Label-Free Characterization of Single Proteins Using Synthetic Nanopores

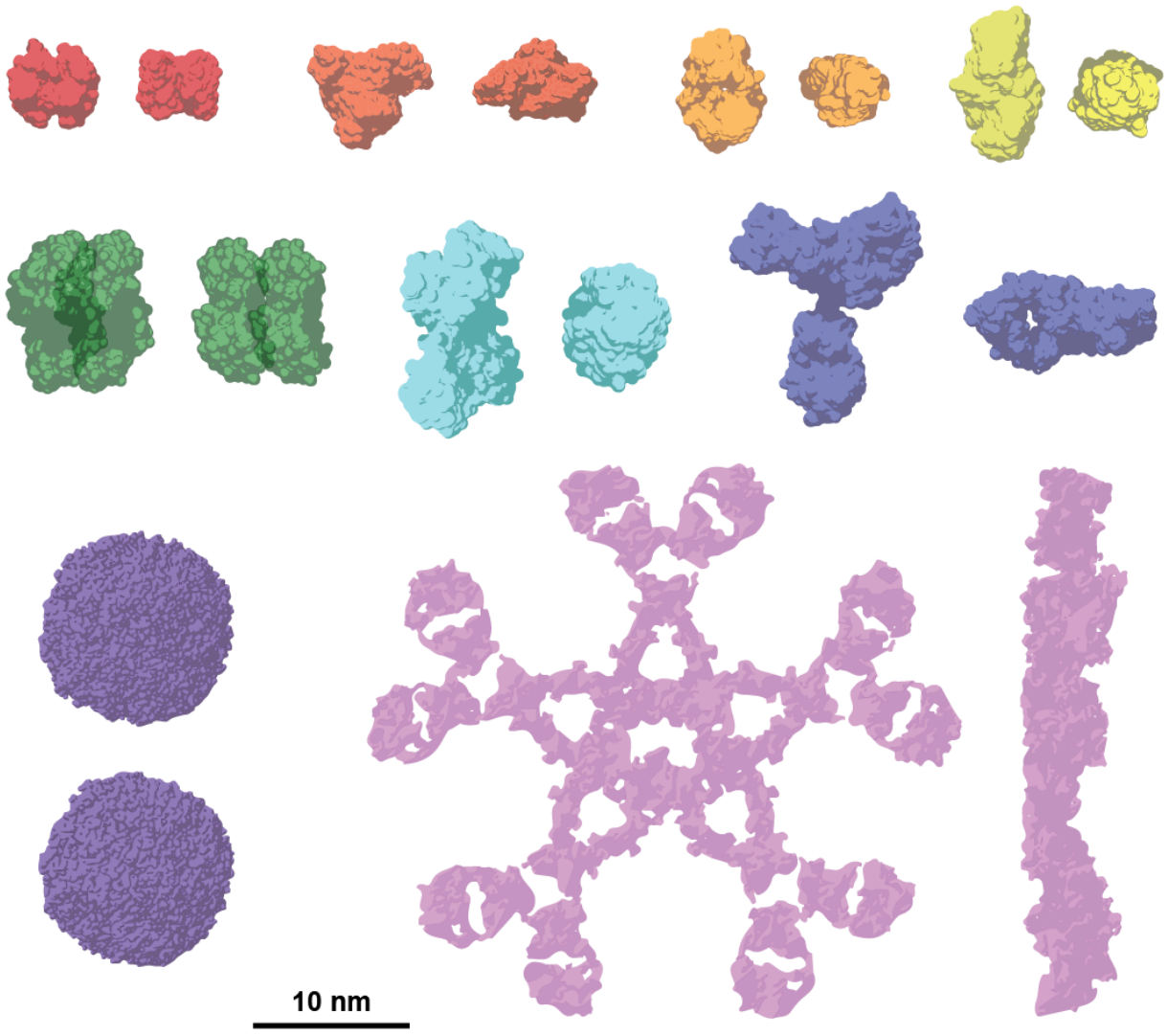
by

Jared Houghtaling

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biomedical Engineering)
in the University of Michigan
2019

Doctoral Committee:

Professor Michael Mayer, Co-Chair
Professor David Sept, Co-Chair
Professor Tim Bruns
Professor Ari Gafni


10 nm

Jared Houghtaling
jaredsh@umich.edu
ORCID iD: 0000-0003-1557-9716

Dedication

To my parents, Gretchen and Scott Houghtaling, for their unwavering support and encouragement, and for instilling in me the confidence to take on challenges – both in research and in life – no matter the size.

To my brother, Zach Houghtaling, for years of lively discussions that helped foster a strong curiosity for all things unknown.

and

To my fiancée, Marjon Tytgat, for her patience, her contagious excitement, and her wondrous ability to shine through the darkest of times.

Acknowledgments

The research described in this thesis was funded by a Graduate Research Fellowship from the National Science Foundation of the United States of America, and by a Rackham Merit Fellowship from the University of Michigan.

Special thanks are extended to the following people, without whom this work would not have been possible:

My research advisor, Dr. Michael Mayer, for his guidance, tremendous enthusiasm for science, and meticulous attention to detail. He has shown me the importance of a rigorous scientific approach, and perhaps even more so, the importance of presenting my research in a way that balances brevity with depth and clarity with style. His decision to move to Switzerland, and my subsequent three and a half years in Fribourg, have been nothing short of life changing and I am profoundly grateful for the experience.

All members of the Mayer research group, both past and present, who have contributed to this work, and who have contributed to a supportive and friendly work environment. In particular:

Cuifeng Ying, Ph.D, for her contributions to Chapters 2, 4, and 5, for her infectious research tenacity, for her help on all things related to physics and optics, and for being a positive and encouraging presence in the lab all hours of the day.

Thomas Schroeder, Ph.D, for being an excellent role model as a Ph.D student, writer, and person. It was an absolute pleasure compiling too many weird facts about insects together.

Simon Finn Mayer for countless adventures, meals, and laughs, but above all, for being a phenomenal friend.

Lennart De Vreede, Ph.D, for his contributions to the noise and fabrication work presented in Chapter 2, and for his comedic relief. There was no friction in our fit.

Aziz Fennouri, Ph.D, for his contributions to Chapter 4, for multiple rides to the hospital, for answering my stream of questions about French, and for his endless invites to anything athletic.

Gogol Guha for putting up with my unorthodox travel itineraries and for the occasional intense tennis match.

Olivia Eggenberger for her contributions to Chapters 2, 3, and 4, and for her support through trying times in Ann Arbor.

Saurabh Awasthi, Ph.D, for his contributions to Chapter 6, and for always having a smile on his face.

Jonathan List, Ph.D, for his contributions to Chapter 1, for being an excellent technical resource for a huge array of scientific questions, and for teaching me a bit of alpine safety.

Peter Nirmalraj, Ph.D, Trevor Kalkus, Louise Bryan, Ph.D, Melissa McGuire, and Jessica Dupasquier for their friendship and insightful scientific discussions.

Stéphane Hess, Pongsatorn Sriboonpeng, and Shuran Xu for their contributions to Chapter 6, their patience with me as a mentor, and their willingness to always ask questions.

Brandon Bruhn, Ph.D, for his mentorship and pioneering research that is featured mainly in Chapter 3 but resonates throughout this thesis. Many components of this work would not have been possible were it not for his contributions and training.

Erik Yusko, Ph.D, for his vast and fundamental contributions to the work presented here, and for the fantastic conversation at Schultzy's on a bleak Seattle evening in 2014 that was instrumental in me joining the group.

Research collaborators and mentors outside of the research group, including:

Prof. Jiali Li, Santoshi Nandivada, Ph.D, and Mitu Achargee for their willingness to fabricate a supply of nanopores that were used in a large number of experiments presented here, and for investing their time to train me on the ion beam sculpting apparatus in Fayetteville, AR.

Prof. Adam Hall, and Osama Zahid, Ph.D, for fabricating nanopores with their helium ion microscope that were used in many experiments throughout this thesis.

Prof. Dave Sept, Prof. Tim Bruns, and Prof. Ari Gafni, who along with Prof. Mayer make up my thesis committee. I thank all committee members for their continued support

and helpful suggestions.

Aghapi Mordovanakis, Ph.D, for his instruction and training on the femtosecond laser setup featured in Chapter 2.

John Mansfield, Ph.D, and the various curators of the Electron Microbeam Analysis Lab (EMAL) at the University of Michigan for their help with the TEM imaging featured in Chapter 2.

Prof. Aleksandra Radenovic, Michael Graf, Ph.D, Sebastian Davis, and other members of the Laboratory of Nanoscale Biology (LBEN) at EPFL for their advice related to using glass capillaries for resistive pulse analyses, and for their willingness to provide information and materials related to resistive pulse sensing.

Those mentors and colleagues who helped shape my earliest years as a researcher and ultimately encouraged me to pursue a Ph.D – namely Prof. Elain Fu, Tinny Liang, and all members of the Yager, Fu, and Lutz lab in the Department of Bioengineering at the University of Washington, as well as Prof. Michael Forrest, who is an extraordinarily gifted instructor and whose lectures on organic chemistry inspired me to follow a path toward a career in academia.

Friends and family outside of my direct academic sphere, including:

The many friends who made Ann Arbor a great place to live and who continue to make it an excellent place to return, especially Daniel Clough, Brian Linhares, and James Day.

Likewise, the many friends in Fribourg and at the Adolphe Merkle Institute who were always up for hiking, skiing, sporting, or apéro-ing. In particular, Jose Vitor who continues to be a great role model and friend, my former Bångstrum bandmates Thomas Moore, Preston Sutton, and Michi Fischer, as well as Pauline Starrenberger and Cloé Brachotte who are largely responsible for my level of French language.

My extended family – grandparents, aunts, uncles, cousins – for their love, support, and advice along the way.

The entire Tytgat family for welcoming me as one of their own, and for always making me feel like I was close to home.

Preface

"A native protein molecule with specific properties must possess a definite configuration, involving the coiling of the polypeptide chain or chains in a rather well-defined way. The forces holding the molecule in this configuration may arise in part from peptide bonds between sidechain amino and carboxyl groups or from sidechain ester bonds or S-S bonds; in the main, however, they are probably due to hydrogen bonds and similar interatomic interactions. Interactions of this type, while individually weak, can by combining their forces stabilize a particular structure for a molecule as large as that of a protein."

– Pauling, L., and Niemann, C. The Structure of Proteins *J. Am. Chem. Soc.* (**1939**): 1860-1867.

Table of Contents

List of Figures

List of Tables

List of Symbols

Symbol	Units	Description
ΔI	A	Non-normalized magnitude of a resistive pulse
$\Delta I/I_0$	unitless	Magnitude of a resistive pulse normalized to baseline current
ρ	Ω m	Resistivity of the electrolyte
l_p	m	Length of the nanopore
r_p	m	Radius of the nanopore
t_b	m	Thickness of the lipid bilayer
l_c	m	Length of the channel adjacent to the nanopore
r_c	m	Radius of the channel adjacent to the nanopore
κ	Ω^{-1}	Conductance of the nanopore
R	Ω	Resistance
R_p	Ω	Resistance of the nanopore
R_c	Ω	Resistance of the channel adjacent to the nanopore
R_a	Ω	Access resistance of the nanopore
t_d	s	Dwell time of the protein in the nanopore during translocation
θ	rad	Angle between an ellipsoid's principle axis and the electric field
γ	unitless	Electrical shape factor
$\gamma_\parallel$	unitless	Electrical shape factor when $\theta = 0$
$\gamma_\perp$	unitless	Electrical shape factor when $\theta = \frac{\pi}{2}$
m	unitless	Length-to-diameter ratio of an ellipsoid of rotation (A/B)
Λ	nm^3	Volume of electrolyte excluded by the protein
μ	D	Dipole moment of the protein

(*continued on next page...*)

(*... continued from previous page*)

D	3.336×10^{-30}Cm	Debyes – unit of measure for the dipole moment
D_R	rad^2 sec^{-1}	Rotational diffusion coefficient of the protein
z	unitless	Net charge of the protein molecule
e	1.60×10^{-19}C	Elementary charge of an electron
η	Pa s	Viscosity of the solution
k_B	1.38×10^{-23}JK^{-1}	Boltzmann constant
T	K	Temperature (295 was used for room temperature)
A_V	6.022×10^{-23}mol^{-1}	Avagadro's number
C	M or mol m^{-3}	Concentration of analyte
$\vec{E}$	V m^{-1}	Electric field
f_c	Hz	Cutoff frequency of an electronic filter
f_{event}	Hz	Frequency of translocation events in a nanopore experiment

Abstract

Molecular diagnosis has proven to be a powerful tool for early detection of neurodegenerative disease, but research in this field is still relatively nascent. In Alzheimer's Disease specifically, levels of microtubule associated protein tau and amyloid-$\beta_{1\text{-}42}$ in cerebrospinal fluid are becoming reliable pathological indicators. The current gold standard for detecting these biomarkers is an enzyme-linked immunosorbent assay, and while this method has a limit of detection on the order of pg mL^{-1}, it lacks the ability to provide information about aggregation extent and structure on a per-protein basis. From a disease standpoint, neurological pathologies are often extremely complex in their biological manifestation, and precise mechanisms for many of these diseases are still being discovered and revised. A thorough understanding of *in situ* structure and properties of neurological disease-related proteins would likely help clarify some of these complicated mechanisms. Resistive-pulse methods may be useful in this effort, as they can determine specific biomarker concentrations and can also unveil multiple physical qualities of single proteins or protein aggregates in an aqueous sample. The latter capability is critical, and could allow for both earlier diagnoses and a stronger mechanistic understanding of neurological disease progression.

The work presented in this dissertation, therefore, represents broad efforts toward developing a nanopore-based system able to characterize amyloids and protein complexes related to neurodegenerative disease. These efforts range from upstream fabrication and characterization of nanopores in synthetic substrates to downstream techniques for optimizing the accuracy and efficiency of analyses on resistive pulses. Single proteins rotating and translating while tethered to the surface of a nanopore provide rich information during transit

through the pore that makes it possible to determine their ellipsoidal shape, volume, dipole moment, charge, and rotational diffusion coefficient in a time frame of just a few hundred microseconds. This five-dimensional protein fingerprint, however, requires chemical modification of each protein and is thus not ideal for studying protein dynamics or transient protein complexes, both of which are relevant when characterizing amyloids. Transitioning to low-noise nanopore substrates and high-bandwidth recordings enables label-free identification and quantification of unperturbed, natively-folded proteins and protein complexes in solution – no chemical tags, tethers, or fluorescent labels are needed. Such a transition is nontrivial; proteins passing uninhibited through the strong electric field inside of a nanopore rotate and translocate rapidly, posing a challenge to time-resolve their various orientations adequately while circumventing adhesion to nanopore walls. Furthermore, during their translocation through the nanopore, untethered, native proteins diffuse laterally, generating asymmetric disturbances of the electric field and larger-than-expected resistive pulse magnitudes. Known as off-axis effects, these latter phenomena add a noise-like element to the electrical recordings. We evaluate, both computationally and experimentally, the influence of such label-free complications on resulting parameter estimates, and place these results in the context of developing future iterations of nanopore-based protein sensors.

In light of the spectacular recent success of nanopore-based nucleic acid sequencing, it is likely that the next frontier for nanopore-based analysis is the characterization of single proteins and, in particular, the characterization of protein aggregates such as amyloids. The experiments and results presented here enable future particle-by-particle analysis of amyloids with nanopores to rapidly reconstruct their heterogeneity in size and shape, both of which are correlated with the neurotoxicity of amyloid samples and are being investigated as biomarkers for neurodegenerative diseases.

Chapter 1: Overview and Concepts

Aggregates of misfolded proteins are associated with several devastating neurodegenerative diseases. These so-called amyloids are therefore explored as biomarkers for diagnosis of dementia and other disorders, as well as for monitoring of disease progression, and assessment of the efficacy of therapeutic intervention. Quantification and characterization of amyloids as biomarkers is particularly demanding because the same amyloid-forming protein can exist in different states of assembly ranging from nanometer-sized monomers to micrometer-long fibrils that interchange dynamically both *in vivo* and in samples from body fluids *ex vivo*. Soluble oligomeric amyloid aggregates, in particular, are associated with neurotoxic effects, and their molecular organization, size, and shape appears to determine their toxicity. The emerging field of nanopore-based analytics, which has recently demonstrated single molecule and single aggregate sensitivity, holds the potential to account for the heterogeneity of amyloid samples and to characterize these particles – rapidly, label-free, and in aqueous solution – with regard to their size, shape, and abundance. This introductory chapter describes the concept of nanopore-based resistive pulse sensing, reviews previous work in amyloid analysis, and discusses limitations and challenges that will need to be overcome to realize the full potential of nanopore-based amyloid characterization on a single-particle level.

1.1 Background

Neurodegenerative disorders like Alzheimer's disease and Parkinson's disease, as well as other medical conditions like Type II diabetes, are often associated with the presence and activity of toxic protein aggregates known as amyloids.[1–6] The incidence, or probability of occurrence, of many of these diseases increases with age.[5] Today, more than 30 million people worldwide suffer from dementias linked to amyloids, and the World Health Organization predicts that this figure may exceed 100 million by the year 2050 as life expectancies increase.[7] While these conditions are devastating with regard to patient suffering and impacts on family members and caregivers, neurological amyloidoses are also responsible for more than USD 500 billion in worldwide annual costs.[7] To reduce patient suffering as well as the associated economic burden, research groups across scientific disciplines have investigated strategies to better understand and interfere with the transition from soluble monomeric proteins and peptides into soluble oligomers and, eventually, into insoluble amyloid fibrils and plaques that are hallmarks of neurodegenerative disease.[8, 9]

Amyloid-forming proteins often undergo an aggregation process analogous to crystal formation where the generation of a "seed" is the rate-limiting step for the assembly of a large ordered structure.[10, 11] Treatments that are currently in clinical trial target these seeds and their precursors in order to redirect or disrupt downstream aggregate formation at an early stage before irreversible nerve cell damage occurs.[12] Additionally, there is evidence that the size, shape, and concentration of oligomeric amyloid aggregates determines their toxicity to neurons,[13, 14] and patients with certain amyloidoses have elevated counts of these oligomeric species in their cerebrospinal fluid (CSF).[15, 16] Findings like these suggest that soluble amyloid aggregates may be valuable biomarkers for predicting or monitoring disease progression and may also help to assess the efficacy of therapeutic intervention. There is thus a need for a characterization technique that provides a comprehensive profile of the individual amyloid particles in bodily fluids like CSF, blood, tears, saliva, or urine.[17, 18]

While a range of techniques is currently being used to characterize amyloids and their

Table 1.1. Comparison of different characterization techniques for amyloid samples. In the resources required and expertise required columns, one star corresponds to minimal resources, and three stars correspond to extensive resources. In the continuous measurement column, "Y", or yes, represents continuous measurements, and "N", or no, represents a snapshot measurement. Adapted with permission.[19] Copyright 2018, Wiley-VCH.

Characterization Technique	Type of Technique				Applicable to these Species of Amyloid					Structural Resolution					Resources Required	Expertise Required	Continuous Measurements
	Single Mol.	Correlation	Imaging	Ensemble	Monomer	Small Oligo.	Large Oligo.	Protofibril	Fibril	Atomic	Secondary	Tertiary and Quarternary	Morphology	Assembly Size Distrib.			
Biological Nanopores	+				+	+							+	+	* *	* *	Y
Synthetic Nanopores	+					+	+	+					+	+	* *	* *	Y
Ion mobility spectroscopy-mass spectroscopy (IMS-MS) [1,4,6]	+				+	+	+							+	* *	* *	N
Dynamic Light Scattering [6]		+			+	+	+	+						+	*	* *	Y
Fluorescence correlation spectroscopy (FCS)		+	+		+	+	+	+						+	* *	* * *	Y
Transmission Electron Microscopy [1,2,3,5,6]			+			+	+	+	+			+	+		* * *	* * *	N
Scanning Tunnelling Microscopy [5,6]			+			+	+	+	+			+	+		* * *	* * *	N
Atomic Force Microscopy [1,3,5,6]			+			+	+	+	+			+	+		* *	* * *	Y
Super Resolution Microscopy			+			+	+	+	+			+	+		* *	* *	Y
Birefringence [2,6]			+						+		+				*	*	N
Fluorescent Binders [1,2,3,5,6]			+	+				+	+		+				*	*	Y
Fluorescence Spectroscopy [2,5,6]				+	+	+	+	+	+		+	+			*	*	Y
Fluorescence Microscopy [1,6]			+		+	+	+	+	+		+	+			*	* *	Y
Solution State NMR [1,4,5,6]				+	+	+	+			+					* * *	* * *	Y
Solid State NMR [1,3,5,6]				+				+	+	+					* * *	* * *	N
X-Ray Absorption spectroscopy [6]				+	+	+	+	+	+	+					* * *	* * *	N
Circular Dichroism Spectroscopy [5,6]				+	+	+	+	+			+				*	* *	Y
Fourier Transform Infrared Spectroscopy [2,3,5,6]				+	+	+	+	+	+		+				* *	* *	Y
Neutron Scattering [5, 6]				+				+	+		+				* * *	* * *	N
X-Ray Fiber Diffraction [1,2,3,5,6]				+				+	+		+				* * *	* * *	N
Electron Spin Resonance [6]				+	+	+	+	+	+			+			* *	* *	N
Hydrogen Deuterium Exchange [6]				+	+	+	+	+	+			+			* *	* *	N
Limited Proteolysis [4,6]				+	+	+	+	+	+			+			* *	* *	N
Gel Electrophoresis [1,4,6]				+	+	+	+							+	*	*	N
PICUP [4,6]				+	+	+	+							+	*	*	N
Size Exclusion Chromatography [1,4,5,6]				+	+	+	+	+						+	*	*	Y
Analytical Ultracentrifugation [6]				+	+	+	+	+						+	* *	* *	N

aggregation processes (Table 1.1), amyloid samples remain extremely demanding to analyze and none of the established techniques meets all demands of an ideal analysis method.[20–22] For instance, amyloid aggregates are challenging analytes to characterize because they are heterogeneous in size and structure, they rearrange, interchange dynamically and grow over time, and they adhere to various surfaces including tubing and microvials.[23, 24] Methods attempting to characterize ensembles of these heterogeneous aggregates obscure potentially important physical differences between individual macromolecular assemblies.[25] Conversely, techniques that label or chemically modify individual amyloids vary in their sensitivity and specificity to different morphologies or chemical structures and the modification itself may alter the sample.[26] Long and extensive sample preparation processes can bias amyloid populations toward stable species and may destroy potentially important transient complexes.[27] Nonetheless, current approaches to characterize amyloids provide a range of relevant parameters of amyloid aggregates. For instance, methods like transmission electron microscopy (TEM), atomic force microscopy (AFM), and mass spectrometry (MS) create "macromolecular snapshots", providing structural information about proteins and aggregates at a single time point.[22] Other techniques using fluorescent dyes, tracer molecules such as 18F-florbetapir, or fluorescently labeled antibodies track aggregation processes *in vivo* or *in vitro* to probe the dynamics and formation rates of different complexes.[28] Some methods focus on defining and detecting amyloid biomarkers that indicate disease predisposition or progression.[15] It is common to perform various analysis approaches in parallel in order to determine structural and chemical characteristics of an amyloid sample, though such combinations often require significant financial and technical resources. An ideal technique for amyloid characterization would combine these functionalities to provide rich, rapid, and robust information about single amyloid particles in high throughput and in a clinical setting without the need for expensive equipment or technical expertise.

One emerging technique that may meet several of these demands is resistive pulse-based nanopore sensing, as it is capable of characterizing individual unlabeled particles in aqueous

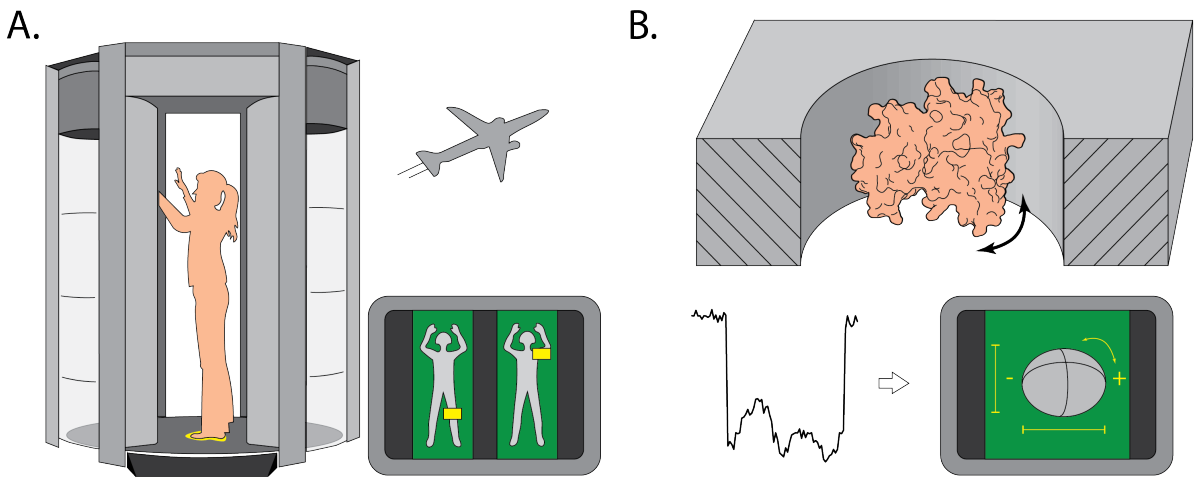


Figure 1.1. Cartoon introduction to nanopore sensing. A) The main principle of single particle resistive pulse analysis is similar to security scanners at airports: the interrogation volume accommodates only one particle (here, one person) at a time. B) Scanning of particle properties occurs during its residence time in this volume, determined properties can be related back to each particular particle because only one is present.

solution. The method was originally developed in the late 1940s for applications on the microscale such as counting and characterizing biological cells.[29–31] Resistive pulse experiments have now made it possible to probe nanoscale analytes including small molecules,[32, 33] metal ions,[34] polynucleotides,[35] nanoparticles,[36, 37] proteins,[38–41] and amyloids.[40, 42–51] The application of resistive pulse sensing to protein characterization emerged less than 15 years ago and is not as developed as the established methods listed in Table 1.1, but it combines attractive capabilities that make it a potentially powerful tool for studying amyloids. Some of these benefits, such as the characterization of shapes, volumes, diffusion coefficients, and electrical and mechanical properties of individual proteins and protein complexes, are appealing for fundamental biophysical studies. Other advantages, including the ability to extract resistive pulses from single unlabeled molecules and to perform analyses of those resistive pulses in real time, may be clinically useful (shown visually later in Figure 6.1).[52] This chapter focuses specifically on the application of resistive pulse sensing to amyloid-related protein analytes and discusses the concepts and challenges of this application.

1.2 Principles of Resistive Pulse Sensing

The concept of resistive pulse sensing traces back to the invention of the Coulter counter for blood cells.[53] Briefly, if two electrolyte-filled reservoirs are connected by a small channel, a difference in electrical potential between the two reservoirs generates a current through the channel. This ionic current is constant at a constant potential difference, but when an insulating particle passes from one reservoir to the other through the channel (*i.e.* through the sensing volume), it transiently displaces conducting electrolyte and reduces the current to produce a resistive pulse.[36, 54, 55] With regard to the other macromolecules in solution, the method can also be thought of as a transient purification, as it interrogates one particle at a time from bulk solution; this concept is illustrated in Figure 1.1. Due to the strong electric field and concomitant fast electrophoretic motion of particles in the pore, the probability of finding two macromolecules in the small sensing volume at the same time is very low, especially when the average duration between particle capture is at least 100-fold longer than the average duration of the resistive pulses (see Section 4.4.5 for calculations). In scenarios where particle concentrations are relatively high ($\sim$1 mM or greater) and each particle dwells within the pore for a relatively long duration ($\sim$1 ms or greater), the probability of multiple-occupancy events increases and may lead to rare events that must be excluded from analysis.[56] For protein analysis by resistive pulse sensing, the protein concentrations are, however, typically in the micromolar range or below and the dwell times are typically shorter than 1 ms. "Continuous" resistive-pulses, like those generated when long strands of nucleic acids pass end-to-end through a sensing volume, have a particular set of intricacies that are reviewed elsewhere.[57–61] Here, we highlight the resistive pulses produced by discrete particles, which typically have lengths shorter than the length of the nanopore, as these are relevant for the sensing and characterization of amyloid oligomers and short protofibrils.

For resistive pulse sensing at any scale, sensing volumes must be appropriately sized to their target analyte.[62] A particle will not produce a detectable signal if its volume is more than 1000-fold smaller than the sensing volume, and it will not translocate through a

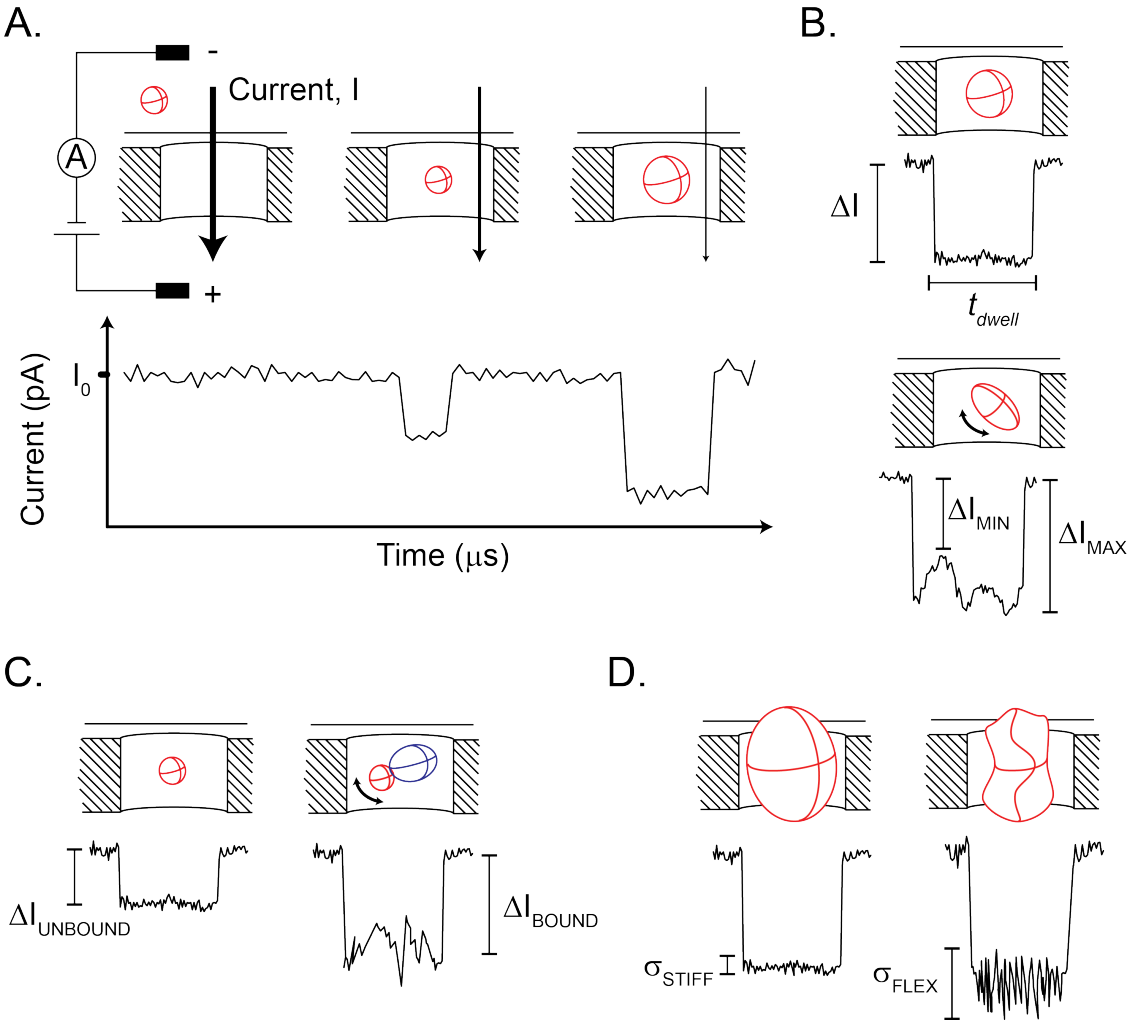


Figure 1.2. Fundamental concepts of resistive pulse sensing. A) An electrical potential applied across the pore creates a constant ionic current, and particles passing through the sensing volume produce resistive pulses proportional to their volume. B) Translocation of spherical particles through cylindrical nanopores produces square-shaped resistive pulses, and the duration of translocation events is proportional to their electrophoretic mobility. Brownian rotation of non-spherical particles modulates the ionic current through the sensing volume depending on their orientation within the pore. The minimum (ΔI_{min}) and maximum (ΔI_{max}) blockade values can be used to estimate shape, dipole moment, and rotational diffusion coefficient of the particle. C) Two or more particles bound together produce a larger resistive pulse (ΔI_{bound}) than sequential translocations of individual particles ($\Delta I_{unbound}$). The fraction of resistive pulses from bound to unbound proteins is related to binding affinities. D) Particle structures can fluctuate when confined or compressed within a sensing volume depending on mechanical properties of the molecules or molecular complexes. Less-flexible particles (σ_{stiff}) produce resistive pulses with smaller amplitude fluctuations than flexible particles (σ_{flex}). Adapted with permission.[19] Copyright 2018, Wiley-VCH.

sensing volume that it cannot enter due to steric constraints. Particles that do not conform to these size restrictions can still provide indirect information about the system,[63] but they cannot be analyzed directly within the sensing volume. If the sensing volume is sufficiently large to allow a non-spherical particle to rotate relatively freely, that particle will produce a unique resistive pulse signature that depends not only on its volume, but also on its shape and relative orientation to the electric field (Figure 1.2).[39, 62] Rotations during transit through the pore cause fluctuations within the resistive pulse that originate from different orientations of a three-dimensional shape in the electric field. The physical basis for this orientation dependence was first explored by Golibersuch[62] and others,[29, 64] who found that rotations of disk-shaped red blood cells passing through a cylindrical channel generated characteristic resistive pulses with distinct minima and maxima. The minimum values (ΔI_{min}) corresponded to the cell in its edgewise orientation relative to the channel axis, and the maximum values (ΔI_{max}), which were about 2-fold greater in amplitude than the minimum, corresponded to the cell in its crosswise orientation relative to the channel axis (Figure 1.2-B). Fricke [65] and others [66] quantified these effects with a physical descriptor, the electrical shape factor γ, which depends on the particle's ellipsoidal shape and orientation within a cylindrical sensing volume, and is directly proportional to resistive pulse amplitude. The electrical shape factor is also valid on the nanoscale, though nanometer-sized proteins and particles rotate at a rate that is several orders of magnitude faster than cells and thus require high-bandwidth recording electronics and strategies like surface anchoring [38] to slow their rotation in order to fully resolve their characteristic fluctuations. We recently took advantage of these fluctuations in order to approximate the shape of an individual particle translocating through a nanopore, and present those results in Chapters 3 and 4.[39, 67] Furthermore, we used the particle's bias toward certain orientations during its transit through the strong electric field inside of the nanopore (several MV m^{-1}) to estimate the net dipole moment of the particle. The most probable speed at which the particle transitions between these orientations is proportional to its bulk rotational diffusion coefficient, and the

amount of time that a particle occupies the sensing volume, also referred to as its dwell or residence time, is a function of the particle's lateral diffusion coefficient and is inversely proportional to its net charge. Measured simultaneously, these five parameters – ellipsoidal shape, volume, charge, rotational diffusion coefficient, and dipole moment – define a high-content multidimensional "fingerprint" of a single particle that helps to characterize and discriminate between particles in a mixture.[39]

Resistive pulse sensing can also probe the conformational variability of particles.[41, 68, 69] When a particle is sterically constrained within a small sensing volume, the resistive pulse associated with that particle's translocation provides information about the particle's conformational variability.[41, 68, 69] In the case of proteins, this information appears as fluctuations in amplitudes between multiple resistive pulses (Figure 1.2-D) and has been attributed to differences in secondary structural composition (*e.g.* β-sheet to α-helix ratio).[41, 68, 69] Resistive pulse experiments, in principle, may also make it possible to monitor the interactions of particles or amyloid-forming molecules with soluble species. For instance, our group used a nanopore to monitor immunoprecipitation and to determine the binding affinity of an antibody to the surface of a virus particle by relating resistive pulse amplitudes over time to the number of binding sites occupied on the particle. In these approaches, each binding event increased the aggregate volume, and thus the magnitude of the resistive pulse, by a constant increment.[70–72] Likewise, Si *et al.* showed that denatured proteins produce different resistive pulses than their native protein counterparts, and that the differences correspond to unfolding processes.[39, 73] Each of these applications of resistive pulse sensing provides different information about individual particles in aqueous solution and about their time-dependent changes in response to various stimuli.

1.3 Detection and Characterization of Amyloids with Biological Nanopores

Biological nanopores range from ion channels to assemblies of pore-forming toxins[74, 75] and their biological function is to facilitate or regulate the passage of polar molecules, ions, or peptides across cell membranes.[76] With regard to resistive pulse sensing, biological nanopores are appealing as their sensing volumes are often defined with atomic precision and they can be produced in large quantities through bacteria fermentation or cell culture followed by purification.[77] Due to their extremely small sensing volumes in pore lumens with diameters smaller than 4 nm (Figure 1.3), biological nanopores are particularly well-suited to sensing small analytes, including proteins smaller than 30 kDa as well as single-stranded and double-stranded DNA or RNA.[78–80] Furthermore, several of these protein-based nanopores can be engineered by site-directed mutagenesis to integrate desired properties such as specific binding sites or residues that modify the electrostatic landscape in the pore lumen.[32, 81–84]

1.3.1 Resistive Pulses Provide Information about the Kinetics of Enzyme Cleavage of Amyloid-Forming Peptides

Two key research developments facilitated the nanoscale application of resistive pulse counting with biological nanopores: inhibiting the voltage-dependent nature of ion channel proteins,[33, 84, 85] and limiting the rate at which biomolecules of interest transit the lumen of the pore.[82, 86] These developments enabled seminal work by Berzrukov and Kasionowicz, who used the pore from α-toxin (*Staphylococcus aureus*) to measure protonation rates and to discriminate between protons and deuterons.[33, 84, 85] Kasianowicz *et al.* were the first to demonstrate the detection of individual polynucleotide molecules using the pore α-hemolysin,[87] and Berzrukov *et al.* showed that dwell times of certain polyethylene glycol (PEG) molecules were more than 1000-fold longer than values predicted by 1D diffusion be-

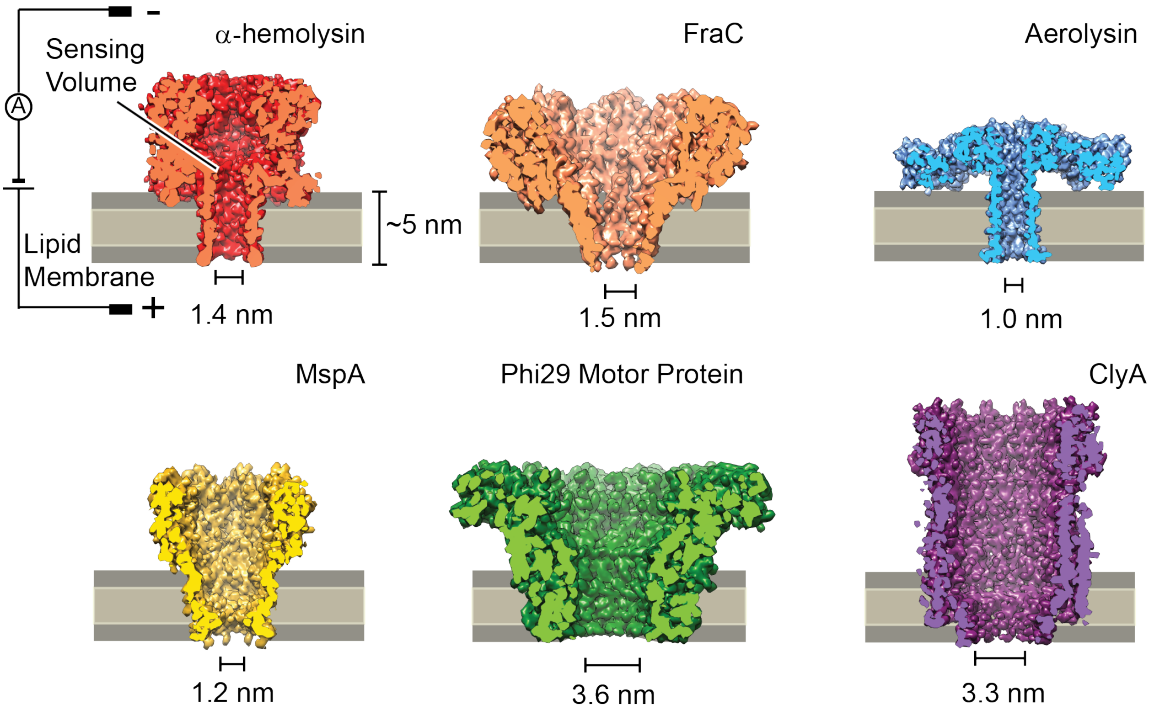


Figure 1.3. Examples of biological nanopores used for sensing biomolecules. Cross-sections show their unique high-resolution structures and sizes that connect two electrolyte-filled chambers across a lipid membrane. As shown for α-hemolysin, protein pores insert into lipid bilayer membranes (grey) and conduct a constant ionic current if a voltage is applied across the membrane. Scale bars show the narrowest constriction of each sensing volume and were measured from high-resolution structures using Chimera software. Protein cartoons were generated using Chimera software with the following codes from the protein data bank: α-hemolysin (7AHL), MspA (1UUN), FraC (4TSY), Phi29 Motor Protein (1JNB), Aerolysin (5JZT) , and ClyA (2WCD). Adapted with permission.[19] Copyright 2018, Wiley-VCH.

cause the molecules reversibly bind within the nanopore;[82] a formal theory later proposed by Lubensky and Nelson helped to explain these polymer-pore interactions.[88] This body of research established biological nanopores as resistive pulse sensors, and biological nanopores have since made it possible to detect not only polymers, but also small molecules,[56] and amyloid-forming peptides.

The first report of characterizing an amyloid-related peptide with biological nanopores focused on the amyloid-β (Aβ) peptide segment A$\beta_{10\text{-}20}$.[89] (Note here that Aβ refers to a peptide involved in Alzheimer's disease that typically contains between 37 and 43 amino acids, and we denote the particular segment of Aβ using subscripted numbers.) By modifying the interior of an α-hemolysin pore to contain additional aromatic binding regions, Zhao *et al.* prolonged the residence times of A$\beta_{10\text{-}20}$ and other peptides rich in aromatic residues. They related residence times and resistive pulse amplitudes of the peptides to binding affinities within the nanopore, and demonstrated that nanopores could be used to determine the presence or absence of various small peptides in a mixture.[43] This setup also allowed Zhao and colleagues to monitor trypsin-catalyzed cleavage of A$\beta_{10\text{-}20}$. Cleavage of proteins like amyloid precursor protein (APP) by β- and γ-secretases plays a role in downstream amyloid formation and thereby possibly influences diseases like Alzheimer's and Parkinson's.[90] These results demonstrated that nanopores can provide information about the kinetics of enzymatic cleavage of amyloids without the need to label the amyloids.[43]

1.3.2 Interactions between Amyloid-Forming Peptides and Molecules that Modulate Aggregation

Experiments with biological nanopores can investigate interactions between small molecules and amyloid-forming peptides. Hai-Yan Wang *et al.* studied the influence of aggregation promotors like β-cyclodextrin and aggregation inhibitors like Congo red on the aggregation kinetics of Aβ. To this end, the authors followed aggregation over time by monitoring the frequencies of resistive pulses, which are related to particle concentration, and the amplitudes

of resistive pulses, which are related to particle size.[44] A similar assay revealed the influence of copper (Cu^{2+}) ions on the conformation of A$\beta_{1\text{-}16}$ peptide and the rate at which it formed oligomers. Human A$\beta_{1\text{-}16}$ had stronger and longer-lasting interactions with Cu^{2+} compared to the rat variant of A$\beta_{1\text{-}16}$, likely because of a single amino acid difference (human HIS-13 versus rat ARG-13).[91] Wang *et al.* proposed that HIS-13 plays a role in metal-induced aggregation and could therefore be a potential therapeutic target.[44, 92] Information about aggregation behavior of specific peptide sub-sequences may be useful in the context of therapeutic approaches to inhibit aggregation, and it can reveal structure-function relationships for the full-sequence peptide. For instance, the A$\beta_{25\text{-}35}$ peptide has a β-sheet structure and forms aggregates, but its inverted sequence A$\beta_{35\text{-}25}$ takes on a random coil structure and does not exhibit neurotoxicity.[93] Hu *et al.* found that A$\beta_{25\text{-}35}$ produced large current blockages due to its extended β-sheets and its translocation events became less frequent with time, suggesting that it was aggregating into complexes too large to enter the sensing volume of the α-hemolysin nanopore.[47] Solutions of A$\beta_{35\text{-}25}$, on the other hand, generated smaller and shorter blockage events than those of A$\beta_{25\text{-}35}$, and frequencies of A$\beta_{35\text{-}25}$ translocation events were consistent over time, suggesting that the β-sheet motif could be detected using a nanopore and likely plays a role in aggregation.[47]

1.3.3 Investigations of Amyloid-Forming Peptides other than Aβ

Most detection and characterization of amyloids using biological nanopores has focused on the Aβ peptide. There are more than 30 amyloid-forming peptides or proteins that are associated with human disorders, including prion protein (PrP) in the spongiform encephalopathies, huntingtin protein in Huntington's disease, and α-synuclein (α-Syn) in Parkinson's disease.[5] One reason that Aβ has been popular in research with biological nanopores thus far is its small (4.5 kDa) molecular weight; many other amyloid-forming proteins are too large in their natively folded conformation to fit through the small confines of a biological nanopore. Some proteins, like α-Syn, are natively unfolded and can pass through the

nanopore as a single strand,[94] while large globular proteins can be denatured in a solution of 5M Guanidinium HCl and passed through the sensing volume in a mostly unfolded state.[95] Jeremy Lee's research group employed this denaturation strategy and reported distributions of current blockades for different prion proteins, as well as for Aβ and α-Syn.[95–98] Despite these approaches, the large size of some amyloid-forming proteins or their aggregates remains a major challenge to their characterization with biological pores, especially when the native structure of the protein or amyloid particle is of interest. At present, no biological pore can accommodate an intact amyloid oligomer larger than 4 nm. Some research groups have begun to engineer protein-based nanopores with larger sizes than natural pores,[99, 100] while others have chosen to fabricate and use synthetic nanopores in different sizes and materials.

1.4 Detection and Characterization of Amyloids with Synthetic Nanopores

Synthetic nanopores with custom diameters in the range from 1 nm to 100 nm facilitate the detection and analysis of larger biomolecules including natively-folded proteins[101] and double-stranded DNA.[102–104] Typically, these sensing volumes are fabricated by generating holes with nanometer-scale diameters in thin (less than 100 nm) insulating membranes (Figure 1.4). Manufacturing techniques for nanopore formation include dielectric breakdown,[105], TEM drilling,[106] helium ion microscope drilling,[107] capillary shrinking,[108] and gold particle heating,[109] while substrates range from silicon,[110] silica (glass),[111] silicon nitride,[102] MoS_2,[112] HfO_2,[113] and graphene.[114] Synthetic nanopores of all sizes and materials still suffer from two critical drawbacks: proteins tend to adhere to the nanopore substrate and hence clog the pore, or – when proteins do not adhere to the substrate – they transit the nanopore too quickly such that the majority of them cannot be detected by conventional electrical recording equipment.[38, 115]

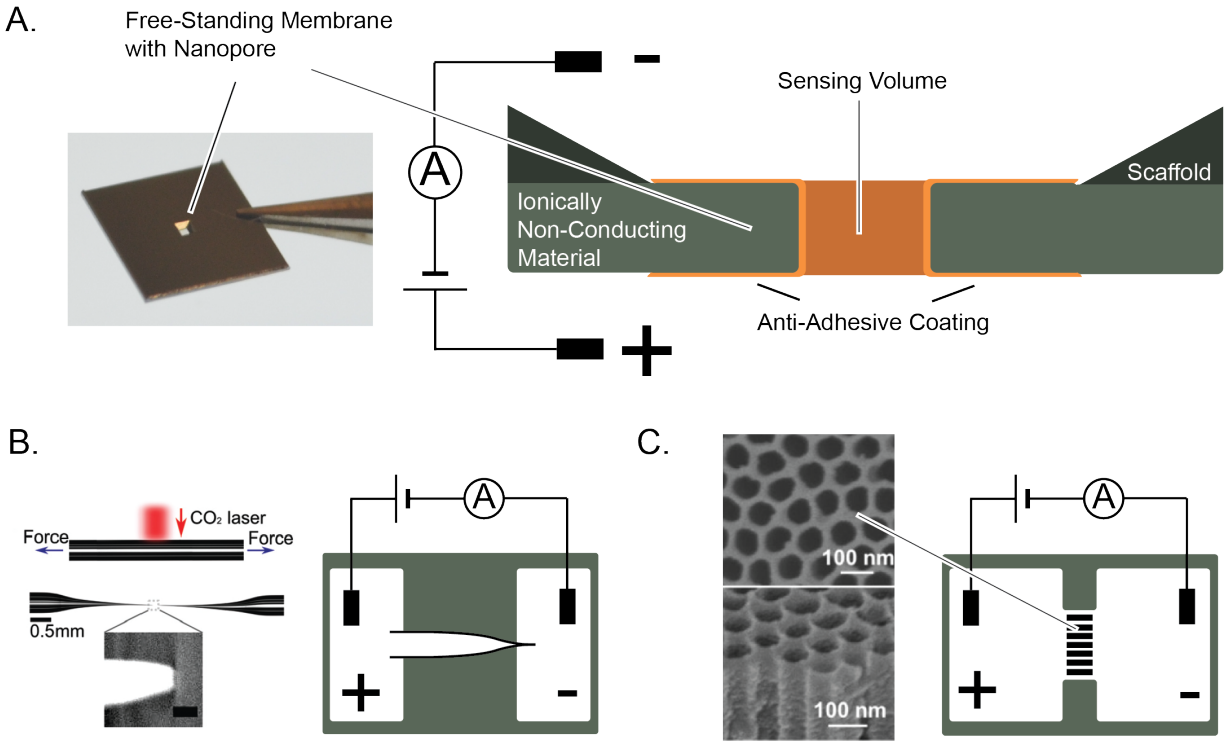


Figure 1.4. Examples of synthetic nanopores for characterizing biomolecules. A) Image of a silicon scaffold supporting a free-standing silicon nitride membrane, and cross-section of a nanopore in a free-standing membrane with an anti-adhesive coating. B) Glass capillary tubes can be locally heated and mechanically stretched to terminate in hollow tips with nanometer diameters. The terminal tips then act as a sensing volume that connects two electrolyte-filled reservoirs. C) Scanning electron microscopy (SEM) images of a membrane containing parallel nanochannels used to connect two reservoirs. The membrane was functionalized with amyloid-forming peptides, and the ionic current through the membrane slowly declined as peptides aggregated and occluded the channels. Adapted with permission.[19] Copyright 2018, Wiley-VCH.

1.4.1 Lipid Bilayer Coatings Allow First Measurements of Amyloids with Synthetic Nanopores

In response to these problems, several research groups in the synthetic nanopore field have focused on creating anti-adhesive coatings.[38, 116] One such coating introduced by our group – fluid lipid bilayers – prevents unwanted adhesion and slows protein transit through the nanopore by anchoring proteins to activated lipids.[117] To this end, we coated a silicon nitride nanopore with a supported lipid bilayer in order to characterize A$\beta_{1\text{-}40}$ peptide in the first application of a synthetic nanopore to an amyloid-related protein.[38, 40] This work quantified the formation of four distinct A$\beta_{1\text{-}40}$ aggregates over the course of a 72-hour aggregation period: spherical oligomers, short protofibrils, long protofibrils, and amyloid fibers. The event frequency of each species, which is a measure of their abundance, reflected the extent of aggregation; spherical oligomers gradually became less frequent as they grew in size while resistive pulses from protofibrils and mature fibers increased in frequency over time.

Coatings derived from materials other than lipid bilayers can also reduce unwanted adhesion. Rui *et al.* investigated α-Syn aggregation using synthetic nanopores coated with polysorbate 20 (Tween-20). This work characterized four different oligomeric species, and investigated the impact of small unilamellar vesicles (SUVs) containing certain lipids on the rates of aggregation of α-Syn.[50] Similarly, Giamblanco *et al.* functionalized nanopores with polyethylene glycol (PEG-5k) chains in order to study the aggregation kinetics and fibril sizes of amyloid particles comprised of lysozyme and other model proteins.[51]

1.4.2 Fibril-Forming Proteins Demonstrate Assembly Processes in Nanopores

Because of difficulties with the preparation or analysis of amyloid samples,[23] several groups have chosen to investigate proteins like lysozyme or bovine serum albumin (BSA) that readily

aggregate into amyloid-like fibrils but are not necessarily pathogenic to humans. In an attempt to develop a system that relates a population of resistive pulse amplitudes and durations to a concentration profile of amyloid (proto)fibrils, Martyushenko *et al.* monitored the aggregation process of lysozyme into long fibrils using glass nanocapillaries.[46] The authors performed the experiments at pH 2.0 in order to prevent adhesion to the glass substrate and determined a distribution of aggregates that they compared with results from simulations.[46] Balme *et al.* expanded upon this research with lysozyme fibrils and focused on the effects of protein adhesion to the surface of a silicon nitride nanopore; these authors mitigated adhesion by treating the nanopore with concentrated sulfuric acid directly before experiments. The authors then extracted distinct populations of lysozyme oligomers from the distribution of resistive pulse amplitudes and correlated those particular populations with individual monomer additions.[48]

1.4.3 Investigations of Prion Protein with Synthetic Nanopores

Li *et al.* took advantage of the large volumes of synthetic nanopores to characterize and compare a range of commonly available proteins like BSA and IgG1 antibody as well as human prion protein (PrP).[45] This report represented the first measurements of native PrP with a synthetic nanopore. The experiments were hampered by transient protein adhesion to the glass substrate, but nonetheless revealed differences in dwell times and pulse amplitudes between standard proteins and PrP; the authors related these differences to protein structure and aggregation processes.[45] In general, synthetic nanopores enable direct monitoring of changes in the populations of oligomeric species and of large natively-folded monomers (Figure 1.5). Recent years indicate a shift toward synthetic nanopores for amyloid sensing, and we expect this trend to continue with the development of low-noise recording setups,[118] high-bandwidth recording equipment,[119] and advanced surface coating technologies[38, 120] that may selectively bind amyloids. Because of their single molecule sensitivity and broad size range, synthetic nanopores may ultimately provide insight into the ways in which

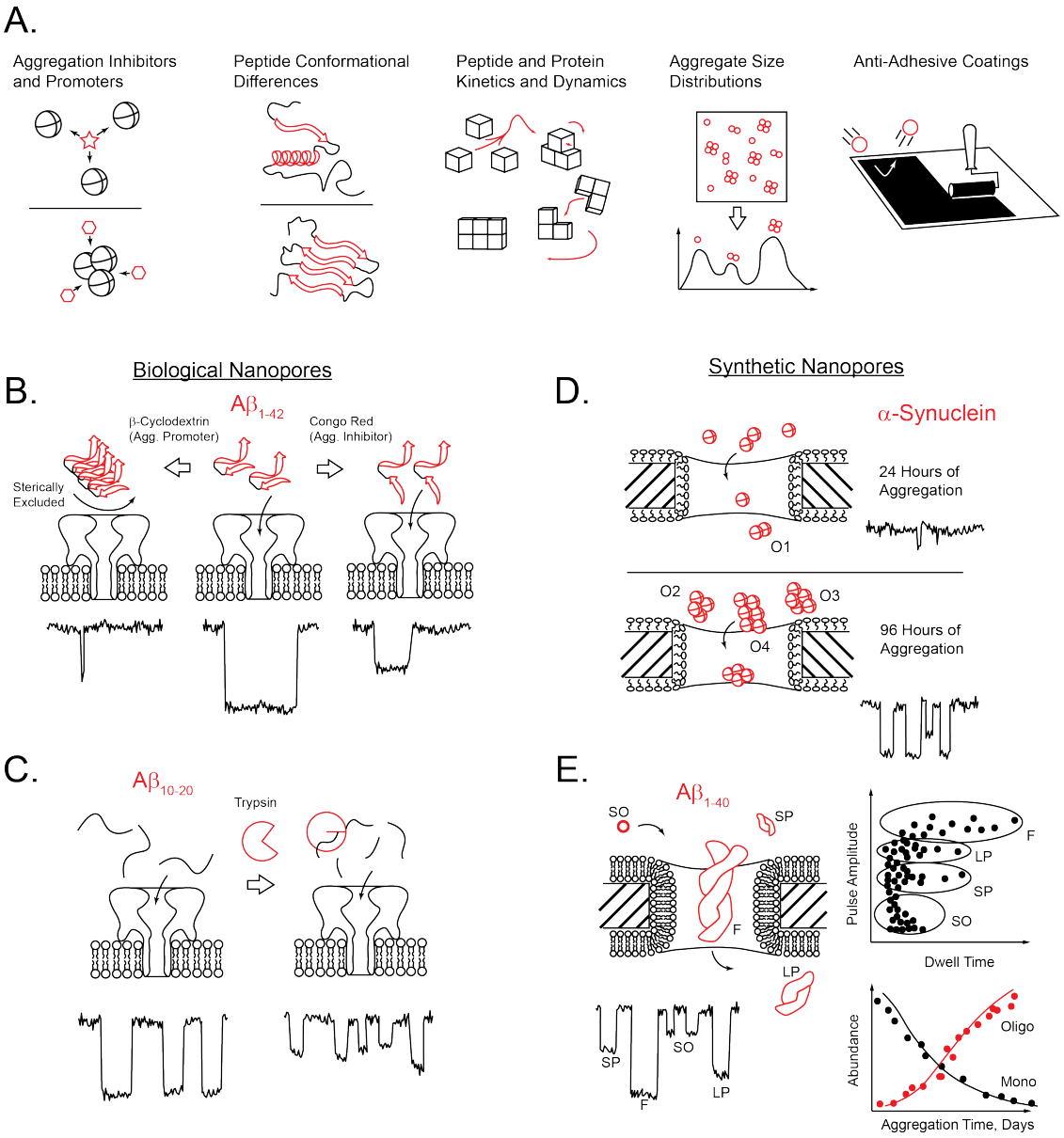


Figure 1.5. Graphical summary of prior nanopore work characterizing amyloid particles. A) Biological nanopores have been used to investigate interactions of amyloid forming peptides with aggregation promotors or inhibitors as well as toevaluate conformational differences between peptides. Synthetic nanopores have determined size distributions of aggregates over time, by employing a variety of anti-adhesive coatings. Both biological and synthetic pores have been used to determine aggregation rates of amyloids. B) The peptide $A\beta_{1\text{-}42}$ produces brief spikes in the presence of an aggregation promoter because the protein is sterically excluded from entering the pore, and generates small, short-lived events in the presence of an aggregation inhibitor.[44] C) As $A\beta_{10\text{-}20}$ is enzymatically cleaved by trypsin, it produces smaller and shorter resistive pulses that correspond to the turnover rate and the length of the Aβ fragments.[43] D) The protein α-Syn aggregates over the course of 96 hours; aggregates can be grouped into four major phenotypes (O1, O2, O3, and O4).[50] E) The peptide $A\beta_{1\text{-}40}$ goes through an aggregation process where it forms small spherical oligomers (SO), short protofibrils (SP), long protofibrils (LP), and finally fibrils (F) over time. Each of these species can be detected and characterized using a lipid bilayer coated synthetic nanopore.[40] Adapted with permission.[19] Copyright 2018, Wiley-VCH.

these amyloid-forming proteins aggregate as well as into their structure.

1.5 Challenges and Motivation

Given the development and commercialization success of nanopore-based DNA and RNA sequencing over the past 20 years,[56, 121–125] it is clear that resistive pulse sensing provides exciting opportunities as a bioanalytical method on the nanoscale. Applications of the technique to protein-based analytes, however, have not yet fully realized this potential. For example, one of the most compelling aspects of resistive pulse sensing – analyzing an individual resistive pulse to determine the physical characteristics of the unique particle that produced it – has yet to be fully exploited on amyloid targets. All of the work in Table 1.2 measured resistive pulses resulting from the translocations of single amyloid particles but performed subsequent analyses on populations of resistive pulses. Furthermore, these studies typically reduced resistive pulses to two quantities, amplitude and dwell time, before clustering those data into groups to generate high-level comparisons about aggregation rates and distributions of aggregate sizes. These analyses produced insights into processes of amyloid aggregation and the size of aggregates, but they overlooked rich information about relevant physical properties of individual amyloid particles such as shape, dipole moment, or conformational variability. In order to take full advantage of this detailed single-amyloid information in a way that may have clinical usefulness, nanopore sensors must first overcome several challenges summarized in Table 1.3.

Diagnostic characterization of a patient's amyloid profile with nanopores requires investigation of complex biological solutions like blood or CSF. A fundamental challenge of applying label-free single molecule techniques to such samples is the ability to discriminate between a few analytes of interest and a large concentration of background molecules. Purification techniques like filtration, size exclusion, or affinity chromatography can remove most of these background molecules, but they prolong analyses and add complications that can limit use-

Table 1.2. List of the studies that characterized amyloid-forming peptides and proteins or amyliod aggregates with nanopores. Adapted with permission.[19] Copyright 2018, Wiley-VCH.

Protein / Peptide	Title	Nanopore Style	Nanopore Details	Unique Aspect(s)	Year	REF
Amyloid-β (10-20)	Real-Time Monitoring of Peptide Cleavage Using a Nanopore Probe	Biological	α-hemolysin	Monitored enyzmatic rate indirectly	2009	Guan *et al.*
Amyloid-β (10-20)	Study of Peptide Transport through Engineered Protein Channels	Biological	α-hemolysin	Modified interior α-hemolysin nanopore to extend dwell times of peptides containing aromatic residues	2009	Guan *et al.*
Amyloid-β (1-40)	Controlling protein translocation through nanopores with bio-inspired fluid walls	Synthetic	Silicon Nitride, Lipid Bilayer Coated	Coated synthetic nanopores with lipid bilayers for the first time, and recorded the translocation of Aβ aggregates	2011	Yusko *et al.*
Amyloid-β (1-42)	Nanopore Analysis of β-Amyloid Peptide Aggregation Transition Induced by Small Molecules	Biological	α-hemolysin	Compared aggregation kinetics of Aβ in the presence of aggregation inhibitors and promotors	2011	Long *et al.*
Amyloid-β (1-40, 1-42), α-Synuclein, Prion Protein (PrP) Human and Bovine	Nanopore analysis: An emerging technique for studying the folding and misfolding of proteins	Biological	α-hemolysin	Investigated a large group of amyloid-forming proteins with a nanopore, denatured larger proteins to fit	2012	Madampage *et al.*
Amyloid-β (1-40)	Single-Particle Characterization of Aβ Oligomers in Solution	Synthetic	Silicon Nitride, Lipid Bilayer Coated	Monitored the aggregation of Aβ and classified four species of aggregates	2012	Yusko *et al.*
Amyloid-β (1-16) , Human and Rat	Investigation of Cu2+ Binding to Human and Rat Amyloid Fragments Aβ (1−16) with a Protein Nanopore	Biological	α-hemolysin	Related interactions between peptides and metal ions to specific amino acid differences between peptide variants	2013	Luchian *et al.*
α-Synuclein	Analysis of a single a-synuclein fibrillation by the interaction with a protein nanopore	Biological	α-hemolysin	Analyzed α-Synuclein with a protein nanopore, quantified aggregate-promoting interactions with lipid bilayers	2013	Wang *et al.*
Prion Protein (PrP)	Single Protein Molecule Detection by Glass Nanopores	Synthetic	Glass Nanocapillary	Compared size distribution and translocation times of PrP with other model proteins	2013	Li *et al.*
Lysozyme	Nanopore analysis of amyloid fibrils formed by lysozyme aggregation	Synthetic	Glass Nanocapillary, Low pH	Correlated the length of amyloid fibrils with residence times within a nanocapillary, verified results with simulations	2015	Martyushenko *et al.*
Amyloid-β (25-35, 35-25)	Single molecule study of initial structural features on the amyloidosis process	Biological	α-hemolysin	Observed differences in conformation and aggregation rate of two identically sized peptides	2016	Long *et al.*
Lysozyme, Avidin and IgG	Influence of Adsorption on Proteins and Amyloid Detection by Silicon Nitride Nanopore	Synthetic	Silicon Nitride, Acid Pre-Cleaning	In-depth analysis of the impact of adhesion on the dwell times of three amyloid-forming proteins. Correlation of blockades with monomer additions	2016	Balme *et al.*
Amyloid-β (1-40)	A novel device of array nanochannels integrated electrochemical detector for detection of amyloid β aggregation and inhibitor screening	Synthetic	Nanochannel Array, IR Measurements	Used a parallel array of nanopores coated with Aβ to monitor the aggregation kinetics of Aβ in solution based on increased resistance	2016	Wang et al.
α-Synuclein	Intrinsic and membrane-facilitated α-synuclein oligomerization revealed by label-free detection through solid-state nanopores	Synthetic	Silicon Nitride, Tween20 Coated	Measured the aggregation kinetics of α-Syn, determined four distinct oligomeric species, monitored aggregation in the presence of different lipid membranes	2016	Hu *et al.*
Lysozyme, β-Lactoglobulin, BSA	Detection of protein aggregate morphology through single antifouling nanopore	Synthetic	Silicon Nitride, PEG-Coated	Used PEG coated pore, and incorporated ellipsoidal shape analysis to estimate fibril lengths	2018	Giambalnco *et al.*

fulness. For instance, the presence of interfaces as well as changes in pH or ionic strength during these procedures may influence the amyloid aggregation state in the sample. Direct analysis of complex samples without purification is possible when employing target-specific detection labels.[126] Amit Meller's group simultaneously monitored optical and electrical signals of a strand of fluorescently labeled DNA transiting a nanopore,[127] and a similar approach may allow for selective resistive pulse analyses in complex protein samples. Optical methods can even replace electrical measurements to monitor ionic current through a nanopore, as has been shown with calcium-flux sensing on nanopore arrays.[68, 128–131] The challenge is, however, to collect a sufficient number of photons during the short-lived dwell times (μs) of proteins through nanopores. Solutions may emerge from sensing volumes themselves, as they can also be engineered to interact specifically with a target analyte. Binding sites designed inside or around biological nanopores enhance detection of target molecules through transient binding,[132] and synthetic nanopores with fluid lipid bilayer coatings can concentrate specific molecules around the surface of the pore by incorporating lipid anchors with binding sites into the coating.[38] But even when applying techniques to improve specificity, resistive pulse sensing with a single nanopore is still inherently a serial process and profiling the individual molecules in a non-purified mixture may require long recording times. This limitation can be addressed through parallelization as shown by Oxford Nanopore Technologies with their recent nanopore-based DNA sequencing devices that record data from hundreds of nanopores independently and simultaneously.[133] Meanwhile, novel integrated CMOS current amplifiers combined with nanopore chips with low electrical capacitance will continue to improve the signal-to-noise ratio of high-bandwidth current recordings and provide more detailed and accurate insights from the translocations of single particles than the electrical setups currently available.[119] Fast and high-fidelity data acquisition requires robust data processing, and improvements in recording equipment have prompted a trend toward applying machine learning algorithms to resistive pulse-based data, including deep learning by neural networks.[134, 135] Amyloid characterization with

Table 1.3. Challenges and their potential solutions for characterizing amyloid particles using nanopores. Adapted with permission.[19] Copyright 2018, Wiley-VCH.

Specific Challenges of Characterizing Amyloids with Nanopores	Potential Solution(s) to Challenges
General to All Nanopore Designs	*General to All Nanopore Designs*
– Distinguish amyloid particles from other macromolecules in solution.	– Amyloid-specific binders, combine with optical techniques.
– A range of nanopores in an array format is necessary to characterize the entire range of amyloid species from monomer to fiber.	– Chambers engineered for parallel electrical recordings.
– Long protofibrils and mature fibrils may not enter the pore or may be sterically prohibited from rotating.	– Increase size distribution of nanopores.
– High salt concentration of recording electrolytes affects aggregation kinetics of amyloid-forming peptides and proteins.	– Perform experiments at physiologic ion concentrations.
– Specialized technique that requires specialized instrumentation and expertise.	– Automate technique in a commercial format.
Specific to Biological Nanopores	*Specific to Biological Nanopores*
– Amyloid-forming proteins and, in particular, amyloid aggregates are too large to translocate through the biological nanopores that are currently available.	– Engineer larger biological nanopores.
– Capacitive current noise from bilayer capacitance.	– Separate chambers with a synthetic polymer.
– Fragility of lipid membranes to biological samples.	– Separate chambers with a synthetic polymer.
– Need to prepare lipid membrane before nanopore insertion.	– Separate chambers with a synthetic polymer.
– Need to reconstitute biological pore into lipid membrane.	– Pre-insert biological nanopore and engineer pore to remain stable for long periods of time.
Specific to Synthetic Nanopores	*Specific to Synthetic Nanopores*
– Strong, non-specific interctions with the nanopore wall can lead to the artifacts in the characterization of amyloids and to clogging of the pore.	– Generate surface coatings with desired properties.
– Poor capacitive properties in silicon substrates.	– Fabricate nanopores in glass (silica) substrates.
– Limited access to high-quality nanopores with desired characteristics like dimensions, geometry, surface chemistry, and amenity to surface coating.	– Improve fabrication and characterization methods for synthetic nanopores.

nanopores will also benefit from further development and optimization of sensing volumes. While biological pores are currently limited to diameters less than 4 nm, engineered protein pores[99,100] as well as DNA origami channels[136] might extend the range of potential analyte sizes. Novel coating strategies taking inspiration from nature [38, 137] can overcome unwanted adhesion issues for synthetic pores, while fabrication techniques like dielectric breakdown [105] or laser-assisted nanopore formation [138] can quickly produce single-use pores without the need for sophisticated equipment.

Few, if any, techniques can quickly identify, quantify, and characterize individual unlabeled proteins or protein aggregates in a complex aqueous sample.[139] Because nanopores can probe multiple physical parameters of individual particles in solution, we suggest that

they are compelling candidates for an analytical platform technology that makes it possible to detect and characterize amyloid aggregates. We hope that the studies summarized here represent the initial steps toward a rapid and robust amyloid characterization platform using nanopores. If solutions to the challenges above can be incorporated into a single device, we propose that nanopore-based single particle analysis has the potential to improve the diagnosis of neurodegenerative diseases. Ultimately, nanopore-based amyloid characterization may enable monitoring of neurodegenerative disease progression using microliter volumes of patient samples in a rapid, low-cost, and broadly accessible format that can be applied routinely and longitudinally to an ever increasing aging population. The insights gained from such population-based monitoring may help to accelerate the development of new therapies against those diseases.

References

[1] J. Hardy and D. J. Selkoe. “The amyloid hypothesis of Alzheimer’s disease: progress and problems on the road to therapeutics”. *science* 297.5580 (2002), pp. 353–356.

[2] F. Chiti and C. M. Dobson. “Protein misfolding, functional amyloid, and human disease”. *Annu. Rev. Biochem.* 75 (2006), pp. 333–366.

[3] T. P. Knowles and M. J. Buehler. “Nanomechanics of functional and pathological amyloid materials”. *Nature nanotechnology* 6.8 (2011), p. 469.

[4] C. Haass and D. J. Selkoe. “Soluble protein oligomers in neurodegeneration: lessons from the Alzheimer’s amyloid -peptide”. *Nature reviews Molecular cell biology* 8.2 (2007), p. 101.

[5] D. Eisenberg and M. Jucker. “The amyloid state of proteins in human diseases”. *Cell* 148.6 (2012), pp. 1188–1203.

[6] C. Ballard et al. “Alzheimer’s disease”. *Lancet (London, England)* 377.9770 (2011), pp. 1019–1031.

[7] W. H. Organization. *Dementia: a public health priority.* World Health Organization, 2012.

[8] J. J. Yerbury et al. “Walking the tightrope: proteostasis and neurodegenerative disease”. *Journal of neurochemistry* 137.4 (2016), pp. 489–505.

[9] T. L. Spires-Jones, J. Attems, and D. R. Thal. “Interactions of pathological proteins in neurodegenerative diseases”. *Acta neuropathologica* 134.2 (2017), pp. 187–205.

[10] M. E. Nelson. “Target detection, image analysis, and modeling”. *Electroreception*. Springer, 2005, pp. 290–317.

[11] J. T. Jarrett and P. T. Lansbury Jr. “Seeding one-dimensional crystallization of amyloid: a pathogenic mechanism in Alzheimer’s disease and scrapie?” *Cell* 73.6 (1993), pp. 1055–1058.

[12] X. Han and G. He. “Toward a Rational Design to Regulate -Amyloid Fibrillation for Alzheimers Disease Treatment”. *ACS chemical neuroscience* 9.2 (2018), pp. 198–210.

[13] P. Prangkio, E. C. Yusko, D. Sept, J. Yang, and M. Mayer. “Multivariate analyses of amyloid-beta oligomer populations indicate a connection between pore formation and cytotoxicity”. *PLoS One* 7.10 (2012), e47261.

[14] S. Pellegrino et al. “-Hairpin mimics containing a piperidinepyrrolidine scaffold modulate the -amyloid aggregation process preserving the monomer species”. *Chemical science* 8.2 (2017), pp. 1295–1302.

[15] M. Hölttä et al. “Evaluating amyloid- oligomers in cerebrospinal fluid as a biomarker for Alzheimers disease”. *PLoS One* 8.6 (2013), e66381.

[16] O. Hansson et al. “Levels of cerebrospinal fluid -synuclein oligomers are increased in Parkinsons disease with dementia and dementia with Lewy bodies compared to Alzheimers disease”. *Alzheimer’s research & therapy* 6.3 (2014), p. 25.

[17] E. Bagyinszky et al. “Role of inflammatory molecules in the Alzheimer’s disease progression and diagnosis”. *Journal of the neurological sciences* 376 (2017), pp. 242–254.

[18] B. Olsson et al. “CSF and blood biomarkers for the diagnosis of Alzheimer’s disease: a systematic review and meta-analysis”. *The Lancet Neurology* 15.7 (2016), pp. 673–684.

[19] J. Houghtaling, J. List, and M. Mayer. “Nanopore-Based, Rapid Characterization of Individual Amyloid Particles in Solution: Concepts, Challenges, and Prospects”. *Small* 14.46 (2018), p. 1802412.

[20] Q. Chen, Y.-Y. Liu, C. Xue, Y.-L. Yang, and W.-M. Zhang. “Energy self-sufficient desalination stack as a potential fresh water supply on small islands”. *Desalination* 359 (2015), pp. 52–58.

[21] S. J. C. Lee, E. Nam, H. J. Lee, M. G. Savelieff, and M. H. Lim. “Towards an understanding of amyloid- oligomers: characterization, toxicity mechanisms, and inhibitors”. *Chemical Society Reviews* 46.2 (2017), pp. 310–323.

[22] H. Li et al. “Amyloids and protein aggregationanalytical methods”. *Encyclopedia of Analytical Chemistry: Applications, Theory and Instrumentation* (2006).

[23] M. G. Zagorski et al. “[13] Methodological and chemical factors affecting amyloid peptide amyloidogenicity”. *Methods in Enzymology*. Vol. 309. Amyloid, Prions, and Other Protein Aggregates. Academic Press, 1999, pp. 189–204.

[24] B. J. Alper and W. K. Schmidt. “A capillary electrophoresis method for evaluation of A proteolysis in vitro”. *Journal of neuroscience methods* 178.1 (2009), pp. 40–45.

[25] C. G. Glabe. “Structural classification of toxic amyloid oligomers”. *Journal of Biological Chemistry* 283.44 (2008), pp. 29639–29643.

[26] A. N. Klein et al. “Optimization of D-peptides for A monomer binding specificity enhances their potential to eliminate toxic A oligomers”. *ACS chemical neuroscience* 8.9 (2017), pp. 1889–1900.

[27] M. R. Nilsson. “Techniques to study amyloid fibril formation in vitro”. *Methods* 34.1 (2004), pp. 151–160.

[28] V. L. Villemagne, V. Doré, S. C. Burnham, C. L. Masters, and C. C. Rowe. “Imaging tau and amyloid- proteinopathies in Alzheimer disease and other conditions”. *Nature Reviews Neurology* 14.4 (2018), pp. 225–236.

[29] R. W. DeBlois and C. P. Bean. “Counting and sizing of submicron particles by the resistive pulse technique”. *Review of Scientific Instruments* 41.7 (1970), pp. 909–916.

[30] H. Bayley and C. R. Martin. “Resistive-pulse sensing from microbes to molecules”. *Chemical Reviews* 100.7 (2000), pp. 2575–2594.

[31] C. D. Ahrberg, J. M. Lee, and B. G. Chung. “Poisson statistics-mediated particle/cell counting in microwell arrays”. *Scientific reports* 8.1 (2018), p. 2438.

[32] L.-Q. Gu, O. Braha, S. Conlan, S. Cheley, and H. Bayley. “Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter”. *Nature* 398.6729 (1999), p. 686.

[33] J. J. Kasianowicz et al. “Analytical applications for pore-forming proteins”. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1858.3 (2016), pp. 593–606.

[34] M. Ali et al. “Metal ion affinity-based biomolecular recognition and conjugation inside synthetic polymer nanopores modified with ironterpyridine complexes”. *Journal of the American Chemical Society* 133.43 (2011), pp. 17307–17314.

[35] W. Shi, A. K. Friedman, and L. A. Baker. “Nanopore sensing”. *Analytical chemistry* 89.1 (2016), pp. 157–188.

[36] W.-J. Lan, D. A. Holden, B. Zhang, and H. S. White. “Nanoparticle transport in conical-shaped nanopores”. *Analytical chemistry* 83.10 (2011), pp. 3840–3847.

[37] R. Vogel, W. Anderson, J. Eldridge, B. Glossop, and G. Willmott. “A variable pressure method for characterizing nanoparticle surface charge using pore sensors”. *Analytical chemistry* 84.7 (2012), pp. 3125–3131.

[38] E. C. Yusko et al. “Controlling protein translocation through nanopores with bio-inspired fluid walls”. *Nature Nanotechnology* 6.4 (2011), pp. 253–60.

[39] E. C. Yusko et al. “Real-time shape approximation and fingerprinting of single proteins using a nanopore”. *Nature Nanotechnology* 12 (2017), pp. 360–367.

[40] E. C. Yusko et al. “Single-particle characterization of A oligomers in solution”. *ACS nano* 6.7 (2012), pp. 5909–5919.

[41] P. Waduge et al. *Nanopore-Based Measurements of Protein Size, Fluctuations, and Conformational Changes*. research-article. 2017.

[42] X.-d. Zhao, Q.-l. Meng, L.-z. Yu, and M.-l. Li. “Composition and bioactivity of secretion from Stauronematus compressicornis (Fabricius)”. *Forestry Studies in China* 11.2 (2009), pp. 122–126.

[43] Q. Zhao, R. S. S. de Zoysa, D. Wang, D. A. Jayawardhana, and X. Guan. “Real-time monitoring of peptide cleavage using a nanopore probe”. *Journal of the American Chemical Society* 131.18 (2009), pp. 6324–6325.

[44] H.-Y. Wang, Y.-L. Ying, Y. Li, H.-B. Kraatz, and Y.-T. Long. “Nanopore analysis of -amyloid peptide aggregation transition induced by small molecules”. *Analytical chemistry* 83.5 (2011), pp. 1746–1752.

[45] W. Li et al. “Single protein molecule detection by glass nanopores”. *ACS nano* 7.5 (2013), pp. 4129–4134.

[46] N. Martyushenko, N. A. Bell, R. D. Lamboll, and U. F. Keyser. “Nanopore analysis of amyloid fibrils formed by lysozyme aggregation”. *Analyst* 140.14 (2015), pp. 4882–4886.

[47] Y.-X. Hu et al. “Single molecule study of initial structural features on the amyloidosis process”. *Chemical Communications* 52.32 (2016), pp. 5542–5545.

[48] S. Balme et al. *Influence of Adsorption on Proteins and Amyloid Detection by Silicon Nitride Nanopore*. research-article. 2016.

[49] C. Wang et al. “A novel device of array nanochannels integrated electrochemical detector for detection of amyloid aggregation and inhibitor screening”. *Electrochemistry Communications* 66 (2016), pp. 25–28.

[50] R. Hu et al. “Intrinsic and membrane-facilitated -synuclein oligomerization revealed by label-free detection through solid-state nanopores”. *Scientific reports* 6 (2016), p. 20776.

[51] N. Giamblanco et al. “Detection of protein aggregate morphology through single antifouling nanopore”. *Sensors and Actuators B: Chemical* 260 (2018), pp. 736–745.

[52] V. Sauvage et al. “Early minion nanopore single-molecule sequencing technology enables the characterization of hepatitis B virus genetic complexity in clinical samples”. *PloS one* 13.3 (2018), e0194366.

[53] W. H. Coulter. *Means for counting particles suspended in a fluid*. Google Patents, 1953.

[54] D. S. Talaga and J. Li. “Single-molecule protein unfolding in solid state nanopores”. *Journal of the American Chemical Society* 131.26 (2009), pp. 9287–9297.

[55] K. J. Freedman et al. “Chemical, thermal, and electric field induced unfolding of single protein molecules studied using nanopores”. *Analytical chemistry* 83.13 (2011), pp. 5137–5144.

[56] J. W. Robertson et al. “Single-molecule mass spectrometry in solution using a solitary nanopore”. *Proceedings of the National Academy of Sciences* 104.20 (2007), pp. 8207–8211.

[57] F. Haque, J. Li, H.-C. Wu, X.-J. Liang, and P. Guo. “Solid-state and biological nanopore for real-time sensing of single chemical and sequencing of DNA”. *Nano today* 8.1 (2013), pp. 56–74.

[58] M. Wanunu. “Nanopores: A journey towards DNA sequencing”. *Physics of life reviews* 9.2 (2012), pp. 125–158.

[59] B. M. Venkatesan and R. Bashir. “Nanopore sensors for nucleic acid analysis”. *Nature nanotechnology* 6.10 (2011), p. 615.

[60] H. P. J. Buermans and J. T. Den Dunnen. “Next generation sequencing technology: advances and applications”. *Biochimica et Biophysica Acta (BBA)-Molecular Basis of Disease* 1842.10 (2014), pp. 1932–1941.

[61] S. Howorka and Z. Siwy. “Nanopore analytics: sensing of single molecules”. *Chemical Society Reviews* 38.8 (2009), pp. 2360–2384.

[62] D. C. Golibersuch. “Observation of Aspherical Particle Rotation in Poiseuille Flow via the Resistance Pulse Technique”. *Biophysical Journal* 13.3 (1973), pp. 265–280.

[63] J. J. Kasianowicz, S. E. Henrickson, H. H. Weetall, and B. Robertson. “Simultaneous Multianalyte Detection with a Nanometer-Scale Pore”. *Analytical Chemistry* 73.10 (2001), pp. 2268–2272.

[64] W. R. Smythe. “Flow around a spheroid in a circular tube”. *The Physics of Fluids* 7.5 (1964), pp. 633–638.

[65] H. Fricke. “The Electric Permittivity of a Dilute Suspension of MembraneCovered Ellipsoids”. *Journal of Applied Physics* 24.5 (1953), pp. 644–646.

[66] S. Velick and M. Gorin. “The electrical conductance of suspensions of ellipsoids and its relation to the study of avian erythrocytes”. *The Journal of general physiology* 23.6 (1940), p. 753.

[67] M. Mayer and E. Yusko. *Nanopore-based determination of protein charge, shape, volume, rotational diffusion coefficient, and dipole moment*. Google Patents, 2017.

[68] A. Ivankin, S. Carson, S. R. Kinney, and M. Wanunu. “Fast, label-free force spectroscopy of histoneDNA interactions in individual nucleosomes using nanopores”. *Journal of the American Chemical Society* 135.41 (2013), pp. 15350–15352.

[69] M. Langecker et al. “Nanopores suggest a negligible influence of CpG methylation on nucleosome packaging and stability”. *Nano letters* 15.1 (2014), pp. 783–790.

[70] J. D. Uram, K. Ke, A. J. Hunt, and M. Mayer. “Submicrometer PoreBased Characterization and Quantification of AntibodyVirus Interactions”. *Small* 2.89 (2006), pp. 967–972.

[71] J. D. Uram, K. Ke, A. J. Hunt, and M. Mayer. “LabelFree Affinity Assays by Rapid Detection of Immune Complexes in Submicrometer Pores”. *Angewandte Chemie International Edition* 45.14 (2006), pp. 2281–2285.

[72] J. D. Uram and M. Mayer. “Estimation of solid phase affinity constants using resistive-pulses from functionalized nanoparticles”. *Biosensors and Bioelectronics* 22.7 (2007), pp. 1556–1560.

[73] W. Si and A. Aksimentiev. “Nanopore Sensing of Protein Folding”. *ACS Nano* 11.7 (2017), pp. 7091–7100.

[74] M. D. Peraro and F. G. v. d. Goot. “Pore-forming toxins: ancient, but never really out of fashion”. *Nature Reviews Microbiology* 14.2 (2016), pp. 77–92.

[75] A. Fennouri, S. F. Mayer, T. B. H. Schroeder, and M. Mayer. “Single channel planar lipid bilayer recordings of the melittin variant MelP5”. *Biochimica et Biophysica Acta (BBA) - Biomembranes* 1859.10 (2017), pp. 2051–2057.

[76] J. Zlatanova and K. van Holde. “Single-Molecule Biology: What Is It and How Does It Work?” *Molecular Cell* 24.3 (2006), pp. 317–329.

[77] A. L. Demain and P. Vaishnav. “Production of recombinant proteins by microbes and higher organisms”. *Biotechnology advances* 27.3 (2009), pp. 297–306.

[78] G. Huang, K. Willems, M. Soskine, C. Wloka, and G. Maglia. “Electro-osmotic capture and ionic discrimination of peptide and protein biomarkers with FraC nanopores”. *Nature Communications* 8.1 (2017), p. 935.

[79] C. Wloka et al. “Label-Free and Real-Time Detection of Protein Ubiquitination with a Biological Nanopore”. *ACS nano* 11.5 (2017), pp. 4387–4394.

[80] D. Wendell et al. “Translocation of double-stranded DNA through membrane-adapted phi29 motor protein nanopores”. *Nature Nanotechnology* 4 (2009), p. 765.

[81] M. Soskine, A. Biesemans, M. De Maeyer, and G. Maglia. “Tuning the size and properties of ClyA nanopores assisted by directed evolution”. *Journal of the American Chemical Society* 135.36 (2013), pp. 13456–13463.

[82] S. M. Bezrukov, I. Vodyanoy, R. A. Brutyan, and J. J. Kasianowicz. “Dynamics and Free Energy of Polymers Partitioning into a Nanoscale Pore”. *Macromolecules* 29.26 (1996), pp. 8517–8522.

[83] S. Howorka, S. Cheley, and H. Bayley. “Sequence-specific detection of individual DNA strands using engineered nanopores”. *Nature Biotechnology* 19 (2001), p. 636.

[84] J. J. Kasianowicz and S. M. Bezrukov. “Protonation dynamics of the alpha-toxin ion channel from spectral analysis of pH-dependent current fluctuations.” *Biophysical Journal* 69.1 (1995), pp. 94–105.

[85] S. M. Bezrukov and J. J. Kasianowicz. “Current noise reveals protonation kinetics and number of ionizable sites in an open protein ion channel”. *Physical Review Letters* 70.15 (1993), pp. 2352–2355.

[86] J. J. Kasianowicz, J. W. Robertson, E. R. Chan, J. E. Reiner, and V. M. Stanford. “Nanoscopic Porous Sensors”. *Annual Review of Analytical Chemistry* 1.1 (2008), pp. 737–766.

[87] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer. “Characterization of individual polynucleotide molecules using a membrane channel”. *Proceedings of the National Academy of Sciences* 93.24 (1996), pp. 13770–13773.

[88] D. K. Lubensky and D. R. Nelson. “Driven polymer translocation through a narrow pore”. *Biophysical Journal* 77.4 (1999), pp. 1824–1838.

[89] Q. Zhao, D. A. Jayawardhana, D. Wang, and X. Guan. “Study of peptide transport through engineered protein channels”. *The Journal of Physical Chemistry B* 113.11 (2009), pp. 3572–3578.

[90] R. A. Sharples et al. “Inhibition of -secretase causes increased secretion of amyloid precursor protein C-terminal fragments in association with exosomes”. *The FASEB Journal* 22.5 (2008), pp. 1469–1478.

[91] L. Hong et al. “Quantification of the binding properties of Cu2+ to the amyloid beta peptide: coordination spheres for human and rat peptides and implication on Cu2+-induced aggregation”. *The Journal of Physical Chemistry B* 114.34 (2010), pp. 11261–11271.

[92] C. C. Curtain et al. “Metal ions, pH, and cholesterol regulate the interactions of Alzheimer’s disease amyloid- peptide with membrane lipid”. *Journal of Biological Chemistry* 278.5 (2003), pp. 2977–2982.

[93] G. Olivieri et al. “NAcetyllcysteine protects SHSY5Y neuroblastoma cells from oxidative stress and cell cytotoxicity: effects on amyloid secretion and tau phosphorylation”. *Journal of neurochemistry* 76.1 (2001), pp. 224–233.

[94] H.-Y. Wang, Z. Gu, C. Cao, J. Wang, and Y.-T. Long. “Analysis of a single -synuclein fibrillation by the interaction with a protein nanopore”. *Analytical chemistry* 85.17 (2013), pp. 8254–8261.

[95] C. A. Madampage, O. Tavassoly, C. Christensen, M. Kumari, and J. S. Lee. “Nanopore analysis: An emerging technique for studying the folding and misfolding of proteins”. *Prion* 6.2 (2012), pp. 116–123.

[96] R. I. Stefureac, C. A. Madampage, O. Andrievskaia, and J. S. Lee. “Nanopore analysis of the interaction of metal ions with prion proteins and peptides”. *Biochemistry and Cell Biology* 88.2 (2010), pp. 347–358.

[97] O. Tavassoly and J. S. Lee. “Methamphetamine binds to synuclein and causes a conformational change which can be detected by nanopore analysis”. *FEBS letters* 586.19 (2012), pp. 3222–3228.

[98] O. Tavassoly, J. Kakish, S. Nokhrin, O. Dmitriev, and J. S. Lee. “The use of nanopore analysis for discovering drugs which bind to -synuclein for treatment of Parkinson’s disease”. *European journal of medicinal chemistry* 88 (2014), pp. 42–54.

[99] A. Henning-Knechtel, J. Knechtel, and M. Magzoub. “DNA-assisted oligomerization of pore-forming toxin monomers into precisely-controlled protein channels”. *Nucleic acids research* 45.21 (2017), pp. 12057–12068.

[100] E. Spruijt, S. E. Tusk, and H. Bayley. “DNA scaffolds support stable and uniform peptide nanopores”. *Nature nanotechnology* (2018), p. 1.

[101] A. Oukhaled et al. “Dynamics of Completely Unfolded and Native Proteins through Solid-State Nanopores as a Function of Electric Driving Force”. *ACS Nano* 5.5 (2011), pp. 3628–3638.

[102] J. Li et al. “Ion-beam sculpting at nanometre length scales”. *Nature* 412.6843 (2001), pp. 166–169.

[103] J. Li, M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko. “DNA molecules and configurations in a solid-state nanopore microscope”. *Nature materials* 2.9 (2003), p. 611.

[104] A. Han et al. “Sensing protein molecules using nanofabricated pores”. *Applied Physics Letters* 88.9 (2006), p. 093901.

[105] H. Kwok, K. Briggs, and V. Tabard-Cossa. “Nanopore fabrication by controlled dielectric breakdown”. *PloS one* 9.3 (2014), e92880.

[106] A. J. Storm, J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker. “Fabrication of solid-state nanopores with single-nanometre precision”. *Nature Materials* 2.8 (2003), pp. 537–540.

[107] J. Yang et al. “Rapid and precise scanning helium ion microscope milling of solid-state nanopores for biomolecule detection”. *Nanotechnology* 22.28 (2011), p. 285310.

[108] L. J. Steinbock, J. F. Steinbock, and A. Radenovic. “Controllable shrinking and shaping of glass nanocapillaries under electron irradiation”. *Nano letters* 13.4 (2013), pp. 1717–1723.

[109] L. J. de Vreede, A. van den Berg, and J. C. Eijkel. “Nanopore fabrication by heating Au particles on ceramic substrates”. *Nano letters* 15.1 (2015), pp. 727–731.

[110] J. B. Heng et al. “The detection of DNA using a silicon nanopore”. *IEEE International Electron Devices Meeting 2003*. 2003, pp. 32.2.1–32.2.4.

[111] M.-H. Lee et al. “A Low-Noise Solid-State Nanopore Platform Based on a Highly Insulating Substrate”. *Scientific Reports* 4 (2014), p. 7448.

[112] A. B. Farimani, K. Min, and N. R. Aluru. “DNA base detection using a single-layer MoS2”. *ACS nano* 8.8 (2014), pp. 7914–7922.

[113] J. Larkin et al. “Slow DNA Transport through Nanopores in Hafnium Oxide Membranes”. *ACS Nano* 7.11 (2013), pp. 10121–10128.

[114] Z. S. Siwy and M. Davenport. “Nanopores: Graphene opens up to DNA”. *Nature nanotechnology* 5.10 (2010), p. 697.

[115] C. Plesa et al. *Fast Translocation of Proteins through Solid State Nanopores*. rapid-communication. 2013.

[116] L. T. Sexton et al. “An adsorption-based model for pulse duration in resistive-pulse protein sensing”. *Journal of the American Chemical Society* 132.19 (2010), pp. 6755–6763.

[117] M. Mayer, E. Yusko, and J. Yang. *Controlling translocation through nanopores with fluid wall*. Google Patents, 2016.

[118] A. Balan, C.-C. Chien, R. Engelke, and M. Drndi. “Suspended Solid-state Membranes on Glass Chips with Sub 1-pF Capacitance for Biomolecule Sensing Applications”. *Scientific Reports* 5 (2015), p. 17775.

[119] J. K. Rosenstein, M. Wanunu, C. A. Merchant, M. Drndic, and K. L. Shepard. “Integrated nanopore sensing platform with sub-microsecond temporal resolution”. *Nature methods* 9.5 (2012), p. 487.

[120] A. Ananth et al. “Reversible Immobilization of Proteins in Sensors and Solid-State Nanopores”. *Small* 14.18 (2018), p. 1703357.

[121] E. A. Manrao et al. “Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase”. *Nature biotechnology* 30.4 (2012), p. 349.

[122] A. S. Mikheyev and M. M. Tin. “A first look at the Oxford Nanopore MinION sequencer”. *Molecular ecology resources* 14.6 (2014), pp. 1097–1102.

[123] J. E. Reiner, J. J. Kasianowicz, B. J. Nablo, and J. W. F. Robertson. “Theory for polymer analysis using nanopore-based single-molecule mass spectrometry”. *Proceedings of the National Academy of Sciences* 107.27 (2010), p. 12080.

[124] S. Kumar et al. “PEG-Labeled Nucleotides and Nanopore Detection for Single Molecule DNA Sequencing by Synthesis”. *Scientific Reports* 2 (2012), p. 684.

[125] C. W. Fuller et al. “Real-time single-molecule electronic DNA sequencing by synthesis using polymer-tagged nucleotides on a nanopore array”. *Proceedings of the National Academy of Sciences* 113.19 (2016), pp. 5233–5238.

[126] E. E. Nesterov et al. “In vivo optical imaging of amyloid aggregates in brain: design of fluorescent markers”. *Angewandte Chemie International Edition* 44.34 (2005), pp. 5452–5456.

[127] G. V. Soni et al. “Synchronous optical and electrical detection of biomolecules traversing through solid-state nanopores”. *Review of Scientific Instruments* 81.1 (2010), p. 014301.

[128] S. Liu, A. R. Hawkins, and H. Schmidt. “Optofluidic devices with integrated solid-state nanopores”. *Microchimica Acta* 183.4 (2016), pp. 1275–1287.

[129] S. Huang, M. Romero-Ruiz, O. K. Castell, H. Bayley, and M. I. Wallace. “High-throughput optical sensing of nucleic acids in a nanopore array”. *Nature Nanotechnology* 10.11 (2015), pp. 986–991.

[130] B. N. Anderson et al. “Probing Solid-State Nanopores with Light for the Detection of Unlabeled Analytes”. *ACS Nano* 8.11 (2014), pp. 11836–11845.

[131] T. Gilboa and A. Meller. “Optical sensing and analyte manipulation in solid-state nanopores”. *Analyst* 140.14 (2015), pp. 4733–4747.

[132] S. Krishnan et al. “Molecular transport through large-diameter DNA nanopores”. *Nature communications* 7 (2016), p. 12787.

[133] T. Tucker, M. Marra, and J. M. Friedman. “Massively parallel sequencing: the next big thing in genetic medicine”. *The American Journal of Human Genetics* 85.2 (2009), pp. 142–154.

[134] K. Misiunas, N. Ermann, and U. F. Keyser. “QuipuNet: convolutional neural network for single-molecule nanopore sensing”. *Nano letters* (2018).

[135] R. Y. Henley et al. “Electrophoretic Deformation of Individual Transfer RNA Molecules Reveals Their Identity”. *Nano letters* 16.1 (2016), pp. 138–144.

[136] M. Langecker et al. “Synthetic lipid membrane channels formed by designed DNA nanostructures”. *Science* 338.6109 (2012), pp. 932–936.

[137] T. B. H. Schroeder, J. Houghtaling, B. D. Wilts, and M. Mayer. “It’s Not a Bug, It’s a Feature: Functional Materials in Insects”. *Advanced Materials* (2018).

[138] R. An et al. “Ultrafast laser fabrication of submicrometer pores in borosilicate glass”. *Optics Letters* 33.10 (2008), pp. 1153–1155.

[139] A. E. Herr. “Disruptive by design: a perspective on engineering in analytical chemistry”. *Analytical chemistry* 85.16 (2013), pp. 7622–7628.

Chapter 2: Fabrication of Nanopores

Synthetic nanopores with diameters from 20-50 nm in freestanding silicon nitride (SiN_x) membranes are useful for single-molecule studies of globular biological macromolecules and protein complexes. These pores can be fabricated to a desired size using many different methods on a broad array of different substrate materials. One limitation to using synthetic nanopores, as discussed in Chapter 1, is that proteins adhere non-specifically to synthetic surfaces; our group has circumvented this issue by coating the surface of these nanopores with a lipid bilayer. Lipid bilayers have qualities that are essential for the protein characterization presented throughout this thesis, but their ability form stable and reproducible coatings in this context depends on factors including lipid composition, nanopore geometry, surface chemistry, surface roughness, and others. Simply creating a nanopore of a desired size is no guarantee that it will be amenable to lipid coating. In this chapter, I discuss my efforts to fabricate nanopores using multiple techniques, and I evaluate the resulting pores with regard to their noise properties, their performance in nanopore recordings, and their ability to support stable lipid bilayer coatings.

2.1 Introduction to Nanopore Fabrication

The critical component of any successful nanopore-based protein sensing experiment is a suitable nanopore. Ideally, this pore would be a perfectly cylindrical hole with a well-characterized geometry, while also maintaining long-term stability under rigorous cleaning, and would be fabricated within a low-noise, low-capacitive substrate having surface qualities amenable to a lipid-bilayer coating. Recent interest in nanoscale techniques has brought about a wide variety of approaches to fabricate such pores including but not limited to; dielectric breakdown,[1], TEM drilling,[2] helium ion microscope drilling,[3] capillary shrinking,[4] and gold particle heating,[5] while substrates range from silicon,[6] silica (glass),[7] silicon nitride,[8] MoS2,[9] HfO2,[10] and graphene [11] (Figure 2.1). Despite this diversity, most successful protein experiments presented in this thesis were performed using nanopores made with an ion beam sculpting technique developed by Li *et al.*,[8] whereby a low-intensity, feedback-controlled helium ion beam induces material migration of silicon nitride, effectively closing a large nanopore ($\geq$ 100 nm diameter) to a nanopore of desired size. This process results in a relatively low-noise pore that can be successfully coated with a lipid bilayer.[12] More importantly, the robust structure of ion beam sculpted nanopores allows them to withstand upwards of 20 experimental setups and cleanings – in comparison, we have observed that pores made with other methods tend to change size or shape after only several experimental attempts.[13] Note that pore stability has important implications with regard to experimental accuracy and reproducibility (see Chapter 3 for details). Nevertheless, due to the timescale and complexity of the ion beam sculpting process, and given that other nanopore substrates like fused silica exhibit superior noise characteristics, we have investigated several alternative fabrication techniques and present the results of those investigations throughout this chapter.

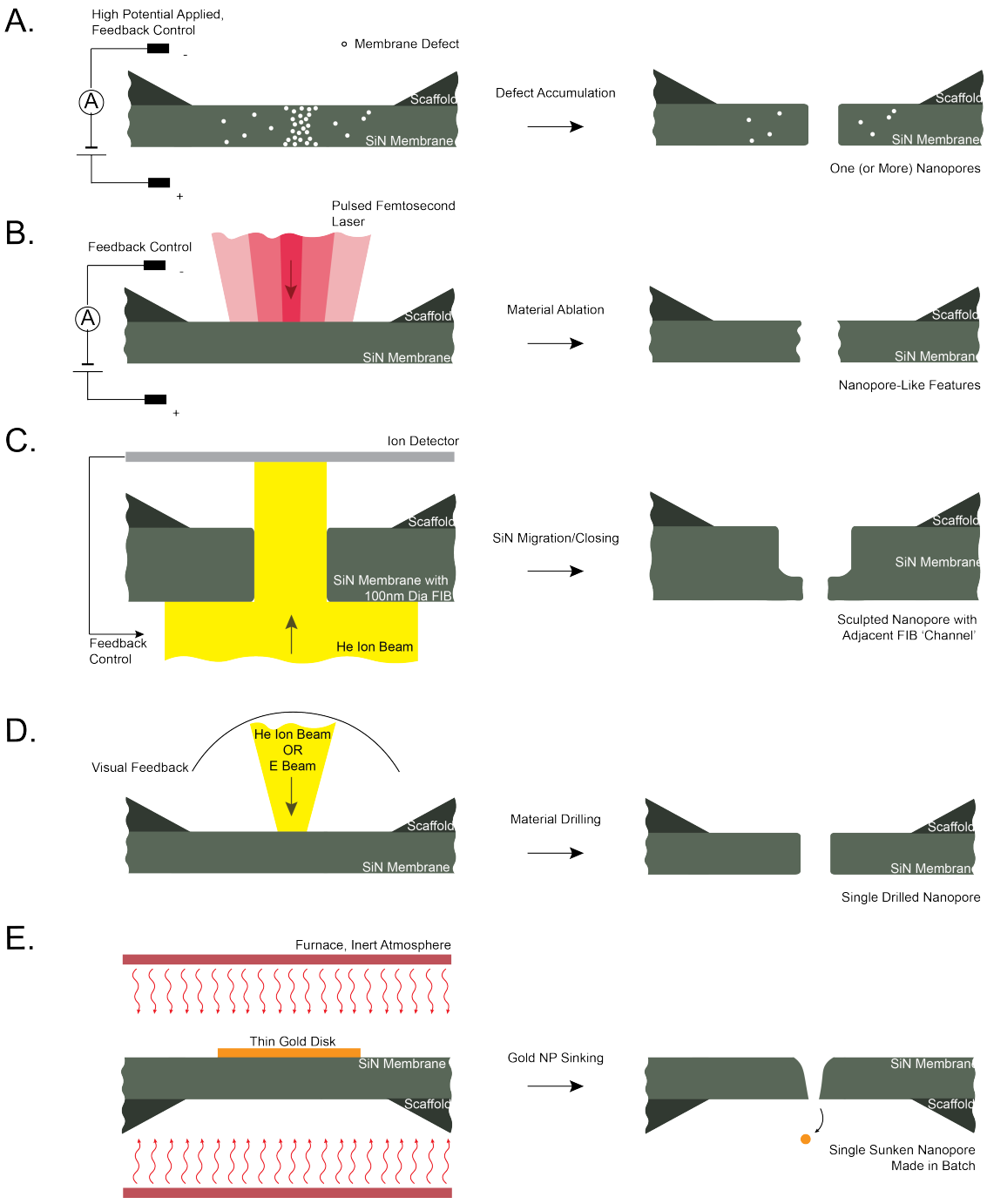


Figure 2.1. Schematic of techniques to fabricate nanopores in synthetic substrates. A) Dielectric breakdown, whereby a strong electrical potential is applied across a freestanding SiN_x membrane in order to accumulate defects which ultimately converge into a nanopore (or multiple nanopores).[1] B) Femtosecond laser ablation, where a femtosecond pulsed laser is focused onto a freestanding membrane and pulses with feedback-control from ionic current through the membrane. C) Ion beam sculpting, whereby helium ions are directed through a thick ($\sim$300 nm) freestanding membrane with a $\sim$100 nm diameter pore, inducing material migration and effectively closing the existing pore to a desired size.[8] D) Electron or Ion Beam drilling, in which a beam of high energy particles drills a hole of desired size – confirmed visually – in a freestanding membrane.[2, 3] E) Gold nanoparticle sinking, in which a gold disk is patterned on a freestanding membrane, and once heated up to $\sim 1000°C$, the gold sinks through the SiN_x, leaving a nanopore behind. Note that this fabrication method can be carried out in batch, rather than serially like the other methods here. [5]

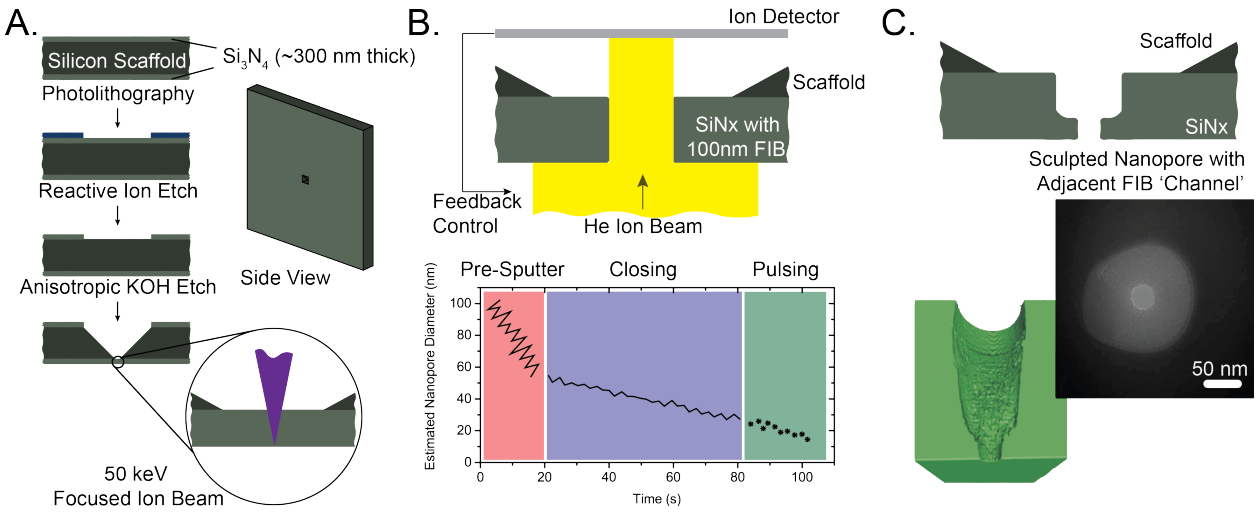


Figure 2.2. Process for fabricating nanopores by ion beam sculpting. A) We generate a freestanding membrane on a silicon scaffold using standard wafer processing techniques, and then irradiate that membrane with a gallium ion beam to produce a channel with a diameter of approximately 100 nm. B) We direct helium ions through through this channel in a three step process: first, we use a high beam energy to close the pore rapidly (*i.e.* pre-sputter) under manual control, then we reduce the beam energy to close the pore slowly under automated feedback, and finally we pulse the beam for precise control of final nanopore diameter. C) The resulting nanopore has a unique geometry with an adjacent channel – green tomographic cross-section adapted from Rollings.[15]

2.2 Ion Beam Sculpting of Single Nanopores in SiN_x Membranes

In 2001, Li *et al.* introduced the method of feedback-controlled, ion beam sculpting,[8] and then employed this technique to fabricate a nanopore with a diameter of 5 nm, which they used to observe translocations of double stranded DNA.[14] Our group later coated these nanopores with a lipid bilayer and demonstrated that the functionalized pores have a variety of appealing properties for protein characterization.[12] These initial works inspired an ongoing collaboration between the Li and Mayer research groups, and during my time as a doctoral researcher under Prof. Michael Mayer, I made several trips to the University of Arkansas in order to fabricate nanopores using this technique. Figure 2.2 shows details of the fabrication method, as well as images of the resulting nanopores.

The fabrication technique begins with a silicon chip containing a small (100 μm × 100 μm) and thin (~300 nm) freestanding SiN_x window, which is produced on a wafer-scale

using standard etching and photolithography methods (Figure 2.2-A). A focused ion beam (FIB) is then used to drill a pore between 70 and 120 nm in diameter completely through the freestanding window, after which the pore is irradiated with low-intensity beam of helium ions to induce SiN_x migration under ion flux-based feedback, effectively closing the FIB hole to a desired diameter (20-40 nm for our purposes). The closing procedure itself consists of three stages, outlined in Figure 2.2-B: 1) an initial pre-sputter step, where the beam current is set to 15 mA under user-controlled feedback (*i.e.* the pore begins to close quickly and the user manually switches off the beam), 2) a secondary closing step, where the beam current is reduced to 200 μA to allow for gradual closing with computer-automated feedback, and 3) a final pulsing step, where the user applies individual pulses 100 ms in length at a beam current of 15 mA for precise diameter control. Once the nanopore appears to have a desired size based on ion-flux measurements, it is placed in a furnace set to $\sim 700^\circ$ C under an inert atmosphere in order to anneal the migrated SiN_x material, and then imaged using a TEM (Figure 2.2-C) to verify pore dimensions. The resulting nanopore has a unique geometry with the nanopore adjacent to a larger channel, as shown in the green tomographic cross-section from Ryan Rollings (Figure 2.2-C).[15] We derive a method to properly account for this geometry in the analysis of resistive pulses in Chapter 4.

Regarding the success rates of these nanopores, we have observed that one in three ion beam sculpted nanopores can be used for protein characterization experiments. There are three typical causes for failure: 1) the measured baseline current is lower than expected ($I_{baseline} < 0.9 * I_{expected}$) based on the known nanopore geometry prior to coating the nanopore with a lipid bilayer, 2) the chip does not coat with a lipid bilayer based on the predicted increase in resistance after coating (see Section 2.4 below for more details), or 3) the baseline current after coating is too noisy to detect translocation events. For the work presented in Chapter 3, we obtained the expected baseline current in 73% of attempts, successfully coated the pore in 37% of attempts (cumulative success rate = 27%), and achieved sufficiently low noise for recording after successfully coating the pore in 46% of attempts

(cumulative success rate = 12%). These statistics indicate that roughly 1 in 10 experiments yields a measurement, on average; however, it is worth stressing that the success rate is highly dependent on the nanopore being used. For example, an ion beam sculpted nanopore that supports a stable lipid bilayer can typically be coated repeatedly, and that particular nanopore may enable upwards of 20 successful experiments before either growing in size or clogging irreversibly.

2.3 Preliminary Results of Femtosecond Laser Ablation of Nanopores

Long-pulse ($\geq$ 10 ps) and continuous-wave (cw) lasers interact with material *via* linear absorption, directly heating a region of interest within some absorption depth, and then heating adjacent regions at a rate dependent on the thermal conductivity of the material, ultimately leading to material ablation.[16] One problem with using these kinds of lasers to produce nanopores is that their spot sizes are typically on the order of hundreds of nanometers, so ablated features are often too large (*e.g.* $\geq$ 100 nm in diameter) for use in experiments aiming to characterize protein.[17] Material absorption of ultrafast pulses ($\leq$ 1 ps), on the other hand, is nonlinear and plasma mediated. The absorption intensity of these types of lasers should, in theory, be strongly concentrated around the center of the laser spot and decay rapidly moving outward, making it possible to fabricate features much smaller than the laser spot size.[16]

In attempt to generate single nanopores in a rapid and reproducible way, we employed a diode-pumped Nd:glass chirped-pulse amplifation laser system (Intralase) to generate 600 fs pulses with a wavelength of 1053 nm. These pulses were focused through a KTP crystal to clean their intensity profile before they passed into a Axiovert 200M inverted microscope with a 0.65 NA Achroplan air objective. We mounted substrate chips containing freestanding SiN_x windows (Norcada, 5 $\times$ 5 mm frame or 3 mm disk, 250 $\times$ 250 μm window with

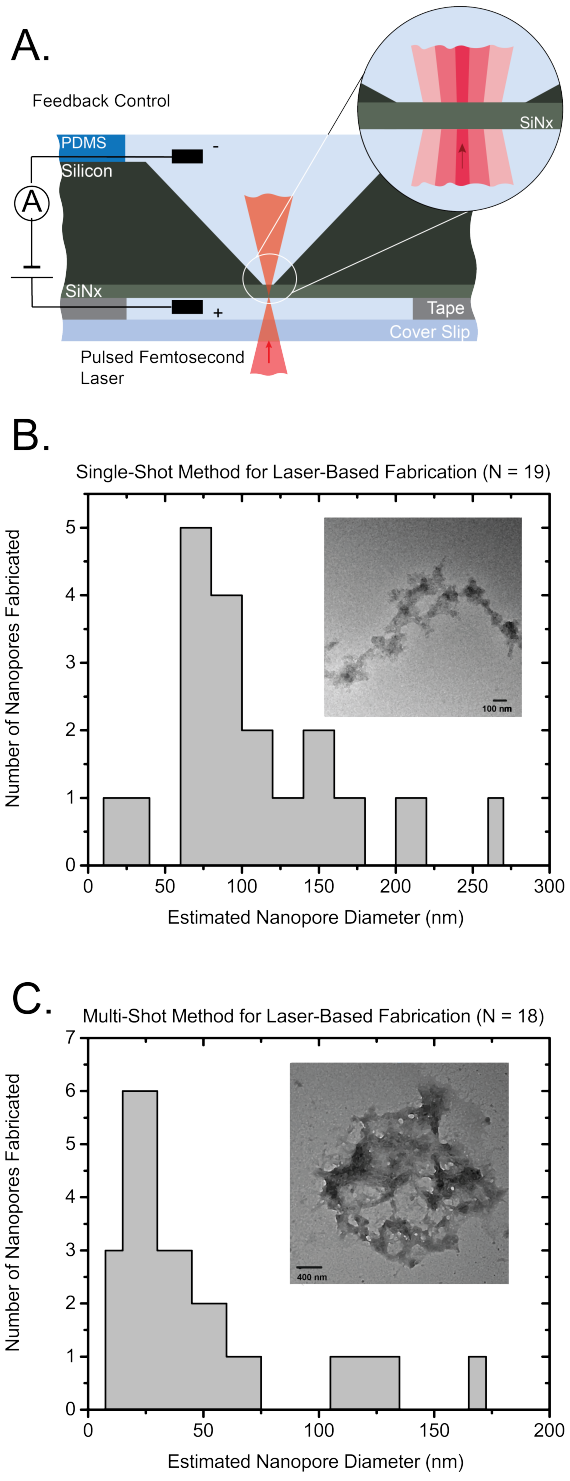


Figure 2.3. Setup schematic and size distributions of nanopores fabricated *via* femtosecond laser. A) Cross-section view of a silicon chip on the microscopy stage, with fluidic channels above and below the membrane accessed by Ag/AgCl electrodes, and the membrane secured to a glass coverslip using double-sided tape. B) Histogram showing the nanopore diameters estimated from the conductance across the membrane (see equation 2.1), with an inset showing a TEM image of a feature generated using the single-shot approach. C) Histogram showing the nanopore diameters estimated from the conductance across the membrane (see equation 2.1), with an inset showing a TEM image of a feature generated using the multi-shot approach.

30 nm thickness) as shown in Figure 2.3-A on the microscope stage and positioned the window laterally with high precision using a piezoelectric positioner.[18] We applied a 100 mV potential across the defect-free membrane, and continuously monitored the current using a Kiethley picoammeter connected to Ag/AgCl electrodes immersed in 2 M KCl filling the compartments above and below the chip.

We employed two different methods for fabricating nanopores with the femtosecond laser: single-shot, and multi-shot. In the single-shot approach, we set the laser potential to 7.4 mV, released a single pulse, and checked the measured current. If the membrane remained intact (*i.e.* the current was 0 pA), we moved the laser focal position by 0.4 μm, incremented the potential by 0.1% and repeated the procedure (Figure 2.3-B). In the multi-shot approach, we set the laser potential to 4.8 mV and released up to 10 pulses with at least a 1 second delay in between at a single location before moving to a new location 1 μm away and repeating the process (Figure 2.3-C). We fabricated approximately 20 "nanopores" using each of these methods, and report their theoretical diameters based upon their conductance after fabrication in Figure 2.3-B,C. Subsequent TEM imaging showed, however, that rather than creating a single well-defined nanopore, the laser – in both fabrication methods – was ablating a large region of the window and creating non-cylindrical nanoscale features that conducted ionic current (Figure 2.3-B,C insets). After confirming this result on several TEM images, we decided to discontinue our efforts to fabricate nanopores with this approach.

2.4 Formation of Nanopores with Dielectric Breakdown

While current nanopore fabrication processes using ion or electron beam drilling can produce nanopores with 1-nm precision,[2, 3, 8, 11] these methods often require advanced skill and specialized instrumentation, limiting broad access. An alternate approach to nanopore fabrication is controlled breakdown (CBD),[1, 19] a stochastic process involving defect accumulation in an insulating membrane under high external electric fields (0.5 to 1 V/nm).

Kwok *et al.* demonstrated that when the density of accumulated defects reaches a critical value, a channel with sub-nanometer diameter opens in the freestanding membrane, the resulting nanopore fills with electrolyte and produces in a measurable increase in current (Figure 2.4).[1] After this initial breakthrough event, reduction of the electric field can be used to slowly enlarge the nanopore to a desired diameter.[1, 20] During the CBD process, monitoring the ionic conductance provides an estimate of the diameter of the nanopore at a time resolution of milliseconds. The equation below gives the relationship between the ionic conductance, G, of a cylindrical nanopore and its radius, R_p (m).[1, 21]

$$G = \sigma \left(\frac{l_p}{\pi R_p^2} + \frac{1}{2R_p} \right)^{-1} \tag{2.1}$$

Here σ (S m^{-1}) is the conductivity of the electrolyte solution and l_p (m) is the length of the nanopore.

Figure 2.4-B shows a controlled breakdown process with an applied voltage of 0.6 V nm^{-1} during the initial breakdown phase (18 V across a SiN_x membrane with thickness of 30 nm), and 0.3 V nm^{-1} during the enlargement phase. The observed stepwise increase in current suggests either fast, discrete enlargement processes during pore growth or the formation of additional breakdown paths. The TEM image in Figure 2.4-C illustrates an example of a nanopore chip after fabrication of a large nanopore by controlled breakdown. The image also shows the unintended formation of a second nanopore in the field of view, which is a result that we observed in preliminary efforts to fabricate pores by dielectric breakdown.

Although TEM images provide a direct visualization of nanopore quantity and size, it can be time consuming and difficult to locate a randomly-generated, nanometer-scale pore in a large area of a SiN_x membrane (*e.g.* 250 μm × 250 μm for most chips used here for dielectric breakdown fabrication) littered with debris – this was also an issue when imaging the laser-fabricated pores in Section 2.3. In addition, TEM imaging precludes the immediate use of the nanopore for resistive pulse recording directly after CBD formation.[1] Hence, to

determine whether more than one nanopore formed in the membrane, we took advantage of geometric considerations of controlled pore shrinkage upon coating the nanopores with a lipid bilayer.[12] Assuming a cylindrical nanopore has radius R_p and length l_p, a lipid bilayer coating reduces this radius to $R_p - t_b - W_L$ and extends the pore length to $l_p + 2t_b + 2W_L$ (Figure 2.4-D). Here $t_b = 3.7$ nm is the length of a lipid bilayer made from 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC), and $W_L = 1.2$ nm is the average thickness of the interstitial water layer between the lipid bilayer and the nanopore wall.[12] Hence, the pore's conductance after lipid bilayer coating changes from above equation to:

$$G = \sigma \left(\frac{l_p + 2t_b + 2W_L}{\pi (R_p - t_b - W_L)^2} + \frac{1}{2(R_p - t_b - W_L)} \right)^{-1} \tag{2.2}$$

Using the equation above, we calculated the expected conductance of a single cylindrical nanopore after lipid bilayer coating as a function of the conductance before coating (red curve in Figure 2.4-D). This reduction in conductance after surface coating is a function of nanopore geometry as well. [23–25] In an attempt to evaluate the effect of non-cylindrical shapes, we determined the expected conductance changes before and after lipid bilayer coating for different nanopore geometries by finite element simulations (see Chapter 5 for more details). A cone shape and a double-cone shape (both with an angle of 45°) had conductance values that were within 10% of the theoretically expected conductance after coating a perfectly cylindrical nanopore (Figure 2.4-D). This result means that deviations of more than 10% from the theoretically expected conductance after coating do not necessarily indicate non-cylindrical pore geometries, but rather indicate the formation of multiple pores or imperfect lipid coatings.

Based on our previous experience with coating nanopores,[12, 26] we expect single nanopores with stable lipid bilayer coatings to have conductance values within 20% of the theoretically expected conductance. When evaluating nanopores formed by controlled breakdown, we observed that only ∼20% of nanopores met the criterion of deviating by less than ±20% from the expected conductance after coating. For the nanopores outside of this threshold,

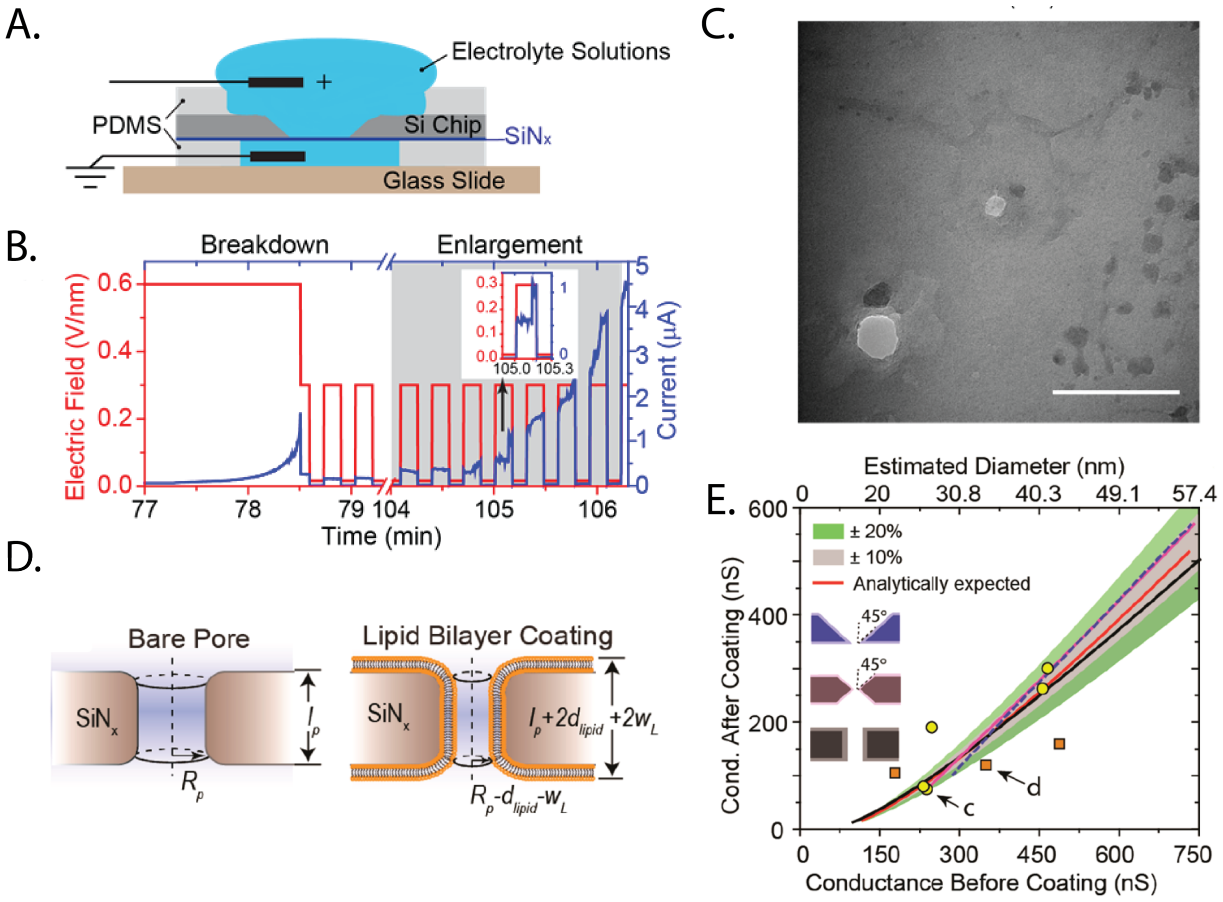


Figure 2.4. Overview of controlled dielectric breakdown technique. A) Schematic drawing of the setup used for controlled breakdown (not to scale); the electrolyte solution was 2 M $LiCl_2$ at pH 8.0. B) Graph of the electric field applied across a 30-nm thick SiN_x membrane and the corresponding current measured during the nanopore formation process and subsequent period of pore enlargement. Once breakdown occurred, as indicated by a sudden increase in current (>150 nA), feedback control decreased the voltage from 0.6 V nm^{-1} for breakdown to 0.3 V nm^{-1} for enlargement. During the enlargement phase of the experiment, the current occasionally changed in steps (see inset), indicating possible formation of additional pores.[1] C) TEM image of the SiN_x membrane indicating undesirable formation of a second nanopore. For this experiment, we used custom-ordered chips with a freestanding SiN_x membrane of 10 μm × 10 μm to reduce the time required to scan the entire window by TEM. Scale bar is 100 nm. D) Schematic of expected changes in nanopore geometry after lipid bilayer coating (not to scale). E) Plot of nanopore conductance after lipid bilayer coating as a function of the conductance before coating calculated by finite element simulations for a single cylindrical nanopore (black curve), a single cone-shaped pore (blue dashed curve), and a single double-cone-shaped pore (violet curve). The red curve represents the analytically expected conductance change of a single cylindrical nanopore based on Eq. 2.1 and Eq. 2.2; the green region represents ± 20% deviation from the red curve, and the grey region represents ± 10% deviation from the red curve. The yellow dots and orange squares are the conductance values before and after coating of different chips used for Ca^{2+} fluorescence detection. Yellow dots indicate single nanopores; orange squares indicate multiple nanopores. To convey the estimated nanopore size, the upper x-axis represents the estimated diameter assuming a single cylindrical nanopore derived from the conductance based on Eq. 2.1. Adapted with permission.[22] Copyright 2018, American Chemical Society.

we attribute their large conductance values after coating to either imperfect coating or the existence of multiple nanopores in the membrane. While it is straightforward to understand why a partial coating or a coating with defects would lead to larger than theoretically predicted conductance after coating, it may seem counterintuitive that the formation of additional pores may lead to larger than expected conductance after coating. Indeed, coating two pores with a lipid bilayer would lead to a smaller conductance after coating than coating a single pore with equivalent pre-coating conductance. The situation is, however, different if one, several, or all pores that may have formed during a controlled breakdown process have diameters smaller than 16 nm. Anecdotal evidence in our lab over the last five years showed that we were never able to coat nanopores with diameters smaller than 16 nm with a lipid bilayer.[12, 26] One plausible reason for this observation is that supported lipid bilayers are limited in their geometric conformations and cannot conform to high-curvature surfaces due to surface tension (see Section 2.6.1).[27–29] For instance, if there are multiple pores in the membrane with diameters smaller than 16 nm, then these pores would contribute to the conductance before coating, but because the supported lipid bilayer would not form within these pores, the total measured conductance after coating would be larger than theoretically predicted by Eq. 2.2. On the other hand, a smaller-than-expected conductance after coating might indicate one of two different scenarios: 1) the nanopore(s) are sufficiently small that liposome fusion forms pore-spanning lipid bilayers, or 2) as mentioned above, multiple nanopores with diameters larger than 16 nm could successfully coat with a lipid bilayer, leading to a smaller conductance than coating a single nanopore with corresponding pre-coating conductance.

Initial attempts to fabricate nanopores with this technique demonstrated that controlled breakdown carries the risk of generating multiple pores in the membrane. We later identified two strategies for minimizing this risk and forming single nanopores: accelerating defect accumulation within a confined membrane area before breakdown, and reducing the magnitude of the electric field applied during enlargement. By applying a focused laser beam on the

SiN_x membrane during the breakdown process and by decreasing the electric field strength during pore enlargement, we tripled the success rate for generating single nanopores with diameters larger than 20 nm. We have also presented evidence [22] that laser-induced local heating contributes to the increased rate of pore formation by controlled breakdown. As with the ion beam sculpted nanopores, a subsequent annealing step at $\sim 700^\circ$ C improved the success rate of coating these nanopores with a lipid bilayer, which is critical to quantify the translational and rotational dynamics of single proteins, and to make such characterization accessible to a broader community. Annealing also reduced the current noise of nanopore chips in general (*i.e.* also without a coating and slowed undesirable enlargement of pore diameters during cleaning and recording compared to chips that were not annealed).[15] One additional exciting aspect of generating large nanopores without the need for FIB or electron beam equipment is that it may further accelerate the development of nanopore-based, single-particle and single-molecule interrogation approaches.

2.5 Fabrication of Nanopores in Fused Silica Substrates and Comparison of Different Nanopore Platforms

The strong (MV m^{-1}) electric field in the nanopore moves charged particles and proteins through the pore on nanosecond to millisecond time scales; the temporal resolution of resistive pulse recordings is, therefore, one of the most important and limiting parameters of the technique. This temporal resolution is limited by recording electronics as well as by the need for low-pass filtering in order to minimize current noise such that resistive pulses can be distinguished from the noise.[30] Specifically, resistive pulse sensing with nanopores suffers from four major contributions to noise, namely flicker noise, thermal noise, dielectric noise and amplifier noise.[30–32] Smeets *et al.* proposed that flicker noise originates from surface charges and nanobubbles present at the nanopore wall; this source of noise influences the lowest frequencies in electrical recordings.[32] The factors contributing to thermal

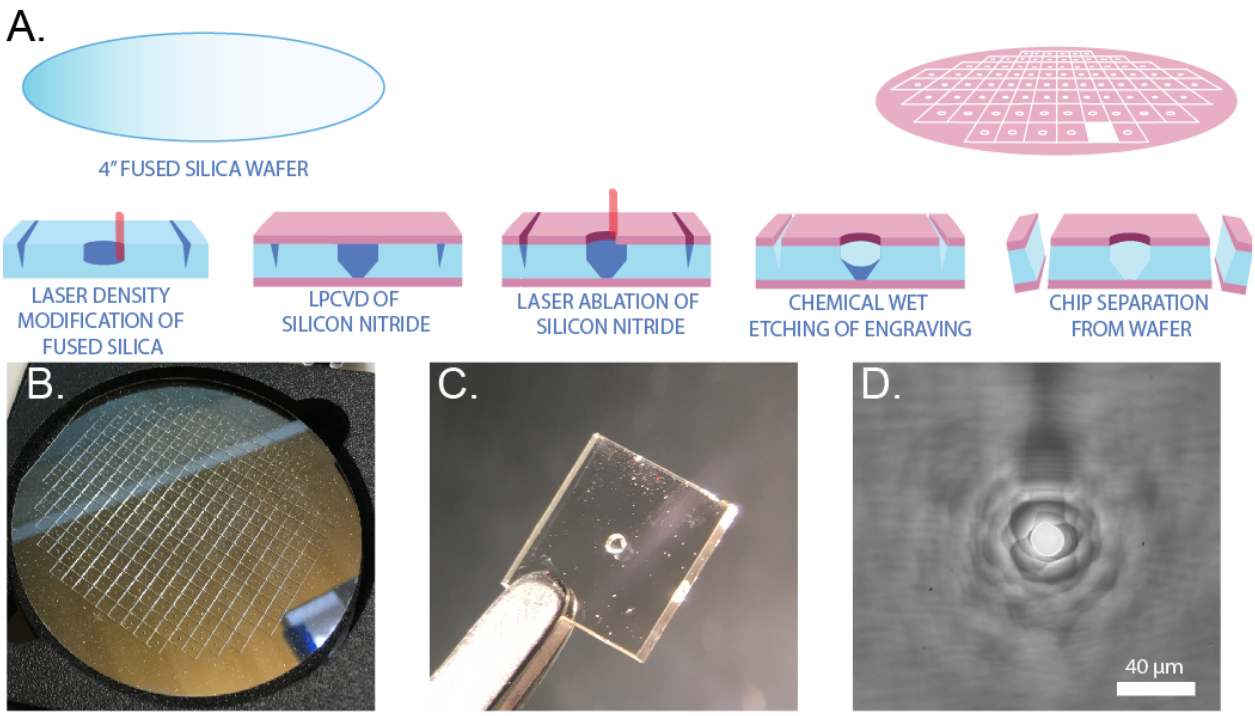


Figure 2.5. Schematic showing wafer-scale fabrication process for fused silica substrates. A) Fabrication scheme for fused silica chips with a SiN_x membrane by laser density modification of fused silica followed by accelerated etching of the laser exposed parts of the chips. B) Photograph of a fused silica wafer with a diameter of 100 mm showing wafer scale fabrication of 322 chips. C) Photograph of a 4 mm $\times$ 4 mm fused silica chip with a 500 µ-thick frame and the opening of the cavity with diameter of 500 µm. D) Optical microscopy image of the transparent, freestanding SiN_x membrane at the bottom of the etched cavity. Adapted with permission.[37] Copyright 2019, IOP Publishing.

noise are the electrolyte concentration and nanopore size and affect the full recording bandwidth.[31, 33] The main noise sources that become increasingly important for high bandwidth recordings are the dielectric noise and amplifier noise. While amplifier noise results from amplifier design, electrode connections, shielding and grounding of the setup as well as from the presence and choice of electrolyte solution, dielectric noise results from the capacitance of the nanopore chip and the recording setup.[30] Therefore, in order to resolve globular macromolecules passing through nanopores on a time scale of microseconds and shorter, high-bandwidth recordings are required, which begin at 50 kHz and ideally approach 1 to 10 MHz.[34–36]

In order to limit the noise levels and to maximize the signal-to-noise ratio (SNR) at high recording bandwidth, it is therefore important to reduce the electrical capacitance

of the nanopore chips. Chips with nanopores for resistive pulse sensing usually consist of a cavity in a support material from one side, which leads to a freestanding membrane containing a nanopore. These freestanding membranes are commonly prepared by thin-film deposition techniques such as low-pressure chemical vapor deposition (LPCVD) or plasma enhanced chemical vapor deposition (PECVD), while the cavities are typically defined using photolithography and formed using dry and wet etching (see Figure 2.5-A) [36, 38–42] or by a membrane transfer technique.[7, 43–45] Most chips used for nanopore recordings are silicon based.[7, 36, 38, 40, 42] Due to the semi-conducting electrical properties of silicon and the presence of thin, insulating SiO_2 or SiN_x layers on many silicon support structures, silicon chips can, however, have a large capacitance, leading to large current noise at high recording bandwidth.[30, 46] Nonetheless, silicon remains often the material of choice due to its established manufacturability. Various methods can reduce capacitive noise from silicon chips. For example, deposition of relatively thick SiO_2 insulation layers in-between the silicon and the SiN_x membrane layer or on all exposed silicon surfaces can greatly reduce capacitance.[47] Shekar *et al.* showed that placing recording amplifiers directly on the chip with the nanopore greatly reduced noise, in part, because the approach minimized capacitive contributions from the electrical recording setup and, in part, because the intrinsic noise of this custom-made amplifier was very low.[36] This approach enabled measurements at bandwidths up to 10 MHz while maintaining adequate SNR to resolve the translocation of ssDNA through nanopores filled with a solution of 3 M KCl.[36] Polymeric insulator coatings deposited on the membrane layer after chip fabrication can also reduce capacitance and dielectric noise. Examples of insulating coatings on chip surfaces include deposition of polyimide nanospheres or thin layers of painted polydimethylsiloxane (PDMS).[38, 42, 43, 48] Alternatively, several groups have shown a significant improvement in the SNR during translocation experiments with silica-based chips or glass nanopipettes. These materials are completely insulating such that their capacitance is in the range of a few picoFarad and below.[7, 32, 38, 39, 49, 50]

In an attempt to reduce capacitance and recording noise, we manufactured chips by 3D patterning of a fused silica wafer with a femtosecond-pulsed laser in combination with LPCVD of a SiN_x layer and a subsequent chemical wet etching step at an approximately 200-fold accelerated etch rate in the laser-exposed regions of the wafer (Figure 2.5). Unlike in Section 2.3 above where we ablated the freestanding membrane, here the laser was applied to the thick fused silica substrate in order to increase its rate of etching. The use of a writing step with a femtosecond-pulsed laser combined with subsequent chemical wet etching for the creation of sub-micrometer features recently became available as a manufacturing technique for high volume production of micro-structured glass.[51–55] We compared the noise levels of the resulting fused silica chips with silicon chips and with previously published work at bandwidths approaching 1 MHz, and Figure 2.6 shows an example equivalent circuit for estimating the capacitance of a layered, silicon-based nanopore chip.

The capacitance of the chip architecture shown in Figure 2.6 can be estimated by treating the freestanding membrane region and substrate region of the chip as capacitors in parallel using the following equation:

$$C_{Tot} = \frac{1}{1/C_1 + 1/C_2} + \frac{1}{1/C_3 + 1/C_4 + 1/C_5 + 1/C_6} \tag{2.3}$$

where

$$C = \varepsilon_0 \varepsilon_{med} \frac{Area}{Thickness} \tag{2.4}$$

As in Table 2.1, ε_0 is the vacuum permitivity, and ε_{med} is the dielectric constant of the medium. Using this approach, we estimated capacitance values for chip architectures reported in the literature, and compared those estimates to capacitance values that we measured experimentally for a variety of different nanopore chips in Table 2.1.

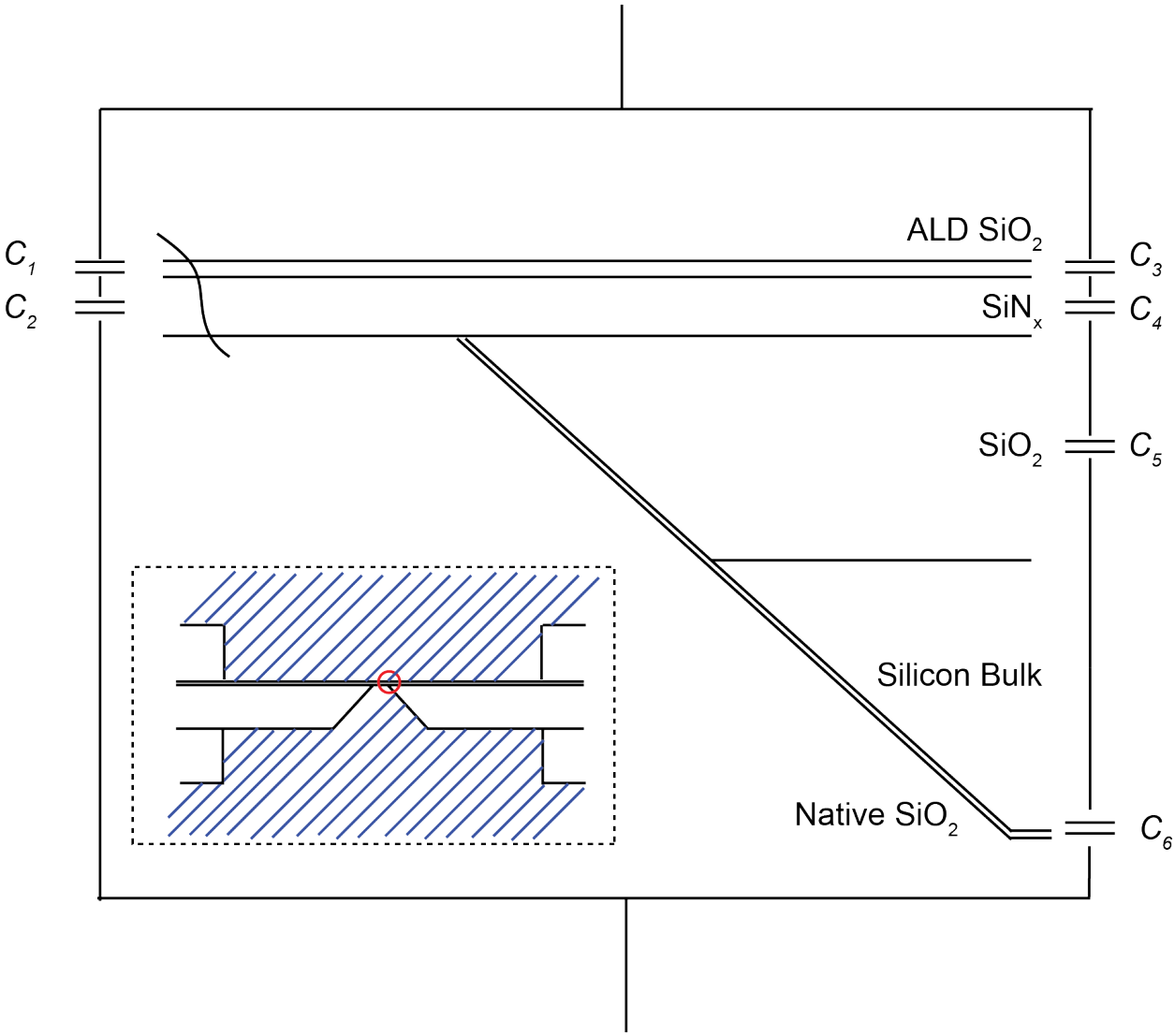


Figure 2.6. Schematic showing cross-section of a synthetic nanopore and the equivalent circuit for estimating capacitance. Each material layer of the chip can be treated as a unique capacitor (C_1 - C_6), and the overall theoretical capacitance can be evaluated using equation 2.3.

Table 2.1. Comparison of noise and capacitance values across different synthetic nanopore platforms. Adapted with permission.[37] Copyright 2019, IOP Publishing.

Chip Materials	Material of Freestanding Membrane	Nanopore Fabrication Method	Membrane Thickness	Membrane Size	Nanopore Diameter	Amplifier	Capacitance with Nanopore [a]	RMS Noise	Estimated Thermal Noise (I_T) [b]	Reference
Fused silica / a-Si	SiNx	Electron Beam	5 nm	500 nm × 500 nm	1.5 nm	Axopatch 200B	n.s.	12.6 pA at 10 kHz	n.a.	Lee et al. Sci Reports 2014
Si / SiNx / PDMS	SiNx	Electron Beam	30 nm	~50 µm	6.0 nm	Axopatch 200B	n.s.	3 pA at 10 kHz	1.2 pA at 10 kHz	Tabbard Cossa et al. Nanotech 2007
Pyrex / a-Si / SiNx / PDMS	SiNx	Electron Beam	20 nm	5 µm × 5 µm	27 nm	A-M systems 2400	n.s.	4 pA at 10 kHz	4.8 pA at 10 kHz	Pitchford et al. ACS Nano 2015
Si / SiO2 / SiNx / Cyanoacrylate / glass	SiNx	Electron Beam	10 nm	150 nm × 150 nm	6.3 nm	Chimera VC100	1.5 pF**	133.3 pA at 1 MHz	19.4 pA at 1 MHz	Balan et al. Nano Lett 2014
Fused silica / SiNx / Graphene	Graphene	Electron Beam	0.34 nm	Ø 2.1 µm	3 nm	Chimera VC100	< 1 pF*	110 pA at 1 MHz	21.6 pA at 1 MHz	Balan et al. Sci Reports 2015
SiO2 / Si / SiO2 / SiNx / Silicone	SiNx	Electron Beam	4 nm	50 nm × 50 nm	2.5 nm	CMOS	10 pF*	23.2 pA at 200 kHz 125.7 pA at 1 MHz	9.8 pA at 200 kHz 21.9 pA at 1 MHz	Shekar et al. Nano Lett 2016
SiO2 / Si / SiO2 / SiNx	SiNx	Electron Beam	10 nm	50 µm	3.5 nm	CMOS	6 pF*	12.9 pA at 100 kHz 155 pA at 1 MHz	3.4 pA at 100 kHz 10.9 pA at 1 MHz	Rosenstein et al. Nat Methods 2012
Si / SiO2 / Si3N4 / PDMS	Si3N4 / polymide bead coating	Dielectric Breakdown	10 nm	500 nm × 500 nm	2 nm	Chimera VC100	50 pF**	580 pA at 1 MHz	6.9 pA at 1 MHz	Goto et al. Sci Reports 2016
Si / SiNx / PDMS	SiNx	Dielectric Breakdown	10 nm	50 µm × 50 µm	5.4 nm	Chimera VC100	n.s.	62 pA at 250 kHz	10 pA at 250 kHz	Karau et al. ACS Sensors 2018
Fused silica capillary	n.a.	Pipette Pulling	n.a.	n.a.	n.s.	Custom	n.a.	50 pA at 1 MHz	n.a.	Fraccari et al. Nanoscale 2016
Fused silica / SiNx	SiNx	Dielectric Breakdown	30 nm	Ø 20 µm	25 nm	Chimera VC100	0.75 pF**, 2.2 pF*	11.8 pA at 10 kHz 99.6 pA at 250 kHz 316.2 pA at 1 MHz	5.9 pA at 10 kHz 29.5 pA at 250 kHz 59.1 pA at 1 MHz	De Vreede et al. Nanotech 2019
Fused silica / SiNx	SiNx	Dielectric Breakdown	30 nm	Ø 20 µm	25 nm	Axopatch 200B	0.75 pF** , 2.2 pF*	18.7 pA at 10 kHz	5.9 pA at 10 kHz	De Vreede et al. Nanotech 2019
Silicon / SiNx	SiNx	Dielectric Breakdown	30 nm	9 µm × 9 µm	26 nm	Axopatch 200B	213 pF*	34.6 pA at 10 kHz	6.1 pA at 10 kHz	De Vreede et al. Nanotech 2019
Silicon / SiNx	SiNx	Ion Beam Sculpting	275 nm	25 µm × 25 µm	29 nm	Axopatch 200B	440 pF*	17.4 pA at 10 kHz	6.9 pA at 10 kHz	Li et al. Nature 2001
Silicon / SiNx	SiNx	Ion Beam Sculpting	275 nm	25 µm × 25 µm	32 nm	Axopatch 200B	720 pF*	28.6 pA at 10 kHz	7.2 pA at 10 kHz	Li et al. Nature 2001
Silicon / SiNx	SiNx	Helium Ion Microscope	30 nm	150 µm × 150 µm	23 nm	Axopatch 200B	884 pF*	32.9 pA at 10 kHz	5.5 pA at 10 kHz	Marshall et al. Scanning 2012
Silicon / SiNx / SiO2	SiNx	Electron Beam	30 nm	25 µm × 25 µm	37 nm	Axopatch 200B	34.4 pF*	73.8 pA at 10 kHz	8.0 pA at 10 kHz	Wanunu et al. Nat Nano 2010
Pyrex / SiNx	SiNx	Dielectric Breakdown	20 nm	Ø 2 µm	40 nm	Axopatch 200B	6.9 pF*	23.1 pA at 10 kHz	9.3 pA at 10 kHz	Lee et al. Sci Reports 2014

[a] Measured capacitance (*) or theoretically estimated capacitance (**) using $C = \varepsilon_r \varepsilon_0 \frac{A}{d}$

[b] The thermal noise(I_T) was estimated using $I_T(f_c) = \sqrt{\frac{4kTc_1 f_c}{R}}$

n.s. not specified

n.a. not applicable

From Table 2.1, it is clear that nanopores fabricated in fused silica substrates outperform those in silicon substrates. In combination with recent developments in integrated CMOS current amplifiers, these low-noise chips could further improve the information content and accuracy of estimates of multiple parameters calculated from the resistive pulses of individual globular biomolecules such as proteins.[26, 40] This improved information content and accuracy has important implications for future investigations that are discussed in detail in Chapter 6.

One additional benefit of fused silica substrates is that they exhibit no laser-induced electrical noise upon illumination. In contrast, measurements with silicon chips showed strong noise upon laser illumination in agreement with a previous report.[43] Therefore, fused silica substrates are well suited for nanopore-based experiments that perform electrical recording and optical monitoring in parallel.[43, 44] Approaches that combine optical with electrical measurements have shown promise in recent applications for DNA sequencing [56, 57] but have thus far been limited by optical and thermal noise under laser illumination.

It is important to note that there are limitations on the resolution of this fabrication method. The femtosecond pulsed laser makes it possible to modify the fused silica to within a distance of 25 μm from the SiN_x thin film (see Figure 2.7). If the laser focal spot is placed closer than 25 μm from the SiN_x film, it can damage or ablate the SiN_x film and render it inadequate as an insulating freestanding membrane material, similar to what is shown in Figure 2.3-B,C. We therefore had to leave a small gap between the laser-patterned fused silica and the interface between the fused silica and the SiN_x to prevent unwanted damage of the SiN_x. Reducing the diameter of the freestanding membrane would require: 1) improving the precision (sub-micrometer) of z-axis control of the laser focal spot, or 2) further optimizing the protocol for KOH etching, possibly by implementing a feedback measurement of electrical capacitance. Another possible solution, for example, would be to use deep reactive ion etching in place of KOH during the etching step. Both of these approaches may be excellent extensions for future optimization of this fabrication method.

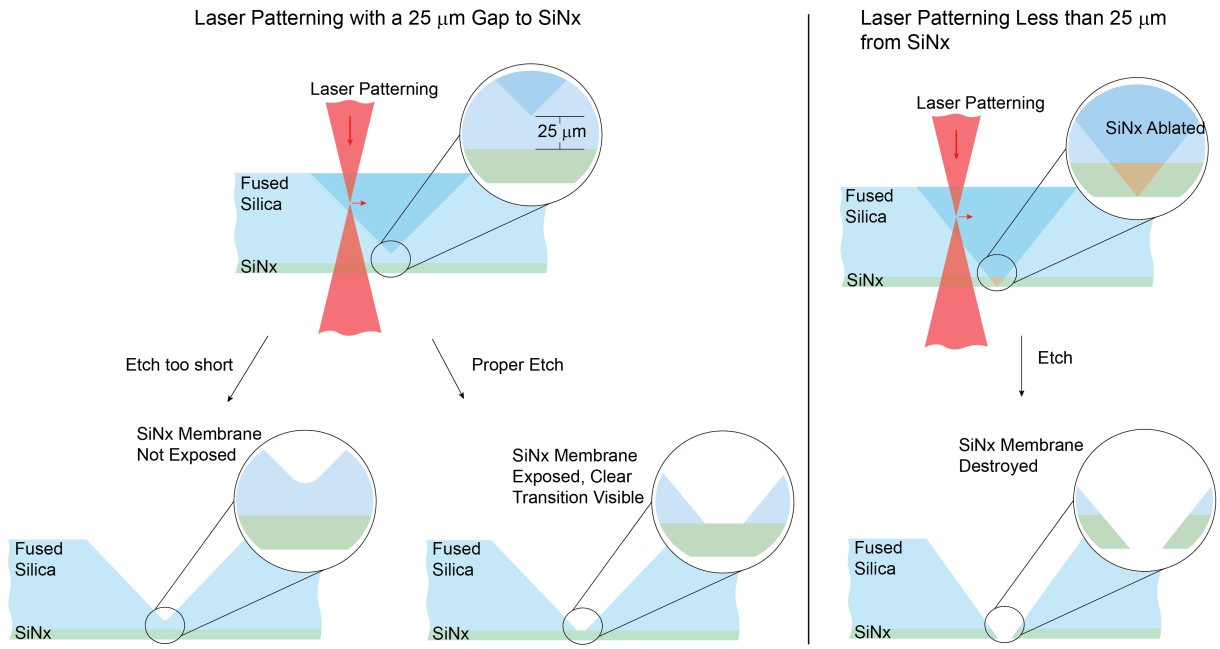


Figure 2.7. (LEFT) Schematic showing laser patterning within 25 μm of the SiN_x membrane. If the subsequent KOH etch is too short (lower left), we achieve a rounded cone in the fused silica which renders the chip useless. Under the proper etching conditions, we expose an approximately circular area of the SiN_x membrane (lower center). (RIGHT) Schematic showing laser patterning too near to the SiN_x membrane, resulting in ablation of SiN_x. After etching the fused silica, there is no longer a free standing SiN_x membrane (lower right).

Looking forward, the fabrication process we report here also has potential benefits extending beyond producing nanopores, as it can generate other three-dimensional nanoscale features such as fluidic channels, mixers and reaction chambers on the surface or within the transparent bulk material of fused silica.[51–55] Advanced combinations of these lab-on-a-chip features with nanopore detectors, produced in a batch format, may ultimately enable the next generation of nano- and microscale devices with possible applications in low-cost diagnostics, point-of-care devices, fundamental biophysics studies, and implantable measurement systems. Moreover, in the context of recent work on large area nanopore arrays [5, 58] the direct writing method presented here may contribute to the field of energy research where such arrays in membranes function as part of energy storage or conversion devices.[59]

2.6 Supplementary Notes and Figures

2.6.1 Automation of FRAP Analyses for Bilayer Characterization

In the context of nanopore-based protein sensing, lipid bilayer coatings carry several benefits; they make it possible to finely tune nanopore diameters of an individual pore between experiments, they prevent unwanted protein adhesion to the nanopore substrate, they can promote analyte specificity through the addition of lipid-conjugated functional groups, and they can slow both the translation and rotation of proteins through lipid anchoring.[12] But why is this important? Nanopore dimensions are critical to translocation analyses, and a single nanopore is useful for a particular range of protein sizes (*e.g.* a 20 nm diameter nanopore can typically accommodate proteins less than 100 kDa, while a 30 nm nanopore is suited to larger sizes). Modifying the lipid coating composition can lead to 1.4 nm of dynamic diameter adjustment, which extends the range of potential protein targets accessible to a well-characterized nanopore, thus improving analysis accuracy and reducing fabrication demand. Additionally, the non-stick nature of lipid coated pores means they rarely clog during an experiment, and their bilayers can be stripped off and regenerated 20 or more times in some cases. Again, this improves inter-experiment versatility and nanopore robustness. The ability to anchor protein to the bilayer, and to target particular proteins with this anchoring, also has important implications for nanopore-based protein experiments. It vastly increases the translocation event frequency (*e.g.* the number of pore-transiting protein molecules detected per second) of a desired protein at a given concentration; bulk (unanchored/low-affinity) protein requires more than a 300-fold greater concentration to translocate at the same frequency as anchored protein.[12] This reduces experimental time and improves general data quality by limiting the influence of false noise-based events. Anchoring also bolsters the information gathered from each individual translocation event, as it slows protein translation and rotation, leading to better time resolution of those actions. We exploit many of these advantages in Chapters 3 and 4 of this thesis.

When attempting to coat nanopores with lipid bilayers, we find that certain pores are more amenable to lipid bilayer coatings than others. One possible reason for the discrepancy in ease of coating between different pores may be due to heterogeneities in the sharpness of the radii of curvature of the corners at the brims of the pores. While we typically observe relatively low success rates when coating nanopores with a lipid membrane, nanopores with diameters less than 20 nm [22] and those with sharp edges rarely support stable lipid coatings. This same phenomenon of curvature-dependent bilayer formation appears both in research on phospholipid vesicles[27] and in applications of lipid membranes to nanoparticles.[28, 29, 60] We therefore tested if lipid membrane compositions that form more flexible coatings with regard to membrane curvature, when compared to those composed of cylindrically shaped lipids, would be able to conform more readily to sharp corners and thereby coat nanopores with extremely small radii of curvature. Figure 2.8 shows an example of a curvature-tolerant lipid composition containing lipids with a cone-shaped geometry (1-oleoyl-2-hydroxy-sn-glycero-3-phosphocholine or LysoPC), with a complementary inverted cone shape (1,2-dioleoyl-sn-glycero-3-phosphoethanolamine or DOPE), and with 1,2-dioleoyl-sn-glycero-3-phosphocholine (DOPC), a cylindrically-shaped lipid.[29]

Curvature-tolerant lipid bilayers did not improve coating success rates, and in fact, we observed that they significantly increased the noise in current recordings. These elevated noise levels putatively stem from the diversity of lipid shapes present in the bilayer. Given that membranes containing conical lipids are prone to forming curved regions,[61] membrane undulations along the pore walls are likely more pronounced when the membrane is formed from a mixture of curvature-tolerant lipids compared to membranes comprised of homogeneously-shaped lipids.

In addition to the curvature-tolerant lipid composition, we evaluated more than ten other lipid mixtures based on their baseline stability, noise properties, viscosity, and their ease of coating. The viscosity in particular is relevant as it determines the extent of "slowing" of a protein anchored within the lipid membrane. We determined the effective viscosity (*i.e.* the

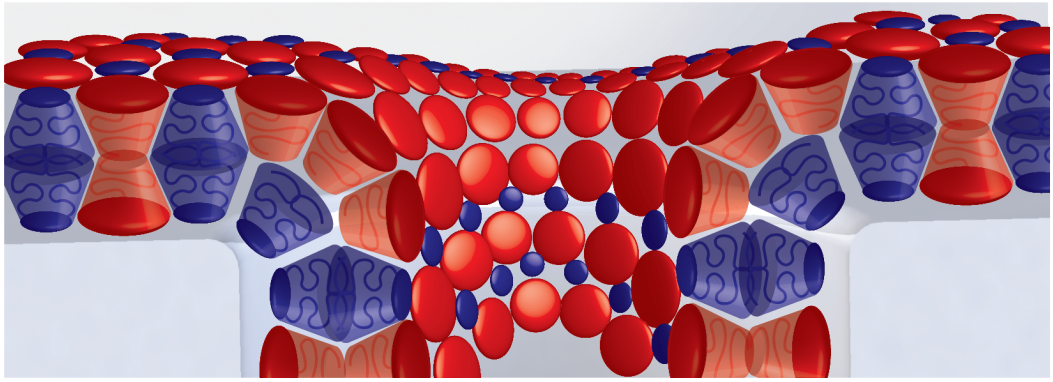


Figure 2.8. Cartoon showing a lipid bilayer coating with complementary conical lipids. We expect this lipid composition to conform to sharp corners more readily by re-arranging such that the head groups with smaller surface area (blue) align to the surface interface, and the head groups with the larger surface area (red) align on the solution interface.

diffusion coefficient of the lipids within the membrane) of each different composition using Fluorescent Recovery After Photobleaching (FRAP) analysis, and designed a MATLAB script and associated GUI to automate the calculations. Figure 2.9 outlines the procedure of determining the lipid diffusion coefficient from a FRAP experiment.

In the context of resistive pulse-based sensing, and based upon the results from all lipid compositions tested, we ultimately recommend lipid coatings comprised of pure POPC with up to 20 mol% cholesterol, or coatings comprised of Archaea-inspired lipids with up to 10-20 area% POPC. Archaea-inspired lipid compositions enabled the longest anchored translocation times of all the lipids tested while maintaining stable, low-noise baselines, and are best suited for experiments that seek to confine proteins in the nanopore for as long as possible.

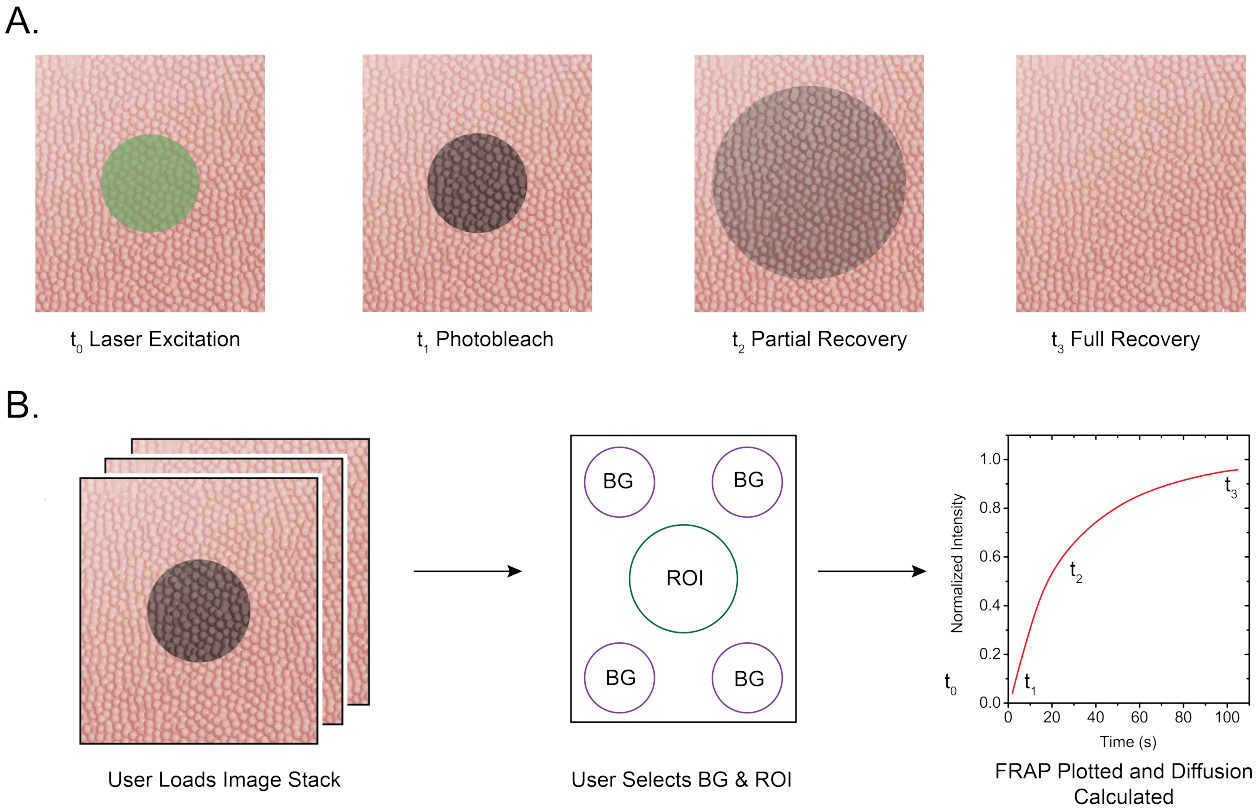


Figure 2.9. Process used for Fluorescent Recovery After Photobleaching (FRAP) measurements. A) We direct a laser at a small region of the membrane in order to photobleach fluorescently active lipid headgroups, and then allow those headgroups to diffuse throughout the membrane (in the absence of light), taking photographs at consistent time intervals. B) A custom MATLAB program reads in these image stacks, allows the user to select regions for the photobleached region of interest (ROI) and adjacent background regions (BG), and then performs a normalized intensity analysis to determine the rate of diffusion of the photobleached lipids into the surrounding bilayer.

References

[1] H. Kwok, K. Briggs, and V. Tabard-Cossa. “Nanopore fabrication by controlled dielectric breakdown”. *PloS one* 9.3 (2014), e92880.

[2] A. J. Storm, J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker. “Fabrication of solid-state nanopores with single-nanometre precision”. *Nature Materials* 2.8 (2003), pp. 537–540.

[3] J. Yang et al. “Rapid and precise scanning helium ion microscope milling of solid-state nanopores for biomolecule detection”. *Nanotechnology* 22.28 (2011), p. 285310.

[4] L. J. Steinbock, J. F. Steinbock, and A. Radenovic. “Controllable shrinking and shaping of glass nanocapillaries under electron irradiation”. *Nano letters* 13.4 (2013), pp. 1717–1723.

[5] L. J. de Vreede, A. van den Berg, and J. C. Eijkel. “Nanopore fabrication by heating Au particles on ceramic substrates”. *Nano letters* 15.1 (2015), pp. 727–731.

[6] J. B. Heng et al. “The detection of DNA using a silicon nanopore”. *IEEE International Electron Devices Meeting 2003*. 2003, pp. 32.2.1–32.2.4.

[7] M.-H. Lee et al. “A Low-Noise Solid-State Nanopore Platform Based on a Highly Insulating Substrate”. *Scientific Reports* 4 (2014), p. 7448.

[8] J. Li et al. “Ion-beam sculpting at nanometre length scales”. *Nature* 412.6843 (2001), pp. 166–169.

[9] A. B. Farimani, K. Min, and N. R. Aluru. “DNA base detection using a single-layer MoS2”. *ACS nano* 8.8 (2014), pp. 7914–7922.

[10] J. Larkin et al. “Slow DNA Transport through Nanopores in Hafnium Oxide Membranes”. *ACS Nano* 7.11 (2013), pp. 10121–10128.

[11] Z. S. Siwy and M. Davenport. “Nanopores: Graphene opens up to DNA”. *Nature nanotechnology* 5.10 (2010), p. 697.

[12] E. C. Yusko et al. “Controlling protein translocation through nanopores with bio-inspired fluid walls”. *Nature Nanotechnology* 6.4 (2011), pp. 253–60.

[13] D. Stein, J. Li, and J. A. Golovchenko. “Ion-beam sculpting time scales”. *Physical review letters* 89.27 (2002), p. 276106.

[14] J. Li, M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko. “DNA molecules and configurations in a solid-state nanopore microscope”. *Nature materials* 2.9 (2003), p. 611.

[15] R. C. Rollings. “The geometry and sensitivity of ion-beam sculpted nanopores for single molecule DNA analysis” (2013).

[16] Z. Qian et al. “Pulsetrain-burst mode, ultrafast-laser interactions with 3D viable cell cultures as a model for soft biological tissues”. *Biomedical optics express* 5.1 (2014), pp. 208–222.

[17] R. S. Marjoribanks et al. “Ablation and thermal effects in treatment of hard and soft materials and biotissues using ultrafast-laser pulse-train bursts”. *Photonics & Lasers in Medicine* 1.3 (2012), pp. 155–169.

[18] B. R. Bruhn et al. “Dual-pore glass chips for cell-attached single-channel recordings”. *Lab on a Chip* 14.14 (2014), pp. 2410–2417.

[19] K. Briggs et al. “Kinetics of nanopore fabrication during controlled breakdown of dielectric membranes in solution”. *Nanotechnology* 26.8 (2015), p. 084004.

[20] E. Beamish, H. Kwok, V. Tabard-Cossa, and M. Godin. “Precise control of the size and noise of solid-state nanopores using high electric fields”. *Nanotechnology* 23.40 (2012), p. 405301.

[21] M. Wanunu. “Nanopores: A journey towards DNA sequencing”. *Physics of life reviews* 9.2 (2012), pp. 125–158.

[22] C. Ying et al. “Formation of Single Nanopores with Diameters of 2050 nm in Silicon Nitride Membranes Using Laser-Assisted Controlled Breakdown”. *ACS Nano* (2018).

[23] A. T. Kuan, B. Lu, P. Xie, T. Szalay, and J. A. Golovchenko. “Electrical pulse fabrication of graphene nanopores in electrolyte solution”. *Applied physics letters* 106.20 (2015), p. 203109.

[24] C. M. Frament, N. Bandara, and J. R. Dwyer. “Nanopore surface coating delivers nanopore size and shape through conductance-based sizing”. *ACS applied materials & interfaces* 5.19 (2013), pp. 9330–9337.

[25] Y. N. D. Bandara, J. W. Nichols, B. Iroshika Karawdeniya, and J. R. Dwyer. “Conductance-based profiling of nanopores: Accommodating fabrication irregularities”. *Electrophoresis* 39.4 (2018), pp. 626–634.

[26] E. C. Yusko et al. “Real-time shape approximation and fingerprinting of single proteins using a nanopore”. *Nature Nanotechnology* 12 (2017), pp. 360–367.

[27] S. H. White. “Small phospholipid vesicles: internal pressure, surface tension, and surface free energy”. *Proceedings of the National Academy of Sciences* 77.7 (1980), pp. 4048–4050.

[28] Y. Roiter et al. “Interaction of nanoparticles with lipid membrane”. *Nano letters* 8.3 (2008), pp. 941–944.

[29] S. Mornet, O. Lambert, E. Duguet, and A. Brisson. “The formation of supported lipid bilayers on silica nanoparticles revealed by cryoelectron microscopy”. *Nano letters* 5.2 (2005), pp. 281–285.

[30] J. D. Uram, K. Ke, and M. Mayer. “Noise and bandwidth of current recordings from submicrometer pores and nanopores”. *ACS nano* 2.5 (2008), pp. 857–872.

[31] C. Wen et al. “Generalized Noise Study of Solid-State Nanopores at Low Frequencies”. *ACS Sensors* 2.2 (2017), pp. 300–307.

[32] R. M. M. Smeets, U. F. Keyser, N. H. Dekker, and C. Dekker. “Noise in solid-state nanopores”. *Proceedings of the National Academy of Sciences* 105.2 (2008), pp. 417–421.

[33] S. Pud et al. “Self-Aligned Plasmonic Nanopores by Optically Controlled Dielectric Breakdown”. *Nano Letters* 15.10 (2015), pp. 7112–7117.

[34] C. Plesa et al. *Fast Translocation of Proteins through Solid State Nanopores*. rapid-communication. 2013.

[35] A. J. Storm et al. “Fast DNA Translocation through a Solid-State Nanopore”. *Nano Letters* 5.7 (2005), pp. 1193–1197.

[36] S. Shekar et al. “Measurement of DNA Translocation Dynamics in a Solid-State Nanopore at 100 ns Temporal Resolution”. *Nano Letters* 16.7 (2016), pp. 4483–4489.

[37] L. de Vreede et al. “Wafer scale fabrication of fused silica chips for low-noise recording of resistive pulses through nanopores”. *Nanotechnology* (2019).

[38] V. Tabard-Cossa, D. Trivedi, M. Wiggin, N. N. Jetha, and A. Marziali. “Noise analysis and reduction in solid-state nanopores”. *Nanotechnology* 18.30 (2007), p. 305505.

[39] A. Balan et al. “Improving Signal-to-Noise Performance for DNA Translocation in Solid-State Nanopores at MHz Bandwidths”. *Nano Letters* 14.12 (2014), pp. 7215–7220.

[40] J. K. Rosenstein, M. Wanunu, C. A. Merchant, M. Drndic, and K. L. Shepard. “Integrated nanopore sensing platform with sub-microsecond temporal resolution”. *Nature methods* 9.5 (2012), p. 487.

[41] Y. Goto, I. Yanagi, K. Matsui, T. Yokoi, and K.-I. Takeda. “Integrated solid-state nanopore platform for nanopore fabrication via dielectric breakdown, DNA-speed deceleration and noise reduction”. *Scientific reports* 6 (2016), pp. 31324–31324.

[42] P. Karau and V. Tabard-Cossa. “Capture and Translocation Characteristics of Short Branched DNA Labels in Solid-State Nanopores”. *ACS Sensors* 3.7 (2018), pp. 1308–1315.

[43] W. H. Pitchford et al. “Synchronized Optical and Electronic Detection of Biomolecules Using a Low Noise Nanopore Platform”. *ACS Nano* 9.2 (2015), pp. 1740–1748.

[44] G. V. Soni et al. “Synchronous optical and electrical detection of biomolecules traversing through solid-state nanopores”. *Review of Scientific Instruments* 81.1 (2010), p. 014301.

[45] A. Balan, C.-C. Chien, R. Engelke, and M. Drndi. “Suspended Solid-state Membranes on Glass Chips with Sub 1-pF Capacitance for Biomolecule Sensing Applications”. *Scientific Reports* 5 (2015), p. 17775.

[46] Y. Li et al. “Harnessing Plasmon-Induced Ionic Noise in Metallic Nanopores”. *Nano Letters* 13.4 (2013), pp. 1724–1729.

[47] V. Dimitrov et al. “Nanopores in solid-state membranes engineered for single molecule detection”. *Nanotechnology* 21.6 (2010), p. 065502.

[48] Z. Roelen, J. A. Bustamante, A. Carlsen, A. Baker-Murray, and V. Tabard-Cossa. “Instrumentation for low noise nanopore-based ionic current recording under laser illumination”. *Review of Scientific Instruments* 89.1 (2018), p. 015007.

[49] N. Di Fiori et al. “Optoelectronic control of surface charge and translocation dynamics in solid-state nanopores”. *Nature nanotechnology* 8.12 (2013), p. 946.

[50] R. L. Fraccari et al. “High-speed detection of DNA translocation in nanopipettes”. *Nanoscale* 8.14 (2016), pp. 7604–7611.

[51] J. D. Uram, K. Ke, A. J. Hunt, and M. Mayer. “LabelFree Affinity Assays by Rapid Detection of Immune Complexes in Submicrometer Pores”. *Angewandte Chemie International Edition* 45.14 (2006), pp. 2281–2285.

[52] J. D. Uram, K. Ke, A. J. Hunt, and M. Mayer. “Submicrometer PoreBased Characterization and Quantification of AntibodyVirus Interactions”. *Small* 2.89 (2006), pp. 967–972.

[53] Y. Bellouard et al. “Femtoprint: A femtosecond laser printer for micro- and nano-scale systems”. *Conference on Lasers and Electro-Optics 2012 (2012), paper ATu3L.3*. Optical Society of America, 2012, ATu3L.3.

[54] R. An et al. “Ultrafast laser fabrication of submicrometer pores in borosilicate glass”. *Optics Letters* 33.10 (2008), pp. 1153–1155.

[55] Sima Felix et al. “Three-dimensional femtosecond laser processing for lab-on-a-chip applications”. *Nanophotonics* 7.3 (2018), p. 613.

[56] J. Larkin, M. Foquet, S. W. Turner, J. Korlach, and M. Wanunu. “Reversible Positioning of Single Molecules inside Zero-Mode Waveguides”. *Nano Letters* 14.10 (2014), pp. 6023–6029.

[57] A. Ivankin et al. “Label-Free Optical Detection of Biomolecular Translocation through Nanopore Arrays”. *ACS Nano* 8.10 (2014), pp. 10774–10781.

[58] D. Xia et al. “Rapid fabrication of solid-state nanopores with high reproducibility over a large area using a helium ion microscope”. *Nanoscale* 10.11 (2018), pp. 5198–5204.

[59] C. Liu et al. “An all-in-one nanopore battery array”. *Nature Nanotechnology* 9.12 (2014), pp. 1031–1039.

[60] M. Deserno and W. M. Gelbart. “Adhesion and wrapping in colloid- vesicle complexes”. *The Journal of Physical Chemistry B* 106.21 (2002), pp. 5543–5552.

[61] C. Hamai, T. Yang, S. Kataoka, P. S. Cremer, and S. M. Musser. “Effect of average phospholipid curvature on supported bilayer formation on glass by vesicle fusion”. *Biophysical journal* 90.4 (2006), pp. 1241–1248.

Chapter 3: 5-D Fingerprinting of Proteins

This chapter demonstrates that analysis of protein rotation and translocation inside nanopores can reveal the ellipsoidal shape, volume, charge, rotational diffusion coefficient, and dipole moment of single proteins simultaneously and in real time. This multiparametric information makes it possible to identify and quantify single proteins and protein complexes in a mixture and provides instantaneous, low-resolution shape information without the need for protein purification. Moreover, the kHz to MHz bandwidth of current recordings complements existing methods for structure determination, which provide higher resolution but are limited to gathering static information, in that it can interrogate and monitor protein dynamics and transient protein-protein complexes in solution. The work presented in this chapter represents a fundamental foundation on which subsequent chapters are built, and focuses largely on my personal contributions to a large project carried out by multiple students over multiple years.[1]

3.1 Introduction to Five-Dimensional Protein Fingerprinting

Methods to characterize, identify, and quantify unlabeled, folded proteins in solution on a single-molecule level do not currently exist.[2] If available, such methods would simplify routine protein analysis, enabling rapid and ultra-sensitive biomarker detection,[3], and allowing the analysis of personal proteomes. Furthermore, methods providing low-resolution approximations of shape could help to reveal the conformation of transient protein complexes or large assemblies that are not accessible by electron microscopy, NMR spectroscopy, X-ray crystallography, or small-angle X-ray scattering.[4] In this chapter, we demonstrate that interrogation of single proteins or protein-protein complexes during their passage through the electric field inside of a nanopore enables the determination of ellipsoidal approximations of their shape, as well as their volume, charge, rotational diffusion coefficient, and dipole moment.

Dipole moment has been largely neglected as a protein descriptor. Despite the pioneering work by Debye and Oncley,[5] the usefulness of this parameter for protein identification and its importance for concentrated protein solutions are often overlooked, and existing experimental methods – as opposed to computational methods – for determining protein dipole moments are cumbersome and unable to measure dipole moments on a single-molecule level. We propose, however, that dipole moment provides a powerful dimension for label-free protein analysis because its magnitude is widely distributed among different proteins, with absolute values ranging from 0 to 4,000 Debye.[6] Dipole moment may, therefore, be comparable to protein size for its usefulness in identification and would likely exceed the usefulness of protein charge, which has values that are less distributed and range from -40e to +40e. Moreover, the pharmaceutical industry is increasingly recognizing the importance of dipole moment for antibody formulations,[7] in part because subcutaneous injection of highly concentrated solutions of monoclonal antibodies, which are the fastest growing class of therapeutics, can be impractical due to high viscosity and aggregation resulting from dipole alignment.[7, 8] Measurements of antibody dipole moments could therefore provide

a criterion to select early candidates in the drug discovery process and reduce development costs.[8]

Interrogating single protein particles during their passage through a pore is simple in principle, [9, 10] and is explained in detail in Chapter 1. Of note for this chapter, the small volume of the nanopore transiently separates single proteins from other macromolecules in solution (see also Figure 1.1), enabling the interrogation and interpretation of the rotational dynamics of one protein at a time. For this reason, time-dependent modulations of ionic current as a single protein passes through a nanopore (Figure 3.1-D) can, under appropriate conditions, relate uniquely to the time-dependent molecular orientation of that protein as well as its ellipsoidal shape, volume, charge, rotational diffusion coefficient, and dipole moment. Several groups have recently considered, in qualitative terms, the effect of a protein's [11–17] or nanoparticle's shape when analyzing distributed ΔI signals [18] as well as the effect of a protein's dipole on its translocation through an α-hemolysin pore in the presence of an AC field.[19] The work presented here develops the theory for a quantitative understanding of the dependence of measured ΔI values on the ellipsoidal shape, dipole, and rotational dynamics of a protein inside a nanopore and makes it possible to estimate the volume, approximate shape, rotational diffusion coefficient, and dipole moment of non-spherical proteins in real time. We suggest that this ability to measure five parameters simultaneously on single proteins in real time has fundamental implications. For instance, analyzing individual proteins one-by-one may inherently mean that these proteins do not need to be purified for determining their approximate shape or the other four parameters (see Chapter 4 for more discussion). This consequence would be a paradigm shift compared to existing methods for determining the shape or structure of proteins, which either require purified, concentrated, or crystallized protein samples or cannot examine protein dynamics.

The main concept underlying the analysis introduced in this chapter is that rotation of a single non-spherical object during translocation through a cylindrical nanopore[20] modulates the current reduction through the pore and that these modulations can be used to

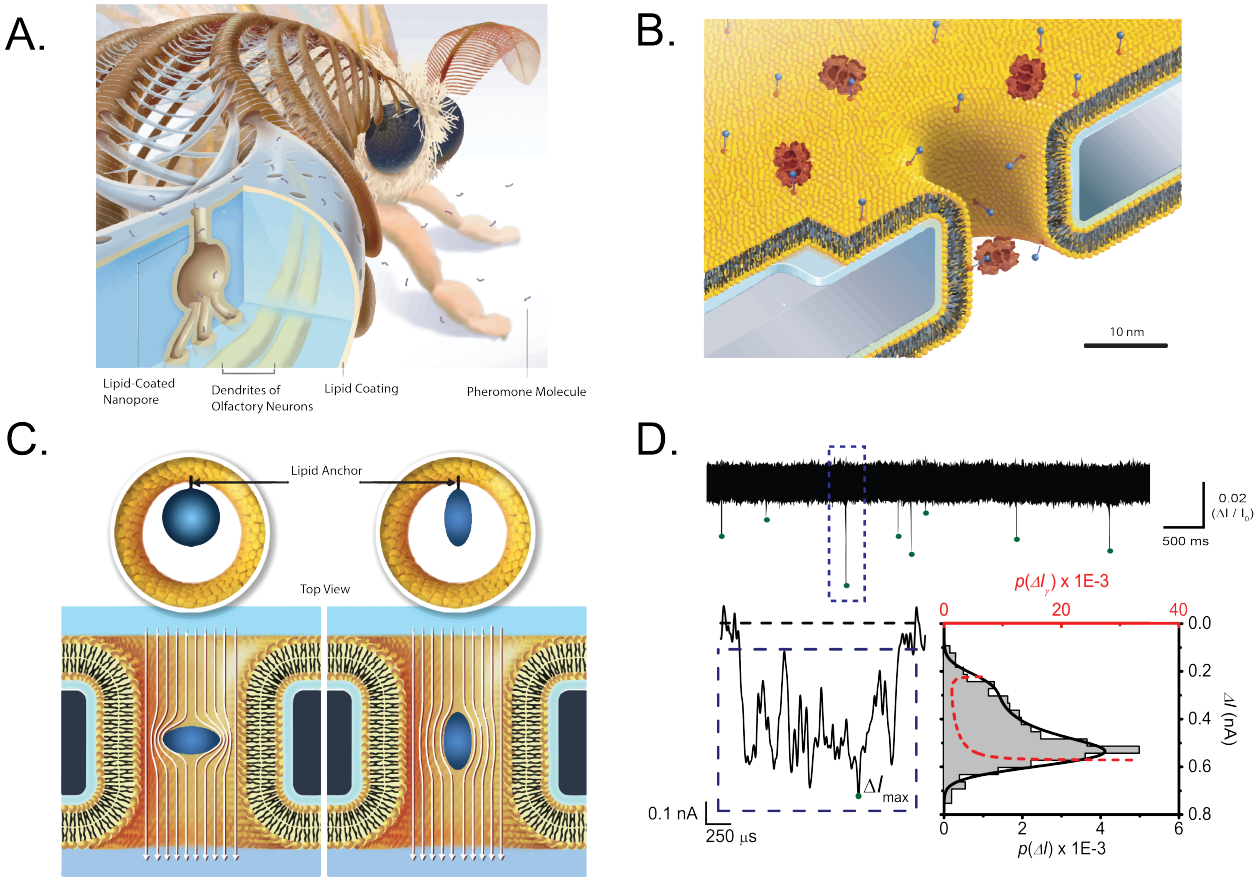


Figure 3.1. Lipid bilayer coatings facilitate tethering and allow for shape determination. A) Drawing showing a cross-section through a sensillum in the antenna of the silk moth *Bombyx mori*. These sensilla facilitate the capture, pre-concentration, and translocation of specific pheromones to the dendrites of olfactory neurons, allowing for tremendous sensitivity and specificity. Adapted with permission.[13] Copyright 2011, Springer. B) Scaled cartoon showing the binding of streptavidin protein to the head groups of lipids functionalized with the small molecule biotin. Adapted with permission.[13] Copyright 2011, Springer. C) Top and side views of a nanopore showing the two extreme orientations of an ellipsoidal particle transiting a nanopore. Adapted with permission.[1] Copyright 2017, Springer. D) Example current trace showing resistive pulses, with a "long" pulse outlined with a dashed blue line. Current values can be plotted as a histogram and subsequently analyzed with a convolution fitting procedure to estimate physical protein parameters.

determine the orientation, approximate shape, and volume of the object in the pore (Figure 3.1-D). Golibersuch [21] and others[22, 23] demonstrated both theoretically on ideal spheroids and experimentally on red blood cells that a crosswise orientation of an oblate (lentil-shaped object) or prolate (rugby ball-shaped object) distorts the electric field along a tube more dramatically than a lengthwise orientation (Figure 3.1-C). In the context of current recordings through a nanopore, this means that the particle-induced blockade of current, ΔI, is maximal when the spheroidal particle is in its extreme crosswise orientation and minimal in the extreme lengthwise orientation; orientations between these two extremes induce intermediate current reductions. The center box labeled "Convolve and Solve" in Figure 3.2 shows that all possible orientations of the protein within the pore are not equally probable; instead, the two orientations corresponding to electrical shape factors γ_{MIN} and γ_{MAX} describing extreme lengthwise and crosswise orientations are more likely than the other orientations because of the manner that the orthogonal projection of a non-spherical particle rotating within a nanopore varies with that particle's orientation relative to the electric field, described by the angle θ. The U-shaped probability distribution that describes this phenomenon implies that randomly rotating ellipsoidal proteins translocating through a nanopore should produce a distribution of ΔI values with two maxima, one corresponding to ΔI_{MIN} and one to ΔI_{MAX}. In contrast, spherical proteins, with a γ value of 1.5 that is independent of orientation (*e.g.* $\gamma_{MIN} = \gamma_{MAX}$), should produce normal distributions of ΔI values, as the probability distribution describing their orientations within the pore resembles a Dirac-δ function rather than a U-shaped distribution.

In addition to these orientation-dependent effects, the particle's volume and shape also affect the extent of electric field line distortion (Figure 3.1-C). For example, when comparing two oblates of equal volume in a cross-wise orientation, the particle that deviates most from a perfect sphere (*i.e.* the flatter oblate) distorts the field lines more dramatically than the rounder object. Conversely, in a lengthwise orientation, the flatter ($m << 1$) of these two particles distorts the field lines less dramatically than the rounder object. In other words,

particles with increasingly non-spherical shapes produce a more extreme ratio between the current blockage in their crosswise versus lengthwise orientation. For translocation of two particles with the same ellipsoidal shape but different volume, both the orientation-dependent minimal and maximal current reductions have larger magnitudes for the larger particle compared to the smaller one. Therefore, the magnitude of the current reduction depends on particle volume, while the ratio between the minimal and maximal current reduction depends on particle shape. The dependence of ΔI values on the shape and orientation of translocating particles has generally been neglected in nanopore-based protein characterization, thereby introducing uncertainty in measurements of volume for particles that are not perfect spheres. Considering these shape-dependent effects, as proposed here, will likely increase the accuracy of nanopore-based particle characterization, as most particles and proteins are not perfect spheres (see, for example, Figure 3.3, 4.12).

3.2 Results of Resistive-Pulse Analyses of Lipid-Tethered Proteins

In order to obtain time-resolved values of ΔI from the translocation of single proteins, we slowed down translocation by tethering proteins to a lipid anchor that was embedded in the fluid lipid bilayer coating of the nanopores, either by incorporating lipids into the bilayer with headgroups containing primary amines to react with a bi-functional NHS PEG-based linker, by incorporating lipids into the bilayer with headgroups containing a protein binding complement (Figure 3.1-B), or by investigating proteins that contain a molecule which actively incorporates into a lipid bilayer (*e.g.* GPI-AchE).[1, 13] In this way, the speed of protein translocation was dominated by the approximately 100-fold higher viscosity of the lipid coating compared to that of the aqueous electrolyte. In addition, we maximized the possibility that the proteins could rotate and sample all orientations in the nanopore by employing long and flexible PEG tethers.[1] Finally, the lipid coating minimized non-specific

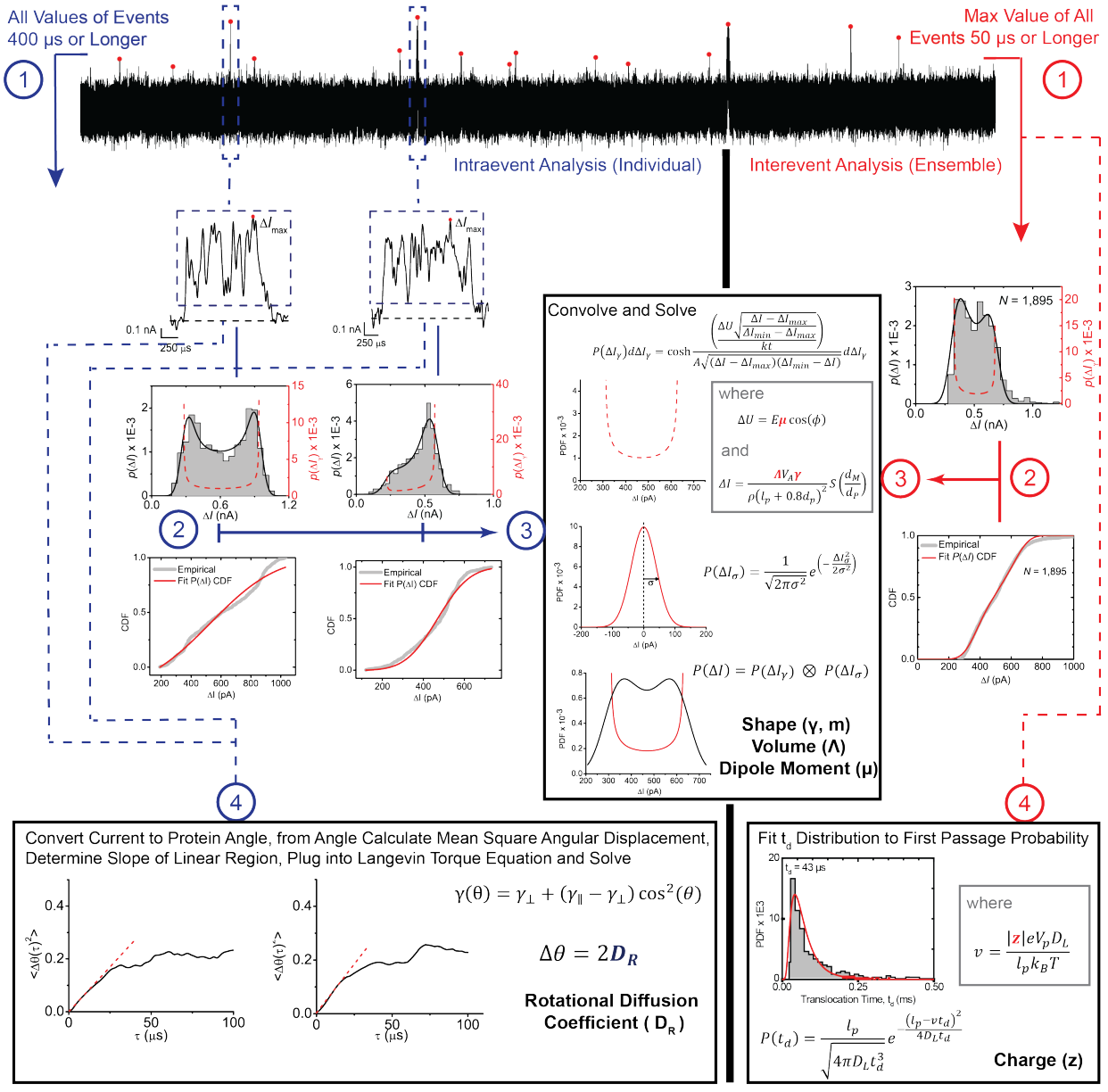


Figure 3.2. Process flow of the analysis of resistive pulses, showing both population-based and individual event proceudres. The procedure of analyzing resistive pulses begins with a peak finding algorithm first introduced by [24], which extracts resistive pulses from baseline noise using a threshold of $5\times$ the standard deviation of the baseline. A single maximum value is taken from *all* events longer than 50 µs or longer (red pathway), while all measured values are taken from events longer than 400 µs. These values are then used to estimate five different protein parameters, labeled in bold font, from a single experiment. Adapted with permission.[1] Copyright 2017, Springer.

interactions between proteins and the pore wall,[13] thus enabling extraction of quantitative data on Brownian rotational and translational dynamics of proteins while they were in the pore. For instance, we took advantage of the resulting translocation times to determine the net charge of all ten proteins and found a strong correlation between the charge from nanopore experiments and reference values for the charge of each protein (Pearson correlation coefficient r = 0.95, see Figure 3.3-C).

To determine the approximate shape and volume of proteins, we developed two strategies based on the theory developed by Fricke[25] Velick and Gorin,[26] and Golibersuch[21] (Figure 3.2 red and blue pathways). Both strategies approximate the shape of proteins as an ellipsoid of rotation and have different strengths and weaknesses. The first strategy estimates shape and volume from distributions of maximum ΔI values from many translocation events that were obtained from a pure protein solution. In other words, only the single maximum value of ΔI from each resistive pulse is used for analysis. Maximum ΔI values have been employed in almost all nanopore-based resistive pulse analyses of protein volume to date combined with the assumption of a perfectly spherical particle shape (i.e. $\gamma = 1.5$, $m = 1$), thereby foregoing the opportunity to evaluate protein shape. In contrast, Golibersuch showed by examining red blood cells that maximum ΔI values could also be used to approximate the shape of particles.[21] Here, we applied this concept for the first time to proteins transiting nanopores. An advantage of using maximum ΔI values to estimate protein shape and volume is that the ratio between the extreme values of current reduction, ΔI_{MIN} and ΔI_{MAX}, is relatively insensitive to deviations in pore geometry from a perfect cylinder. A disadvantage of this approach is that ellipsoidal shape and volume cannot be determined from a single translocation event because only the maximum ΔI value from each translocation event is analyzed and thus many translocations are required to sample all possible electrical shape factors (see Section 4.4.4 for more discussion). It is important to note that population-based analysis **depends** on inadequate sampling of protein orientations within the pore in order to capture the full distribution of translocation values. In other words, if we

were to perform population-based analysis only on translocation events of proteins that adequately sampled all orientations within the nanopore, we would select only the value for the maximum orientation of that protein, vastly overestimating its volume and underestimating its shape.

Determining the shape and volume of spheroids from distributions of maximum ΔI values proceeds in three steps (Figure 3.2 red pathway numbers 1-3). First, an algorithm detects resistive pulses from the translocation of hundreds to thousands copies of the same protein and determines the maximum amplitude of the current modulation, ΔI, with respect to the baseline current for each pulse (Figure 3.2 red number 1). As predicted theoretically, the resulting distribution of maximum ΔI values is bimodal for non-spherical proteins (Figure 3.2). Second, in order to circumvent binning effects encountered with probability distributions, the experimentally determined distribution of ΔI values is converted to an empirical cumulative density function, CDF (Figure 3.2 red number 2), and fit iteratively with an equation that describes the variation in ΔI due to rotation of proteins with non-spherical shape (Figure 3.2 red number 3). We refer to this equation as the convolution model because it accounts for broadening of the ΔI distribution due to convolution of the true signal with noise (Figure 3.2 center box, more details in Chapter 4) and for bias toward either the crosswise or lengthwise orientation as a result of the electric-field-induced torque on the protein's dipole moment.[5] The bias in a distribution of maximum ΔI values, however, may also be affected by other factors than the dipole moment (as discussed in Chapter 4), which are all accounted for by the same fitting parameter. The values of ΔI_{MIN} and ΔI_{MAX} returned by the fitting procedure reflect the two extreme orientations of the protein (Figure 3.1-C). Third, based on the direct proportionality between ΔI and γ, and the geometric relationship between γ and the length-to-diameter ratio m of a spheroid (see equations in Chapter 4), we determine the shape and volume that agree best with the experimental distribution of ΔI values for the protein.

Figure 3.3-D shows the spheroidal approximation of the shape of ten different proteins

compared to the respective crystal structure for each protein, illustrating that this analysis yields excellent estimates of protein shape, particularly for proteins that closely resemble a spheroid. Figure 3.3-A,B for instance, shows that the volume and length-to-diameter ratio values agree well with the expected reference values; the average deviation of both parameters is less than 20% (see Supporting Information of [1]). These results also show that two proteins with a similar molecular weight and volume but different shape are clearly distinguishable by this analysis; for instance, compare the ellipsoids determined for the IgG_1 antibody and GPI-AChE in Figure 3.3-D.

Independent from these experimental results, we confirmed the accuracy of this approach for ellipsoidal shape and volume determination using simulated data that was generated from the theory of biased one-dimensional Brownian diffusion and convolved with current noise. Fitting the simulated data with the convolution model, just as with the experimental data, returned values of length-to-diameter ratio and volume that were in excellent agreement with the input parameters (see Supporting Information of [1]).

Compared to other methods for determining the shape and volume of proteins in aqueous solution such as solution-state NMR spectroscopy, analytical ultracentrifugation, and dynamic light scattering, the nanopore-based approach is faster (seconds to minutes), requires smaller sample volumes (1 μL) and lower protein concentrations (pM to nM), and may see improved performance as the size of proteins or protein complexes increases due to the concomitant potential increase in signal-to-noise ratio. While the resolution of shape is significantly lower than that of NMR spectroscopy for small proteins (<80 kDa), it is higher than the resolution of analytical ultracentrifugation and dynamic light scattering. In addition, although the limited time-resolution of currently available amplifiers requires tethering proteins to the lipid coating (a reaction that occurs *in situ* on the nanopore chip), the nanopore-based approach does not require extensive modification of pure proteins by isotope labeling as it is the case for protein NMR spectroscopy.

As opposed to this first strategy, which analyzes maximum ΔI values from many translo-

cation events, the second strategy estimates the length-to-diameter ratio and volume of proteins from individual resistive pulses by analyzing all current values from the beginning to the end of single translocation events, $\Delta I(\mathrm{t})$, in a stand-alone manner (Figure 3.2 blue pathway numbers 1-3). This analysis relies on a single translocating protein to rotate and sample virtually all orientation-dependent γ values such that the resulting single-event, or intra-event, ΔI distribution reveals ΔI_{MIN} and ΔI_{MAX} and thereby the protein's spheroidal shape approximation and its volume from an individual translocation event. The advantage of this strategy, in addition to estimating length-to-diameter ratio and volume from the translocation of a single protein, is that it can also determine the protein's rotational diffusion coefficient and dipole moment from individual resistive pulses based on orientation-dependent modulations in current over time. In fact, estimates of all four parameters can be determined and updated in real time as a single protein travels through the pore. The disadvantage of this simultaneous multiparametric analysis of single molecules is that the analysis is limited to resistive pulses with durations of at least 400 μs to ensure that each protein resides sufficiently long in the pore to sample the full range of electrical shape factors (under the conditions used in this work, approximately 10% of events exceeded this threshold). We chose this duration based on the mean-square angular displacement equation that predicts a protein will sample all possible orientations in less than 400μs, on average, as long as its rotational diffusion coefficient exceeds 3,000 $\mathrm{rad}^2\ \mathrm{s}^{-1}$, which was the case for the tethered proteins examined here (see Chapter 4, specifically Figure 4.2). Other disadvantages of this analysis include that it is more sensitive to deviations of the pore geometry from a perfect cylinder than the multi-event analysis of maximum ΔI values (see Section 3.4.3) and that the analysis of individual resistive pulses is associated with relatively high uncertainty as with other single-molecule measurements.

We find that the individual-event distributions from translocations of individual proteins retain their key features (*e.g.* minimal and maximal ΔI values) although the current recordings are smoothed due to filtering (see, again, Chapter 4). The median protein ellipsoidal

shapes obtained from this analysis are in reasonable agreement with their crystal structure,[1] although the analysis of maximum ΔI values yielded more accurate estimates. With regard to the robustness of each stand-alone single molecule measurement, more than half of all measurements yielded values of the length-to-diameter ratio and of volume that were within $\pm$ 35% of the median value. Based on this result and the expectation that further improvements are possible, we propose that intra-event analysis has the potential to yield good estimates of ellipsoidal shape and volume of single proteins from individual translocation events. Moreover, this strategy of analyzing individual-event ΔI distributions introduces, to the best of our knowledge, the only existing method for estimating, in real time, the ellipsoidal shape and volume of single protein molecules in solution. Shape and volume determination on a single particle level is particularly advantageous for analysis of samples with large heterogeneity in size and shape (*e.g.* amyloids, see Chapter 1); ensemble methods such as dynamic light scattering are poorly suited for such samples as they blur potentially important heterogeneities within a population into a single output value. Other techniques for analyzing the shape and volume of single proteins such as cryo-electron microscopy and atomic force microscopy either require freezing or surface immobilization that fixes the orientation of the proteins; therefore, these methods are not particularly well suited to tracking protein dynamics.

Figure 3.2 shows that monitoring the time-dependent modulations of ΔI while a single particle moves through a nanopore makes it possible to measure its rotational diffusion coefficient, D_R, by tracking its rotation over short time scales and therefore over small fluctuations in angle. We carried out this analysis in three steps by transforming the intra-event current signal into an angle (*i.e.* orientation) versus time curve, calculating the mean-square angular displacement over various time intervals, τ, and fitting its initial slope with a model for rotational diffusion about a single axis (Figure 3.2 Bottom Left Box). Figure 3.3-F shows that the most-probable D_R values for tethered proteins obtained from many individual-event analyses of single resistive pulses were strongly correlated with the expected

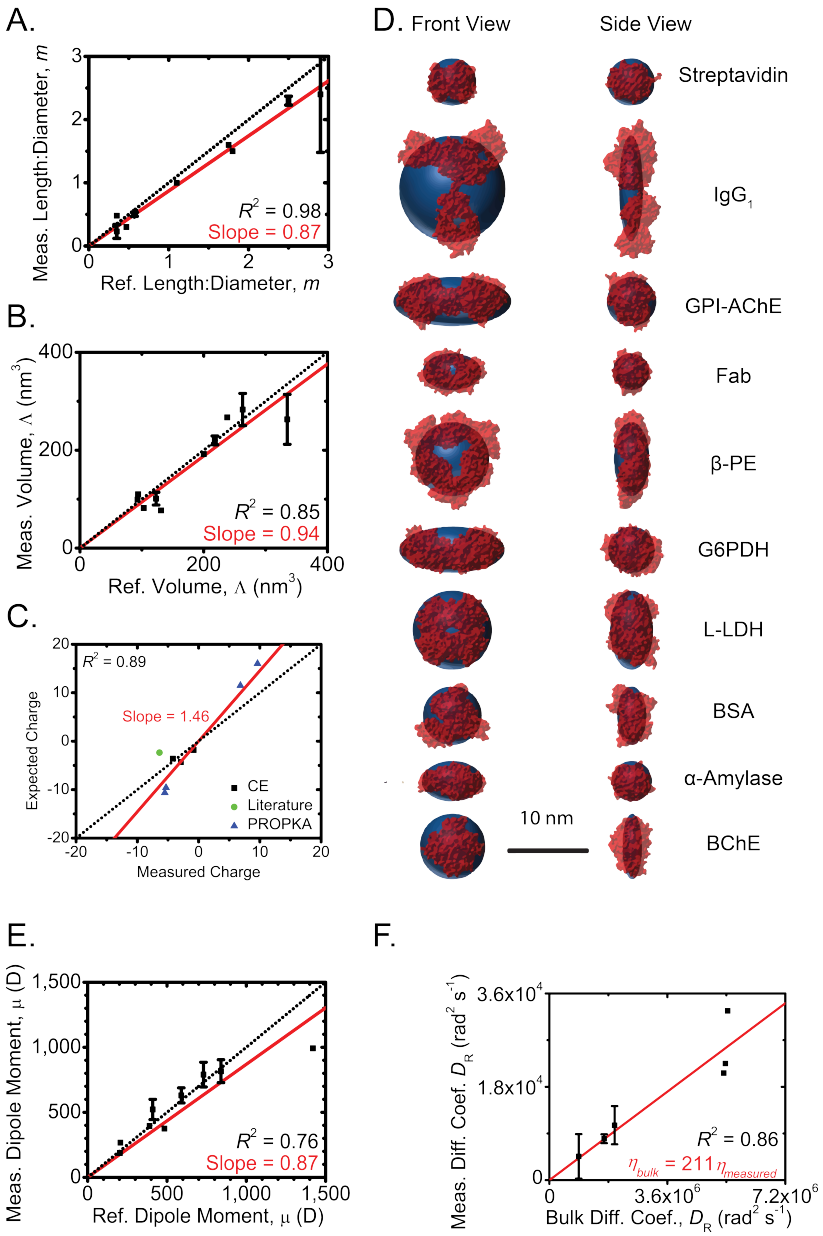


Figure 3.3. Main results of 5D analysis across all proteins. A) Comparison of the measured length-to-diameter ratios, m, of all proteins determined through population-based analyses with expected reference values. B) Comparison of the measured volumes, Λ, of all proteins determined through population-based analyses with expected reference values. C) Comparison of the measured net charge of all proteins determined through population-based analyses (Figure 3.2 red pathway number 4) with expected reference values. D) Comparison of the approximate ellipsoidal shape of ten proteins as determined by analysis of resistive pulses (blue ellipsoids) with crystal structures from the Protein Databank in red (streptavidin, 3RY1; bovine serum albumin (BSA), 3V03; anti-biotin antibody fragment (FAB), 1F8T;α-amylase, 1BLI; anti-biotin immunoglobulin G1, 1HZH; glucose-6-phosphate dehydrogenase (G_6PDH), 4EM5; L-Lactate dehydrogenase (L-LDH), 2ZQY; butyrylcholineesterase (BChE), 1P01; β-phycoerythrin, 3V57; and GPI-anchored acetylcholinesterase, 3LII). E) Comparison of the measured most-probable dipole moment of all proteins determined through individual-event analyses (Figure 3.2 blue pathway numbers 1-3) with expected reference values. F) Comparison of the measured most-probable rotational diffusion coefficient of all proteins determined through individual-event analyses (Figure 3.2 blue pathway number 4) with expected reference values. Adapted with permission. [1] Copyright 2017, Springer.

values of D_R in bulk solution (Pearson's r = 0.93). As expected, the presence of the lipid tether and close proximity of the proteins to the bilayer coating reduced D_R significantly; this tether-induced attenuation of rotation was consistent with an apparent viscosity increase by a factor of 211 compared to the viscosity in bulk solution (see Supporting Information of [1]). This value is in excellent agreement with fluorescence polarization measurements of GPI-anchored AChE by Yuan and Axelrod, which revealed that the rotational diffusion coefficient of tethered AChE is 199 times smaller than its expected value in bulk solution.[27] For analyzing the rotational dynamics of proteins in real time as presented here, this tether-induced reduction of D_R was beneficial as it enabled changes in protein orientation to be resolved in time (see Chapter 4 for a direct comparison of time-resolution of tethered and untethered proteins).

With regard to the robustness of these measurements, we found that, on average, the relative standard deviation of the most probable value of D_R from distributions of measured single molecule values was 46% from experiment-to-experiment or day-to-day; however, as is typical for many single molecule measurements, the variation from event-to-event was large with a mean absolute deviation of 403%.

To the best of our knowledge, this approach is the fastest method (sub-millisecond) for estimating the rotational diffusion coefficient of single proteins in solution, albeit with considerable uncertainty at this initial stage of the technology; it is also the only non-fluorescent method to determine D_R.[7] While the requirement for tethering proteins precludes direct determination of the bulk value of D_R by this approach, the good correlation shown in Figure 3.3-F demonstrates that bulk D_R values can be estimated from the measured D_R values of tethered proteins.

Monitoring the rotational dynamics of proteins at long time scales and hence over large changes in angle shows theoretically and experimentally that proteins with a dipole moment do not rotate randomly when they experience the MV m^{-1} electric field intensity inside the pore; instead, the proteins undergo biased Brownian rotation due to electric-field-induced

torque on their dipole moment.[5] Quantifying this bias in orientation by fitting the intra-event ΔI distribution from an individual resistive pulse with the convolution model made it possible to calculate a protein's dipole moment by considering the potential energy landscape of a dipole in an electric field (see Supporting Information of [1]). In this analysis, the fitting parameter μ of the convolution model is equivalent to the dipole moment and therefore yields its magnitude. In contrast, in the population-based analysis of maximum ΔI values, the same parameter encompasses additional factors including inadequate sampling of protein orientation, as discussed before, and hence precludes estimation of dipole moment.

Figure 3.2-E shows that the most probable values of dipole moment from this nanopore-based analysis agree well with expected values; the average deviation is less than 25%. With regard to the robustness of this method from experiment-to-experiment or day-to-day: the relative standard deviation of the most probable value from distributions of measured single molecule values was 12% and compares well with dielectric impedance spectroscopy measurements; however, as is typical for many single molecule measurements, the variation from event-to-event was large with a mean absolute deviation of 227%.

While the uncertainty in each stand-alone single molecule measurement of dipole moment will have to be reduced in order to realize the full potential of this approach (see Chapter 6 for details), this technique introduces the first experimental method for determining the dipole moment of individual proteins in solution. To this end it exploits a fundamental advantage of single molecule techniques, namely that statistical fluctuations of one particle are easier to interpret and to compare with theoretical models than it would be of an ensemble of particles. An additional advantage of this single-particle analysis is that it can estimate dipole moments in real time and requires only pico- to nanomolar concentrations of proteins. In contrast, the standard method for measuring dipole moment, dielectric impedance spectroscopy, requires micromolar protein concentrations and significantly larger sample volumes.[7]

To extend the technique beyond biophysical characterization of pure protein solutions, we applied it to the identification and quantification of different proteins in a well-defined

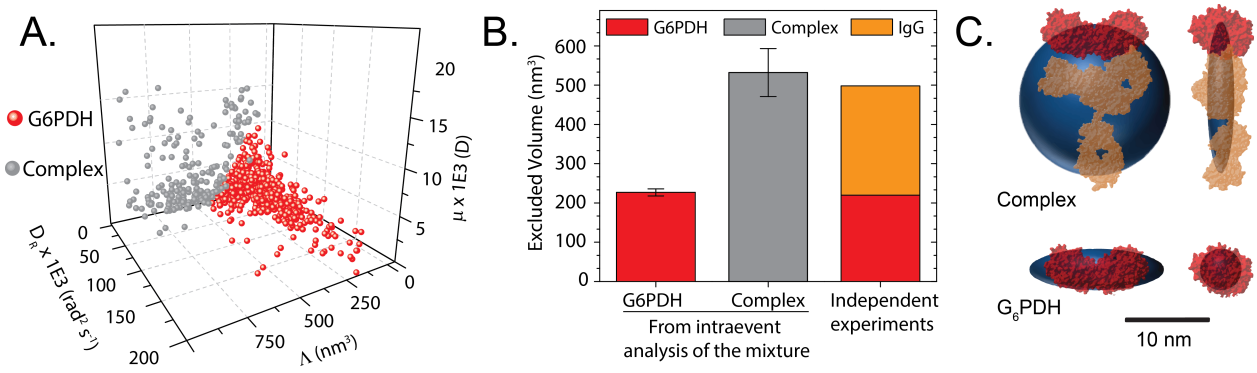


Figure 3.4. Results of experiment with heterogeneous protein mixture. A) Values for the volume, rotational diffusion coefficient, and dipole moment determined from individual events. The k-means clustering algorithm in MATLAB classified single events as corresponding to G_6PDH (red points) or G_6PDH / IgG_1 (grey points). This individual-event classification indicated that 28% of events were associated with the complex, which is approximately the same proportion of events as determined by population-based analysis. B) The volume of G_6PDH and G_6PDH / IgG_1 complex from clustering individual-event analyses agree well with those determined for the pure proteins. C) The volume and ellipsoidal shape approximation of G_6PDH and G_6PDH / IgG_1 complex, determined through individual-event analysis and shown in the same fashion as the proteins in Figure 3.3-D. Adapted with permission.[1] Copyright 2017, Springer.

mixture (Figure 3.4). We repeated our original experiment with glucose-6-phosphate dehydrogenase (G_6PDH), but this time added a polyclonal anti-G6PDH IgG_1 antibody to the recording buffer after cross-linking. This procedure resulted in a mixture of three distinct species; anchored G_6PDH, anchored G_6PDH bound to IgG_1, and unanchored IgG_1 in bulk solution. Because the lipid coating prevents adhesion, we assumed that nearly all (99.9+%) of the translocation events resulting from freely translocating IgG_1 would fall short our 400 μs threshold for individual event analysis and thus were discarded. We analyzed the remaining "long" events individually, and used a clustering algorithm (Figure 3.4-A) to classify each one as G_6PDH (red) or G_6PDH/IgG_1 complex (grey). This analysis produced excellent estimates of both species (Figure 3.4-B,C), and the experiment serves as a proof-of-concept for future nanopore-based analysis of heterogeneous protein mixtures.

3.3 Conclusion and Future Prospects of Five-Dimensional Analyses

The work presented here extends the potential of nanopore-based DNA sequencing to five-dimensional characterization and fingerprinting of proteins and protein complexes. Unlike standard bulk methods, this technique interrogates individual proteins one-at-a-time by taking advantage of the molecular scale volume of the nanopore. This zeptoliter volume (10^{-21} L) temporarily separates individual proteins from other proteins in the bulk solution and inherently forms a focal point for measuring protein-induced changes in ionic conductance with exquisite sensitivity. Hence, only the protein residing in the nanopore modulates the electrical signal. This arrangement, together with the lipid coating that minimizes non-specific interactions and slows down the translocation and rotation of lipid-anchored proteins, enables interrogation of the translational and rotational dynamics of single proteins and uses those dynamics to determine their ellipsoidal shape, volume, charge, rotational diffusion coefficient, and dipole moment. Additionally, we showed that this approach has advantages in distinguishing a protein from its complex with another protein in a binary mixture.

Based on the success in commercializing nanopore-based DNA sequencing,[28–30] we predict that improvements to the approach introduced here will increase the potential of nanopore-based protein characterization. For instance, the individual-event analysis likely suffers from deviations in the pore geometry from a perfect cylinder. These irregularities, which are a consequence of the current state-of-the-art fabrication methods, affect the local resistance along the lumen of the pore and hence affect the precision with which the maximum and minimum ΔI value can be determined. Alternative fabrication methods may produce pores that are almost perfectly cylindrical and should therefore minimize possible artifacts from this source of error, as discussed in Chapter 2, but to take advantage of the five-dimensional characterization presented here, these pores must also be amenable to a lipid bilayer coating. In addition, the recent development of integrated CMOS current

amplifiers,[31] which can be produced in parallel to record from hundreds of nanopores simultaneously while reaching at least ten-times higher bandwidth and three-times higher signal-to-noise ratio compared to the amplifier used in this work,[31] will increase the throughput and improve the precision and accuracy of determining the rotational dynamics of proteins on their journey through the pore. Such fast amplifiers may eliminate the need for tethering proteins to lipid anchors [32] while their improved signal to noise ratio combined with the recent development of low-noise nanopore chips[33] will likely reduce the uncertainty in each determined parameter (see Chapter 2,4).[31, 34] Furthermore, computational approaches that can model proteins with shapes more complex than simple spheroids (see Chapter 6) may increase the resolution of shape determination, while the capability to monitor current modulations with MHz bandwidths may enable measurements of transient changes in protein conformation and folding as well as the ability to estimate the shape of short-lived protein complexes whose structure and dynamics are not easily accessible by existing techniques.

We suggest that the ability to measure five parameters simultaneously on single proteins in real time, including parameters that can otherwise not be obtained on a single molecule level, has transformative potential for the analysis and quantification of proteins as well as for the characterization of nanoparticle assemblies. For instance, fast protein identification and quantification in complex mixtures is an unsolved problem.[3] Despite its tremendous capabilities, mass spectrometry has currently limited throughput and is not broadly applicable to meet demand for routine protein analysis.[2, 3] Two-dimensional gel electrophoresis remains one of the most important techniques for analyzing complex protein samples, but its reproducibility is limited, and the method is slow and semi-quantitative. We propose that multi-dimensional analysis and fingerprinting of single proteins in nanoscale volumes may be one alternative. The work presented here is only a first step in this direction; if improvements similar to the ones made in nanopore-based DNA sequencing can be realized, we think it has the potential to replace methods such as 2-D gel electrophoresis by providing additional protein descriptors, improved quantification, increased sensitivity, reduced analysis time, and

lower cost. Such a capability may ultimately make it feasible to characterize and monitor an individual's proteome with significant implications for personalized medicine. Multiparameter protein characterization on a single molecule level may also reveal biochemically- or clinically-relevant static or dynamic heterogeneities, such as sub-populations of phosphorylated proteins, that are often hidden in ensemble measurements.[35] Moreover, real-time identification of single proteins might ultimately enable single molecule sorting in a fashion analogous to cell sorting.

Finally, this chapter focused on one of the most relevant and challenging applications of nanoscale shape approximation, namely the characterization of single proteins. The same approach may, however, apply to particles such as DNA origami, synthetic nanoparticles, and nanoparticle assemblies, whose characterization on a single particle level is important because they are typically more heterogeneous than proteins and because their charge, shape, volume, and dipole moment affect their assembly characteristics and function.

3.4 Supplementary Notes and Figures

3.4.1 Protein-Based Calibration of Nanopore Length

One question that arises when considering the analysis used to determine parameters from resistive pulses of proteins is: can we use the known size and ellipsoidal shape of a protein to calibrate geometric parameters of a particular nanopore, and then use that same nanopore with its calibrated geometric parameters to improve estimates of physical parameters of other proteins? To do this experiment, we calibrated a nanopore chip by translocating streptavidin (SA) protein whose molecular shape can be approximated as spherical (i.e. shape $m \approxeq 1$ and electrical shape factor $\gamma \approxeq 1.5$) and then we determined the volume of BSA using the same pore with a newly calibrated length. Before calibration, we used our usual approach for determining l_p from the conductance of the uncoated nanopore using equation 2.1. From this standard approach, we determined the following geometry for this particular pore: r_p = 12.8 nm (from the TEM image of this pore using the area-equivalent pore radius $r_p = (r_1 * r_2) * 0.5$) and l_p = 16 nm from its conductance before coating by solving for total resistance of the system. The standard, uncalibrated analysis with this pore determined a volume of 77 nm^3 for SA and in a second experiment, after Piranha cleaning, the same pore using the same r_p and l_p, determined a volume of 104 nm^3 and a length-to-diameter ratio m of 0.51 for BSA. Next, we calibrated the pore by keeping r_p constant and adjusting l_p so that the determined volume for SA was equal to 95 nm^3, which is the average of all reported volumes we could find for SA (range 94-105 nm^3). With the resulting calibrated, new $l_{p,eff}$ value of 19.3 nm, we then Piranha cleaned this pore again and ran a new translocation experiment with BSA. The results of the calibrated analysis for BSA with r_p = 12.8 nm and $l_{p,eff}$ = 19.3 nm now were length-to-diameter ratio m = 0.51 and volume Λ = 122 nm^3. For comparison, the average reported value we could find for BSA in the literature is 121.5 nm^3. In summary, in this particular experiment, our uncalibrated, standard analysis underestimated the volume of SA by 19% and underestimated the volume of BSA by 14%,

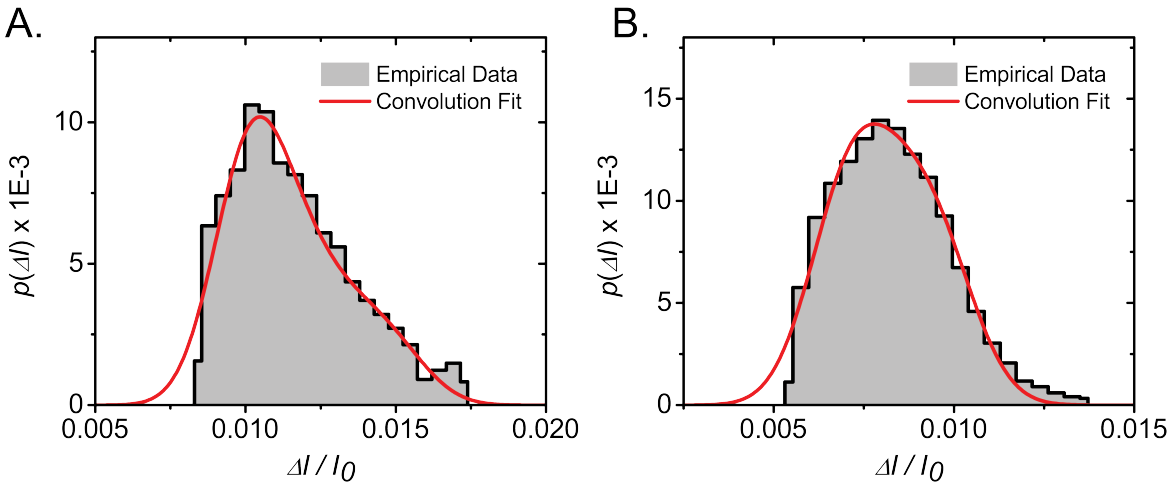


Figure 3.5. Results of calibration experiment for Streptavidin and BSA. A) Distribution of maximum current blockade values (population-based) for BSA protein, (grey) and fit from the convolution model after calibration. (red) B) Distribution of maximum current blockade values (population-based) for streptavidin (grey) and associated fit (red). In the convolution fit, we varied the nanopore length, l_p and assigned a constant length-to-diameter ratio $m = 1$ until we achieved the reference volume of 95 nm^3. We applied this new length, $l_{p,eff}$, to analyze the data for BSA shown in (A) in order to improve our estimate for volume. Adapted with permission.[1] Copyright 2017, Springer.

while the same pore after obtaining $l_{p,eff}$ from calibration with SA, determined the exact average of the reference volumes published for BSA.

This result shows that, not surprisingly, a pore calibration improves the accuracy of determining protein volumes. This analysis also shows, however, that a pore whose area-equivalent diameter is known from TEM images (as is the case for all pores that we used throughout Chapter 3) and whose length is determined from its conductance (as we do with each pore before coating with a lipid bilayer), is able to determine protein volumes within an average uncertainty of $\pm$ 20% without calibration (consistent with this result, we stated earlier in this chapter that the average error in length-to-diameter ratio m and volume is smaller than $\pm$ 20%). This result is again unsurprising once the cylindrical shape of the sensing zone of the pore is established because the difference between the determination of l_p by our standard conductance measurement compared to determination of $l_{p,eff}$ by calibration is that our standard method could be viewed as a pseudo one-point calibration by measuring baseline current through the nanopore (which could also be thought of as translocating a

particle with zero volume) compared to real one-point calibration that measures ΔI from the translocation of a spherical, lipid-tethered particle with well-known volume such as the SA protein.

Here, we followed the protocol of previous groups who used the volume of BSA in its natively folded state, implicitly assuming that BSA translocates in its native state at applied potentials below 0.3 V. We used an electrolyte with a similar ionic strength and made the same assumption that proteins transit in their natively folded state by making sure to limit the applied potentials to $\sim$ 0.12 V and below. As mentioned previously, the majority of nanopore-based protein characterization work has made the same assumption of natively folded proteins and obtained good agreement between their determined protein volumes and reference volumes of the folded state. Therefore, the standard approach without calibration, presented throughout this chapter and outlined above, led to reasonable agreement between estimated volumes of 11 proteins and associated reference values. The error bars in Figure 3.3-A show this agreement with the variations for measuring the same protein either on different days with the same pore or with different pores on different days. Without calibration, the method leads to an average uncertainty of $\pm$ 20%. While the calibration is hence not required for good agreement with reference values, the new calibration experiment shown above provides the opportunity to further improve the accuracy of the approach.

Calibrated and uncalibrated nanopore-analyses both have advantages and disadvantages. We value the fact that straight-forward physics equations based Ohm's law combined with knowledge of the pore shape (Figure 2.2-C), area-equivalent diameter (from TEM images), and length (from IV curves of the uncoated pore), make it possible to determine protein volumes on average within $\pm$ 20% of reference values without the need for calibration, because this approach circumvents an extra step. On the other hand, the advantage of calibration is that it improves estimation accuracy; its disadvantage is that it requires an extra step/experiment, which further limits the already-limited lifetime of nanopores that are repeatedly cleaned with stringent Piranha solution.

3.4.2 Evaluation of Size-Distribution of Gold Nanoparticles

We also attempted to calibrate a nanopore using synthetic nanoparticles, out of a concern that proteins may be denaturing and are thus unsuitable for such measurements. To do so, we first ordered 20 nm non-functionalized NanoXact Silica nanoparticles from the company nanoCompsix (Lot JEA0218) with the following specifications: diameter (TEM): 22.5 ± 2.8 nm; coefficient of variation (CV): 12.8%; and zeta potential -23 mV.

We then suspended an appropriate concentration of these particles in our standard 2 M KCl recording electrolyte and analyzed this suspension by DLS. Unfortunately the DLS results showed that these particles aggregate at pH 7.4 although at this pH most of the silanol groups on the silica are expected to be deprotonated (because their pKa is $\leq$ 2). Consistent with aggregation, our attempt to characterize these particles with a lipid-coated nanopore resulted in the absence of translocation peaks at both polarities of the applied 0.1 V potential. Aggregation of nanoparticles in aqueous solutions with high ionic strength such as our recording electrolyte has been observed by many other groups before.[1]

We then obtained gold nanoparticles coated with a thiol-PEG5000 terminated with a carboxylic acid group. Figure 3.6 shows a TEM image of these particles as well as their size distribution by TEM (red) in comparison to the size distribution we determined from resistive pulses through a lipid coated nanopore using our standard, uncalibrated analysis (grey). DLS indicated that these particles did not aggregate in our 2 M KCl electrolyte.

The most striking result from Figure 3.6 is that the most frequently observed diameters determined by TEM analysis and nanopore analysis are between 14 and 15 nm and therefore very similar with both techniques. In both cases, we assumed a spherical shape for the comparison of sizes. This excellent agreement in size, once again, shows that our standard approach of determining the length of the sensing zone of the nanopore by its conductance together with a TEM image of the pore yields size estimates that are within ± 20% of the expected reference values; as argued before, calibration was not necessary for this type of accuracy.

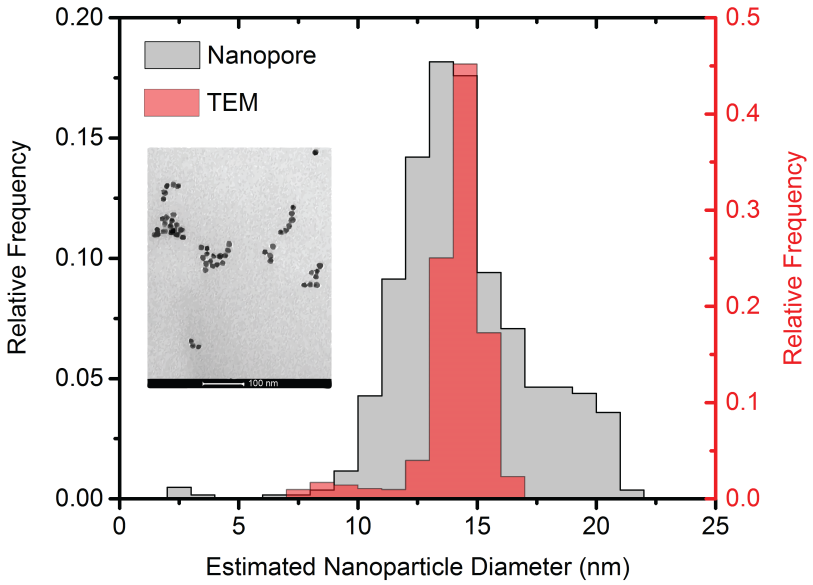


Figure 3.6. Results of size distribution analysis of gold nanoparticles using both nanopore and TEM-based analyses. The grey histogram represents an estimated nanoparticle diameter determined from individual-event analysis of translocation events, assuming a length-to-diameter ratio of $m = 1$, while the red histogram represents the equivalent diameter determined from area-based measurements of TEM images (inset). Note that while the distribution shapes are inherently different, the peak values show very good agreement.

The second interesting result is that the distribution from the nanopore analysis is somewhat wider than the distribution from TEM analysis. To reveal the exact reason for this difference is in our view sufficiently complex for a stand-alone publication in its own right. Speculative reasons are: 1) the assumption of a spherical particle shape leads to uncertainty of resistive-pulse based volume determination of those particles that deviate from a perfect sphere as Golibersuch [21] and we (above) have argued. Because close inspection of the TEM images in the inset of Figure 3.6 illustrates that most particles are not perfect spheres, this uncertainty affects the determined size distribution by nanopore analysis due to its volumetric measurement principle, while the analysis of the TEM image is not affected to the same extent since it is straightforward to envelope a particle on a 2D image with a circle. 2) Another important consideration is that these PEG-coated gold nanoparticles translocated the pore in an untethered fashion. This consideration means that, in contrast to all our protein translocation experiments in this chapter, nanoparticles were not confined to a distance

of maximally ∼1.5 nm from the wall of the nanopore but rather untethered particles were able to sample all distances from the pore walls during their translocation. This is a critical difference that we describe in detail in Chapters 4 and 5. Qin *et al.* demonstrated in an excellent article that off-axis positions of spherical particles during translocations can induce up to 15% variations in the magnitude in ΔI if the diameter of the particles is approximately half the diameter of the pore, as was the case for the experiment shown in Figure 3.6.[11] In other words, if a particle undergoes Brownian motion from the center of the nanopore to the wall of the nanopore during its translocation, then the same particle can result in ΔI values that differ by ∼15%. This effect widens the distribution of nanopore-based size determination in cases when particles are untethered. On the other hand, in our case of lipid tethered proteins, the maximum distance of proteins from the wall is limited by the length of the extended tether, which we estimate to be ∼1.5 nm, while the minimal distance from the wall is likely 0.5 to 1 nm due to the thickness of the water layer near the headgroup of phospholipid bilayers (see [13]). Nonetheless, even when considering the unrealistic case that tethered proteins would somehow be able to sample all distances from the pore wall including the center position of the pore, Figure 5 of the Qin *et al.* article predicts that the maximum ΔI variations will be smaller than 7% because of the small ratio between the protein diameters and the pore diameters that we used in this chapter.

The reason why we discuss the effect of off-axis translocation of untethered particles here in such detail is that we would expect this effect to distort the accurate determination of the shape of untethered translocating particles by our current analysis approach (in Chapter 4, however, we determine that, fortuitously, effects of lowpass filtering counteract these effects). We predict that any fluctuation in ΔI caused by off-axis diffusion during the translocation of a single particle would be interpreted as the rotation of a non-spherical particle and would result in a non-spherical shape even for perfectly spherical particles. In contrast, as illustrated Figures 3.7 and 3.8, the translocation of the lipid-tethered spherical proteins ferritin or streptavidin does not result in fluctuations of the ΔI signal beyond the

baseline current noise, showing that such off-axis effects are negligible for lipid-tethered proteins that are significantly smaller than the pore diameter as is typically the case in our experiments. This discussion is continued in great detail in subsequent chapters, and these "free translocation" experiments with gold nanoparticles represent the first step toward – and, in some ways, a motivation for – the extensive work on freely translocating proteins presented in Chapter 4.

3.4.3 Evaluation of Nanopore Cylindrical Geometry

In order to address concerns about the geometry of the nanopores used throughout this chapter, we conducted additional experiments specifically designed to shed light on the validity – or lack thereof – of approximating our nanopores by a cylindrical nanopore shape. Specifically, we considered the following hypothesis: if many copies of a lipid-tethered, perfect sphere translocate through a nanopore that is significantly longer than the diameter of the sphere, then we expect these translocations to produce square-shaped resistive pulses **if** the nanopore is shaped like a cylinder. The shape of these pulses will have a relatively fast rise from the baseline current to its plateau peak value and then a relatively fast drop back to the baseline due to relative magnitude of access resistance compared to pore resistance. We anticipate the shape of the resistive pulses to be symmetric with respect to the beginning and end portion of the resistive pulse. This expectation is confirmed by the finite element simulation that we carried out in a cylindrical nanopore in the Supporting Information of Yusko *et al.* [1] showing a symmetric peak at its beginning and end with a flat part at its maximum. In contrast, a nanopore with a pronounced hourglass-shape would mean that a significant time-fraction of the resistive pulse would grow in amplitude as the pore would gradually become smaller at its entrance and then a significant fraction of the pulse would become smaller as the particle exits the gradually bigger exit of the pore. An hour-glass pore shape would lead to a relatively small fraction of the resistive pulse being flat at its maximum. Finally, a conical pore shape would lead to an asymmetric profile of the resistive

pulse with either a fast rise and gradual drop for pores with a small entrance and large exit or vice versa for the opposite pore shape. A conical pore shape would not lead to a flat maximum of the peak shape but rather to a tip-shaped resistive pulse and such a conical pore shape would lead to an asymmetric IV curve before coating the pore.

Figure 3.7 shows results for individual translocation events of lipid-tethered ferritin, a large protein with perfectly spherical shape (see Protein Data Bank: 1HRS) and Figure 3.8 shows the results for translocation of lipid-tethered streptavidin, a smaller protein that is also approximately spherical. As can be appreciated both from the individual translocation events and the average of 20 resistive pulses obtained from both proteins, respectively, the resistive-pulse shape, in both cases, is almost perfectly square with a relatively large and flat plateau section. This pulse shape confirms that the sensing zone of our nanopores can indeed be approximated by a perfect cylinder. As we have argued before, any other pore shape with significant deviations from a cylinder would have introduced errors into parameter estimates relative to reference values and we would not have been able to obtain good agreement of five measured parameters across ten different proteins.

3.4.4 Analysis of Chemical Protein Denaturation in a Nanopore

In an attempt to demonstrate that proteins transited the pore in a native state, we performed experiments where we chemically denatured protein prior to analysis, and compared the results from translocation events with denatured protein to those of the same native protein. As reported previously, the apparent volume of BSA proteins increases significantly (in our experiment by 60%) upon denaturation, presumably because of a looser conformation and increased solvent-exposed area compared to native BSA. Similarly, the shape of the BSA changes dramatically upon denaturation from an m value of $\sim$0.5 to $\sim$0.3.

Interestingly the normalized ΔI distribution for BSA from the experiment with 0 M urea and with 1 M urea are similar, indicating that BSA maintains its native conformation in the presence of 1 M urea. This result suggests that, under the experimental conditions and

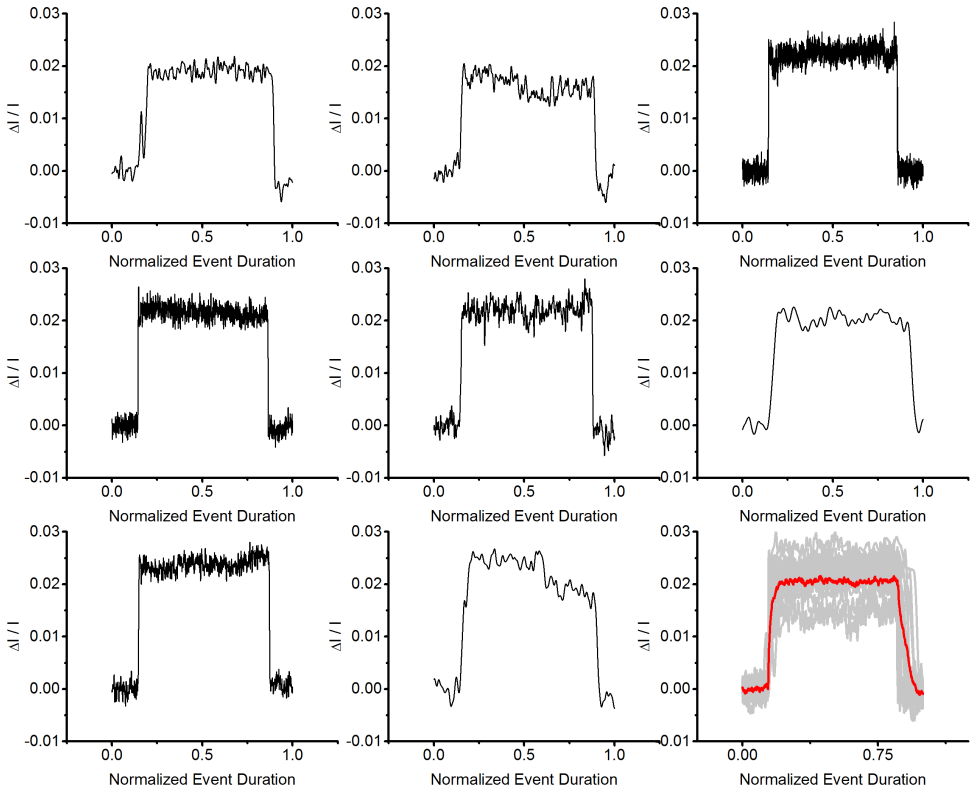


Figure 3.7. Evaluation of cylindrical nanopore geometry using long events of ferritin. Resistive pulses were normalized with regard to the event duration such that all events had a relative duration from 0 to 1. The panel in the lower right corner shows the overlay of 20 resistive pulses (grey) with their average current shown in red. The other panels show 8 examples of these 20 pulses, which were selected based on a random number generator. The important result from these experiments is that translocation of a spherical protein leads to square-shaped resistive pulses indicating a uniform electric field in the sensing zone of the pore and hence confirming that the shape of the sensing zone can be approximated by a cylinder. Note also that for spherical proteins, the magnitude of the current noise during the translocation event is similar to the magnitude of the current noise of the baseline near this event. This result is in stark contrast to the significant current modulations during translocation events of non-spherical proteins (see for instance Figure 3.1-D, 4.1-E,F), which are the result of time-variant shape- and orientation-dependent electrical shape factors γ while non-spherical proteins translocate and rotate through the pore. Using our standard, uncalibrated individual-event analysis approach, we determined a volume of 705 nnm^3 for ferritin, which compares well to the expected volume of 775 nm^3. Adapted with permission.[1] Copyright 2017, Springer.

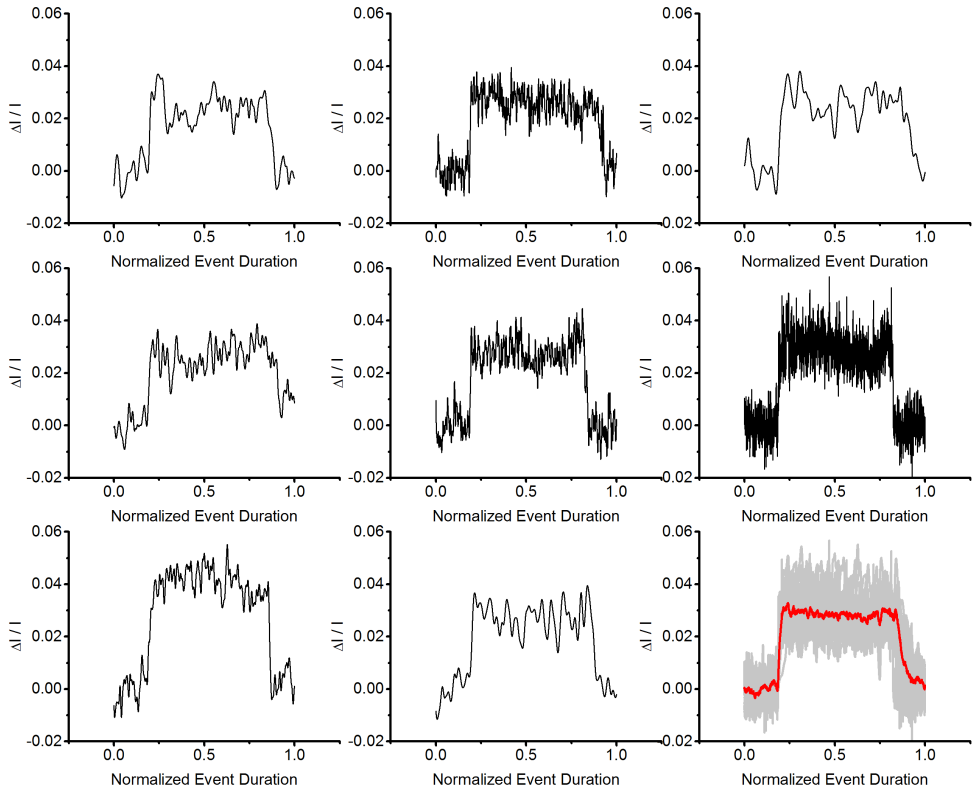


Figure 3.8. Evaluation of cylindrical pore geometry using long events of streptavidin. Resistive pulses were normalized with regard to the event duration such that all events had a relative duration from 0 to 1. The panel in the lower right corner shows the overlay of 22 resistive pulses (grey) with their average current shown in red. The other panels show 8 examples of these 22 pulses, which were selected by a random number generator. As in Figure RL3, the important result from these experiments is that translocation of a spherical protein leads to square-shaped resistive pulses indicating a uniform electric field in the sensing zone of the pore and hence confirming that the shape of the sensing zone can be approximated well by a perfect cylinder. Using our standard individual-event analysis approach, we determined a volume of 106 nm^3 for streptavidin, which compares well to the expected volume of 95 nm^3. Adapted with permission.[1] Copyright 2017, Springer.

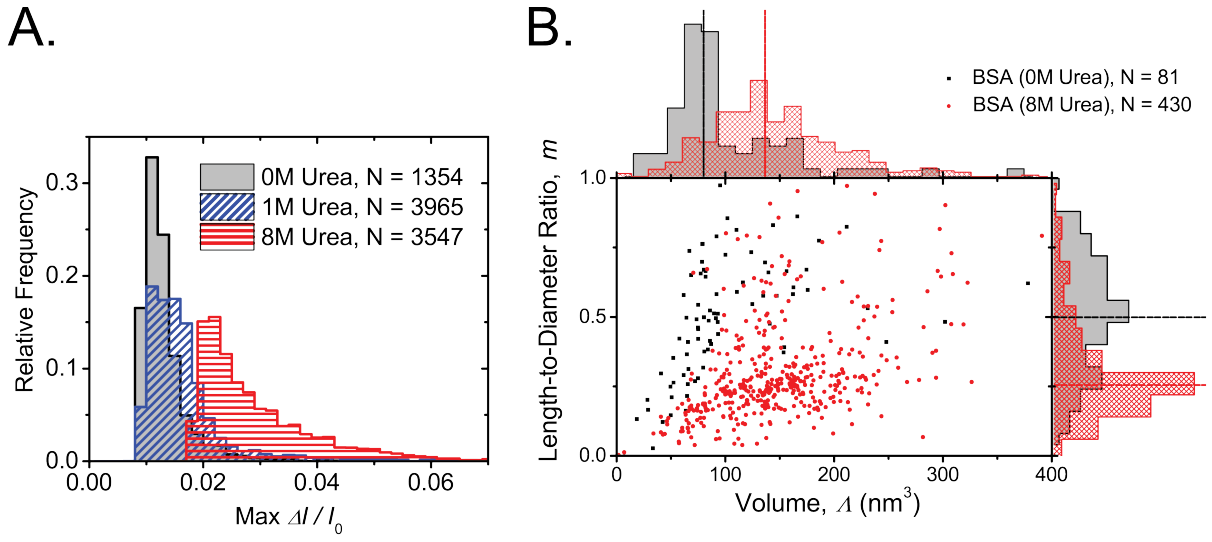


Figure 3.9. Evaluation of a chemical denaturant on protein shape and volume. A) Results from population-based analyses, comparing histograms of maximum ΔI values from the translocations of natively folded proteins (grey) to those from chemically denatured proteins (blue and red). Note that in 8M urea, we observe a large increase in resistive pulse amplitude, which we expect to represent a heterogenous mixture of globular expanded and unfolded proteins and multimeric protein complexes. B) Results from individual-event analyses, with length-to-diameter ratios and volumes plotted in black for native proteins, and in red for proteins that have been chemically denatured. Again, looking at the associated normalized histograms at the top of the plot, we see that the most common volume of the denatured proteins is approximately a dimer with a wide spread, and the length-to-diameter ratio tends toward a more flattened disk shape after denaturation.

electric fields we used, proteins appear to maintain their native conformation even in the presence of low concentrations of chemical denaturant. This result supports the idea that under the experimental conditions we used, proteins translocate through the pores in their native conformation, as we have argued before and as our results indicate for eleven different proteins. In the presence of 4 and 8 M urea, however, we observe a significant shift in apparent protein volume. In fact, we observe only a small overlap in the histogram between the data from 0 M urea and 8 M urea, indicating that either 1) almost all BSA molecules have lost their native conformation in the presence of 8 M urea and are translocating through the nanopore in an expanded and flattened globular disk, or 2) cysteine residues typically buried within BSA become more exposed upon unfolding, and the protein forms semi-unfolded dimers. As reported before, the change in dwell times upon denaturation was relatively small.

While it is not the focus of the present chapter, we were at first surprised that the shape of BSA flattened (*i.e.* became less spherical) after denaturation compared to its native state. Naïvely, we would have expected the shape of denatured BSA to correspond to a loosely packed, random-coiled sphere. Importantly, however, BSA contains 17 disulfide bonds in its native state, which makes it plausible that upon unfolding with 8 M urea in the absence of a reducing agent to break S-S bonds, these disulfide bonds persist and hold the coil together in a flattened conformation with 8 interconnects and one free cysteine.[36] Speculative alternative or additional reasons for a non-spherical shape of denatured monomeric BSA in the pore may be the consequence of 1) interactions between exposed hydrophobic amino acid residues with the bilayer coating, leading to flattening of the random coil or 2) confinement of the denatured coil within 1.5 nm from the pore wall due to the lipid tether, 3) depletion forces between the bilayer coating and the denatured, loose protein conformation due to depletion of counterions in the gap between lipid bilayer surface and the denatured protein, or 4) electric field induced distortions in the loose conformation. An alternative and simpler explanation to these four is that the protein is forming dimers or oligomers, which has been observed for BSA.[36] Regardless of the possible explanations for our findings, the purpose of the experiment shown in Figure 3.9 was to demonstrate that our method is able to resolve significant differences in volume and shape of BSA before and after denaturation and Figure 3.9 indeed provides strong evidence to this effect. We are currently investigating the effects of denaturation on the ellipsoidal shapes of other model proteins and anticipate that we will be able to build upon this preliminary finding.

3.4.5 Investigation of Initialization Values for Dipole Moment

When we analyze distributions of current blockade values using our convolution model (described in more detail in Chapter 4), we need to initialize four fitting parameters, namely $(\Delta I/I_0)_{min}$ as the 10^{th} percentile of the input data, $(\Delta I/I_0)_{max}$ value as the 90^{th} percentile of the data, the spread of the noise σ as the standard deviation of the current trace, and the

protein's permanent dipole moment. In order to investigate the effects of the initialization of dipole moment in particular, we re-performed the analysis of all nine non-spherical proteins examined in this chapter, initializing each individual-event fit with either the reference value for the dipole moment of the protein, or the average value (550 D) across a broad sample of proteins in the Protein Databank. We present the values in Figure 3.10.

From this analysis, we determined that the method of dipole initialization in the convolution fit has a small but measurable effect the accuracy of resulting estimates. Importantly, this analysis shows that initializing with an approximately average dipole moment (550 D) can produce accurate estimates of protein dipole moment, and is therefore applicable in a scenario where the dipole moment of a particular protein is unknown.

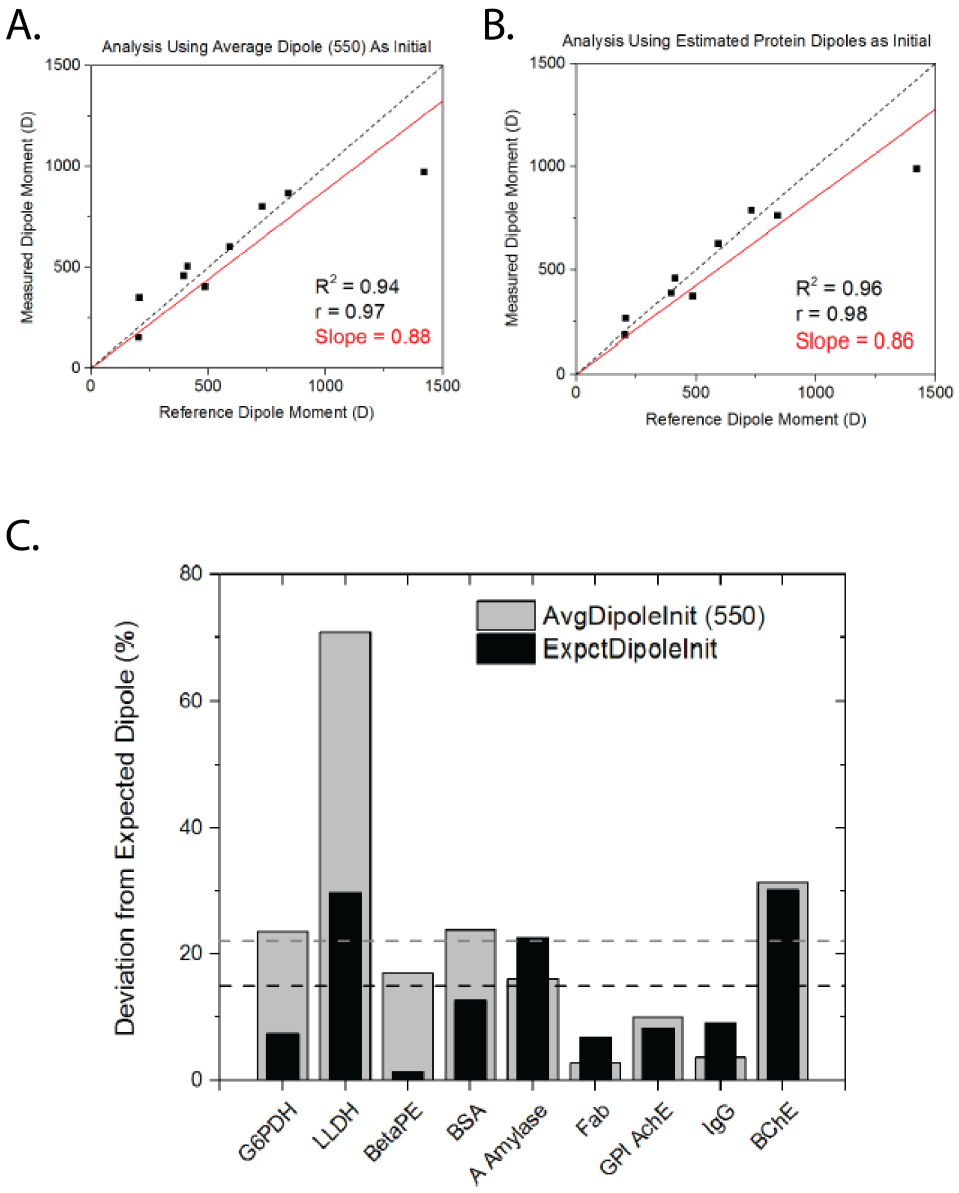


Figure 3.10. Evaluation of initialization parameters for dipole moment in the convolution fitting procedure. A) Comparison of the measured most-probable dipole moment of all proteins determined through individual-event analyses (Figure 3.2 blue pathway numbers 1-3) with expected reference values, initializing the dipole moment as 550 D. B) Comparison of the measured most-probable dipole moment of all proteins determined through individual-event analyses (Figure 3.2 blue pathway numbers 1-3) with expected reference values, initializing the dipole moment as the associated reference value of the protein. C) Comparison of the percent error in estimated dipole moment versus reference dipole moment between the two different initialization styles across all nine non-spherical proteins. The average error of each initialization mode is shown as a dashed line of the respective color (black or grey).

References

[1] E. C. Yusko et al. “Real-time shape approximation and fingerprinting of single proteins using a nanopore”. *Nature Nanotechnology* 12 (2017), pp. 360–367.

[2] P. Picotti and R. Aebersold. “Selected reaction monitoring–based proteomics: workflows, potential, pitfalls and future directions”. *Nature methods* 9.6 (2012), p. 555.

[3] A. E. Herr. “Disruptive by design: a perspective on engineering in analytical chemistry”. *Analytical chemistry* 85.16 (2013), pp. 7622–7628.

[4] A. Sali, R. Glaeser, T. Earnest, and W. Baumeister. “From words to literature in structural proteomics”. *Nature* 422.6928 (2003), p. 216.

[5] J. L. Oncley. “The Investigation of Proteins by Dielectric Measurements.” *Chemical Reviews* 30.3 (1942), pp. 433–450.

[6] C. E. Felder, J. Prilusky, I. Silman, and J. L. Sussman. “A server and database for dipole moments of proteins”. *Nucleic acids research* 35.suppl_2 (2007), W512–W521.

[7] S. N. Singh, S. Yadav, S. J. Shire, and D. S. Kalonia. “Dipole-Dipole Interaction in Antibody Solutions: Correlation with Viscosity Behavior at High Concentration”. *Pharmaceutical Research* 31.9 (2014), pp. 2549–2558.

[8] D. M. Ecker, S. D. Jones, and H. L. Levine. “The therapeutic monoclonal antibody market”. *MAbs*. Vol. 7. 1. Taylor & Francis. 2015, pp. 9–14.

[9] A. Han et al. “Sensing protein molecules using nanofabricated pores”. *Applied Physics Letters* 88.9 (2006), p. 093901.

[10] C. Dekker. “Solid-state nanopores”. *Nature Nanotechnology* 2.4 (2007). WOS:000245920900010, pp. 209–215.

[11] Z. Qin, J. Zhe, and G.-X. Wang. “Effects of particle’s off-axis position, shape, orientation and entry position on resistance changes of micro Coulter counting devices”. *Measurement Science and Technology* 22.4 (2011), p. 045804.

[12] D. Fologea, B. Ledden, D. S. McNabb, and J. Li. “Electrical characterization of protein molecules by a solid-state nanopore”. *Applied physics letters* 91.5 (2007), p. 053901.

[13] E. C. Yusko et al. “Controlling protein translocation through nanopores with bio-inspired fluid walls”. *Nature Nanotechnology* 6.4 (2011), pp. 253–60.

[14] C. Raillon et al. “Nanopore detection of single molecule RNAPDNA transcription complex”. *Nano letters* 12.3 (2012), pp. 1157–1164.

[15] G. V. Soni and C. Dekker. “Detection of nucleosomal substructures using solid-state nanopores”. *Nano letters* 12.6 (2012), pp. 3180–3186.

[16] N. Di Fiori et al. “Optoelectronic control of surface charge and translocation dynamics in solid-state nanopores”. *Nature nanotechnology* 8.12 (2013), p. 946.

[17] E. C. Yusko et al. “Single-particle characterization of A oligomers in solution”. *ACS nano* 6.7 (2012), pp. 5909–5919.

[18] I. Nir, D. Huttner, and A. Meller. “Direct sensing and discrimination among Ubiquitin and Ubiquitin chains using solid-state nanopores”. *Biophysical journal* 108.9 (2015), pp. 2340–2349.

[19] R. I. Stefureac, C. A. Madampage, O. Andrievskaia, and J. S. Lee. “Nanopore analysis of the interaction of metal ions with prion proteins and peptides”. *Biochemistry and Cell Biology* 88.2 (2010), pp. 347–358.

[20] J. Li et al. “Ion-beam sculpting at nanometre length scales”. *Nature* 412.6843 (2001), pp. 166–169.

[21] D. C. Golibersuch. “Observation of Aspherical Particle Rotation in Poiseuille Flow via the Resistance Pulse Technique”. *Biophysical Journal* 13.3 (1973), pp. 265–280.

[22] R. W. DeBlois, E. E. Uzgiris, D. H. Cluxton, and H. M. Mazzone. “Comparative measurements of size and polydispersity of several insect viruses”. *Analytical biochemistry* 90.1 (1978), pp. 273–288.

[23] R. W. DeBlois and C. P. Bean. “Counting and sizing of submicron particles by the resistive pulse technique”. *Review of Scientific Instruments* 41.7 (1970), pp. 909–916.

[24] D. Pedone, M. Firnkes, and U. Rant. “Data Analysis of Translocation Events in Nanopore Experiments”. *Analytical Chemistry* 81.23 (2009), pp. 9689–9694.

[25] H. Fricke. “The Electric Permittivity of a Dilute Suspension of MembraneCovered Ellipsoids”. *Journal of Applied Physics* 24.5 (1953), pp. 644–646.

[26] S. Velick and M. Gorin. “The electrical conductance of suspensions of ellipsoids and its relation to the study of avian erythrocytes”. *The Journal of general physiology* 23.6 (1940), p. 753.

[27] Y. Yuan and D. Axelrod. “Subnanosecond polarized fluorescence photobleaching: rotational diffusion of acetylcholine receptors on developing muscle cells”. *Biophysical journal* 69.2 (1995), pp. 690–700.

[28] J. J. Kasianowicz, E. Brandin, D. Branton, and D. W. Deamer. “Characterization of individual polynucleotide molecules using a membrane channel”. *Proceedings of the National Academy of Sciences* 93.24 (1996), pp. 13770–13773.

[29] E. A. Manrao et al. “Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase”. *Nature biotechnology* 30.4 (2012), p. 349.

[30] D. Branton et al. “The potential and challenges of nanopore sequencing”. *Nature Biotechnology* 26.10 (2008). WOS:000259926000029, pp. 1146–1153.

[31] J. K. Rosenstein, M. Wanunu, C. A. Merchant, M. Drndic, and K. L. Shepard. “Integrated nanopore sensing platform with sub-microsecond temporal resolution”. *Nature methods* 9.5 (2012), p. 487.

[32] C. Plesa et al. *Fast Translocation of Proteins through Solid State Nanopores*. rapid-communication. 2013.

[33] M.-H. Lee et al. “A Low-Noise Solid-State Nanopore Platform Based on a Highly Insulating Substrate”. *Scientific Reports* 4 (2014), p. 7448.

[34] J. D. Uram, K. Ke, and M. Mayer. “Noise and bandwidth of current recordings from submicrometer pores and nanopores”. *ACS nano* 2.5 (2008), pp. 857–872.

[35] A. R. De Pascalis et al. “Binding of ferredoxin to ferredoxin: NADP+ oxidoreductase: the role of carboxyl groups, electrostatic surface potential, and molecular dipole moment”. *Protein Science* 2.7 (1993), pp. 1126–1135.

[36] I. Rombouts et al. “Formation and reshuffling of disulfide bonds in bovine serum albumin demonstrated using tandem mass spectrometry with collision-induced and electron-transfer dissociation”. *Scientific reports* 5 (2015), p. 12210.

Chapter 4: Label-Free Characterization of Proteins

This chapter demonstrates that high-bandwidth current recordings in combination with low-noise silicon nitride nanopores make it possible to determine the molecular volume, approximate shape, and dipole moment of single native proteins in solution without the need for labeling or other chemical modifications of these proteins. The analysis is based on current modulations caused by the translation and rotation of single proteins through a uniform electric field inside of a nanopore. We applied this technique to nine proteins and show that these measured protein parameters agree well with reference values, but only if the nanopore walls were coated with a non-stick fluid lipid bilayer. One potential challenge with this approach is that an untethered protein is able to diffuse laterally while transiting a nanopore, which generates increasingly asymmetric disruptions in the electric field as it approaches the nanopore walls. These "off-axis" effects add an additional noise-like element to the electrical recordings, which can be exacerbated by non-specific interactions with pore walls that are not coated by a fluid lipid bilayer. We performed finite element simulations to quantify the influence of these effects on subsequent analyses. Examining the size, approximate shape, and dipole moment of unperturbed, native proteins in aqueous solution on a single-molecule level in real time while they translocate through a nanopore may enable applications such as monitoring the assembly or disassembly of transient protein complexes based on changes in their shape, volume, or dipole moment.

4.1 Introduction to Label-Free Nanopore Sensing

Recent advances in single-molecule methods, including resistive pulse sensing with nanopores, have made it possible to interrogate the physical characteristics of individual proteins and other biomolecules in aqueous solution.[1–4] Other established biophysical techniques such as atomic force microscopy (AFM) can determine the mechanical properties of proteins in an aqueous sample [5–7], while optical techniques like fluorescence correlation spectroscopy (FCS) [8–10] and fluorescence resonance energy transfer (FRET) [11–14] reveal spatial and temporal information about single proteins that can provide insight into their structure or interactions with other molecules in solution. In each of these methods, proteins are often subjected to either chemical or physical modifications for optimal analysis on a single-molecule level. These preparation steps can alter intrinsic protein properties and subsequent measurements may not be representative of native proteins in their physiological environments.[5, 9, 11]

Nanopore-based, resistive pulse sensing is a single-molecule method capable of investigating the physical and structural properties of individual proteins and protein complexes.[15–27] This technique can characterize up to 100 particles per second, in electrolyte-rich aqueous solution, as they individually transit the zeptoliter confines of a nanopore.[28, 29] As shown in Figure 4.1, each particle disrupts the flow of ions to an extent that is proportional to its volume, shape, and relative orientation to the electric field, while its dwell time (t_d) within the pore corresponds to its net charge and electrophoretic mobility in the applied electric field.[30, 31] Resistive pulse-based, nucleic acid sequencing has made notable progress recently by combining biological nanopores with complimentary enzymes that ratchet nucleotide strands through the pore one base at a time.[32–40] Protein characterization with nanopores, however, is not as advanced as nucleic acid characterization for two key reasons: the size of globular proteins necessitates synthetic nanopores with large diameters that are prone to non-specific adhesion [41], and protein characterization targets mobile particles that transit the sensing zone - in the absence of adsorption - at rates approaching or exceeding

the bandwidth of conventional current-recording systems.[42–45]

In response to these challenges, we and others have designed a variety of nanopore systems that slow the transit speeds of proteins while avoiding clogging of the pores that would otherwise end experiments. For example, Wloka *et al.* engineered a non-stick biological nanopore, Cytolysin A, with a 5.5 nm diameter sensing vestibule above a 3.8 nm diameter pore and used it to detect the attachment of an individual ubiquitin protein (8.6 kDa) to a protein substrate. [27, 47, 48] Waduge *et al.* estimated the sizes and intrinsic flexibilities of proteins as they slowly squeezed through synthetic nanopores with diameters only slightly larger than the proteins themselves [49], while others have investigated protein analytes with a variety of anti-adhesive approaches.[3, 50–57] Increasing the viscosity of the recording buffer (*e.g.* by adding glycerol) also makes it possible to slow the diffusion of proteins, though this approach reduces the conductivity of the recording buffer and thus reduces the amplitude of resistive pulses. In Chapters 2 and 3, we discussed coating synthetic nanopores with a fluid lipid bilayer, which prevents or minimizes non-specific adhesion to the nanopore substrate.[58] This coating can present fluid lipid anchors to slow the diffusion of tethered proteins by taking advantage of the viscosity of the bilayer.[58] We then demonstrated in Chapter 3 that lipid-coated nanopores make it possible to determine the volume, ellipsoidal shape, dipole moment, rotational diffusion coefficient, and charge of proteins simultaneously in aqueous solution and employed this multiparametric fingerprint to categorize populations of protein in a binary mixture.[15] For these applications, however, the protein needed to be attached to the lipid bilayer using either chemical cross-linking or a lipid-functionalized protein binding complement.

In this chapter, we demonstrate that it is possible to determine the ellipsoidal shape, volume, and dipole moment of single untethered and unmodified proteins in aqueous solution as they translocate through a nanopore, driven by electrophoretic force due to their net charge in the electric field. This approach is different than the method we demonstrated previously because here we allow proteins to diffuse freely in solution and no longer slow their

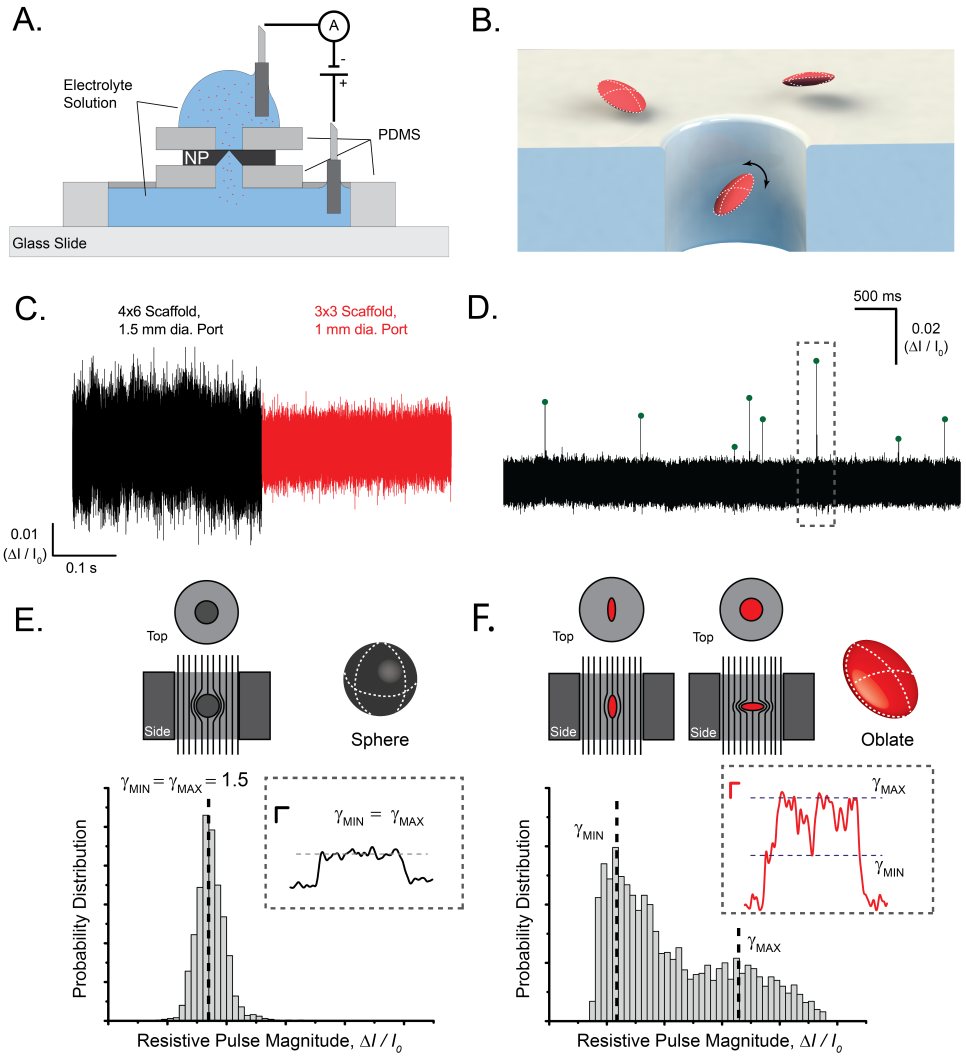


Figure 4.1. Introduction to experiments with untethered proteins. Individual, non-spherical proteins passing untethered through a nanopore create current modulations that contain information about their length-to-diameter ratios, volumes, and dipole moments. A) Schematic cross-section of a setup with a nanopore in a silicon chip and fluid compartments confined by a silicone elastomer (PDMS). B) Cartoon representation of oblate ellipsoids (red) passing through a nanopore in a free-standing silicon nitride membrane of a nanopore chip. C) Baseline current measured across the bare (without lipid coating) nanopore substrate at an applied potential difference of -100 mV and digitally filtered with a 50 KHz Gaussian lowpass filter. Both nanopores are approximately 25 nm in diameter. The RMS current noise with the 3 $\times$ 3 scaffolds was a factor of 1.8 lower than with the previously used configuration. D)Example current trace with a duration of 10 s, digitally low-pass filtered at 50 kHz, with maximum values of all resistive-pulse events shown as green dots, and a long event with a duration greater than 150 μs signified by a dashed grey box. E,F) Probability distribution of current values within a single resistive pulse as a function of particle length-to-diameter ratio and orientation during its translocation through the nanopore. Inset: original current-versus-time traces of single resistive pulses from the translocation of a streptavidin protein (black) and an IgG protein (red). These traces were digitally low-pass filtered at 10 kHz for clarity. Maximum and minimum blockade values corresponding to electrical shape factors γ_{max} and γ_{min} are shown by dashed lines. Scale bars represent 100 μs and 0.01 ($\Delta I/I_0$). Adapted with permission.[46] Copyright 2019, American Chemical Society.

diffusion by tethering them to a lipid anchor in the fluid lipid bilayer. To make this approach possible, we now use nanopore chips with a 3 mm $\times$ 3 mm frame, sandwiched between two layers of polydimethylsiloxane (PDMS) with small access ports to the nanopore (Figure 4.1-A). This experimental design reduced the current noise by 40% at 50 kHz bandwidth and thus increased the signal-to-noise ratio (SNR) at this bandwidth (see Figure 4.1-C,4.14). By digitally low-pass filtering the data at 50 kHz as opposed to 15 kHz, as we had done in previous work, we increased the number of analyzable resistive pulses [42, 59] as well as the temporal resolution from each translocation event. We show here that the number, duration, and bandwidth of resistive pulses are critical for the accuracy of subsequent analyses on those pulses.

We previously discussed the analysis of data from non-spherical particles rotating in an electric field to calculate protein parameters from individual resistive pulses (see Chapter 3);[15] the approach is based on fundamental theory developed by Golibersuch [31], Fricke [60], Velick and Gorin [61], and others.[54, 58, 62–65] Briefly, particles rotate and adopt different orientations relative to the electric field during their passage through the pore (Figure 4.1-B). In the case of simple ellipsoidal particles rotating and translocating through a nanopore, an electrical shape factor, γ, relates the particle's orientation, θ, within the electric field, and its length-to-diameter ratio, m, to the current blockade, $\Delta I/I_0$. A perfectly spherical particle samples only one γ value (equal to 1.5, Figure 4.1-E) during transit. A non-spherical particle can sample all γ values contained between perfectly crosswise (γ_{max}) and perfectly lengthwise (γ_{min}) orientations (Figure 4.1-F), and will do so with a probability for various γ values that can be described by a U-shaped distribution.[31]

In order to quantify protein length-to-diameter ratio, volume, and dipole moment, we determined the particular probability distribution of γ values for a given protein by using an iterative convolution fitting procedure (described in detail in Section 4.4.1).[15] In this procedure, we fit the entire $\Delta I/I_0$ distribution from each individual translocation event with a dwell time greater than 150 μs (Figure 4.1-D, grey box). This approach returned an

approximate ellipsoidal shape, volume, and dipole moment value for tens to hundreds of individual protein translocation events within a single experiment. Characterization of freely translocating proteins based on stand-alone analysis of single events one-by-one is uncommon, as the majority of work characterizing proteins with nanopores extracts parameters like volume from populations of resistive pulses rather than from the individual pulses themselves. Here, we report all values from individual translocation events as well as the median length-to-diameter ratio, median volume, and the most-probable dipole moment determined from distributions of single molecule-based individual event analyses (Table 4.2), and discuss factors that influence the uncertainty of those values. In order to demonstrate that this characterization methodology can be applied to determine a range of protein characteristics, we chose a set of proteins that vary widely in length-to-diameter ratio ($m = 0.14$ to 2.5), volume ($\Lambda = 95$ to 1700 nm^3) and dipole moment ($\mu = 484$ to 1846 D). We demonstrate that, in order to be feasible, the method combines three important characteristics: First, it estimates multiple physical parameters of proteins translocating freely through a synthetic nanopore by combining low-noise nanopores and high-bandwidth recordings. Second, it proceeds in a more straightforward manner than its tethered alternative by circumventing the tethering step and thus provides measurements on unperturbed proteins. And third, it takes advantage of anti-adhesive nanopore coatings that are critical in providing the free translocation and rotation necessary to collect these measurements without artifacts.

4.2 Results and Considerations of Label-Free Protein Analyses

4.2.1 Filtering Attenuates Fluctuations in the Resistive Pulses from Freely Translocating Proteins

Signal bandwidth is critical for recording accurate amplitudes and durations of resistive pulses; inadequate bandwidth can clip the amplitude of the signal or overlook resistive pulses that occur between two sampled points.[66] Moreover, proteins that are not tethered to a lipid bilayer during translocation rotate at a rate that is approximately two orders of magnitude faster than tethered proteins.[15, 67, 68] This difference in rotational diffusion coefficient has important implications when the goal is to resolve differences in protein orientation during a single translocation event, as is necessary, for instance, to determine a bias in orientation that reflects the dipole moment of a protein.

To investigate the extent to which we could resolve different orientations of an untethered protein rotating and translocating freely through a lipid bilayer coated nanopore, we performed random-walk simulations.[15] To do so, we used the rotational diffusion coefficient of the protein to determine the average rotated angle during each time step around a single rotational axis. We then selected a rotational direction (*e.g.* clockwise or counterclockwise) based on orientation-dependent biased diffusion in an electric field, and converted the resulting array of angles to their corresponding $\Delta I/I_0$ values.[15] Figure 4.2 shows ideal representations (*i.e.* without recording noise) of resistive pulses produced by a simulated random walk for an oblate ellipsoid that represents a 150 kDa protein as it translocates through a nanopore. When the protein rotated with a rotational diffusion coefficient that corresponds to a lipid anchored state [15], filtering at 50 kHz retained the large majority of fluctuations as well as their maximum and minimum blockade amplitudes (Figure 4.2-A). This figure panel also shows that in order to sample both the minimum and maximum orientations of the

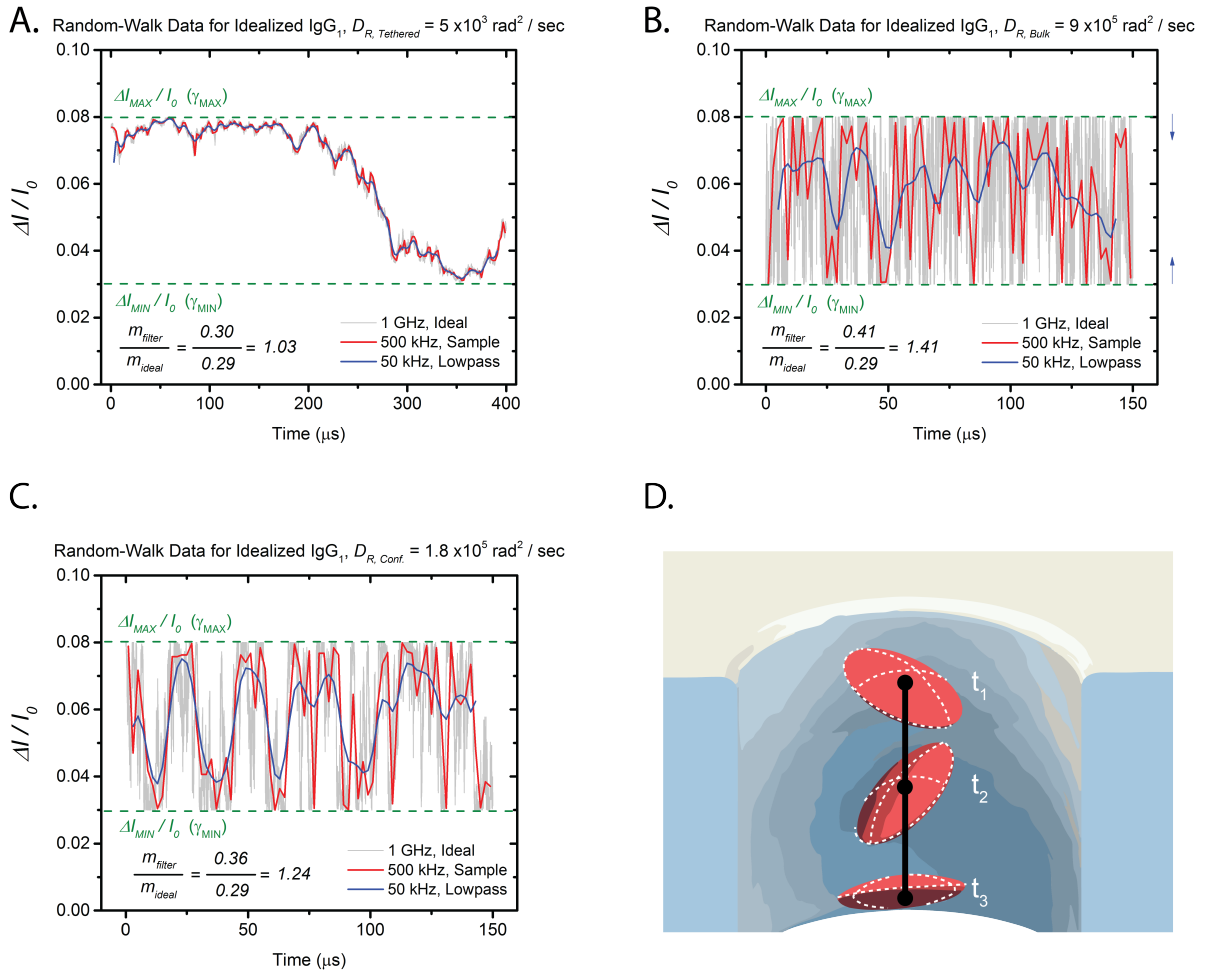


Figure 4.2. Effects of filtering on simulated resistive pulse signals. A) Simulated resistive-pulse trace generated from a random-walk simulation for a 150 kDa protein with a shape that can be approximated by an oblate (*m* = 0.29, Λ = 340 nm^3, μ = 840 D) rotating at a rate of 5 $\times$ 10^3 rad^2 sec^{-1} that corresponds to this protein after tethering it to a lipid bilayer with a lipid anchor as reported by Yusko *et al.* [15] B) Simulated resistive-pulse trace for the same protein as in (A) rotating with a diffusion coefficient of 9 $\times$ 10^5 rad^2 sec^{-1}, estimated for the particle in bulk solution using the software HydroPRO.[69] Blue arrows show the attenuation effect of low-pass filtering. C) Simulated resistive-pulse trace for the same protein as in (A) rotating with a rotational diffusion coefficient of 1.8 $\times$ 10^5 rad^2 sec^{-1} that corresponds to an untethered protein inside of the confinement of a nanopore according to Dix & Verkman.[67] A-C) Grey lines represent each ns time step of the random-walk simulation, red lines show the same trace downsampled at 500 kHz, and the blue lines are the downsampled trace digitally low-pass filtered at 50 kHz. Note that the blue filtered trace is shifted in time to show visual alignment with other traces (see Figure 4.6). Green dashed lines show minimum and maximum $\Delta I/I_0$ values that correspond to electrical shape factors, γ_{min} and γ_{max}. Estimated values of protein length-to-diameter ratio, *m*, are compared between the filtered data (blue) and the ideal random-walk data (grey). The dwell time of the protein in the pore was set to either (A) 400 μs or (B,C) 150 μs. The electric field within the pore was 1.646 MV m^{-1}, corresponding to a nanopore with a length of 38 nm and a diameter of 17 nm connected to an adjacent channel with a length of 275 nm and diameter of 100 nm [70] with an applied potential of 0.1 mV in 2 M KCl solution with a resistivity of 0.046 Ω m. D) Cartoon showing small, time-step rotations of an ellipsoidal particle passing through a nanopore, representing the random-walk simulations shown in (A-C). Adapted with permission.[46] Copyright 2019, American Chemical Society.

particle, however, dwell times of at least 400 μs duration were required.[15] The reason is, that in order to estimate the ellipsoidal shape, m, of a protein accurately, the protein must remain sufficiently long in the pore to sample the orientations that correspond to minimum and maximum blockade values because these two values represent the extreme lengthwise and extreme crosswise orientations of a particle with a given volume and length-to-diameter ratio (see Eqn's 4.10-4.15 for how $(\Delta I/I_0)_{min}$ and $(\Delta I/I_0)_{max}$ influence the quantification of ellipsoidal shape and volume). In contrast, when the protein rotated with a rotational diffusion coefficient that corresponds to bulk solution, a sampling rate of 500 kHz was too slow to completely resolve fluctuations between minimum and maximum orientations, and digital filtering at 50 kHz produced a 41% underestimate of length-to-diameter ratio as represented by the m value (Figure 4.2-B). At a rotational diffusion coefficient that was 5-fold slower than bulk – one that describes the protein rotating untethered but within a confined space [67] – sampling at 500 kHz was sufficient to track the protein rotating between minimum and maximum orientations. Low-pass filtering at 50 kHz, however, still produced a 24% underestimate of protein length-to-diameter ratio under these conditions (Figure 4.2-C).

In general, filtering attenuates the current fluctuations that correspond to different protein orientations during its translocation through the nanopore. This effect depends on the time resolution of the recording and is pronounced when the protein rotates at faster rates, as is the case in Figure 4.2-B. Estimates of length-to-diameter ratio appear to be most influenced by choice of filter cutoff frequency and the rotational diffusion coefficient of the protein (Figure 4.15-A,B). Estimates of volume follow a similar trend to those of length-to-diameter ratio, whereby the estimates are more accurate at higher filter frequencies and slower rotational diffusion coefficients (Figure 4.15-C,D). Estimates of dipole moment, however, appear to be relatively independent of protein rotation rate and filter cutoff frequency (Figure 4.15-E,F), but improve as dwell times increase (Figure 4.11). These trends highlight the usefulness of anchoring proteins to anti-adhesive coatings in order to slow both their rotation as well as their speed of transit through the nanopore; they also highlight the importance of attempts

to increase recording bandwidth and SNR in future resistive pulse-based experiments for protein characterization. For example, according to Figure 4.2-B, the commercially available Chimera VC-100 amplifier, with a bandwidth of ∼1 MHz, should be able to time-resolve the rotation of an untethered protein in the confines of a nanopore with significantly higher fidelity than the Axopatch 200B used here, assuming that the SNR at this bandwidth permits resolving various orientations.[59]

With the results from the simulations in Figure 4.2 in mind, we expected untethered proteins translocating through the confines of a nanopore (analogous to the case illustrated in Figure 4.2-C) to produce estimates of length-to-diameter ratio, m, that were less extreme than predicted and thus closer to a sphere ($m = 1$) than to a flattened or elongated ellipsoid of rotation ($m << 1$ or $m >> 1$). Specifically, data filtered at 50 kHz will not reach the full amplitudes of γ_{min} and γ_{max} (blue curve in Figure 4.2-C) and thus lead to a systematic underestimate of the m value by approximately 25 %.

4.2.2 Individual Translocations through Lipid Bilayer-Coated Nanopores Contain Information about Protein Ellipsoidal Shape, Volume, and Dipole Moment

Figure 4.3 (A,C) shows individual data of experimental length-to-diameter ratio and volume determinations that resulted from individual translocation events of a single protein moving through the nanopore. To allow for sufficient time such that proteins could sample all electrical shape factors, we restricted the analysis to resistive pulses with a duration of at least 150 μs. Determined values for length-to-diameter ratio and volume were in reasonable agreement with reference values: median estimates for the length-to-diameter ratio deviated on average by 35% from reference values across all proteins (Figure 4.3-B), and median volume estimates were within 40% of reference values (Figure 4.3-D). Figure 4.3 also shows that estimates of protein shape, volume, and dipole moment from individual translocation

events varied considerably for each protein. We attribute this variability to the intrinsic uncertainty of single-molecule analyses, to noise in the signal recording [66], to the limited recording bandwidth,[66], to the fundamental limitation of approximating a complex three-dimensional protein shape with a simple ellipsoid (see Figure 4.12 for visual details on the accuracy of this approximation), and to off-axis effects as discussed below.[30] We note that the approach of analyzing single translocation events in a stand-alone fashion as shown in Figure 4.3 (A,C,E) is particularly demanding and thus the median length-to-diameter ratio and volume estimates from these analyses were less accurate than those estimates from population-based analyses, similar to findings from previous work.[15] For instance, population-based analyses performed on the same data revealed protein shapes that were within 20% of reference values for all oblate-shaped proteins, and volume estimates that were strongly correlated with reference values for all nine proteins (slope = 0.989, Pearson's r = 0.99, see Figure 4.5). The reason why we emphasized stand-alone event-by-event analysis in the work presented here is that this approach is required when the long-term goal to analyze mixtures of proteins is to be achieved.

Proteins with a permanent dipole moment do not rotate randomly while passing through the electric field within a nanopore; rather, they experience a Brownian rotation that is biased by torque acting on their dipole from the electric field.[71] The convolution-based model used to fit the data in this work accounted for this bias, and ultimately determined the extent of bias each protein had for its minimum (γ_{min}) and maximum (γ_{max}) electrical shape factors; *i.e.* did the protein sample both maxima with equal probability, or did it sample one orientation and hence electrical shape factor more often than the other? Figure 4.3-F shows the most-probable dipole moment values (μ) determined through log-normal fitting of the distribution of measured dipoles for each protein.[15] These dipole estimates were in good agreement with reference values; most probable dipole moment estimates for each protein deviated on average by less than 20% from their reference values. Event-to-event variability in dipole moment was large, consistent with the data spread in shape and

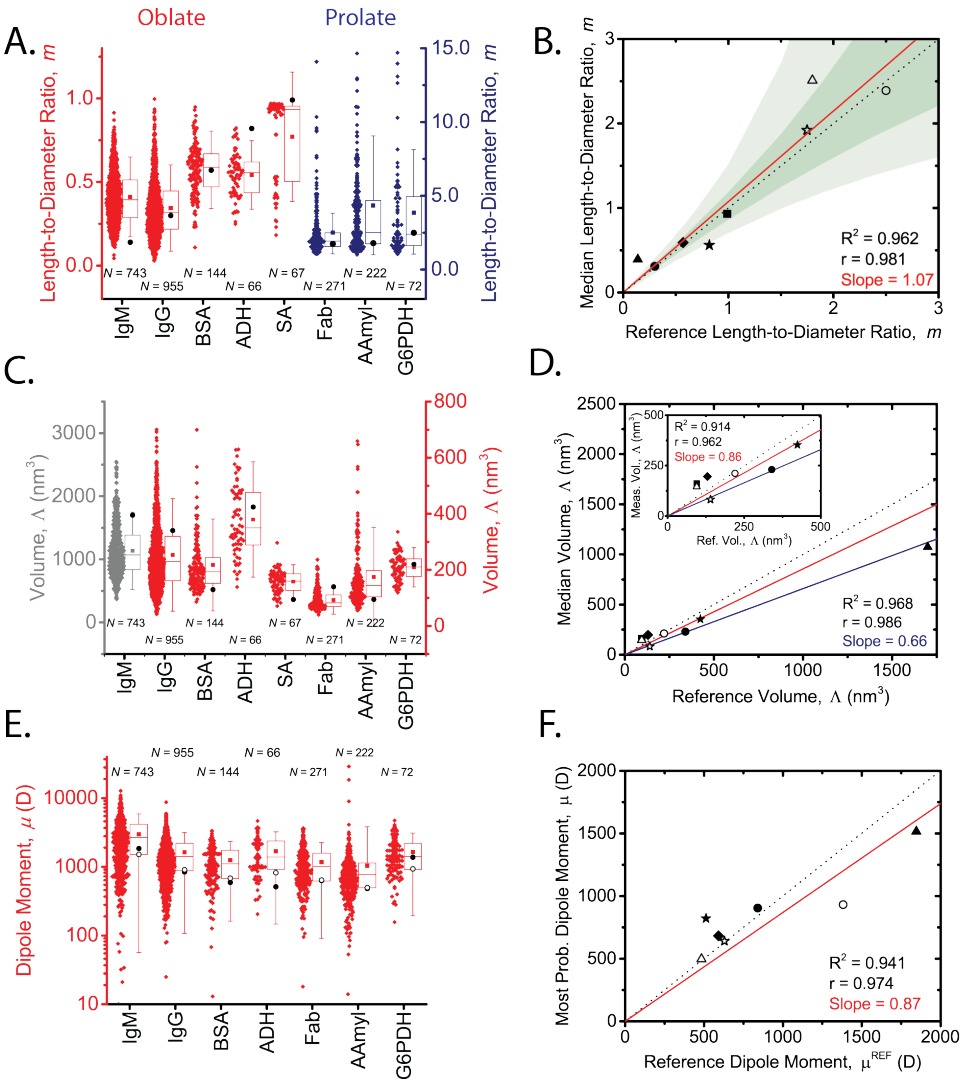


Figure 4.3. Results of individual-event analyses on untethered proteins. A) Length-to-diameter ratios determined from individual resistive pulses from the translocation of various oblate-shaped proteins (red) and various prolate-shaped proteins (blue). B) Comparison of the median values of measured length-to-diameter ratios, *m*, with reference values for each protein. The red line is the linear fit for all proteins in the study, the black dotted line shows the ideal 1:1 agreement, and the dark and light green regions represent estimates of length-to-diameter ratio in the presence of $\pm$10% and $\pm$25% deviations in minimum and maximum blockade values, respectively. C) Volumes of all proteins, with IgM in grey corresponding to the left y-axis in grey, and other proteins in red corresponding to the right y-axis in red. D) Comparison of the median values of the volumes for 8 proteins determined from single event analyses with reference values for these proteins. Inset shows values for proteins with volumes smaller than 500 nm^3. The red line represents a linear fit for only the proteins with volumes smaller than 500 nm^3 and the blue line represents a linear fit through data for all proteins. E) Dipole moment estimates for individual events of all non-spherical proteins. Data are plotted on a log-scale to represent their underlying log-normal distribution, as discussed by Yusko *et al.* [15] A,C,E) Parameter estimates from individual long events are shown as red diamonds, red squares represent mean values, horizontal lines represent median and quartile values, whiskers represent the standard deviation of the values, black filled circles denote reference values, and black open circles denote most-probable values determined through log-normal fitting.[15] F) Comparison of estimated dipole moments with reference values for all 7 non-spherical proteins investigated. The black dotted line represents the ideal 1:1 agreement, and the red solid line is the linear fit. B,D,F) Proteins are plotted as follows: anti-biotin Fab (open stars), α-Amylase (open triangles), G6PDH (open circles), IgM (triangles), anti-biotin IgG (circles), BSA (diamonds), ADH (stars), and Streptavidin (squares). See Table 4.3 for quantitative values. Adapted with permission.[46] Copyright 2019, American Chemical Society.

volume estimates; the estimates of dipole moment for individual BSA events, for example, had a median absolute deviation of 73%. While this level of uncertainty is relatively high, we note that this is the only technique available to estimate the dipole moment of individual unmodified proteins in solution. In our previous work [15], tethering removed one positive charge from each protein when cross-linking them to a lipid anchor in the bilayer by means of a bifunctional N-hydroxysuccinimide (NHS) linker. Removing a charge of a randomly located amine on the surface of the protein inherently distorted that protein's permanent dipole moment and is thus not desirable.[15] Furthermore, the technique presented here generates dipole estimates within a few hundred microseconds as the protein passes through the nanopore, and is compatible with small sample volumes ranging from nano- to microliters as well as protein concentrations ranging from nano- to micromolar. These characteristics are attractive because dipole moments are becoming increasingly important for the rheological properties of concentrated antibody formulations used for subcutaneous administration.[72] Rapid quantification of dipole moments in aqueous solution also provides additional discriminatory power in heterogeneous protein mixtures as dipole moments are distributed broadly between different proteins and show little to no correlation with the volume or shape of proteins.[15]

In previous work on lipid-anchored proteins, we restricted the analyses to individual resistive pulses longer than 400 μs because we could be confident that most tethered proteins would sample all possible orientation-dependent electrical shape factors (γ) within this time frame.[15] Proteins that rotate freely in bulk solution, however, do so at a rate approximately two orders of magnitude faster than tethered proteins inside of a nanopore [15, 68], and thus freely translocating proteins sample all possible γ values on much shorter time scales than tethered proteins. When selecting a threshold for sufficiently long events, we struck a balance between accuracy and sample size (see Section 4.4.4 for details). In other words, we needed to gather current-versus-time data with sufficient duration to produce accurate estimates of protein parameters, but we also needed to collect as many events as possible for analysis

within a standard experimental time frame. Based on theoretical predictions using a model that treats translocations of charged proteins as a biased first-passage-time problem [73], we estimated that the most probable dwell times for the proteins in this study were all shorter than 10 μs. Hence, only a small percentage ($< 1\%$) of the total number of translocation events that we resolved had dwell times longer than 150 μs.[15, 42, 58] Nonetheless, we found that a threshold of 150 μs was the optimal choice for our recording setup combined with a digital Gaussian low-pass filter with a cutoff frequency of 50 kHz, rather than the 15 kHz filter that we used in previous work.[15] This elevated cutoff frequency made it possible to time-resolve a larger fraction of high-frequency protein movement, and to reduce the rise time of the digital filter and thus increase the fraction of each resistive pulse that could be analyzed (Figure 4.2-C). Unexpectedly, the values that we determined for the shape of eight proteins after filtering the data with this low-pass filter were not systematically attenuated as the simulations in Figure 4.2-C predicted; rather, they varied evenly over all proteins (slope = 1.07). Figure 4.4 shows that this better-than-expected agreement between simulation and experiment arose, at least in part, from off-axis effects that influenced resistive pulse amplitudes and acted to shift determined protein shapes toward more extreme values (Figure 4.4).

In order to determine if the analysis approach presented here could be applied in the context of commonly used anti-adhesive coatings other than lipid bilayers, we also performed experiments with Tween-20 coated nanopores (see Figure 4.8), following the protocol of Li *et al.*[74] This coating is attractive because it is more straightforward to prepare than lipid bilayer coatings of high quality. We found that while volume estimates from a Tween-20 coated pore agreed with reference values, the estimates for length-to-diameter ratio were skewed toward an m value of 0.5, and estimates for dipole moment showed a weak correlation with reference values (see Figure 4.8-D). Additionally, the event frequencies and protein dwell times that we observed with Tween-20 coated nanopores did not correspond to freely translocating proteins. For example, we recorded an average event frequency of 1.7 Hz for

Alcohol Dehydrogenase (ADH) at a protein concentration of 80 nM in a Tween-20 coated pore. A theoretical analysis following the approach by Plesa *et al.* [42] predicts a detection frequency of only 0.02 Hz at this protein concentration. The reason why we experimentally observed almost 100-fold more frequent translocations than theoretically expected at the bandwidth of our experiments is that the protein's residence times in the Tween-20 coated pore were significantly longer than predicted. We attribute this observation to nonspecific interactions with the pore walls [45] possibly in combination with electroosmotic flow (EOF) mediated by residual surface charge in nanopores coated with Tween-20.[75] In contrast, when we used nanopores that were coated with a lipid bilayer, the theoretical predictions of event frequencies were in excellent agreement with experimentally observed event frequencies. Specifically, at a concentration of 10 μM ADH, we observed an event frequency of 1.2 Hz using a lipid bilayer coated pore; the theoretical prediction was 1-5 Hz at this concentration. If significant non-specific interactions of the protein with the pore wall would have been present, then our experimentally observed detection frequency would again have been significantly higher than the theoretical prediction, and we would have expected to observe differences between the shortest and longest translocation events for a given protein (see Figure 4.9, 4.10).

While the Tween-20 surface coating could not completely circumvent EOF and adhesive interactions with the pore wall, it did, however, prevent clogging of the nanopore. If we used pores in silicon nitride without any surface coating, we encountered clogging during translocation experiments of Ferritin and IgM proteins at typically employed concentrations between 0.1 and 1 μM which terminated the experiments. Tween-20 also enabled the accurate estimation of protein volumes (see Figure 4.8-C) as reported previously.[44, 74] Adhesive interactions with the nanopore wall, however, led to erroneous quantification of dipole moments and these interactions also led to inaccurate determination of protein shape. Therefore, the results with Tween-20 highlight the critical advantage of lipid bilayer coatings in providing non-stick synthetic nanopores with almost completely suppressed EOF.[76, 77]

4.2.3 Off-Axis Effects Distort Resistive Pulse Magnitudes

The analysis that we used to determine the shape and, in particular, the dipole moment of proteins requires that the protein is able to rotate unperturbed within the nanopore. It follows that the diameter of the nanopore must be larger than the largest dimension of the protein, which inherently means that the protein has space to diffuse laterally within the pore during its translocation.[15] Proteins that translocate untethered, therefore, may be electrically sampled at the center of the pore, at the pore wall, or anywhere in between, limited only by steric hindrance and random diffusion.[30, 44] When a particle transits a nanopore not through the very center but at some radial distance b from the central pore axis, it distorts the electric field within the pore asymmetrically.[30] This asymmetrical disruption produces a larger-than-expected resistive pulse whose magnitude depends on the particle's off-axis distance (b) and on the ratio of particle diameter to pore diameter (Figure 4.4-A,B). This phenomenon, known as off-axis effects, has been studied sporadically in the context of Coulter counters since the 1970s [78, 79]; it has attracted renewed consideration as research groups are beginning to perform finite element simulations on nanopore systems, and as the analysis of resistive pulses from single macromolecules is becoming increasingly sophisticated and information-rich.[30, 80–83]

In order to provide insight into the extent to which off-axis effects influence the parameter estimates of freely translocating non-spherical proteins - as opposed to spherical proteins [30] - we performed finite element simulations of ellipsoidal particles passing through a cylindrical nanopore at various distances from the central axis. Figure 4.4 (C-F) shows the results of these simulations, carried out for relatively extreme oblate-shaped and prolate-shaped particles in their minimum and maximum orientations relative to the electric field. We found that for the range of nanopore and protein sizes that are typically used for protein characterization, and in the most extreme off-axis scenario (prolate-shaped particle, γ_{max}, b_{max}), these effects can distort the resistive pulse magnitudes of non-spherical particles by up to 18%. An untethered protein may therefore sample its maximum blockade orientation

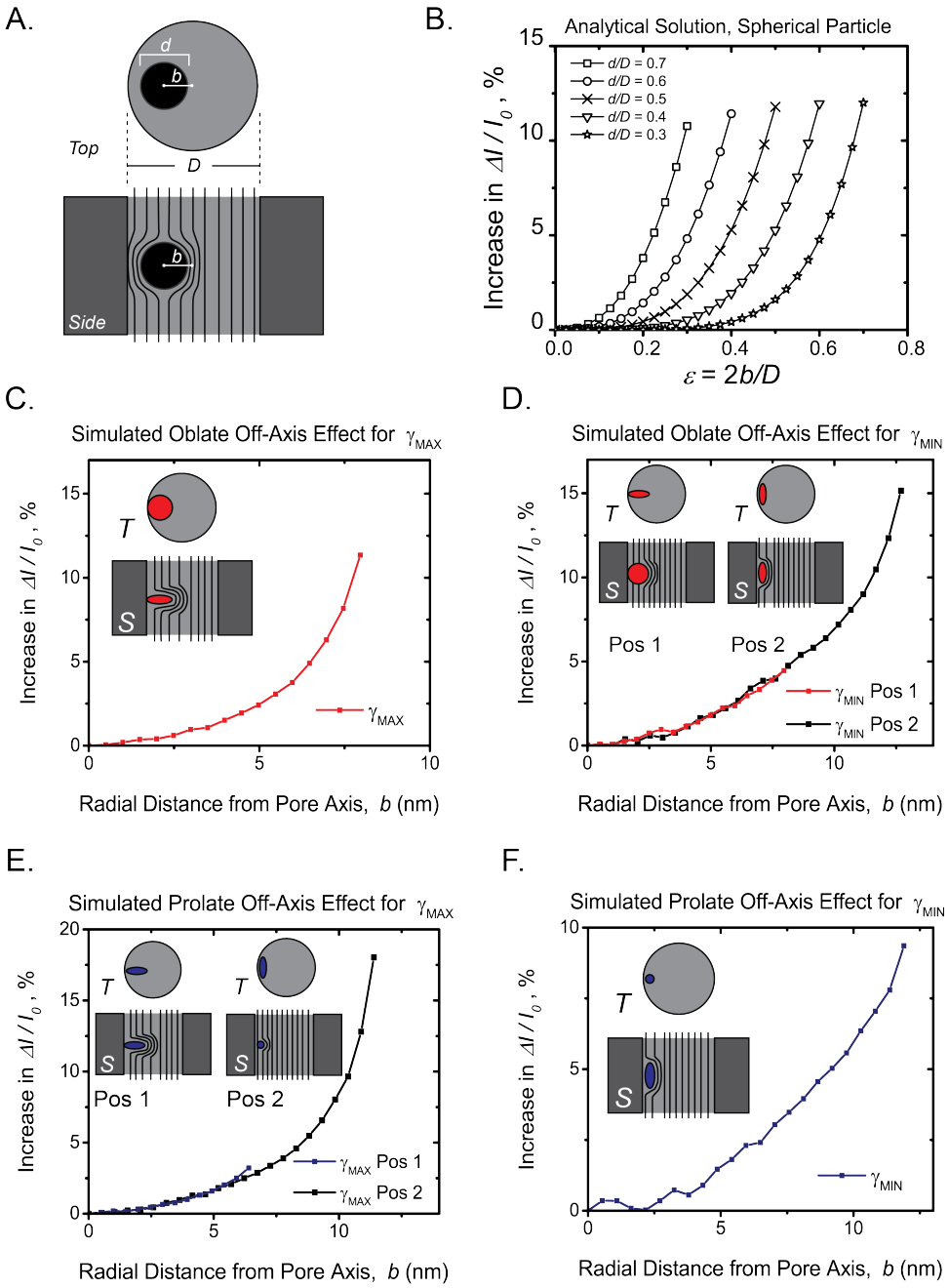


Figure 4.4. Quantified effects of off-axis translocation. A) Schematic of a spherical particle (black) passing through a pore (grey). The diameter of the particle is denoted as *d*, the diameter of the pore is labeled *D*, and the radial distance from the center of the pore is labeled *b*. B) Analytical solutions according to Qin *et al.* showing the magnitude of off-axis effects on the magnitude of the resistive pulse labeled as an increase in $\Delta I/I_0$ in the scenario where a perfectly spherical particle transits a cylindrical nanopore.(19) The x-axis represents ε, which relates the off-axis distance to the pore diameter. C-F) Plots generated using finite element simulations (COMSOL) showing the increase of the magnitude of resistive pulses for both oblate (C-D) and prolate (E-F) proteins at orientations that produce a maximum blockade (C,E) and a minimum blockade (D,F). For the oblate scenario, the simulation was conducted using a particle with a size and shape similar to anti-biotin IgG (m = 0.2, Λ = 275.6 nm^3) passing through a pore with a 30 nm diameter and a 30 nm length. For the prolate scenario, the simulation was conducted using a particle with a size and shape similar to G6PDH (m = 3.0, Λ = 268.4 nm^3) passing through a pore with a 30 nm diameter and a 30 nm length. For non-spherical particles, we defined b as the distance from the central axis of the pore to the center of mass of the particle. Adapted with permission.[46] Copyright 2019, American Chemical Society.

near the pore wall, and its minimum blockade orientation near the central pore axis, leading to a larger difference between maximum and minimum blockade values and thus to a more extreme estimate for protein shape than if the protein transited only through the central pore axis. With regard to the resulting error in shape estimates, this distortion falls within the green shaded regions in Figure 4.3-B and likely contributes to the spread of the data in Figure 4.3-A. We see that prolate-shaped particles are the most susceptible to variation; this result agrees with the uncertainty in the length-to-diameter ratio estimates for prolate-shaped proteins. We defined the threshold for detecting a resistive pulse as $5\times$ the standard deviation of the baseline noise, and thus the magnitude of each resistive pulse can deviate up to 20% simply due to noise from the recording setup. We suggest that the off-axis effect on resistive pulse signals can be viewed as another contributor to the overall noise in the system, with a frequency component related to the lateral diffusion coefficient of the protein within the pore. Hence, off-axis effects contributed to the variability in calculated parameter values - especially for analyses of individual events - but they did not preclude the determination of protein parameters, as evidenced by the agreement of parameter estimates with reference values in Figure 4.3. Furthermore, and fortuitously for this work, off-axis effects, which led to estimates of particle shape that were more extreme than those from experiments without off-axis effects, offset the effects of sampling rate and filtering, which led to estimates of particle shape that were less extreme than those revealed from experiments with adequate bandwidth (*e.g.* greater than 500 kHz). It appears, therefore, that these two opposing effects canceled each other to some extent and led to the good agreement between experimental and reference length-to-diameter ratios reported in Figure 4.3. Based on these arguments, off-axis effects will become more dominant in scenarios with higher bandwidths and lower noise levels than the recording setup used here and will need to be considered in future efforts to improve analysis methods of particles translocating freely and untethered through nanopores.

4.3 Conclusions and Future Prospects of Label-Free Analysis with Nanopores

This work estimates three distinct protein parameters from the resistive pulses generated by single untethered proteins passing through a lipid-coated synthetic nanopore. Among these parameters is dipole moment, which can be quantified on a single-molecule level for unmodified proteins in solution. No other technique has this capability, and given the increasing importance of dipole moments for formulations of monoclonal antibodies [72], this capability may accelerate the development of formulations for subcutaneous administration of therapeutic antibodies, which is the fastest growing class of therapeutics.[84] In addition, dipole moment is an excellent protein descriptor that is orthogonal to protein volume and shape [15], such that simultaneous quantification of these three parameters in sub-millisecond time frames of unmodified proteins in solution may be a first step toward a plug-and-play benchtop protein analysis system that counts and characterizes single proteins. The approach introduced here can likely be optimized; ongoing improvements in SNR will reduce the spread in parameter estimates [85–87], and further increases in recording bandwidth through improved CMOS current amplifiers [59] will increase event capture rates, resolve larger fractions of t_d distributions, and monitor information about protein rotation and shape at smaller time steps.

The work presented here reaffirms the critical importance of anti-adhesive coatings in nanopore-based analyses that rely on translational and rotational dynamics of proteins in an electric field [15, 55, 58, 74, 88, 89], and highlights the need for developing future systems that confine proteins to a single translocation axis rather than allowing them to diffuse laterally within the pore. But perhaps the most promising aspect of the nanopore-based analysis technique presented here is that it has the potential to probe native proteins and protein complexes that are transient in nature, including amyloids and amyloid oligomers without the need for modifications.[3, 4] Interest in characterizing this class of heterogeneous protein

analytes with synthetic nanopores is growing [44, 55, 90], and the single-particle analysis approach that we present here adds additional protein descriptors such as shape and dipole moment to particle volume and charge. Together these parameters may be useful for correlating the physical characteristics like size and shape of various amyloid species with their toxicity [91] as well as for defining and detecting biomarkers that reveal disease progression or the efficacy of therapeutics at early stages of amyloid-induced neurodegenerative diseases.[8, 92]

4.4 Supplementary Notes and Figures

4.4.1 Convolution Fitting Model

The core of the convolution fitting model is a U-shaped probability distribution (derived previously by Yusko *et al.*[15]) that describes the probability of observing some orientation-dependent electrical shape factor, γ, as a protein undergoes biased rotation in an electric field:

$$P(\gamma)d\gamma = \frac{1}{A} cosh \left[\frac{E\mu \left(\frac{\gamma-\gamma_\perp}{\gamma_\parallel-\gamma_\perp} \right)^{\frac{1}{2}}}{k_B T} \right] \left[\frac{1}{\pi[(\gamma-\gamma_\perp)(\gamma_\parallel-\gamma_\perp)]^{\frac{1}{2}}} \right] d\gamma \tag{4.1}$$

and

$$P(\gamma)d\gamma = \frac{1}{A} cosh \left[\frac{E\mu \left(\frac{\gamma-\gamma_\parallel}{\gamma_\perp-\gamma_\parallel} \right)^{\frac{1}{2}}}{k_B T} \right] \left[\frac{1}{\pi[(\gamma-\gamma_\perp)(\gamma_\parallel-\gamma_\perp)]^{\frac{1}{2}}} \right] d\gamma \tag{4.2}$$

where A is a normalization constant for integration

$$A = \int_{\gamma_\parallel}^{\gamma_\perp} cosh \left[\frac{E\mu \left(\frac{\gamma-\gamma_\perp}{\gamma_\parallel-\gamma_\perp} \right)^{\frac{1}{2}}}{k_B T} \right] \left[\frac{1}{\pi[(\gamma-\gamma_\perp)(\gamma_\parallel-\gamma_\perp)]^{\frac{1}{2}}} \right] d\gamma \tag{4.3}$$

and

$$A = \int_{\gamma_\parallel}^{\gamma_\perp} cosh \left[\frac{E\mu \left(\frac{\gamma-\gamma_\parallel}{\gamma_\perp-\gamma_\parallel} \right)^{\frac{1}{2}}}{k_B T} \right] \left[\frac{1}{\pi[(\gamma-\gamma_\perp)(\gamma_\parallel-\gamma_\perp)]^{\frac{1}{2}}} \right] d\gamma \tag{4.4}$$

$\gamma_\parallel$ describes the orientation of the ellipsoid when its singleton axis (ellipsoids of rotation have two axes identical in length and one axis that is different, which we denote here as a singleton axis) is aligned parallel to the electric field; $\gamma_\perp$ describes the ellipsoid's singleton axis aligned perpendicular to the electric field. If the particle/protein is an oblate, then

$\gamma_{\parallel} \propto (\Delta I/I_0)_{max}$, and $\gamma_{\perp} \propto (\Delta I/I_0)_{min}$, or if the particle is a prolate, e.g. $\gamma_{\perp} \propto (\Delta I/I_0)_{max}$, and $\gamma_{\parallel} \propto (\Delta I/I_0)_{min}$. The reason that we have two probability distribution equations is to account for dipole-moment bias (or skew) in either the $\gamma_{\perp}$ direction, or in the $\gamma_{\parallel}$ direction. In practice, we fit the data with both equations and determine which fit to accept by comparing their goodness of fit (e.g. is the data skewed toward $\gamma_{\perp}$ or $\gamma_{\parallel}$?). Because they are proportional, we substitute in $\Delta I/I_0$ values for γ to determine the particular U-shaped distribution for each set of data, as we alluded to in the introduction and as explained in detail by Yusko *et al.*[15] The minimum, or left side of the 'U', corresponds to $(\Delta I/I_0)_{min}$ and the maximum or right side of the 'U', is $(\Delta I/I_0)_{max}$. This probability distribution does not account for noise in the current recording, so we iteratively convolve the U-shaped distribution with a standard Gaussian noise distribution function, shown below:

$$P((\Delta I/I_0)_{\sigma}) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(\Delta I/I_0)_{\sigma}^2}{2\sigma^2}} \tag{4.5}$$

The expression describing the probability of observing any blockade value, $\Delta I/I_0$, is therefore:

$$P(\Delta I/I_0) = P((\Delta I/I_0)_{\gamma}) \otimes P((\Delta I/I_0)_{\sigma}) \tag{4.6}$$

In the iterative convolution, performed using MATLABs lsqcurvefit function, we initialize a total of four parameters: the protein's permanent dipole moment to 550 Debeye, which is an approximate average dipole moment across all proteins [15], $(\Delta I/I_0)_{min}$ as the 10^{th} percentile of the input data, $(\Delta I/I_0)_{max}$ value as the 90^{th} percentile of the data, and the spread σ as the standard deviation of the current trace (we also now constrain this value to be greater than or equal to the standard deviation of baseline noise). It is important to note that while the convolution itself is performed on probability distributions, the fitting function compares cumulative distributions during each iteration to circumvent binning artifacts. As such, the algorithm runs in the following steps: (1) we generate an experimental cumula-

tive distribution from the data that represents the distribution we would like to match with our fitting function, (2) the fitting function performs a convolution between the theoretical U-shaped probability distribution and theoretical Gaussian noise (equations above) substituting only initial parameter estimates, (3) the function then integrates that convolution result to generate a predicted cumulative distribution,(4) after which it compares this predicted cumulative distribution to the experimental cumulative distribution from step 1, (5) it then adjusts the input parameters accordingly, and (6) iterates again until the experimental and predicted cumulative distributions align within some defined tolerance (Termination tolerance on the function value, or 'TolFun', and Step Tolerance, or 'TolX' both set to 1×10^{-9}, and minimum change in variables for finite-difference gradients, or 'DiffMinChange' set to 0.5). As expected, this fitting procedure produces four values corresponding to the initial input values; that is, $(\Delta I/I_0)_{min}$, $(\Delta I/I_0)_{max}$, dipole moment (μ), and distribution spread (σ). We can then take the estimates for $(\Delta I/I_0)_{min}$ and $(\Delta I/I_0)_{max}$, (again, either side of the U-distribution), and substitute them into the system of equations outlined by Golibersuch [31], and later described by Yusko *et al.* [15], to determine a corresponding protein shape and volume. Briefly, the relationship between normalized resistive pulse magnitude $\Delta I/I_0$, and the volume of the associated translocating particle traces back to Maxwell's derivation [93]:

$$\frac{\Delta I}{I_0} = -\frac{4\Lambda\gamma}{\pi d_p^2(l_p + 0.8d_p)} \left(\frac{1}{1 - 0.8\left(\frac{d_M}{d_p}\right)^3} \right) \tag{4.7}$$

where Λ is the volume of the particle, γ is the particle's electrical shape factor ($\gamma = 1.5$ for perfect spheres), d_p is the diameter of the pore, l_p is the length of the pore, and d_M is the diameter of the particle. Note that the expression in large brackets on the right represents a correction factor which we set to equal 1, as we performed the experiments using nanopores much larger than the longest dimension of the protein(e.g. $d_M < 2d_p$). Upon closer examination of the left-hand side of the equation, we see that the denominator

represents the volume of the pore with a slight modification of the length of pore, which we define as an effective length, e.g. $l_{eff} = (l_p + 0.8d_p)$. The '4' in the numerator simply accounts for the use of the square of the diameter (rather than the square of the radius) in the denominator. The nanopore chips that we use for these experiments are fabricated using an ion beam sculpting process, giving them a unique geometry where the nanopore sits adjacent to a larger channel, ∼250 nm long and ∼100 nm in diameter.[70] Expanding on previous work, we now account for this channel geometry by first determining the total resistance in the system (note $\Delta I/I_0 = R_0/\Delta R$), and then deriving a new effective length term that incorporates channel dimensions, giving the final form:

$$\frac{\Delta I}{I_0} = \frac{R_{tot}}{\Delta R_{tot}} = \frac{\gamma \Lambda r_c^2}{\pi r_p^2 \left(r_c^2 l_p + \frac{\pi}{2} r_p r_c^2 + r_p^2 l_c + \frac{\pi}{4} r_p^2 r_c \right)} \tag{4.8}$$

where r_c is the radius of the channel, r_p is the radius of the pore, and l_c is the length of the channel. We first rearrange this equation into an aggregate variable, g, to isolate all of the constants:

$$g = \frac{\Delta I/I_0}{\gamma \Lambda} = \frac{r_c^2}{\pi r_p^2 \left(r_c^2 l_p + \frac{\pi}{2} r_p r_c^2 + r_p^2 l_c + \frac{\pi}{4} r_p^2 r_c \right)} \tag{4.9}$$

Then we substitute this variable g into the following system of equations for the well-described depolarization factors n for oblate-shaped particles [31, 64, 94, 95]:

$$n_{min,oblate} = \frac{2n_{max,oblate}}{(2n_{max,oblate} - 1)} \tag{4.10}$$

$$n_{max,oblate} = \left(\frac{m \cdot cos^{-1}(m)}{(1 - m^2)^{\frac{3}{2}}} - \frac{m^2}{1 - m^2} \right)^{-1} \tag{4.11}$$

$$\Lambda_{oblate} = \frac{(\Delta I/I_0)_{max}}{g \cdot n_{max,oblate}} = \frac{(\Delta I/I_0)_{min}}{g \cdot n_{min,oblate}} \tag{4.12}$$

And into the following depolarization factor equations for prolate-shaped particles:

$$n_{min,prolate} = \left(\frac{m^2}{m^2 - 1} - \frac{m \cdot cosh^{-1}(m)}{(m^2 - 1)^{\frac{3}{2}}} \right)^{-1} \tag{4.13}$$

$$n_{max,prolate} = \frac{2n_{min,prolate}}{(2n_{min,prolate} - 1)} \tag{4.14}$$

$$\Lambda_{prolate} = \frac{(\Delta I/I_0)_{max}}{g \cdot n_{max,prolate}} = \frac{(\Delta I/I_0)_{min}}{g \cdot n_{min,prolate}} \tag{4.15}$$

To solve, we substitute in three known values, $(\Delta I/I_0)_{min}$, $(\Delta I/I_0)_{max}$, and g, and determine values for the two unknowns, Λ and m, which represent volume and length-to-diameter ratio, respectively. We determine an oblate solution from all data sets, and some data sets produce both an oblate and a prolate solution. This is an aspect of the fitting procedure that we previously discussed at length [15] and ultimately arises because we distill a rich data array of current measurements into two components:$(\Delta I/I_0)_{min}$ and $(\Delta I/I_0)_{max}$. When these two values differ considerably, e.g. $(\Delta I/I_0)_{min} < \frac{1}{3}(\Delta I/I_0)_{max}$, they can only be de-

scribed by an oblate-like shape, but in the scenario where $(\Delta I/I_0)_{min} > \frac{1}{3}(\Delta I/I_0)_{max}$,their separation could correspond to the rotation of a slightly disk-shaped (oblate) protein, or it could correspond to the rotation of a slightly rod-shaped (prolate) protein, resulting in two possible solutions. This is an intrinsic challenge with this analysis methodology, and we propose that it could be solved in future low-noise implementations of the technique by comparing auxiliary protein parameters like rotational diffusion coefficient and using those to discriminate between the two approximations. We expect to observe differences in rotational diffusion coefficients between oblate-shaped and prolate-shaped ellipsoids, and indeed we did measure differences between the rotational diffusion coefficients of oblates and prolates. [15] Alternatively, it may be beneficial to move away from a pure ellipsoidal model and toward a bead-like model for non-spherical particles to improve the shape resolution and the ability to discriminate between different current traces.

We carry out the procedure for population-based and individual-event analyses in almost exactly the same fashion. The only difference is whether the experimental cumulative distribution that we recreate with the fitting algorithm is generated from a set of maximum blockade values (population-based) or if it is generated from all of the individual data points contained within a single long event (individual-event).

4.4.2 Population-Based Analyses

In population-based analyses, we created a single probability distribution by combining the maximum $\Delta I/I_0$ value of all protein translocation events that had dwell times (t_d) longer than 30 μs (Figure 1-C in the Main Text, green dots). In order to calculate one length-to-diameter ratio and one volume value for a given experiment, we used this approach to estimate ellipsoidal shape and volume values for nine proteins (Figure 4.5, Table 4.3). The volume and length-to-diameter ratio estimates were in good agreement with reference values; the average deviation of length-to-diameter ratios was less than 30% for all proteins (less than 20% for oblate-shaped proteins), and the average deviation of volumes was less than 40% for

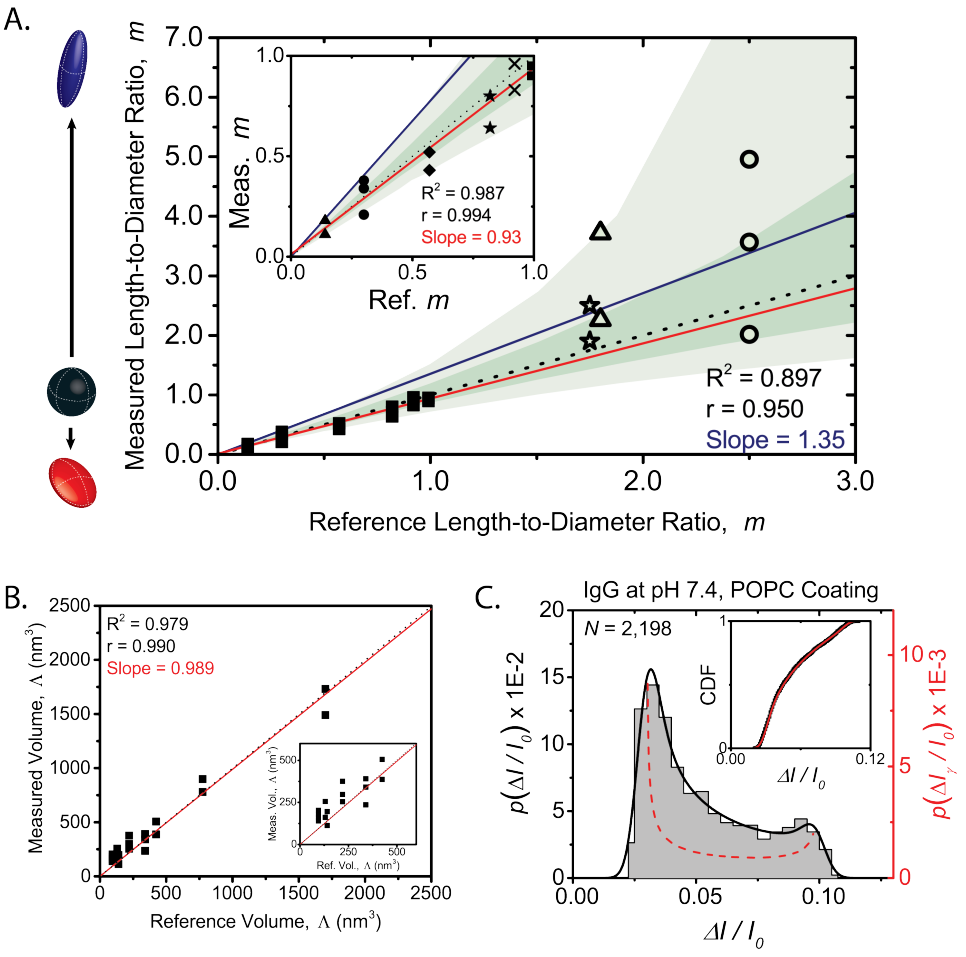


Figure 4.5. Results of population-based analyses for untethered proteins. A) Comparison of estimate and reference length-to-diameter ratio values for the three prolate proteins Fab (open stars), α-Amylase (open triangles), and G6PDH (open circles) and for (INSET) the 6 oblate proteins IgM (triangles), IgG (circles), BSA (diamonds), ADH (stars), Ferritin (X's), and Streptavidin (squares) - tested in this study. Note m = 1 represents spheres. The blue line is the linear fit for all proteins in the study, the black dotted line shows an ideal 1:1 relationship, and the red line is a fit to data from oblate-shaped proteins only. The dark and light green regions represent $\pm$10% and $\pm$25% uncertainty in minimum and maximum blockade values, respectively. B) Comparison of volume estimates (including replicates) with reference values for all proteins examined in the study. The inset corresponds to smaller volume proteins, the red line is the linear fit, and the black dotted line is the 1:1 relation. C) Example probability distribution and cumulative distribution (inset) fits for IgG protein to determine shape and volume (see Section 4.4.1 for details). Refer to Table 4.3 for precise values shown in scatter plots in panels A and B. Adapted with permission.[46] Copyright 2019, American Chemical Society.

all proteins. We performed duplicate or triplicate experiments on each protein, and denote replicates as unique points with the same symbol (Figure 4.5-A,B). As shown in Figure 1 in the main text, non-spherical proteins fluctuate between minimum and maximum blockade orientations while transiting a nanopore. Oblate-shaped proteins, as they deviate further away from a sphere (e.g. become flatter disks), produce large current fluctuations when rotating between minimum and maximum orientations during transit, and thus provide rich information about their approximate shape. The ratio of minimum and maximum blockade values for prolate-shaped proteins, however, levels off asymptotically as the protein becomes closer to an extreme rod-like shape (see Section 4.4.1).[31] Small experimental uncertainties in observed minimum and maximum blockade values lead to larger uncertainties in length-to-diameter estimates for prolate-shaped proteins than for oblate proteins (Figure 4.5-A, green shaded regions). This effect is visible in the spread in m values that we observe for prolate-shaped proteins.

Since Han *et al.* determined the volume of BSA with a synthetic nanopore in 2006 [1], the majority of synthetic nanopore-based protein characterization methods have estimated protein volumes by averaging across large sets of maximum blockade values ($\Delta I/I_0$, or simply ΔI) and assuming spherical protein geometry.[53, 54, 58] This population-based approach has the advantage that it can be carried out on shorter events than individual-event analyses (30 μs vs 150 μs), permitting the calculation of protein shape and volume in experiments where it is difficult to gather a sufficient amount of events longer than 150 μs, as was the case here for Ferritin (see Section 4.4.3 for details on the 30 μs threshold selection). Furthermore, we demonstrated previously with tethered proteins that shape and volume estimation is typically more accurate using population-based methods compared to individual event methods.[15] However, population-based methods rely on samples with only one pure protein in solution. These methods cannot be easily applied to mixtures of different proteins if the proteins are similar in volume (overlapping $\Delta I/I_0$ distributions), limiting their practicality.

4.4.3 Threshold Selection for Population-Based Analyses

We selected 30 μs as the minimum event dwell time for incorporation into the population-based analysis. Previously, we only accepted events longer than 50 μs, as we digitally low-pass filtered the data at 15 kHz. Here, with a filter frequency of 50 kHz, we performed a filter response analysis similar to that demonstrated by Plesa *et al.*[42] Briefly, we added digital square pulses of varying durations to experimentally recorded baseline noise at full bandwidth. For reference, the Axopatch 200B has an intrinsic bandwidth of approximately 57 kHz. We applied a digital low-pass filter with a cutoff frequency of either 50 kHz or 15 kHz to determine the extent of signal truncation at various pulse widths for the respective cutoff frequencies. Figure 4.6 shows these filter responses at the two frequencies. The plots represent an ideal response scenario for a real translocation event; we chose 30 μs rather conservatively to prevent any truncation of experimental data. While reducing this threshold from 50 μs to 30 μs may seem like a marginal change, it increased the fraction of time-resolved events by approximately 60%, on average.

4.4.4 Threshold Selection of Adequate Dwell Times for Individual-Event Analysis

Event duration is a critical parameter with regard to individual-event analysis. We demonstrated previously that the accuracy of parameter estimates improves as a function of event duration; conversely, analyzing individual events with relatively short durations introduces more uncertainty into the determined values. Because we reduced the noise of our experimental setup, increased the recording bandwidth, and analyzed untethered proteins, we were able to select a lower threshold than the 400 μs chosen in previous work. In order to determine which particular threshold to select, we performed an investigation of individual-event analysis for a group of events longer than 400 μs for BSA protein. We performed more than 200 individual convolution model fits on each event, analyzing the first 2 μs, then 4 μs, then

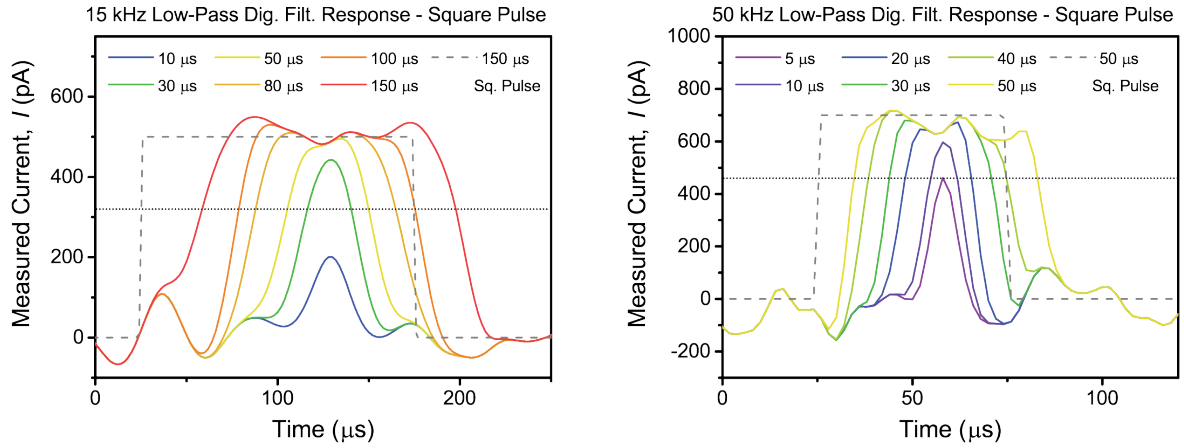


Figure 4.6. Filter responses for determining minimum event duration for population-based analyses. (Left) Data filtered at 15 kHz with pulses ranging from 10 to 150 µs in duration. The square pulse shown in dashed grey represents the longest pulse length (150 µs). (Right) Data filtered at 50 kHz with pulses ranging from 5 µs to 50 µs. The square pulse shown in dashed grey represents the longest pulse length (50 µs). The black dotted line represents the resistive pulse amplitude - 5× the standard deviation of the baseline noise - required for consideration as a translocation event. Note that while all pulses were applied in the same region of the noise trace for each particular cutoff frequency, the 15 kHz pulses were added in a different region of the trace than the 50 kHz pulses, for demonstration purposes. Additionally, we offset the square pulses by 25 µs for visual clarity. Refer to Table 4.3 for precise values shown in scatter plots in panels A and B. Adapted with permission.[46] Copyright 2019, American Chemical Society.

6 µs, and so on until the full length of the event. Figure 4.7 shows the results of this analysis - more specifically, it shows the dependence of event duration on length-to-diameter ratio estimates - for recorded long events of BSA.

4.4.5 Lipid Bilayers Provide Unperturbed, Free Translocation

Chemical tethering, non-specific adhesion, and electroosmotic flow can all slow down protein transit through a nanopore.[75] In the absence of these phenomena, many protein translocation events are too rapid (< 10 µs) to detect using standard recording equipment such as the Axopatch 200B amplifier.[42] Untethered protein experiencing biased diffusion in an electric field requires roughly 300-fold greater concentrations to produce the same frequency of observed events as protein tethered to a lipid bilayer.[42, 58] When we added protein concentrations between 1 nM and 100 nM to the upper fluid compartment of a lipid-coated nanopore, we observed few if any translocation events (f < 0.5 Hz). However, upon increas-

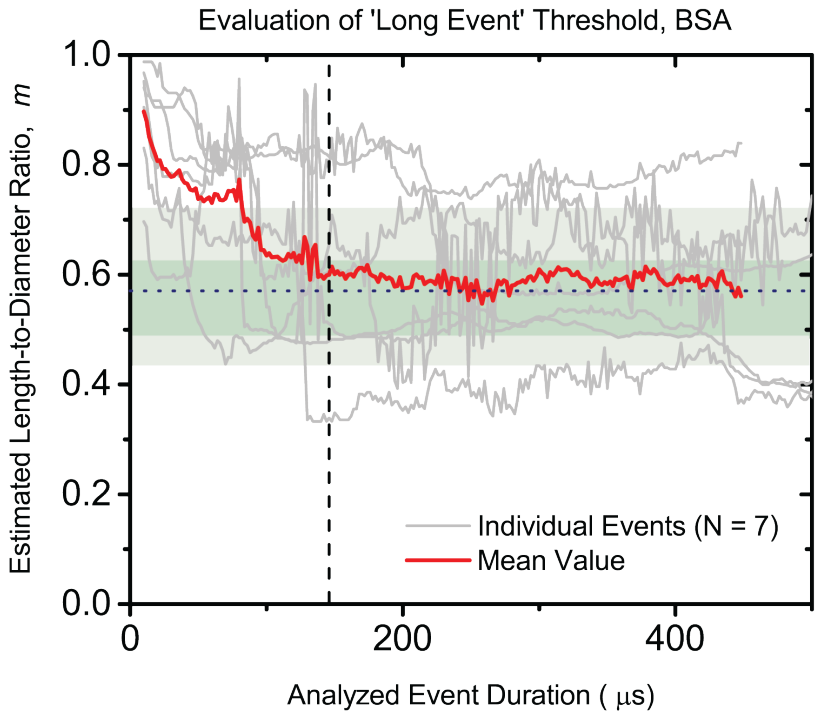


Figure 4.7. Determination of minimum event duration for individual-event analyses. Grey solid lines represent results of analyses for individual events longer than 400 μs of untethered BSA translocations, and each data point (separated by 2 μs) was generated through a unique convolution model fit. The red curve is the average value across all long events displayed. The blue dotted line is the reference shape value for BSA (0.57), and, for reference, the grey dashed line denotes the 150 μs threshold that we selected. Dark green and light green regions ±10% and ±25% uncertainty in minimum and maximum blockade values, respectively. Note that when taken individually, long event length-to-diameter estimates vary considerably around a reference value. When taken in aggregate, their most probable value typically converges to within ±20% of the reference length-to-diameter value. Adapted with permission.[46] Copyright 2019, American Chemical Society.

ing the concentration of protein to 1-10 μM, we observed event frequencies between 1 and 5 Hz, which is in agreement with Schrödinger's first passage probability relationship described more recently in this context by Ling and Ling.[73]

A concern with using relatively high protein concentrations for nanopore experiments is that more than one protein may reside in the pore at a given time, resulting in resistive pulse artifacts.[51] Taking a 100 kDa protein with a translational diffusion coefficient of 30×10^{-12} m^2 sec^{-1}, note that confined diffusion can be 5-fold slower than in bulk [67], a net protein charge of 3e in aqueous solution, in a system with 100 mV applied across a nanopore with a length of 24 nm and diameter of 21 nm after coating with a lipid bilayer, we approximate based on first passage probabilities that only 0.2% of all translocation events will last longer than the ∼20 μs temporal resolution threshold of an Axopatch 200B amplifier. The most probable t_d of a protein in this system is approximately 3.5 μs, so we are detecting only the tail end of this t_d distribution.[42, 73] If we detect three translocation events per second, a rate that was typical for the experiments in this work, then we estimate that approximately 1700 proteins are transiting the pore each second and on average the pore is occupied by any single protein 5.97 ms out of every second, or about 0.6% of the time. It follows that the probability of two proteins residing in the pore simultaneously in this scenario is 0.004%; this situation is thus unlikely to significantly affect the calculated parameters. These theoretical event frequency estimates are also in agreement with the Smoluchowski rate equation, which predicts that for this same system with a 21 nm diameter nanopore and a protein with a diffusion constant of 30×10^{-12} m^2 sec^{-1} at a concentration of 2 μM (12×10^{20} molecules m^{-3}), we would expect 2400 proteins to diffuse passively into the pore each second. As Plesa *et al.* showed [42], proteins with molecular weights in the range of 100 kDa tend to translocate a single nanopore at a frequency that is somewhere between 10% and 80% of what is predicted by Smoluchowski, consistent with the calculations here. Based on a most probable translocation time of 3.5 μs, this simplified calculation reveals that the pore is empty 99.4% of the time, and the probability of multiple proteins residing in the pore

simultaneously is well below 1%, assuming negligible protein-protein and protein-surface interactions.

4.4.6 Experiments with Tween-20 Coated Nanopores

We observed relatively low success rates when attempting to coat synthetic nanopores with lipid bilayers. We hypothesize that the ability of a particular pore to successfully coat is dependent on a combination of factors including its surface chemistry, surface roughness, aspect ratio, radius of curvature, and cleanliness.[41] In an attempt to alleviate some of this difficulty, we coated nanopores with Tween-20, in a similar fashion to that described by Li *et al.*[74] Briefly, we cleaned the nanopore chips using a Piranha solution, as described in Section 4.5.2. Once clean, we air dried the nanopore chips under a flow of nitrogen gas for 60 seconds, and then incubated them in a solution of 0.01% w/v fresh Tween-20 (dissolved in 1 M KCL, 10 mM Tris-HCl, 1 mM EDTA, pH 4.0) for 30 minutes. After incubation, we rinsed the chips with DI water and quickly dried them again with nitrogen gas before placing them into a standard experimental setup (Figure 1-A of the Main Text). In this fashion, we performed replicate experiments on all of the oblate proteins in the study and found that in both population-based and individual-event analyses on experiments with Tween-20 coated nanopores, the length-to-diameter ratio estimates were all skewed toward 0.5 oblate solutions, as shown in Figure 4.8.

4.4.7 Finite Element Simulations

We calculated the blockage current of particles in their off-axis position in Figure 4 of the main text by solving Poisson-Nernst-Planck [96–99] equations using COMSOL Multiplysics 5.3a.

$$\nabla^2\Phi = -\frac{F}{\varepsilon}\sum z_i c_i \tag{4.16}$$

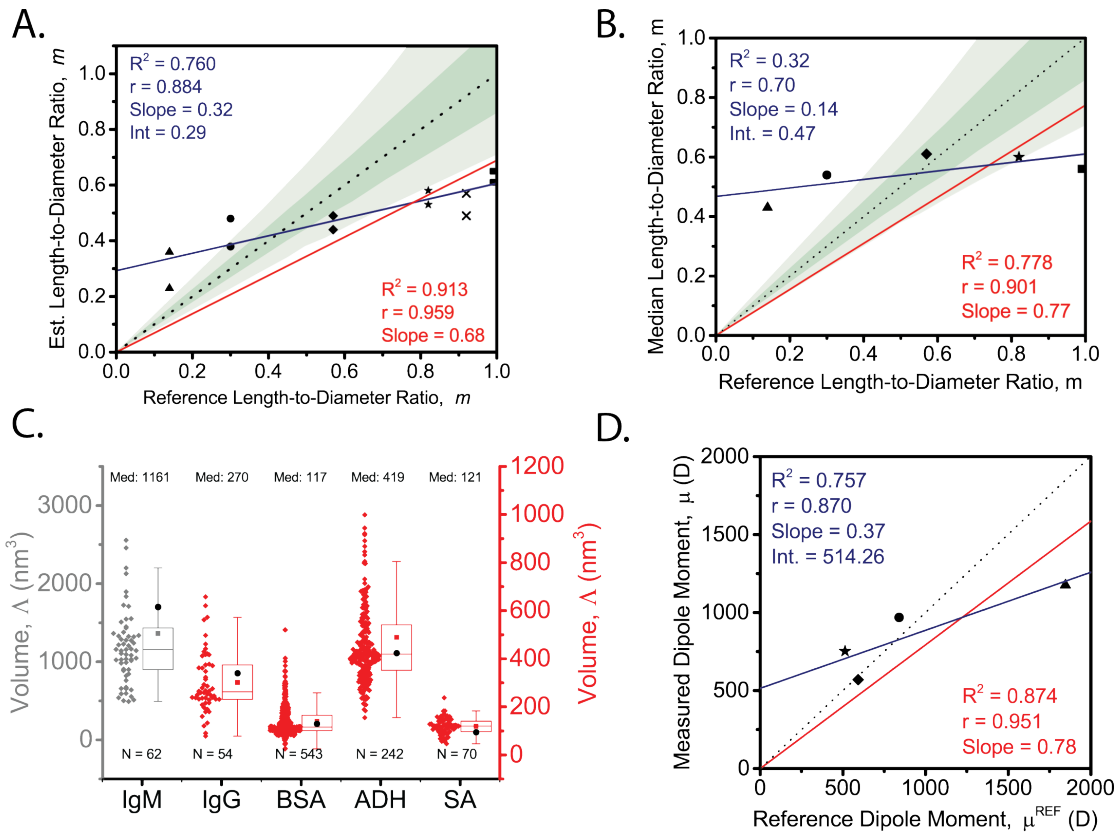


Figure 4.8. Results from experiments with untethered proteins passing through Tween-20 coated pores. (A) and individual-event (B-D) analyses. A. Shape calculations for the oblate proteins IgM (triangles), IgG (circles), BSA (diamonds), ADH (stars), Ferritin (Xs), and Streptavidin (squares) tested with Tween-20 coated pores. The blue dashed line shows an unconstrained linear fit, the red dashed line shows a 0-intercept linear fit, and the black dotted line shows a 1:1 relationship. B. Median length-to-diameter ratios calculated through individual event analysis. A-B. Dark green and light green regions $\pm$10% and $\pm$25% uncertainty in minimum and maximum blockade values, respectively. C. Volumes (nm^3) of all proteins, with IgM in grey and other proteins in red. Reference values are shown as black circles. D. Estimated dipole moments for all proteins investigated with Tween-20 coatings. Black dotted line represents the 1:1 ratio, and red solid line is the linear fit. Adapted with permission.[46] Copyright 2019, American Chemical Society.

Table 4.1. Constants and parameters used for COMSOL Multiphysics simulations. Adapted with permission.[46] Copyright 2019, American Chemical Society.

Constants	Values	Details
zK	1	the charge of K+
zCl	-1	the charge of Cl-
F	96485.3365 C· mol^{-1}	Faraday Constant
ε	80	dielectric constant of the fluidic medium
R	8.31 J· $(mol \cdot K)^{-1}$	gas constant
T	295 K	absolute temperature
DK	1.957×10^{-9} $m^2 \cdot s^{-1}$	diffusion coefficient of K+
DCl	2.032×10^{-9} $m^2 \cdot s^{-1}$	diffusion
Parameters	**Values**	**Details**
Φ	-0.1 V	applied voltage across the nanopore
c_i (K+ or Cl-)	1680 $mol \cdot m^3$	effective concentration of the K+ and Cl- ions according to the measured solution conductivity(30)
l_p	30 nm	length of the nanopore
d_p	30 nm	diameter of the nanopore
σ_{pore}	0 $mC \cdot cm^{-2}$	surface charge density of the lipid bilayer coated nanopore[76]
$\sigma_{particle}$	0 $mC \cdot cm^{-2}$	surface charge density of the particles
Λ	268 nm^3	volume of the particle
m_{oblate}	0.2	length-to-diameter ratio,m, of the oblate
$m_{prolate}$	3	length-to-diameter ratio, m, of the prolate

$$J_i = -D_i \nabla c_i - \frac{z_i F}{RT} D_i c_i \nabla \Phi \tag{4.17}$$

where i represents an ionic species (K+ or Cl-, in this case). Table 4.1 lists the constant parameters [100, 101] that we used in the simulations.

We performed off-axis simulations for three different particle orientations to represent γ_{max} and γ_{min}, obtained the open pore current as I_0 with absence of the particle, and reported the corresponding $\Delta I/I_0$ values in Figure 4 of the Main Text. We adjusted the mesh size in the simulation to achieve the best resolution within a reasonable time frame: finer for the nanopore domain, extra fine for the nanoparticles, and normal for other domains. Table 4.1 lists the experimental parameters that we used in the simulation.

4.4.8 Comparison of Resistive Pulses with Different Dwell Times

While we cannot exclude the possibility that a fraction of the long events may originate from artifacts such as wall interactions, we are confident that the *majority* of events arise from a single distribution and that the results from this largest fraction are representative of their respective protein populations. The reasons are:

- The theoretical frequencies of events with a duration longer than 25 μs that we determined using the Smoluchowski rate equation and the biased first passage probability relationship (Section 4.4.3) agreed with the frequencies that we observed experimentally using lipid bilayer-coated nanopores. In contrast, when non-specific interactions between proteins and the pore walls were not eliminated, then these theoretical estimates did not agree with the 100-fold more-frequent-than-predicted translocation events during experiments with Tween-20 coated nanopores, as discussed throughout Chapter 4.

- Estimated values for length-to-diameter ratio, volume, and dipole moment agree well with reference values across all proteins that we tested. Significant non-specific interactions with the pore wall would introduce errors, especially for the determined dipole moment, again as we indeed observed with the Tween-20 coatings.

- When we overlay histograms of measured current values from subpopulations of the shortest 10%, middle 10%, and longest 10% of events of a particular protein, we see excellent overlap, as shown in Figure 4.9. In other words, the shortest 10% of events lead to parameter estimates that are similar (albeit with larger uncertainty) to parameter estimates from the longest 10% of events, again suggesting no significant differences between these populations.

The translocation events that we observed did in fact originate from the tail of the dwell time distribution for each protein, but such a tail is expected in First Passage Time Probability theory,[73] and from a purely theoretical point of view, the proteins in this tail region

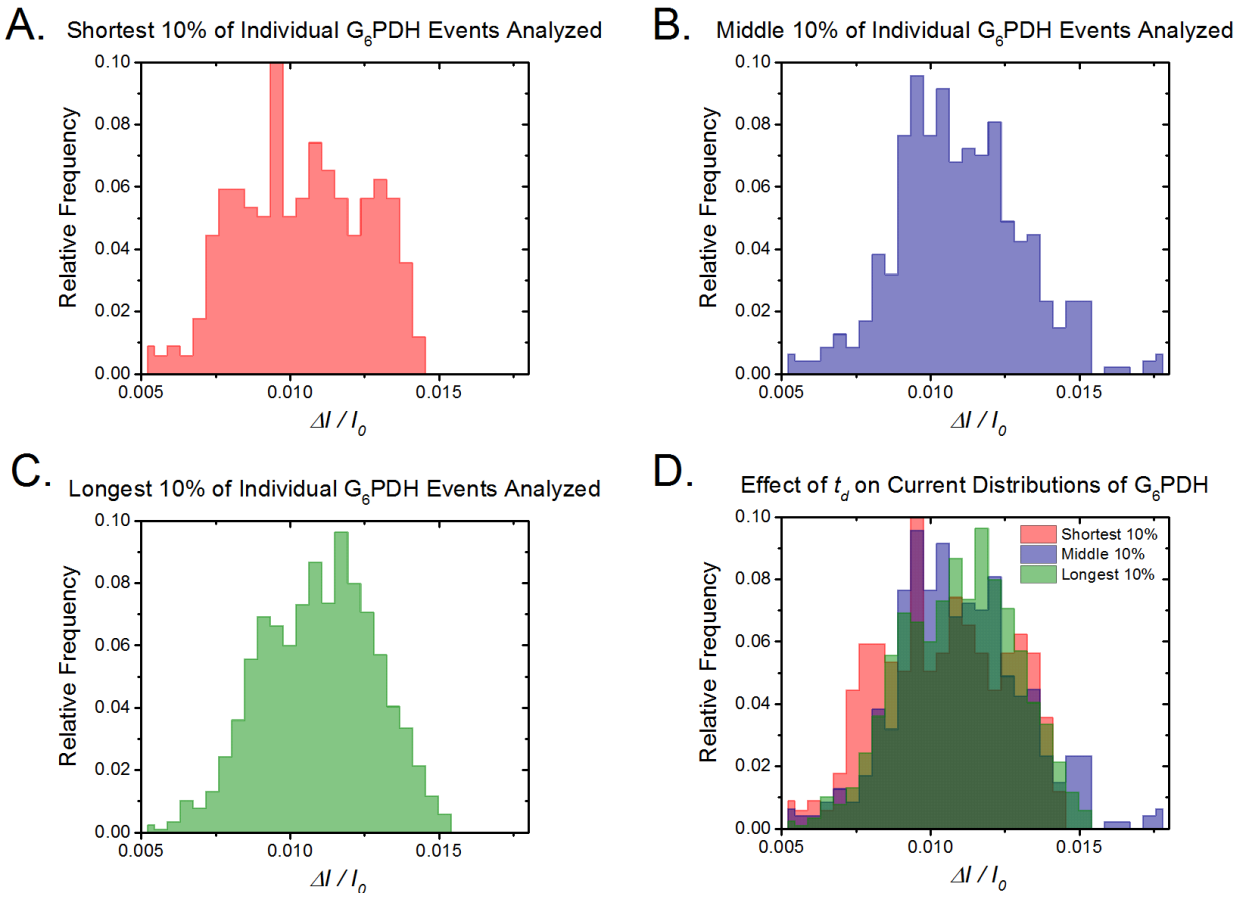


Figure 4.9. Overlay of current blockade distributions for events of G_6PDH with different dwell times. A) Shortest 10% of G_6PDH events with durations longer than 150 μs plotted as an ensemble histogram, B) Middle 10% - between 45th and 55th percentile of G_6PDH events with durations longer than 150 μs plotted as an ensemble histogram. C) Longest 10% of G_6PDH events with durations longer than 150 μs plotted as an ensemble histogram. D) Overlay of histograms, showing relative frequencies and alignment of minimum and maximum blockade values. Adapted with permission.[46] Copyright 2019, American Chemical Society.

do not physically differ from proteins in other regions of the distribution. In other words, the tail is inherent to the passage time distribution even in the absence of non-specific interactions. If significant non-specific interactions were to take place, then the number of events above a threshold dwell time value would be larger than the number predicted by the biased first passage time equation. This is the observation we made with Tween-20 coatings, while lipid bilayer coatings resulted in a good agreement between the theoretically predicted and experimentally observed number of events, therefore providing no indication for a significant fraction of artificially prolonged translocation events.

We demonstrated before [58] that lipid bilayer coatings eliminate non-specific interactions with proteins almost entirely. As discussed above, if a large fraction of the analyzed proteins were non-specifically interacting with the surface of the nanopore, we would expect to observe greater-than-predicted event frequencies, as was the case in the experiments with Tween-20 coatings (see Section 4.4.6). Note that the bilayer coatings largely eliminate electroosmotic flow [58, 76], which may also slow down proteins passing through nanopores coated with Tween-20 depending on the polarity of the applied electric field. The biased first passage time probability relationship, outlined by Ling and Ling [73] assumes no surface interactions or electroosmotic flow, and still shows that long-lived translocation events are expected to occur, albeit with probabilities that decrease with increased dwell times.

In the presence of nonspecific interactions with the membrane, we would expect clogging or issues with nanopore stability during long experiments when using protein concentrations greater than 1 μM. We did not observe these issues when running experiments with lipid bilayer coated nanopores, while uncoated nanopores frequently clog at protein concentrations greater than 1 μM (see Section 4.2).

While it is possible to observe a protein entering and exiting a nanopore, we are unable to track the radial (*i.e.* between pore walls and central axis) motion of a protein while it transits the pore. Proteins that occupy the pore for extended periods of time are more likely to sample all radial positions during transit, and we expect them to diffuse relatively freely

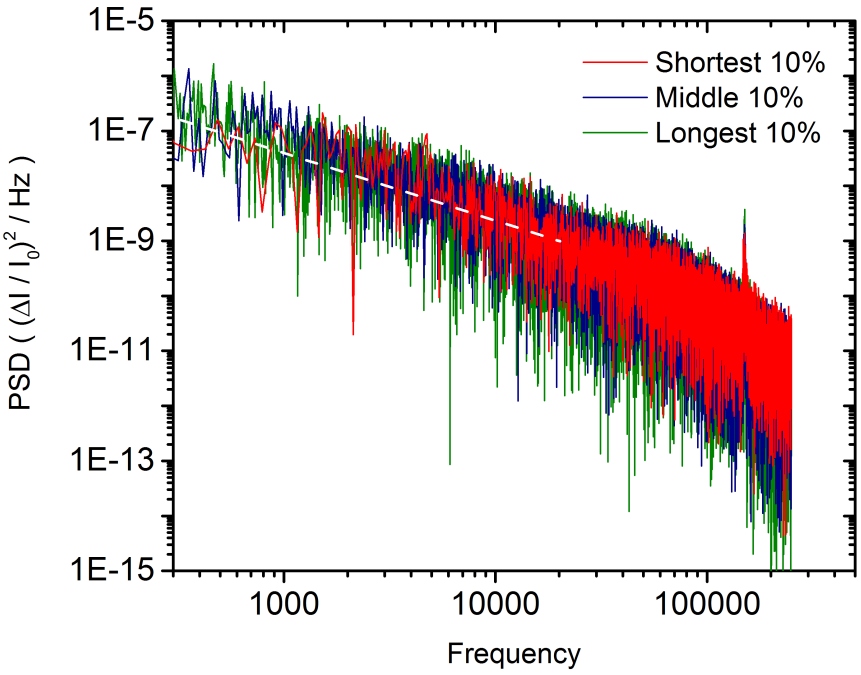


Figure 4.10. Comparison of power spectral densities between translocation events with different durations. Power spectral densities of the recorded current within the shortest 10%, middle 10%, and longest 10% of experimentally recorded translocation events of IgG_1. The white dashed line represents the $1/f$ noise. All data were digitally lowpass filtered at 50 kHz, as described in Section 4.5.2. Adapted with permission.[46] Copyright 2019, American Chemical Society.

between the central axis of the pore and pore walls. We have no evidence for significant interactions between the lipid bilayer coated nanopore and the translocating proteins, as evidenced by our observed event frequencies, agreement between estimated and reference values, and prior work with lipid bilayer-coated nanopores.[15, 58]

As Figure 4.10 shows, we did not observe significant differences in $1/f$ noise related to the length of the event, and we found that the accuracy of parameter estimates tended to improve, rather than worsen, with event length (see Figure 4.7, 4.11), which is consistent with our previous work tethering proteins to the lipid bilayer.[15] For the data plotted in Figure 4.10, we calculated the power spectral densities for the shortest 10%, the middle 10%, and the longest 10% of experimental data for translocations of IgG_1 protein. Note that, in order to determine a meaningful power spectra from a sufficiently large data set, we

concatenated the resistive pulse traces of each subpopulation of events to generate a single "long" current trace for each subpopulation. Consequently, the shortest subpopulation has one order of magnitude fewer current values than the longest subpopulation. Figure 4.10 also shows that the three data sets overlapped completely indicating no detectable difference in $1/f$ noise. Because one component of $1/f$ noise are fluctuations in the surface ionic current,[102] variations in charge density at the surface of the nanopore would likely alter $1/f$ noise levels. The almost complete overlap between the shortest, middle, and longest subpopulations of translocation events in Figure 4.10 provides further evidence that we do not observe significant differences in surface interactions between the shortest and longest translocation events, and therefore that long-lasting events are not artificially prolonged by surface interactions.

In order to determine whether current blockades themselves vary with event length, we examined all translocation events of G_6PDH longer than 150 μs that were suitable for individual-event analysis from a single experiment. It is clear that the three distributions overlap, with approximately the same width and associated ΔI_{MIN} (0.0079) and ΔI_{MAX} (0.0126) values. To generate the distributions, we concatenated the current values of events from each respective region of the total population, and created a histogram for each of the three sets of concatenated values (Figure 4.9 A-C). It is worth noting that the distribution representing the shortest 10% of events contains approximately 15% of the amount of total current data relative to the distribution of the longest 10% of events. The minor differences in histogram appearance may therefore be due to the bias in data quantity, and the longest 10% of events provide a better representation of the protein population (m = 2.56, Λ = 233 nm^3), than the shortest subpopulation (m = 3.93, Λ = 201 nm^3), relative to reference values of length-to-diameter ratio, m = 2.5 and volume, Λ = 222 nm^3 for G_6PDH. Such a finding agrees with previous work on proteins passing through nanopores while tethered to lipid bilayers, where we showed that estimates of protein parameters typically improve with increasing event length.[15]

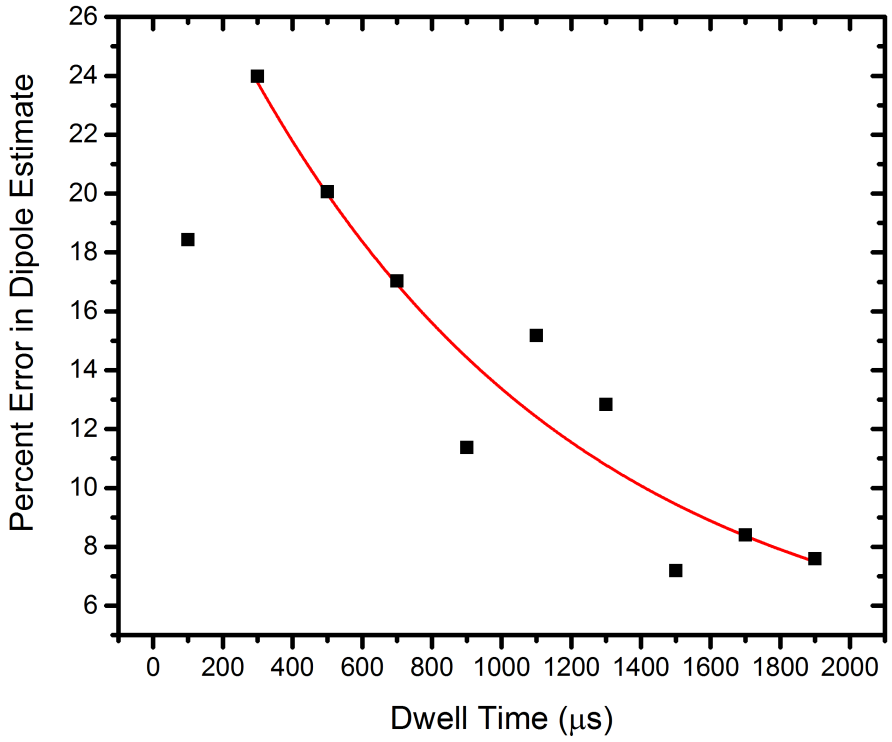


Figure 4.11. Error in estimates of protein dipole moment relative to event length. Data were generated through random-walk simulations, as described previously,[15] for an ellipsoid representing IgG_1 (length-to-diameter ratio m = 0.29, volume Λ = 340 nm^3). We performed a parameter sweep across 60 different conditions (6 different dipole moments 450, 550, 650, 750, 850, and 950 D and 10 different dwell times plotted). We simulated 100 events for each combination of dipole moment and dwell time, combined those events with true experimental noise, filtered them at 50 kHz, fit each event with our convolution model (Section 4.4.1), and then averaged the resulting percent error in dipole estimate across the various subsets of dipole moments (6 for each data point above). We fit the aggregate data plotted above with an exponential decay function because we would expect the percent error to tend asymptotically toward 0 as dwell time approaches infinity. Adapted with permission.[46] Copyright 2019, American Chemical Society.

In Section 4.4.4, we examined the trend of accuracy versus dwell time in order to determine a cutoff threshold for individual event analysis (Figure 4.7), and focused primarily on the effect of event duration on the determination of the length-to-diameter ratio. In light of the distributions plotted in Figure 4.9, we also performed random-walk simulations to better understand the effect of event duration on the quantification of dipole moment. Figure 4.11 shows that the uncertainty of determined dipole moments decreases with increasing event duration, as intuitively expected.

With the discussion above in mind, we assert that the majority of the events that we analyze with individual-event analysis do not originate from a secondary or biased population, but rather they are representative of the entire protein population with regard to all determined protein parameters. In fact, we find that the longest events produce the most accurate protein parameter estimates relative to reference values, which would not be the case if (1) those events resulted from a different protein in solution, or (2) if their duration was influenced by non-specific interactions with the pore walls. As we stated in Section 4.4.4, choosing a dwell-time threshold for individual event analysis longer than 150 μs would increase the accuracy of parameter estimates at the cost of reducing the total number of analyzable events.

4.4.9 Limitations to Ellipsoidal Approximation

There are limits to approximating complex three dimensional shapes as simple ellipsoids, as discussed in Yusko *et al.*[15] To provide insight into these limits, we have overlaid crystal structures of each protein in this study with the ellipsoidal estimate that we determined through convolution analysis and include these cartoons in Figure 4.12.

For many of the proteins, simple ellipsoids appear to capture much of the protein's shape information. However, in the case of proteins like BSA, IgM, and IgG_1, it is clear that ellipsoids fail to account for all three-dimensional complexity. It is important to keep in mind that while the ellipsoidal approximations may provide shape information with limited-

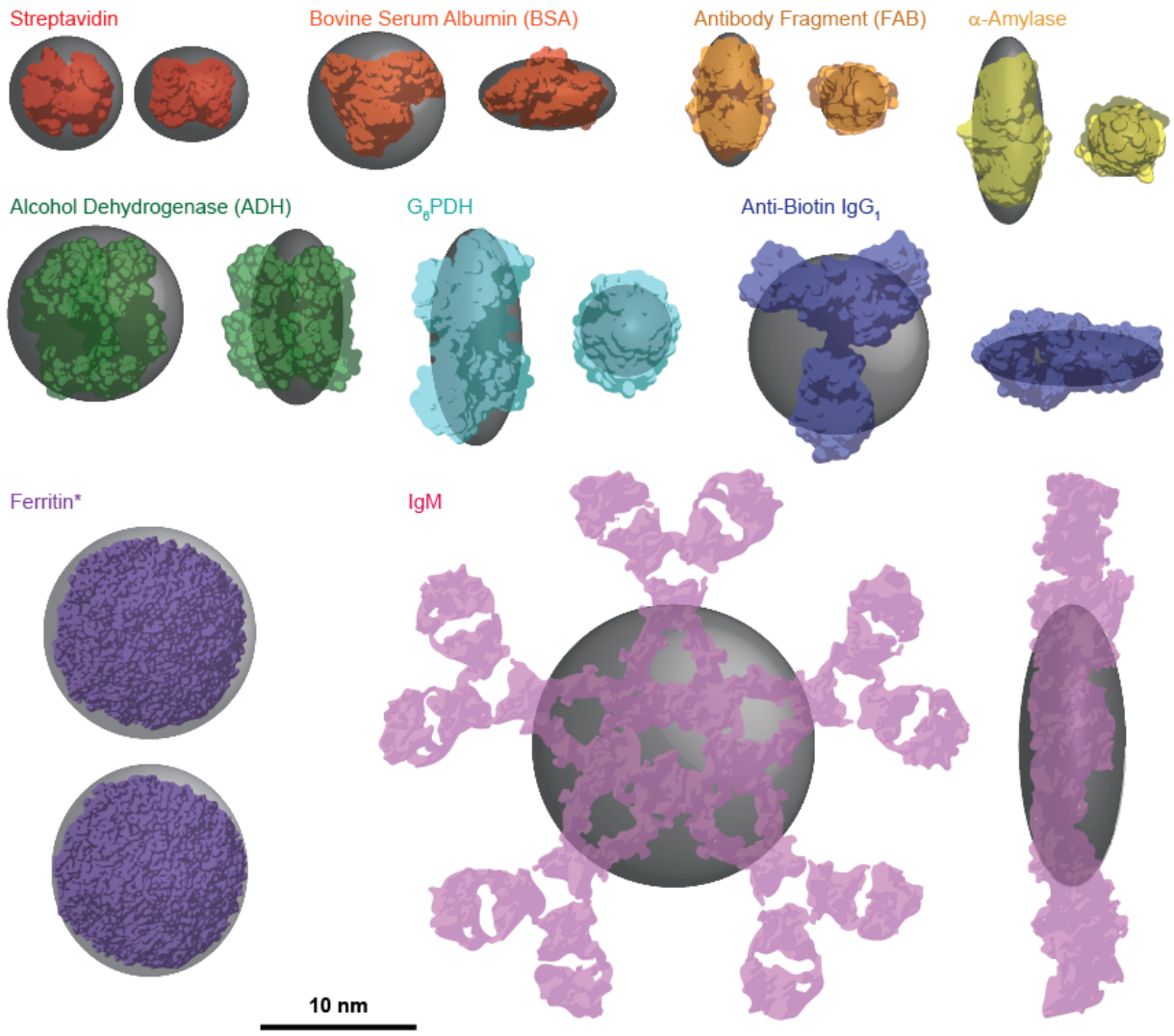


Figure 4.12. Comparison of the approximate shapes and volumes of nine proteins as determined by analysis of resistive pulses (gray spheroids) with crystal structures from the Protein Data Bank in various colors (streptavidin, red, 3RY1; bovine serum albumin (BSA), dark orange, 3V03; anti-biotin antibody fragment (FAB), light orange, 1F8T;α-amylase, yellow, 1BLI; alcohol dehydrogenase (ADH), green, 4W6Z; anti-biotin immunoglobulin G1, blue, 1HZH; ferritin, indigo, 1MFR; human immunoglobulin M antibody (IgM), violet, 2RCJ). The length-to-diameter ratio and volume for Ferritin were calculated using population-based analyses. All proteins and ellipsoidal approximations are shown on the same scale, as indicated by the scale bar. Note that the ellipsoid approximations of molecular shape agree well with most proteins but this simple model has limitations for proteins whose complex shape deviates significantly from perfect ellipsoids, such as BSA, IgG_1, and IgM. Nonetheless, even for these three proteins, the true molecular shape resembles the ellipsoidal approximation more closely than a perfect sphere. Adapted with permission.[46] Copyright 2019, American Chemical Society.

resolution, very few methods make it possible to determine structural information about single, unlabeled, natively folded proteins in solution, and we demonstrated previously that estimating ellipsoidal shape was critical in discriminating between multiple proteins in a mixture.[15] As the nanopore community moves toward low-noise platforms, we expect the information content of each resistive pulse signal to increase, and this rich information content will make it possible to improve upon the ellipsoidal model.

4.4.10 Investigation of Protein Purity

We always acquired pure protein with the highest possible quality and purity directly from the supplier (see Section 4.5.1 for more details), although it is true that there remains some level of impurity in commercially sourced proteins. We have found that for the size range (>40 kDa) of proteins that we investigate in this study, most impurities are typically much smaller than the target protein (*e.g.* buffer salts, glycerol, or other additives). We expect that such impurities will be too small to elicit a 5× threshold resistive pulse when transiting the nanopore, and will not contribute noticeably to resistive pulses above baseline. In the case of low concentrations of impurities by protein fragments or small proteins, we consider it unlikely that they transit at the same time as the target protein in our range of concentrations (see Section 4.4.5 for more details). In order to investigate the purity of all of our proteins, we performed HPLC size exclusion chromatography (SEC) experiments on each protein in this study, and report the chromatograms for each protein in Figure 4.13, and summarize the results in Table 4.2.

4.5 Materials and Methods

4.5.1 Materials

We purchased all phospholipids, namely 1-palmitoyl-2-oleoyl-sn-glycero-3-phosphocholine (POPC) and 1,2-dipalmitoyl-sn-glycero-3- phosphoethanolamine-N-(lissamine rhodamine B

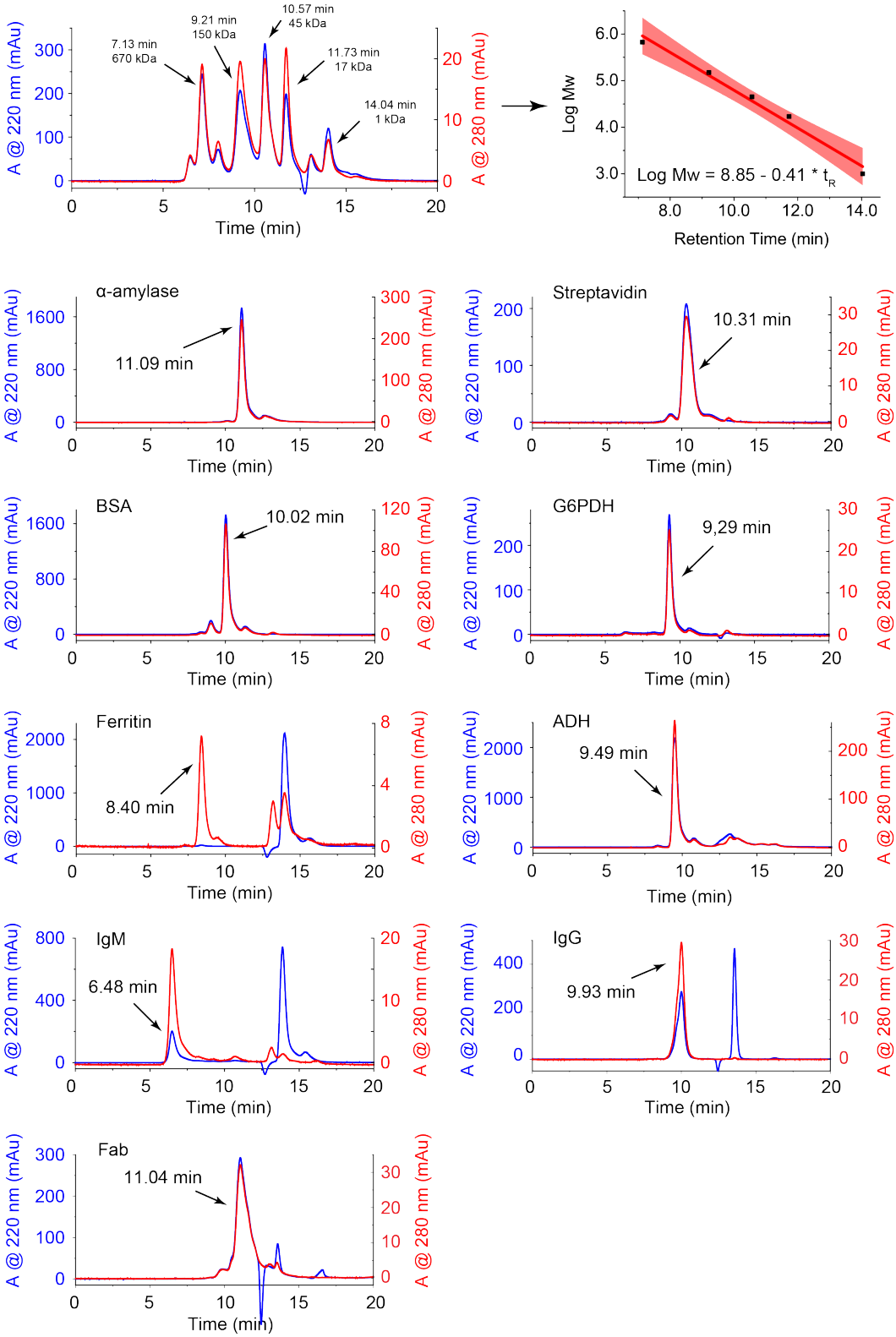


Figure 4.13. Chromatograms from HPLC experiments of proteins. Absorbance values at 220 nm are shown in blue, absorbance values at 280 nm are shown in red. The calibration standards are plotted in the top row (left) with a linear fit of log(MW) as a function of retention time (top row, right) showing a 95% confidence interval shaded in red. Adapted with permission.[46] Copyright 2019, American Chemical Society.

Table 4.2. Summary of results from HPLC analysis of proteins. Results were obtained for nine different commercially available proteins using size exclusion chromatography (SEC). Adapted with permission.[46] Copyright 2019, American Chemical Society.

Protein	Expected Mol. Weight (kDa)	Peak Retention Time (min)	Mol. Weight (kDa)[a]
ADH	82	9.49	100.1 (55.1 – 171.0)
Fab	50	11.04	23.5 (13.7 – 41.7)
IgM	950	6.48	1670.9 (941.2 – 2984.3)b
IgG_1	150	9.93	66.1 (38.9 – 114.8)
G_6PDH	128	9.29	120.7 (65.9 – 215.8)
BSA	66	10.02	61.0 (36.1 – 105.4)
α-Amylase	55	11.09	22.4 (13.1 – 40.9)
Streptavidin	54	10.31	46.5 (27.7 – 80.0)
Ferritin	474	8.40	277.5 (137.4 – 559.8)

[a] *Values in bold estimated from the linear fit in Figure 4.13, values in parentheses show 95% confidence interval of the linear fit at the given point.*
[b] *Range estimated from average percent error across other proteins (43.7% lower, 78.6% upper).*

sulphonyl) (Rh-PE), from Avanti Polar Lipids. Monoclonal anti-biotin IgG1 (B7653), glucose-6-phosphate dehydrogenase (G5885), IgM from human serum (I8260), bovine serum albumin (A7638), α-amylase (A4551), streptavidin from Streptomyces avidinii (85878), alcohol dehydrogenase from saccharomyces cerevisiae (A7011), and ferritin from human liver (F6754) were purchased from Sigma Aldrich. Polyclonal anti-biotin IgG-Fab fragments (800-101-098) were purchased from Rockland Immunochemicals. We investigated the purity of each sample using using size exclusion high performance liquid chromatography (SEC-HPLC), with an Agilent SEC3 column (300 mm, 300 nm pore size, 4.6 mm internal diameter) using 1× PBS + 1 mM EDTA as the running buffer at 0.3 mL min^{-1}, and include the results of this investigation in Figure 4.13, and Table 4.2.

4.5.2 Setup and Experimental

Nanopores were fabricated in a silicon nitride membrane, 275 nm thick, supported by a silicon chip, 3 mm × 3 mm square, using an ion beam sculpting technique (see Section 4.4.1 for details on pore geometry).[70] Experiments using Tween-20 coated pores (see Section 4.4.6) were performed both on these ion beam sculpted nanopores, as well as on nanopores generated using a helium ion microscope drilling technique.[103] Before each experiment, we cleaned nanopores in a freshly mixed Piranha solution containing 3:1 (v/v) concentrated

sulfuric acid and aqueous hydrogen peroxide solution at a temperature of $60 - 70°C$ for at least 30 min, rinsed the chips copiously with deionized water and dried them with nitrogen gas. We then mounted the chips between two pieces of cured PDMS containing ports with a 1 mm diameter for access the nanopore in order to separate *cis* and *trans* electrolyte reservoirs (Figure 4.1-A). The active lead of the headstage of the amplifier was connected to the *cis* compartment, while the ground lead of the headstage was connected to the *trans* compartment. To apply the lipid coatings, we formed supported lipid bilayers with 0.8 mol% of fluorescently labeled lipid, Rh-PE, in 99.2 mol% POPC through fusion of small unilamellar vesicles (SUVs), which were prepared in a buffer containing 150 mM KCl and 10 mM HEPES at pH 7.4 for 20 min before thoroughly rinsing with deionized water and then replacing the solution in both compartments with a recording buffer (2 M KCl) as described previously.[58] To perform the experiments, we placed the experimental setup with the nanopore chip in a Faraday cage, immersing Ag/AgCl pellet electrodes (Warner Instruments) into their respective electrolyte compartments. We applied a constant potential of ± (100 to 150) mV across the nanopore, and then measured the current (500 kHz sampling rate *via* NI PCI 6281 or USB-6361, 100 kHz analog low-pass filter, f_c = 50 kHz digital Gaussian low-pass FIR filter of length L = 2N + 1 with delay N / SR seconds) using an AxoPatch 200B (Molecular Devices) patch-clamp amplifier in voltage-clamp mode ($\beta = 1$) in combination with LabVIEW (National Instruments) software. We defined a resistive pulse by a reduction of the baseline current that exceeded a threshold of 5× the standard deviation of the noise, and marked the beginning and end of that resistive pulse where the current started and returned to within 1× standard deviation of the baseline.[104] We defined the translocation time of each event as the full-width at half-maximum value of that resistive pulse. All peak-finding and analysis procedures were performed using MATLAB (MathWorks) software.

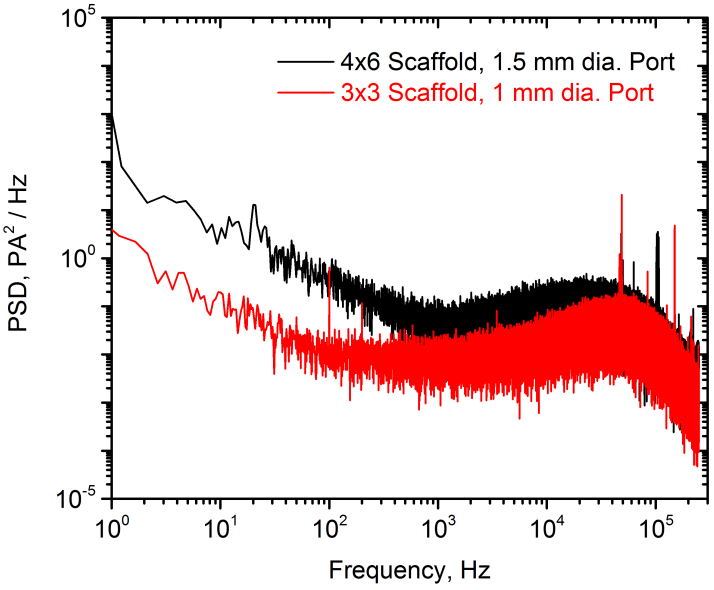


Figure 4.14. Comparison of noise levels of electrical recording setup used in this work (RED) with the setup used previously by Yusko *et al.*[15] (BLACK). Power spectral densities (PSD) are plotted to show the frequency dependence of noise levels in each experimental setup. The Axopatch 200B amplifier was set to β = 1 with its analog lowpass filter set to 100 kHz; the sampling frequency was 500 kHz. Adapted with permission.[46] Copyright 2019, American Chemical Society.

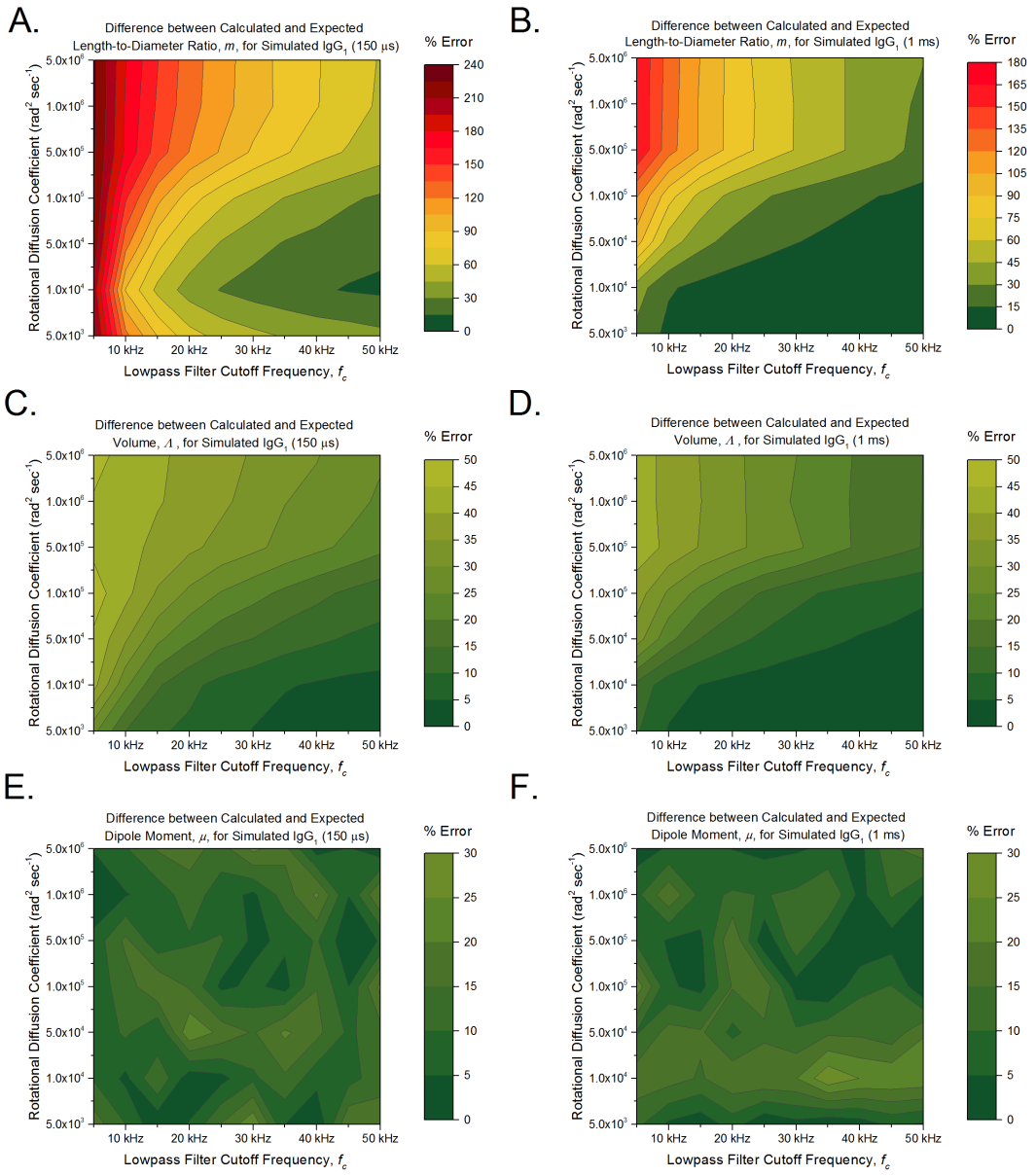


Figure 4.15. Accuracy of protein parameter estimates depends on the protein's rotational diffusion coefficient, on the cutoff frequency, f_c, of the digital filter, and on the event duration. A,B) Contour plot showing the error (%) in estimates for length-to-diameter ratio, *m*, determined from simulated resistive pulses of an oblate-shaped protein ($m = 0.29$, $\Lambda = 340$ nm^3, $\mu = 840$ D). C,D) Contour plot showing the error in estimates of protein volume, Λ, for the same protein as in (A). E,F) Contour plot showing the error in estimates of protein dipole moment, μ, for the same protein as in (A). A-F) All plots display the results of a parameter sweep across 70 different simulation conditions (7 rotational diffusion coefficients, 10 cutoff frequencies) where 100 unique resistive pulses with a duration of 150 μs (A,C,E) or 1 ms (B,D,F) were simulated for each of the 70 conditions and combined with experimentally recorded baseline noise before fitting with the convolution model. Note that both length-to-diameter ratio and volume estimates tend toward larger uncertainties as rotational diffusion coefficient increases and as filter cutoff frequency decreases, while dipole moment shows little to no correlation with either rotational diffusion coefficient or filter frequency. Adapted with permission.[46] Copyright 2019, American Chemical Society.

Table 4.3. Most-probable values determined for freely translocating proteins. Adapted with permission.[46] Copyright 2019, American Chemical Society.

Protein	Volume, Λ (nm^3)			Length-to-Dia. Ratio, m			Dipole Moment, μ (D)	
	Meas. Populat.	Meas. Indiv. (Med.)	Ref.	Meas. Popul.	Meas. Indiv. (Med.)	Ref.	Meas. Indiv. (Most Prob.)	Ref.
IgG1	341 236 391	231	DLS[a]: 391 Theor.[b]: 266 Lit.: 347 ± 15[105]	0.20 0.36 0.37	0.31	0.2-0.5[106, 107]	905, pH 7.4	840[c]
ADH	387 506	353	Theor.: 424	0.72 0.78	0.56	0.82[d]	821, pH 7.4	513[e]
Fab	196 113	82	DLS: 141 Theor.: 97 Lit.: 140[108] 170 ± 31[58]	1.85 3.27	1.92	1.7[d] 1.8[69]	640, pH 7.4	630[c] 550[e]
IgM	1490 1732	1074	Theor.: 1700	0.14 0.09	0.39	0.14[d]	1516, pH 7.4	1846[e]
G6PDH	298 254 376	212	DLS: 214 Theor.: 222	3.60 2.76 4.66	2.39	2.5[d]	932, pH 7.4	1382[f]
BSA	256 161	196	DLS: 143 Theor.: 111 Lit: 109[72] 123[109]	0.38 0.46	0.59	0.57[d]	682, pH 6.1	592[e]
α-Amylase	203 179	145	DLS: 97 Theor.: 89	3.17 4.47	2.51	1.8[d]	497, pH 7.4	484[f]
Streptavidin	140 160	161	DLS: 94 Theor.: 88 Lit.: 94 ± 18[58] 105 ± 3[110]	0.91 0.93	0.93	0.99[d]	–[g]	–[g]
Ferritin	898 779	–[h]	Theor. : 775	0.82 0.97	–[h]	0.92[d]	–[g]	–[g]

(a) Calculated from the hydrodynamic radius measured using dynamic light scattering by Yusko et al.[15]
(b) An estimate of the volume of the hydrated protein determined from the crystal structure of the protein using the software HydroPRO.
(c) Measured using dielectric impedance spectroscopy by Yusko et al.[15]
(d) Estimated from the crystal structure of the protein using Chimera software.
(e) Calculated from the crystal structure of the protein using the software HydroPRO.
(f) Calculated from the crystal structure of the protein using the Weizmann server (http://bioinfo.weizmann.ac.il/dipol/).
(g) Theoretically, we are not able to estimate the dipole moment of a protein that can be approximated by a sphere.
(h) We were not able to collect a sufficient number of long events to perform these analyses.
A table with a similar format and reference values was published in the supplemental material of Yusko et al.[15]

References

[1] A. Han et al. “Sensing protein molecules using nanofabricated pores”. *Applied Physics Letters* 88.9 (2006), p. 093901.

[2] J. Li, M. Gershow, D. Stein, E. Brandin, and J. A. Golovchenko. “DNA molecules and configurations in a solid-state nanopore microscope”. *Nature materials* 2.9 (2003), p. 611.

[3] E. C. Yusko et al. “Single-particle characterization of A oligomers in solution”. *ACS nano* 6.7 (2012), pp. 5909–5919.

[4] J. Houghtaling, J. List, and M. Mayer. “Nanopore-Based, Rapid Characterization of Individual Amyloid Particles in Solution: Concepts, Challenges, and Prospects”. *Small* 0.0 (), p. 1802412.

[5] S. W. Moore, P. Roca-Cusachs, and M. P. Sheetz. “Stretchy proteins on stretchy substrates: the important elements of integrin-mediated rigidity sensing”. *Developmental Cell* 19.2 (2010), pp. 194–206.

[6] R. Afrin, M. T. Alam, and A. Ikai. “Pretransition and progressive softening of bovine carbonic anhydrase II as probed by single molecule atomic force microscopy”. *Protein Science : A Publication of the Protein Society* 14.6 (2005), pp. 1447–1457.

[7] P. Nirmalraj et al. “Nanoelectrical analysis of single molecules and atomic-scale materials at the solid/liquid interface”. *Nature Materials* 13.10 (2014), pp. 947–953.

[8] R. Kayed et al. “Common structure of soluble amyloid oligomers implies common mechanism of pathogenesis”. *Science* 300.5618 (2003). WOS:000182295900040, pp. 486–489.

[9] C. Eggeling et al. “Direct observation of the nanoscale dynamics of membrane lipids in a living cell”. *Nature* 457.7233 (2009). WOS:000263680100047, 1159–U121.

[10] H. Yang et al. “Protein conformational dynamics probed by single-molecule electron transfer”. *Science* 302.5643 (2003). WOS:000185825900035, pp. 262–266.

[11] B. Schuler and W. A. Eaton. “Protein folding studied by single-molecule FRET”. *Current Opinion in Structural Biology* 18.1 (2008), pp. 16–26.

[12] A. A. Deniz et al. “Single-molecule protein folding: diffusion fluorescence resonance energy transfer studies of the denaturation of chymotrypsin inhibitor 2”. *Proceedings of the National Academy of Sciences of the United States of America* 97.10 (2000), pp. 5179–5184.

[13] K. A. Merchant, R. B. Best, J. M. Louis, I. V. Gopich, and W. A. Eaton. “Characterizing the unfolded states of proteins using single-molecule FRET spectroscopy and molecular simulations”. *Proceedings of the National Academy of Sciences* 104.5 (2007), pp. 1528–1533.

[14] E. A. Jares-Erijman and T. M. Jovin. “FRET imaging”. *Nature Biotechnology* 21.11 (2003). WOS:000186320000042, pp. 1387–1395.

[15] E. C. Yusko et al. “Real-time shape approximation and fingerprinting of single proteins using a nanopore”. *Nature Nanotechnology* 12 (2017), pp. 360–367.

[16] A. Ivankin, S. Carson, S. R. Kinney, and M. Wanunu. “Fast, label-free force spectroscopy of histoneDNA interactions in individual nucleosomes using nanopores”. *Journal of the American Chemical Society* 135.41 (2013), pp. 15350–15352.

[17] A. Ananth et al. “Reversible Immobilization of Proteins in Sensors and Solid-State Nanopores”. *Small* 14.18 (2018), p. 1703357.

[18] G. V. Soni and C. Dekker. “Detection of nucleosomal substructures using solid-state nanopores”. *Nano letters* 12.6 (2012), pp. 3180–3186.

[19] F. Rivas et al. “Label-free analysis of physiological hyaluronan size distribution with a solid-state nanopore sensor”. *Nature Communications* 9.1 (2018), p. 1037.

[20] J. J. Kasianowicz and S. M. Bezrukov. “Protonation dynamics of the alpha-toxin ion channel from spectral analysis of pH-dependent current fluctuations.” *Biophysical Journal* 69.1 (1995), pp. 94–105.

[21] J. J. Kasianowicz et al. “Analytical applications for pore-forming proteins”. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1858.3 (2016), pp. 593–606.

[22] S. M. Bezrukov and J. J. Kasianowicz. “Current noise reveals protonation kinetics and number of ionizable sites in an open protein ion channel”. *Physical Review Letters* 70.15 (1993), pp. 2352–2355.

[23] A. Oukhaled et al. “Dynamics of Completely Unfolded and Native Proteins through Solid-State Nanopores as a Function of Electric Driving Force”. *ACS Nano* 5.5 (2011), pp. 3628–3638.

[24] K. J. Freedman et al. “Chemical, thermal, and electric field induced unfolding of single protein molecules studied using nanopores”. *Analytical chemistry* 83.13 (2011), pp. 5137–5144.

[25] N. Martyushenko, N. A. Bell, R. D. Lamboll, and U. F. Keyser. “Nanopore analysis of amyloid fibrils formed by lysozyme aggregation”. *Analyst* 140.14 (2015), pp. 4882–4886.

[26] W. Li et al. “Single protein molecule detection by glass nanopores”. *ACS nano* 7.5 (2013), pp. 4129–4134.

[27] C. Wloka et al. “Label-Free and Real-Time Detection of Protein Ubiquitination with a Biological Nanopore”. *ACS nano* 11.5 (2017), pp. 4387–4394.

[28] S. Howorka and Z. Siwy. “Nanopore analytics: sensing of single molecules”. *Chemical Society Reviews* 38.8 (2009), pp. 2360–2384.

[29] R. Wei, V. Gatterdam, R. Wieneke, R. Tampé, and U. Rant. “Stochastic sensing of proteins with receptor-modified solid-state nanopores”. *Nature Nanotechnology* 7.4 (2012), pp. 257–263.

[30] Z. Qin, J. Zhe, and G.-X. Wang. “Effects of particle’s off-axis position, shape, orientation and entry position on resistance changes of micro Coulter counting devices”. *Measurement Science and Technology* 22.4 (2011), p. 045804.

[31] D. C. Golibersuch. “Observation of Aspherical Particle Rotation in Poiseuille Flow via the Resistance Pulse Technique”. *Biophysical Journal* 13.3 (1973), pp. 265–280.

[32] G. V. Soni et al. “Synchronous optical and electrical detection of biomolecules traversing through solid-state nanopores”. *Review of Scientific Instruments* 81.1 (2010), p. 014301.

[33] T. Gilboa and A. Meller. “Optical sensing and analyte manipulation in solid-state nanopores”. *Analyst* 140.14 (2015), pp. 4733–4747.

[34] A. Ivankin et al. “Label-Free Optical Detection of Biomolecular Translocation through Nanopore Arrays”. *ACS Nano* 8.10 (2014), pp. 10774–10781.

[35] J. J. Kasianowicz, J. W. Robertson, E. R. Chan, J. E. Reiner, and V. M. Stanford. “Nanoscopic Porous Sensors”. *Annual Review of Analytical Chemistry* 1.1 (2008), pp. 737–766.

[36] W. H. Pitchford et al. “Synchronized Optical and Electronic Detection of Biomolecules Using a Low Noise Nanopore Platform”. *ACS Nano* 9.2 (2015), pp. 1740–1748.

[37] A. Balan, C.-C. Chien, R. Engelke, and M. Drndi. “Suspended Solid-state Membranes on Glass Chips with Sub 1-pF Capacitance for Biomolecule Sensing Applications”. *Scientific Reports* 5 (2015), p. 17775.

[38] D. Branton et al. “The potential and challenges of nanopore sequencing”. *Nature Biotechnology* 26.10 (2008). WOS:000259926000029, pp. 1146–1153.

[39] S. Carson and M. Wanunu. “Challenges in DNA motion control and sequence readout using nanopore devices”. *Nanotechnology* 26.7 (2015), p. 074004.

[40] E. A. Manrao et al. “Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase”. *Nature biotechnology* 30.4 (2012), p. 349.

[41] C. Ying et al. “Formation of Single Nanopores with Diameters of 2050 nm in Silicon Nitride Membranes Using Laser-Assisted Controlled Breakdown”. *ACS Nano* (2018).

[42] C. Plesa et al. *Fast Translocation of Proteins through Solid State Nanopores*. rapid-communication. 2013.

[43] C. Dekker. “Solid-state nanopores”. *Nature Nanotechnology* 2.4 (2007). WOS:000245920900010, pp. 209–215.

[44] S. Balme et al. *Influence of Adsorption on Proteins and Amyloid Detection by Silicon Nitride Nanopore*. research-article. 2016.

[45] L. T. Sexton et al. “An adsorption-based model for pulse duration in resistive-pulse protein sensing”. *Journal of the American Chemical Society* 132.19 (2010), pp. 6755–6763.

[46] J. Houghtaling et al. “Estimation of Shape, Volume, and Dipole Moment of Individual Proteins Freely Transiting a Synthetic Nanopore”. *ACS nano* ().

[47] N. S. Galenkamp, M. Soskine, J. Hermans, C. Wloka, and G. Maglia. “Direct electrical quantification of glucose and asparagine from bodily fluids using nanopores”. *Nature Communications* 9.1 (2018), p. 4085.

[48] V. Van Meervelt et al. “Real-Time Conformational Changes and Controlled Orientation of Native Proteins Inside a Protein Nanoreactor”. *Journal of the American Chemical Society* 139.51 (2017), pp. 18640–18646.

[49] P. Waduge et al. *Nanopore-Based Measurements of Protein Size, Fluctuations, and Conformational Changes*. research-article. 2017.

[50] J. Lee and H. Bayley. “Semisynthetic protein nanoreactor for single-molecule chemistry”. *Proceedings of the National Academy of Sciences* 112.45 (2015), pp. 13768–13773.

[51] J. W. Robertson et al. “Single-molecule mass spectrometry in solution using a solitary nanopore”. *Proceedings of the National Academy of Sciences* 104.20 (2007), pp. 8207–8211.

[52] L. T. Sexton et al. “Resistive-pulse studies of proteins and protein/antibody complexes using a conical nanotube sensor”. *Journal of the American Chemical Society* 129.43 (2007), pp. 13144–13152.

[53] A. Han et al. “Label-free detection of single protein molecules and protein protein interactions using synthetic nanopores”. *Analytical chemistry* 80.12 (2008), pp. 4651–4658.

[54] D. Fologea, B. Ledden, D. S. McNabb, and J. Li. “Electrical characterization of protein molecules by a solid-state nanopore”. *Applied physics letters* 91.5 (2007), p. 053901.

[55] N. Giamblanco et al. “Detection of protein aggregate morphology through single antifouling nanopore”. *Sensors and Actuators B: Chemical* 260 (2018), pp. 736–745.

[56] M. Lepoitevin, T. Ma, M. Bechelany, and S. Balme. “Functionalization of single solid state nanopores to mimic biological ion channels: A review”. *Advances in colloid and interface science* (2017).

[57] D. S. Talaga and J. Li. “Single-molecule protein unfolding in solid state nanopores”. *Journal of the American Chemical Society* 131.26 (2009), pp. 9287–9297.

[58] E. C. Yusko et al. “Controlling protein translocation through nanopores with bio-inspired fluid walls”. *Nature Nanotechnology* 6.4 (2011), pp. 253–60.

[59] J. K. Rosenstein, M. Wanunu, C. A. Merchant, M. Drndic, and K. L. Shepard. “Integrated nanopore sensing platform with sub-microsecond temporal resolution”. *Nature methods* 9.5 (2012), p. 487.

[60] H. Fricke. “The Electric Permittivity of a Dilute Suspension of MembraneCovered Ellipsoids”. *Journal of Applied Physics* 24.5 (1953), pp. 644–646.

[61] S. Velick and M. Gorin. “The electrical conductance of suspensions of ellipsoids and its relation to the study of avian erythrocytes”. *The Journal of general physiology* 23.6 (1940), p. 753.

[62] N. Di Fiori et al. “Optoelectronic control of surface charge and translocation dynamics in solid-state nanopores”. *Nature nanotechnology* 8.12 (2013), p. 946.

[63] I. Nir, D. Huttner, and A. Meller. “Direct sensing and discrimination among Ubiquitin and Ubiquitin chains using solid-state nanopores”. *Biophysical journal* 108.9 (2015), pp. 2340–2349.

[64] R. W. DeBlois, E. E. Uzgiris, D. H. Cluxton, and H. M. Mazzone. “Comparative measurements of size and polydispersity of several insect viruses”. *Analytical biochemistry* 90.1 (1978), pp. 273–288.

[65] W. R. Smythe. “Flow around a spheroid in a circular tube”. *The Physics of Fluids* 7.5 (1964), pp. 633–638.

[66] J. D. Uram, K. Ke, and M. Mayer. “Noise and bandwidth of current recordings from submicrometer pores and nanopores”. *ACS nano* 2.5 (2008), pp. 857–872.

[67] J. A. Dix and A. S. Verkman. “Crowding Effects on Diffusion in Solutions and Cells”. *Annual Review of Biophysics* 37.1 (2008), pp. 247–263.

[68] Y. Yuan and D. Axelrod. “Subnanosecond polarized fluorescence photobleaching: rotational diffusion of acetylcholine receptors on developing muscle cells”. *Biophysical journal* 69.2 (1995), pp. 690–700.

[69] J. Garcia de la Torre and B. Carrasco. “Hydrodynamic properties of rigid macromolecules composed of ellipsoidal and cylindrical subunits”. *Biopolymers* 63.3 (2002), pp. 163–167.

[70] J. Li et al. “Ion-beam sculpting at nanometre length scales”. *Nature* 412.6843 (2001), pp. 166–169.

[71] J. L. Oncley. “The Investigation of Proteins by Dielectric Measurements.” *Chemical Reviews* 30.3 (1942), pp. 433–450.

[72] S. N. Singh, S. Yadav, S. J. Shire, and D. S. Kalonia. “Dipole-Dipole Interaction in Antibody Solutions: Correlation with Viscosity Behavior at High Concentration”. *Pharmaceutical Research* 31.9 (2014), pp. 2549–2558.

[73] D. Y. Ling and X. S. Ling. “On the distribution of DNA translocation times in solid-state nanopores: an analysis using Schrödingers first-passage-time theory”. *Journal of Physics: Condensed Matter* 25.37 (2013), p. 375102.

[74] X. Li et al. “Non-sticky translocation of bio-molecules through Tween 20-coated solid-state nanopores in a wide pH range”. *Applied Physics Letters* 109.14 (2016), p. 143105.

[75] M. Firnkes, D. Pedone, J. Knezevic, M. Döblinger, and U. Rant. “Electrically Facilitated Translocations of Proteins through Silicon Nitride Nanopores: Conjoint and Competitive Action of Diffusion, Electrophoresis, and Electroosmosis”. *Nano Letters* 10.6 (2010), pp. 2162–2167.

[76] L. Galla et al. “Hydrodynamic slip on DNA observed by optical tweezers-controlled translocation experiments with solid-state and lipid-coated nanopores”. *Nano letters* 14.7 (2014), pp. 4176–4182.

[77] A. Sischka et al. “Controlled translocation of DNA through nanopores in carbon nano-, silicon-nitride- and lipid-coated membranes”. *Analyst* 140.14 (2015), pp. 4843–4847.

[78] L. I. Berge, T. Jossang, and J. Feder. “Off-axis response for particles passing through long apertures in Coulter-type counters”. *Measurement Science and Technology* 1.6 (1990), p. 471.

[79] W. R. Smythe. “OffAxis Particles in Coulter Type Counters”. *Review of Scientific Instruments* 43.5 (1972), pp. 817–818.

[80] J. Guo, T. S. Pui, A. R. A. Rahman, and Y. Kang. “3D numerical simulation of a coulter counter array with analysis of electrokinetic forces”. *Electrophoresis* 34.3 (2013), pp. 417–424.

[81] P. Hinkle et al. “A hybrid resistive pulse-optical detection platform for microfluidic experiments”. *Scientific Reports* 7.1 (2017), p. 10173.

[82] D. Kozak, W. Anderson, M. Grevett, and M. Trau. “Modeling elastic pore sensors for quantitative single particle sizing”. *The Journal of Physical Chemistry C* 116.15 (2012), pp. 8554–8561.

[83] M. Zhang. “Numerical investigation of hydrodynamic focusing and coulter principle in a microfluidic device” (2014).

[84] D. M. Ecker, S. D. Jones, and H. L. Levine. “The therapeutic monoclonal antibody market”. *MAbs*. Vol. 7. 1. Taylor & Francis. 2015, pp. 9–14.

[85] R. L. Fraccari et al. “High-speed detection of DNA translocation in nanopipettes”. *Nanoscale* 8.14 (2016), pp. 7604–7611.

[86] V. Tabard-Cossa, D. Trivedi, M. Wiggin, N. N. Jetha, and A. Marziali. “Noise analysis and reduction in solid-state nanopores”. *Nanotechnology* 18.30 (2007), p. 305505.

[87] Z. Roelen, J. A. Bustamante, A. Carlsen, A. Baker-Murray, and V. Tabard-Cossa. “Instrumentation for low noise nanopore-based ionic current recording under laser illumination”. *Review of Scientific Instruments* 89.1 (2018), p. 015007.

[88] M. Wanunu and A. Meller. “Chemically modified solid-state nanopores”. *Nano letters* 7.6 (2007), pp. 1580–1585.

[89] Z. Tang et al. “Surface Modification of SolidState Nanopores for StickyFree Translocation of SingleStranded DNA”. *Small* 10.21 (2014), pp. 4332–4339.

[90] N. Giamblanco et al. “Amyloid Growth, Inhibition, and Real-Time Enzymatic Degradation Revealed with Single Conical Nanopore”. *Analytical Chemistry* 90.21 (2018), pp. 12900–12908.

[91] P. Prangkio, E. C. Yusko, D. Sept, J. Yang, and M. Mayer. “Multivariate analyses of amyloid-beta oligomer populations indicate a connection between pore formation and cytotoxicity”. *PLoS One* 7.10 (2012), e47261.

[92] A. R. De Pascalis et al. “Binding of ferredoxin to ferredoxin: NADP+ oxidoreductase: the role of carboxyl groups, electrostatic surface potential, and molecular dipole moment”. *Protein Science* 2.7 (1993), pp. 1126–1135.

[93] J. C. Maxwell. *A treatise on electricity and magnetism*. Vol. 1. Clarendon press, 1881.

[94] J. Hurley. “Sizing Particles with a Coulter Counter”. *Biophysical Journal* 10.1 (1970), pp. 74–79.

[95] J. A. Osborn. “Demagnetizing Factors of the General Ellipsoid”. *Physical Review* 67.11-12 (1945), pp. 351–357.

[96] R. T. Perera, R. P. Johnson, M. A. Edwards, and H. S. White. “Effect of the Electric Double Layer on the Activation Energy of Ion Transport in Conical Nanopores”. *The Journal of Physical Chemistry C* 119.43 (2015), pp. 24299–24306.

[97] J. Liu et al. “Surface Charge Density Determination of Single Conical Nanopores Based on Normalized Ion Current Rectification”. *Langmuir* 28.2 (2012), pp. 1588–1595.

[98] C. Kubeil and A. Bund. “The Role of Nanopore Geometry for the Rectification of Ionic Currents”. *The Journal of Physical Chemistry C* 115.16 (2011), pp. 7866–7873.

[99] Z. Jiang and D. Stein. “Charge regulation in nanopore ionic field-effect transistors”. *Physical Review E* 83.3 (2011), p. 031203.

[100] D. A. MacInnes and A. S. Brown. “The Determination of Activity Coefficients from the Potentials of Concentration Cells with Transference.” *Chemical Reviews* 18.2 (1936), pp. 335–348.

[101] A. S. Brown and D. A. MacInnes. “The determination of activity coefficients from the potentials of concentration cells with transference. I. Sodium chloride at 25”. *Journal of the American Chemical Society* 57.7 (1935), pp. 1356–1362.

[102] C. Wen et al. “Generalized Noise Study of Solid-State Nanopores at Low Frequencies”. *ACS Sensors* 2.2 (2017), pp. 300–307.

[103] J. Yang et al. “Rapid and precise scanning helium ion microscope milling of solid-state nanopores for biomolecule detection”. *Nanotechnology* 22.28 (2011), p. 285310.

[104] D. Pedone, M. Firnkes, and U. Rant. “Data Analysis of Translocation Events in Nanopore Experiments”. *Analytical Chemistry* 81.23 (2009), pp. 9689–9694.

[105] S. W. Schneider, J. Lärmer, R. M. Henderson, and H. Oberleithner. “Molecular weights of individual proteins correlate with molecular volumes measured by atomic force microscopy”. *Pflügers Archiv* 435.3 (1998), pp. 362–367.

[106] A. J. Ozinskas. *In Topics in Fluorescence Spectroscopy, Vol. 4: Probe Design and Chemical Sensing; Lakowicz, JR, Ed*. Plenum Press: New York, 1994.

[107] B. Carrasco et al. “Crystallohydrodynamics for solving the hydration problem for multi-domain proteins: open physiological conformations for human IgG”. *Biophysical chemistry* 93.2-3 (2001), pp. 181–196.

[108] C. A. J. Jr et al. *Immunobiology*. 5th. Garland Science, 2001.

[109] N. El Kadi et al. “Unfolding and refolding of bovine serum albumin at acid pH: ultrasound and structural studies”. *Biophysical journal* 91.9 (2006), pp. 3397–3404.

[110] C. S. Neish, I. L. Martin, R. M. Henderson, and J. M. Edwardson. “Direct visualization of ligandprotein interactions using atomic force microscopy”. *British journal of pharmacology* 135.8 (2002), pp. 1943–1950.

Chapter 5: Computational Evaluation of Off-Axis Effects

Most approaches to analyze resistive pulses assume that the particles that produce them translocate along the central pore axis – *i.e.* equidistant from all pore walls. In reality, particles freely transiting a nanopore also diffuse laterally, and as they approach the pore walls they generate non-uniform distortions in the electric field, resulting in larger-than-expected resistive pulses. In this chapter, we present results from studies of these off-axis effects that we obtained using finite element and random-walk simulations. We varied the size, ellipsoidal shape, and axial position of individual particles, as well as the size of the nanopore. We demonstrated that randomly sampling a translocating nanoparticle at different distances from the central axis leads to overestimates of that particles volume. Additionally, we have shown that minimizing – or better characterizing– off-axis effects is critical to accurately determine a particle's shape; we observed up to 200% deviation in estimates of ellipsoidal shape when off-axis effects were present. We present optimal ratios of nanopore to nanoparticle size for experiments targeting freely translocating particles, and highlight the importance of tethering nanoparticles to the pore walls to mitigate unwanted off-axis effects on the shape determination of those particles. By considering off-axis effects in nanopore sensing, and even possibly correcting for them, we are able to further improve our ability to accurately characterize individual nanometer-sized particles in solution.

5.1 Introduction to Off-Axis Effects

The underlying concepts of resistive pulse-based nanopore sensing have been described now many times throughout this dissertation. Importantly for this chapter, the amplitude, duration, and frequency of resistive pulses provide information about the size, charge, and concentration of the particles passing through the pore.[1, 2] Moreover, as a non-spherical particle adopts different orientations within the nanopore, it produces fluctuations in the measured current that relate to the ellipsoidal shape of the particle, which can be described by an electrical shape factor,γ.[3–6] By comparing the minimum and maximum current values within a resistive pulse, nanopore sensing makes it possible to characterize proteins with different sizes, ellipsoidal shapes, charges, dipole moments, and rotational diffusion coefficients, as described in Chapter 3.[6] This fundamental resistive pulse analysis, however, assumes that the particles translocate through the central axis of a perfectly cylindrical nanopore, and that the particles are able to rotate freely within that nanopore. Such an assumption holds true when proteins are tethered to a lipid anchor within a fluid lipid bilayer coating on the surface of the nanopore, because the lipid tether confines protein translocation to a single (albeit off-center) translocation axis.[6, 7] We demonstrated in Chapter 4, however, that the same multiparametric analysis makes it possible to characterize proteins that are not tethered to the lipid bilayer. When these freely translocating particles diffuse away from the central axis during their transit through the nanopore, they distort the electric field asymmetrically and produce resistive pulse fluctuations that depend on their distance from the central axis,[8–10] as well as on their particular electrical shape factor, γ, relating their ellipsoidal shape to their orientation relative to the electrical field in the nanopore. These distortions lead to larger-than-expected estimates of protein volume and deviations away from reference values for estimates of protein length-to-diameter ratio. Additionally, the fluctuations produced when a spherical particle diffuses laterally within the pore appear similar to that of a rotating non-spherical particle passing through the center of a nanopore. It is therefore important to quantify the effects of off-axis translocation on resistive pulse

measurements, especially in the context of characterizing mixtures of proteins or examining transient changes in a protein population.

In this chapter, we determine the influence of off-axis effects in the context of characterizing ellipsoidal nanoparticles passing through a cylindrical nanopore using finite element simulations (COMSOL Multiphysics 5.2a) in combination with random-walk simulations. To this end, we performed a wide parameter sweep, varying the radial position of individual particles as well as the ratio of particle size to nanopore radius. We further examined the effects of off-axis translocation on the estimation of protein length-to-diameter ratio by considering particles with a fixed volume but with different ellipsoidal shapes and at different radial positions and rotational orientations. The results suggest that 1) off-axis effects become more extreme in scenarios where the diameter of the nanopore is much larger than the diameter of the particle, 2) ellipsoidal particles with length-to-diameter ratios further from a sphere ($m << 1$ or $m >> 1$) illicit larger – and orientation-dependent – off-axis effects than spherical particles of the same volume, and 3) off-axis effects add a noise-like element to resistive pulse recordings that can produce large errors in estimates of length-to-diameter ratio and volume, both in individual-event and in population-based analyses.

5.2 Analysis Methods

We described the methods to analyze resistive pulses produced by non-spherical particles at length in the previous chapters of this thesis.[6] Briefly, resistive pulse-based nanopore sensing monitors ion flux through a sensing volume – here, the nanopore – in the presence and absence of non-conducting particles. The principle is straightforward when considering the electrolyte as a homogeneous conductive medium with resistivity ρ. The resistance of a cylindrical nanopore with diameter of d_p and length of l_p consists of two components: resistance to ions passing through the confines of the pore itself, R_{pore}, and access resistance, R_{access}, of ionic current paths converging to the entrance of the pore.

$$R = R_{pore} + R_{access} = \rho \left(\frac{4l_p}{\pi d_p^2} + \frac{1}{d_p} \right) \tag{5.1}$$

Hence, a particle occupying the nanopore modulates the resistance to $R = \int A^{-1}(x)dx$, where *A(x)* is the area of the electrolyte at some position, x. In principle, each resistive pulse contains information about a single particle's 2-dimensional orthogonal projection on the transverse plane that is normal to the axis of the nanopore. For resistive pulse-based analyses, the electric field is considered to be uniform along the effective nanopore (Figure 5.1). Note that l_p and d_p are the length and diameter of the nanopore, respectively. We relate the volume, Λ, and electrical shape factor, γ, to the magnitude of the current blockade using the following equation, as discussed in Chapter 4:

$$\frac{\Delta I}{I_0} = \frac{4\gamma\Lambda}{\pi d_p^2(l_p + \frac{\pi}{4}d_p)} \tag{5.2}$$

Non-spherical particles have a maximum shape factor, γ_{MAX}, describing their crosswise orientation in the nanopore, and a minimum shape factor, γ_{MIN}, describing their lengthwise orientation in the nanopore, both of which are directly related to current blockade maxima.[3] These non-spherical particles generate resistive pulses that fluctuate between $\Delta I_{min}/I_0 = \frac{4\gamma_{min}\Lambda}{\pi d_p^2(l_p+\frac{\pi}{4}d_p)}$ and $\Delta I_{max}/I_0 = \frac{4\gamma_{max}\Lambda}{\pi d_p^2(l_p+\frac{\pi}{4}d_p)}$. The electric field produces a torque on non-spherical particles that have a permanent dipole moment; this torque biases non-spherical particles toward either their minimum or maximum orientation in the nanopore, and the magnitude of this bias is proportional to the strength of the electric field in the nanopore and the dipole moment of the particle (see Chapters 3 and 4).[6, 11].

This model simplifies the electrolyte solution as homogeneous conducting medium and resolves the particle information steadily, without dynamic analyses. It describes the scenario whereby a particle is located at the center of a large-aspect-ratio (diameter:length) cylindrical nanopore. The mobility of ions in the nanopore is, in practice, non-uniform as ions are unable to diffuse freely near the nanopore wall. In other words, the ohmic medium inside

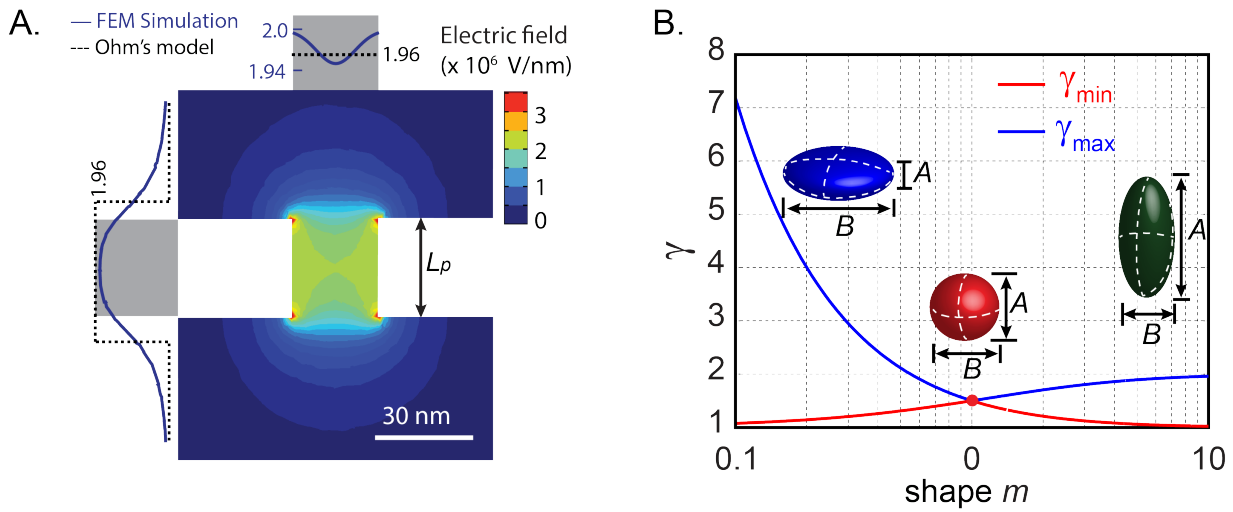


Figure 5.1. Schematic showing the finite element-based determination of the electric field strength in a nanopore. A) The distribution of electric field strength in the nanopore simulations, as well as the electric field strength according to Ohm's model, for comparison. B) The electrical shape factor as a function of a particles length-to-diameter ratio, *m*.

the nanopore is non-homogeneous. This inhomogeneity leads to deviations away from the model when the particle is electrically sampled away from the central axis of the pore.

5.3 Finite Element Analyses of Resistive Pulses Based on the PNP Equations

Poisson-NernstPlanck (PNP) equations are commonly used to express the flow of ionic current through a nanopore.[12–15] The Poisson function describes the distribution of potentials, ϕ, in an electrolyte solution containing ionic species, i, at some concentration c_i and charge of z_i.[16]

$$\nabla^2\Phi = -\frac{F}{\varepsilon}\sum_i z_i c_i \tag{5.3}$$

The Nernst-Planck relationship describes the motion of the ions in a fluid medium under an applied external potential by considering the diffusion of ions, D_i, [17, 18]

$$J_i = -D_i \nabla c_i - \frac{z_i F}{RT} D_i c_i \nabla \Phi \tag{5.4}$$

Where F represents the Faraday constant F = 96485.3365 C mol^{-1}, ε is the dielectric constant of the fluidic medium, $\varepsilon = 80$, R is the gas constant R = 8.31 J mol $^{-1}$ K^{-1}, and T is the absolute temperature, T = 295 K. Solving the PNP equations by using the finite element method, we can determine the ionic current through a nanopore. In comparison with analytical models, numerical simulations of the PNP equations make it possible to incorporate complex parameters including surface charge and curvature, off-axis position of particles, and irregular geometries for both the nanopore and the particles.

Here, we used COMSOL Multiphysics 5.2a to solve the PNP equations. To simulate our experiments for lipid-coated nanopores, we define the nanopore length to be 40 nm (30 nm before coating) and the surface charge of zero on nanopore wall. We applied 0.1 V across the nanopore containing a recording buffer with an ionic strength of 2 M KCl. Because the effective ionic transference number is nonlinearly related to salt concentration,[17, 18] we used a concentration of 1.68 M KCl instead of 2 M according to the measured solution conductivity. Figure 5.1 shows the electric field distribution in an uncoated (*e.g.* no lipid bilayer) 30-nm diameter nanopore determined by solving the PNP equations. We plot the electric field strength across the axis and longitude center of the nanopore, along with the electric field considered in Eq. 2. The simulated electric field is non-uniform on both the lateral and transverse axes, as shown in the grey boxes at the top and left of Figure 5.1.

5.4 Coupled Random-Walk Simulations to Probe Off-Axis Effects on Individual-Event Analysis

We numerically simulated translocation events due to ellipsoidal particles in MATLAB in order to probe the influence of off-axis effects on individual-event analyses. We performed two

types of simulations in parallel, and then combined these simulations with results from the finite element model in order to generate an array of off-axis influenced resistive pulses. We described the first type of simulation, referred to here as a rotational simulation, previously in Chapters 3 and 4.[6] The second type of simulation simply tracks the radial diffusion (*i.e.* between the central pore axis and pore walls) of a protein within a nanopore, which we assume to be unbiased. Input parameters for the rotational simulations included $\Delta I_{MIN}/I_0$, $\Delta I_{MAX}/I_0$, the dipole moment, μ, the rotational diffusion coefficient of the protein, D_R, the pore geometry including length and diameter (assuming a perfect cylinder), the resistivity of the solution, the standard deviation of the baseline noise measured from a true baseline recording, and the dwell time, or t_d, of each particle in the nanopore. Input parameters for the lateral simulations include the 2-D lateral diffusion coefficient of the protein, D_L, the pore geometry including length and diameter (assuming a perfect cylinder), and the dwell time, or t_d, of each particle in the nanopore.

To generate the rotational portion of the resistive pulse signal, we adapted a model developed by Gauthier and Slater for translational motion,[19, 20] and assumed that any bias in rotation of the particle was due to the electric field acting on that particle's permanent dipole moment (assumed here to be along the principle axis of the particle). We simulated discrete 1-ns time steps, in which we varied the angle of the particle's principle axis relative to the electric field by a fixed step size based on the mean squared angular displacement around a single axis:

$$MSAD = \Delta\theta = 2D_R\Delta t \tag{5.5}$$

For each time step, we used the following equation to determine the probability that the particle would rotate in the positive or negative direction:

$$p_{\pm} = \frac{1}{1 + e^{\Delta U/k_BT}} = \frac{1}{1 + e^{\pm E\mu[cos(\theta-\Delta\theta)-cos(\theta+\Delta\theta)]/(2k_BT)}} \tag{5.6}$$

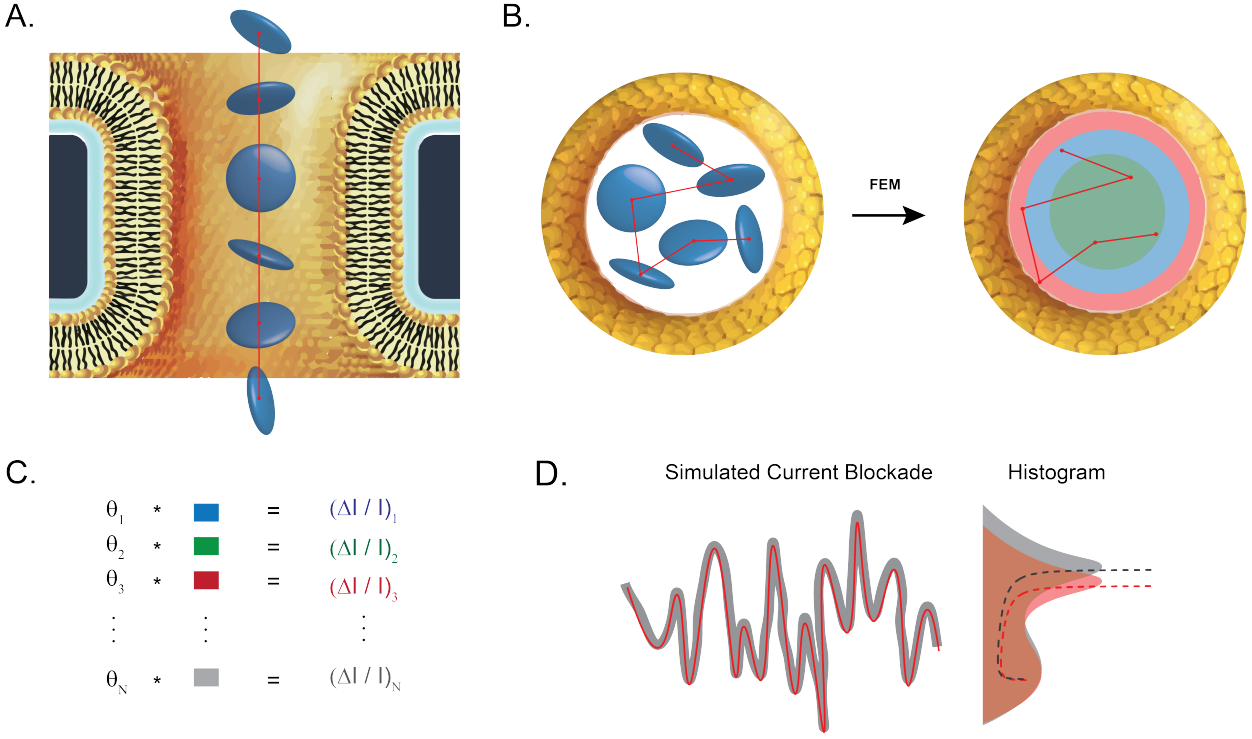


Figure 5.2. Overview of coupled random-walk simulations combined with FEM data. A) Side-view of a lipid-coated nanopore showing an oblate-shaped rotating particle, representing our rotational simulation. B) Top-view of a lipid-coated nanopore showing an oblate-shaped particle diffusing radially, and the discretization of that radial diffusion using our FEM model. C) By combining the rotational diffusion simulation ($\theta_1 : \theta_N$) with the discretized radial diffusion corresponding to different extents of off-axis effects (colored boxes), we can convert each combination of orientation and radial distance to an associated blockade amplitude, $(\Delta I/I_0)_{OffAx}$. D) Simulated current blockade from a single rotation profile combined with many radial diffusion simulations, and associated histograms with U-Shaped convolution fits (see Chapters 3 and 4).

After simulating the biased rotational diffusion of a protein transiting a nanopore, we performed a random lateral diffusion of that same protein relative to the central pore axis,[19, 20] and then combined the two simulations (both with a number of steps that we pre-defined based on an event length of 1 ms) using an error matrix generated from COMSOL describing an expected percent increase in the measured resistive pulse for a particle with a particular length-to-diameter ratio at various distances from the central pore axis. Figure 5.2 gives an overview of this simulation method, and a more thorough description is included in the appendix.

5.5 Results and Discussion

5.5.1 Off-Axis Position Modulates Resistive Pulse Amplitudes

A particle that passes through a nanopore near to the pore wall produces a larger resistive pulse than that same particle passing through the center of the pore. Figure 5.3 demonstrates this phenomenon, showing the passage of an 8-nm diameter dielectric particle through a nanopore along two paths along with their associated current traces. When the particle passes through the nanopore along the central axis, it produces a rounded resistive pulse shape due to gradual changes in access resistance at the entrance and exit of the pore. The off-axis passage, however, yields a square-shaped resistive pulse due to sharp changes in electric field at the corners of the pore. As expected the blockage current for off-axis translocation, ΔI_{off}, is larger than the blockage current from translocation along the central axis, ΔI.

The size of the nanopore relative to the size of the nanoparticle strongly influences the magnitude of off-axis effects. To quantify this influence, we varied the nanopore radius from 8 nm to 26.7 nm for a spherical particle with a radius of 4 nm. Figure 3 shows the distortion of resistive pulse magnitude, or deviation in $\Delta I/I_0$, for different off-axis positions the spherical nanoparticle in this scenario.

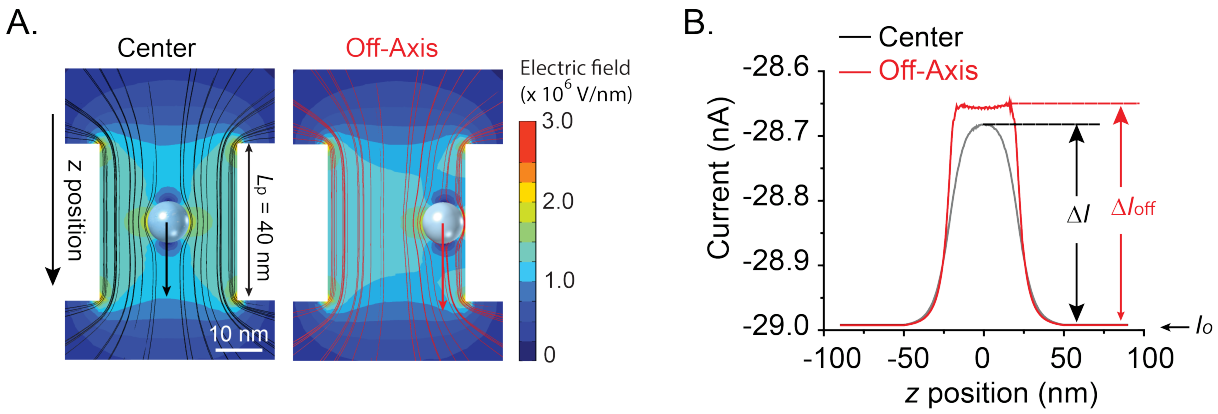


Figure 5.3. Influence of off-axis effects on the electric field distribution within a nanopore. A) Electric field distribution in a nanopore with an 8 nm diameter particle passing through the center (left) or off-axis (right). The fine streamlines represent the electric field. B) The measured current as a function of the particle position when particle transits through the center (black curve) and near the wall (red curve) of the nanopore.

5.5.2 Effects of Off-Axis Translocation on Estimates of Protein Volume and Ellipsoidal Shape from Population-Based Analyses

Most approaches to characterize protein with nanopores analyze populations of resistive pulses (see Chapter 1).[7, 21, 22] In an ideal scenario, a population of identical, perfectly spherical particles would all produce the same resistive pulse amplitude. In practice, ensemble analyses are influenced by the recording noise of the experimental setup,[23–25] which produces an approximately Gaussian distribution of resistive pulse amplitudes. The experimental $\Delta I/I_0$ distribution can be viewed as the convolution of the recording noise and the theoretical resistive pulse amplitude (see Figure 3.2). For a spherical particle, this theoretical amplitude is a Dirac-δ function, while for a non-spherical ellipsoid, the theoretical value is described by a U-shaped distribution that represents the probability of sampling any electrical shape factor, γ, between γ_{MIN} and γ_{MAX}.[3]

Off-axis translocation influences the population-based estimation of protein parameters in two ways: 1) off-axis effects increase $\Delta I/I_0$ values and thus lead to an overestimation of the particle volume by up to 30% depending on the size and ellipsoidal shape of the particle, and 2) the spread of the distribution of $\Delta I/I_0$ for a population of resistive pulses increases, as not

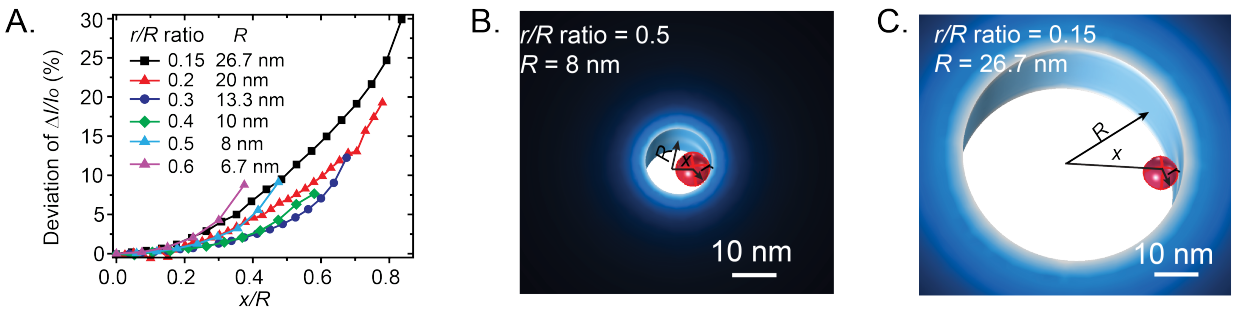


Figure 5.4. Influence of the ratio between particle size and nanopore size on off-axis effects. A)The deviation of $\Delta I/I_0$ as a function of radial position. We simulated the same spherical nanoparticle with a radius, *r*, of 4 nm, passing through different nanopores with radius R. B) Cartoon example of a large r/R ratio, simulating a nanopore with a diameter of 16 nm, and a particle with a diameter of 8 nm. C) Cartoon example of a small r/R ratio, simulating a nanopore with a diameter of 53.4 nm, and a particle with a diameter of 8 nm.

all particles transit the pore at a consistent off-axis position, which can produce large errors in length-to-diameter ratio relative to reference values (Figure 5.5-C-E, grey regions). Both of these consequences of off-axis translocation influence the accuracy of parameter estimates determined through population-based analyses of resistive pulses.

5.5.3 Effects of Off-Axis Translocation on Estimates of Protein Volume and Ellipsoidal Shape from Individual-Event Analyses

Non-spherical particles rotate as they translocate through the nanopore, and these rotations produce characteristic fluctuations in their associated resistive pulse. Figure 5.2-D shows the fluctuations within a resistive pulse caused by different orientations of a non-spherical particle (red), as well as the variation in those fluctuations due to off-axis effects (grey). The ratio of maximum to minimum blockade current, $\Delta I_{max}/\Delta I_{min}$, can be used to estimate the particle's ellipsoidal shape (see Chapter 3,4). When we perform a coupled random-walk simulation, as described above, on a prolate-shaped particle ($m = 2.5$) and on an oblate-shaped particle ($m = 0.3$), we find, somewhat surprisingly, that the oblate particle is associated with larger influence from off-axis effects than the prolate particle. We previously observed that length-to-diameter ratio estimates of prolate-shaped particles tend to be associated with larger uncertainties,[6] and we report a similar finding above for population-based analyses. In

individual-event analyses, however, we observed a ∼5% error between the ideal (*i.e.* without off-axis effects) and measured (*i.e.* with off-axis effects) values for the length-to-diameter ratio in the case of the prolate particle, and a ∼13% error for the oblate particle. For both particles, off-axis effects led to estimates of length-to-diameter ratio that were more extreme (*e.g.* further from a sphere) than expected. We are in the process of performing a more thorough evaluation of the effects of off-axis translocation on individual-event analyses by, again, performing a parameter sweep across a range of particle shapes and sizes relative to the nanopore, and then using the output values from that parameter sweep within our coupled random-walk simulations. From this evaluation, we hope to describe general trends in the influence of off-axis effects on parameter estimates not only of length-to-diameter ratio, but also of volume and dipole moment. We expect that such trends will be useful in improving the accuracy of future efforts to characterize proteins passing uninhibited through synthetic nanopores.

5.5.4 Efforts to Correct for Effects of Off-Axis Translocation in Population-Based Analysis

Similar to the calibration for nanopore geometry presented in Chapter 3, here we performed an experiment characterizing streptavidin protein in a nanopore coated with Tween-20, fit the results with a 2-Gaussian distribution, shifted each respective Gaussian distribution such that streptavidin had a length-to-diameter ratio of 1 (*i.e.* had a spherical shape), and then applied that same shift to experimental data measured for Alcohol Dehydrogenase (ADH). Saleh *et al.* previously reported a similar method for correcting off-axis effects on the microscale.[26] We used the following relationships to manipulate the Gaussian distributions:

$$\mu_Y = a\mu_X + b \tag{5.7}$$

and

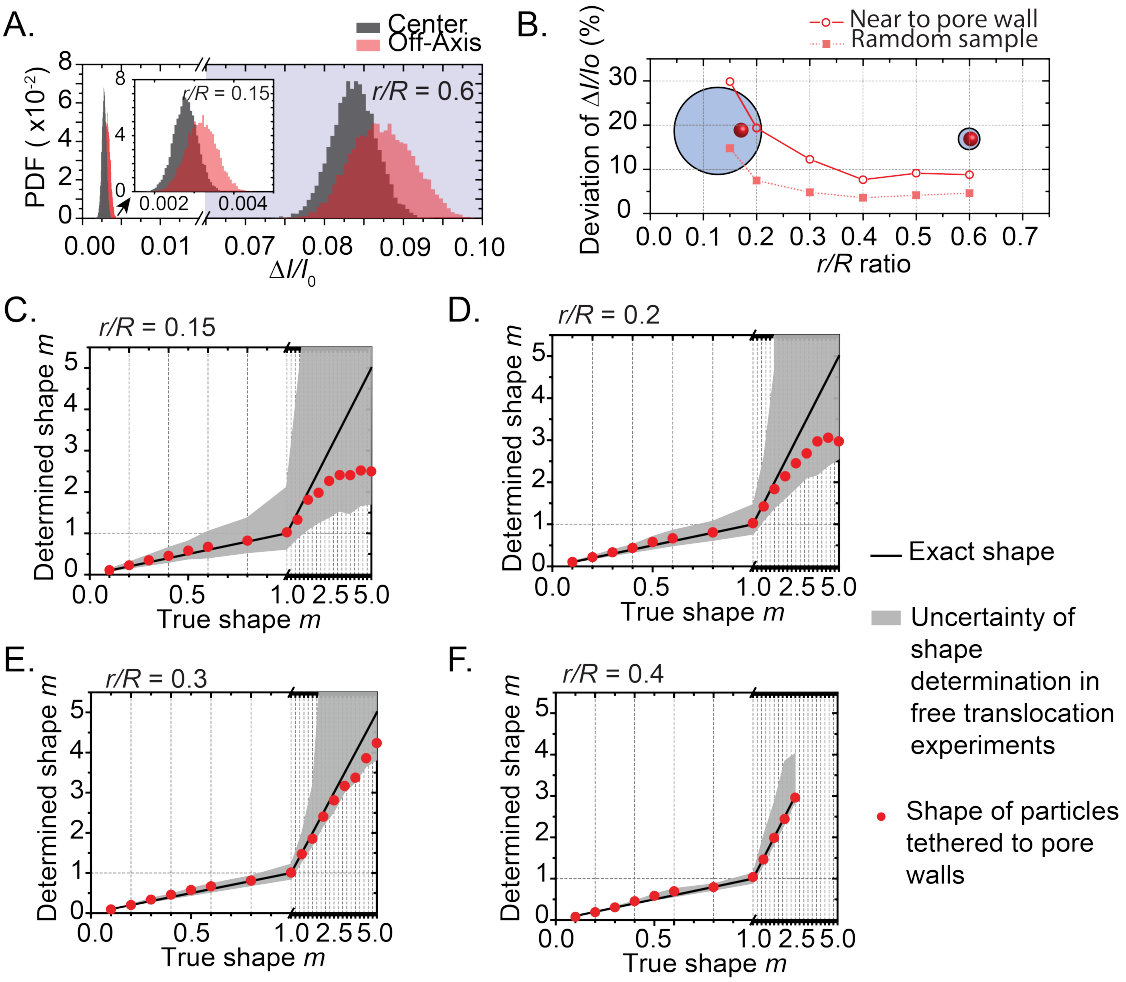


Figure 5.5. Effects of off-axis translocation on population-based analysis of length-to-diameter ratio and volume. A) The probability distribution of $\Delta I/I_0$ values determined by randomly sampling a translocating nanoparticle at different distances from the central axis. We fit the distribution with a probability distribution function (PDF) that is calculated by convolving the $\Delta I/I_0$ with normally distributed noise (25 pA), and the spatial distribution of the particle in the nanopore. B) The deviation of $\Delta I/I_0$ as a function of the r/R ratio. C-F) Determined versus true shapes for different particle-to-pore size ratios, ranging from 0.15 to 0.4. The largest deviations away from true shape values in both tethered and untethered translocations occured when the pore was much larger than the protein and the particle was a relatively extreme prolate shape.

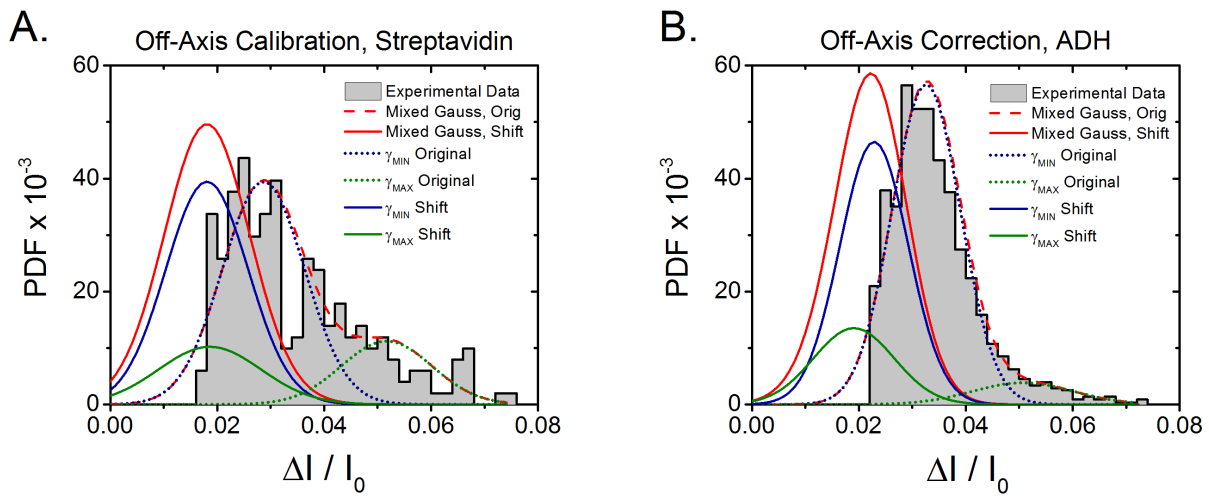


Figure 5.6. 2-Gaussian fit of experimental data to correct for effects of off-axis translocation. A) Histogram of population-based data for streptavidin, with initial 2-gaussian fits shown by dotted lines, and shifted 2-gaussian fits shown with solid lines. B) Histogram of population-based data for ADH, with initial 2-gaussian fits shown by dotted lines, and shifted 2-gaussian fits shown with solid lines.

$$\sigma_Y^2 = a^2\sigma_X^2 \tag{5.8}$$

Where μ represents the mean of each Gaussian distribution, and σ^2 represents the spread, or variance of each distribution. Figure 5.6 shows the results of this attempted correction.

We found that for these two proteins in particular, we were unable to correct for off-axis effects. There are several reasons that could explain the failure of this approach: 1) we have found, as discussed in Chapter 4, that resistive pulses measured on nanopores coated with Tween-20 are likely prolonged due to non-specific adhesion with the nanopore walls [27] and/or electroosmotic flow [16, 28, 29] that acts to slow proteins down during transit. These prolonged events ($t_d > 2ms$) do not allow for accurate estimations of protein parameters, especially with regard to length-to-diameter ratio. 2) Population-based analysis selects only a single, maximum value for each translocation event, and therefore effects of the protein not sampling all orientations within a single event (as is the case for the shortest events that we analyzed here) will confound effects of off-axis translocation in subsequent analyses.

A future method to correct for these effects would be to perform tethered and unteth-

ered experiments of the same protein on the same lipid-bilayer coated nanopore, and then compare the resulting resistive pulse signals from each experiment, looking in particular at frequency components of the signals. If we treat the effects of off-axis translocation as another contributor to the overall noise in nanopore recordings of freely translocating proteins, as proposed in Chapter 4, then we might expect this noise component to be related to the rate of lateral diffusion of the protein within the pore, and such a component could appear in power spectra or autocorrelations of resistive pulses.

5.6 Conclusion

Off-axis effects produce larger-than-expected, population-based estimates of particle volume in resistive pulse experiments, and can result in more than 200% error in population-based estimates of length-to-diameter ratio in extreme scenarios (prolate-shaped particles, small ratio of particle size to nanopore size). Additionally, effects of off-axis translocation influence parameter estimates from individual-event analyses, though this influence appears to be less-pronounced than for population-based analyses. We observed such a result experimentally in Chapter 4, where estimates of length-to-diameter ratio for prolate-shaped proteins had much larger uncertainties in population-based analyses (Figure 4.5) than in individual-event analyses (Figure 4.3). An initial conclusion of this work, therefore, is that estimates from individual events should be emphasized in future work characterizing proteins with nanopores, especially when those proteins are able to diffuse laterally within the nanopore. Moreover, selecting a nanopore with a diameter that is as small as possible while still allowing the particle to fully rotate within the pore minimizes off-axis effects in both population-based and individual-event analyses, however, this choice may increase the amount of time required for the particle to sample both maxima due to effects of confinement within the nanopore. Tethering nanoparticles to fluid coatings on the nanopore walls mitigates the effects of off-axis translocation on estimates of nanoparticle shapes, in particular for particles with shape

values between $m = 0.1$ and $m = 2.0$.[6] By considering off-axis effects in nanopore sensing, and possibly correcting for them,[26] we are further improving the ability to accurately characterize individual nanometer-sized particles free in solution.

References

[1] A. Han et al. “Sensing protein molecules using nanofabricated pores”. *Applied Physics Letters* 88.9 (2006), p. 093901.

[2] C. Dekker. “Solid-state nanopores”. *Nature Nanotechnology* 2.4 (2007). WOS:000245920900010, pp. 209–215.

[3] D. C. Golibersuch. “Observation of Aspherical Particle Rotation in Poiseuille Flow via the Resistance Pulse Technique”. *Biophysical Journal* 13.3 (1973), pp. 265–280.

[4] R. W. DeBlois, E. E. Uzgiris, D. H. Cluxton, and H. M. Mazzone. “Comparative measurements of size and polydispersity of several insect viruses”. *Analytical biochemistry* 90.1 (1978), pp. 273–288.

[5] R. W. DeBlois and C. P. Bean. “Counting and sizing of submicron particles by the resistive pulse technique”. *Review of Scientific Instruments* 41.7 (1970), pp. 909–916.

[6] E. C. Yusko et al. “Real-time shape approximation and fingerprinting of single proteins using a nanopore”. *Nature Nanotechnology* 12 (2017), pp. 360–367.

[7] E. C. Yusko et al. “Controlling protein translocation through nanopores with bio-inspired fluid walls”. *Nature Nanotechnology* 6.4 (2011), pp. 253–60.

[8] Z. Qin, J. Zhe, and G.-X. Wang. “Effects of particle’s off-axis position, shape, orientation and entry position on resistance changes of micro Coulter counting devices”. *Measurement Science and Technology* 22.4 (2011), p. 045804.

[9] W. R. Smythe. “OffAxis Particles in Coulter Type Counters”. *Review of Scientific Instruments* 43.5 (1972), pp. 817–818.

[10] L. I. Berge, T. Jossang, and J. Feder. “Off-axis response for particles passing through long apertures in Coulter-type counters”. *Measurement Science and Technology* 1.6 (1990), p. 471.

[11] J. L. Oncley. “The Investigation of Proteins by Dielectric Measurements.” *Chemical Reviews* 30.3 (1942), pp. 433–450.

[12] R. T. Perera, R. P. Johnson, M. A. Edwards, and H. S. White. “Effect of the Electric Double Layer on the Activation Energy of Ion Transport in Conical Nanopores”. *The Journal of Physical Chemistry C* 119.43 (2015), pp. 24299–24306.

[13] J. Liu et al. “Surface Charge Density Determination of Single Conical Nanopores Based on Normalized Ion Current Rectification”. *Langmuir* 28.2 (2012), pp. 1588–1595.

[14] C. Kubeil and A. Bund. “The Role of Nanopore Geometry for the Rectification of Ionic Currents”. *The Journal of Physical Chemistry C* 115.16 (2011), pp. 7866–7873.

[15] Z. Jiang and D. Stein. “Charge regulation in nanopore ionic field-effect transistors”. *Physical Review E* 83.3 (2011), p. 031203.

[16] L. Galla et al. “Hydrodynamic slip on DNA observed by optical tweezers-controlled translocation experiments with solid-state and lipid-coated nanopores”. *Nano letters* 14.7 (2014), pp. 4176–4182.

[17] D. A. MacInnes and A. S. Brown. “The Determination of Activity Coefficients from the Potentials of Concentration Cells with Transference.” *Chemical Reviews* 18.2 (1936), pp. 335–348.

[18] A. S. Brown and D. A. MacInnes. “The determination of activity coefficients from the potentials of concentration cells with transference. I. Sodium chloride at 25”. *Journal of the American Chemical Society* 57.7 (1935), pp. 1356–1362.

[19] M. G. Gauthier and G. W. Slater. “A Monte Carlo algorithm to study polymer translocation through nanopores. I. Theory and numerical approach”. *The Journal of chemical physics* 128.6 (2008), 02B612.

[20] M. G. Gauthier and G. W. Slater. “A Monte Carlo algorithm to study polymer translocation through nanopores. II. Scaling laws”. *The Journal of chemical physics* 128.20 (2008), 05B619.

[21] D. Fologea, B. Ledden, D. S. McNabb, and J. Li. “Electrical characterization of protein molecules by a solid-state nanopore”. *Applied physics letters* 91.5 (2007), p. 053901.

[22] A. Han et al. “Label-free detection of single protein molecules and protein protein interactions using synthetic nanopores”. *Analytical chemistry* 80.12 (2008), pp. 4651–4658.

[23] J. D. Uram, K. Ke, and M. Mayer. “Noise and bandwidth of current recordings from submicrometer pores and nanopores”. *ACS nano* 2.5 (2008), pp. 857–872.

[24] C. Wen et al. “Generalized Noise Study of Solid-State Nanopores at Low Frequencies”. *ACS Sensors* 2.2 (2017), pp. 300–307.

[25] R. M. M. Smeets, U. F. Keyser, N. H. Dekker, and C. Dekker. “Noise in solid-state nanopores”. *Proceedings of the National Academy of Sciences* 105.2 (2008), pp. 417–421.

[26] O. Saleh and L. Sohn. “Correcting off-axis effects in an on-chip resistive-pulse analyzer”. *Review of scientific instruments* 73.12 (2002), pp. 4396–4398.

[27] L. T. Sexton et al. “An adsorption-based model for pulse duration in resistive-pulse protein sensing”. *Journal of the American Chemical Society* 132.19 (2010), pp. 6755–6763.

[28] A. Sischka et al. “Controlled translocation of DNA through nanopores in carbon nano-, silicon-nitride- and lipid-coated membranes”. *Analyst* 140.14 (2015), pp. 4843–4847.

[29] M. Firnkes, D. Pedone, J. Knezevic, M. Döblinger, and U. Rant. “Electrically Facilitated Translocations of Proteins through Silicon Nitride Nanopores: Conjoint and Competitive Action of Diffusion, Electrophoresis, and Electroosmosis”. *Nano Letters* 10.6 (2010), pp. 2162–2167.

Chapter 6: Outlook

The work presented throughout this thesis has focused on detecting and analyzing *individual* proteins with lipid bilayer coated nanopores. Chapter 1 outlined fundamental principles of nanopore-based protein sensing and described progress in this field related to characterizing amyloids; Chapter 2 discussed efforts to fabricate nanopores in different substrates and the associated challenges with those efforts; Chapter 3 detailed methods for the analysis and five-dimensional characterization of resistive pulses generated by lipid-tethered proteins; Chapter 4 explored the considerations and nuances associated with label-free sensing of the resistive pulses produced by untethered proteins; and finally, Chapter 5 investigated, *via* computational simulation, off-axis effects due to freely translocating proteins. These studies present a glimpse toward the forefront of nanopore sensing; we are currently on the cusp of an interdisciplinary convergence, where biophysical experimentalists are joining forces with bioinformaticians, machine learning experts, biomedical engineers, computational biologists, mathematicians, microfabrication experts, and theoretical physicists in order to realize the full potential of the resistive-pulse sensing technique. In this chapter, I discuss this convergence, and describe the ongoing efforts within our research group – as well as outside of it – aiming to push the limits of nanopore-based protein characterization.

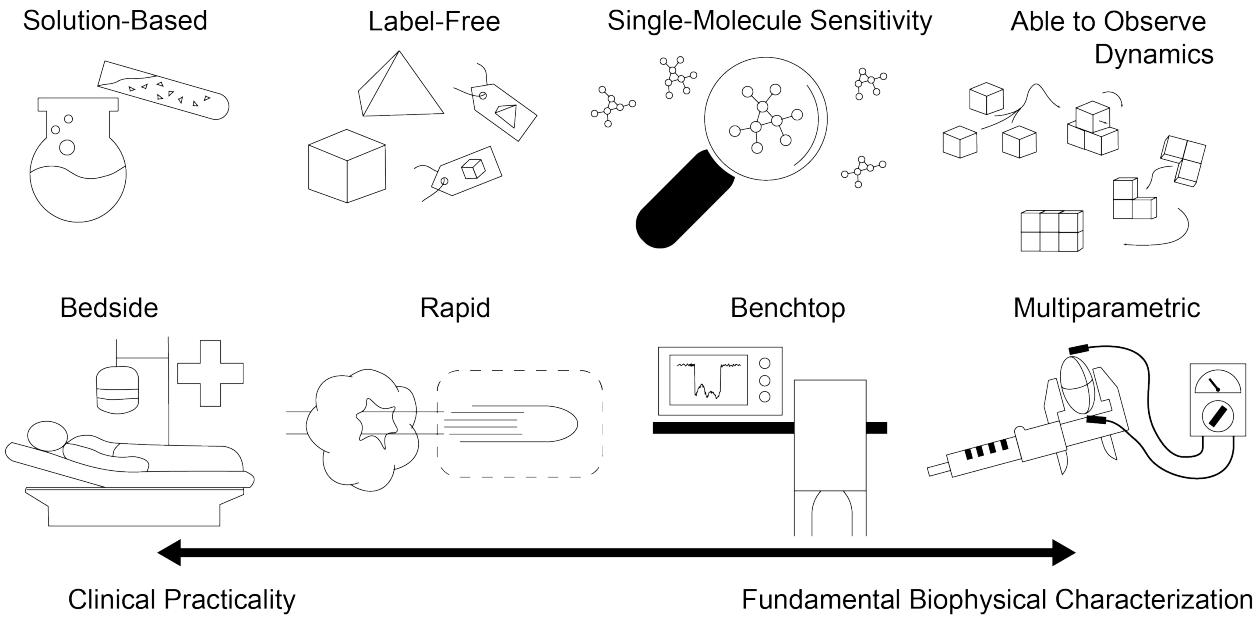


Figure 6.1. Advantages of the nanopore sensing technique. These advantages are further related to their potential usefulness in clinical analysis or fundamental biophysical characterization.

6.1 Avenues for Future Development in Nanopore Sensing

Resistive-pulse sensing with nanopores is a technique with a variety of unique advantages that make it well-suited to both clinical use (see MinION[1]) and fundamental biophysical studies.[2–14] Figure 6.1 highlights some of these advantages and acts as a starting point in the design of future nanopore-based systems and applications requiring such capabilities.

In order for the technique to see success in characterizing proteins on a commercial scale, however, a number of challenges must be met. Proteins are an incredibly diverse class of biomolecules, both in form and in function. Methods to interrogate complex mixtures of proteins, therefore, need to accommodate different shapes, sizes, surface charges, stabilities, and functionalities.[15] Because nanopores are sized to match their target analyte, a single nanopore in a freestanding membrane has a relatively narrow size-range of proteins to which it is sensitive. That is, proteins that are much smaller than the pore will not produce a resistive pulse distinguishable from baseline noise, and those that are too large will either be unable to rotate within the pore or simply not translocate through the pore. There are several different approaches to overcome the geometric issue: 1) reduce the recording noise of

the experimental setup such that a relatively large nanopore becomes sensitive to very small proteins,[16–21] 2) perform experiments with an array of differently sized, independently monitored nanopores that are each sensitive to a portion of the protein mixture, 3) sort protein mixtures into groups based on size and evaluate each group independently using a different nanopore, or 4) denature all proteins and pass them through a small nanopore as a single strand similar to nucleic acid sequencing.[22] Charge is the next challenge; in a diverse protein mixture, some proteins will be highly charged will transit the nanopore much more rapidly than those with fewer net charges.[23] While dwell time can be used here as an additional descriptor for distinguishing proteins (*e.g.* by setting the pH of the buffer we can target proteins with an isoelectric point in a certain range), short-lived resistive pulses have inherently less information content than long-lived pulses and produce larger uncertainties in parameter estimates (see Chapters 3,4). As with geometry, there are several solutions to inequalities in net charge: 1) anchor proteins to a fluid surface coating such that the transit speed of *all* proteins is slowed considerably, though at the cost of chemically modifying the proteins [23] 2) slow *all* proteins by increasing solution viscosity, either by decreasing temperature or adding glycerol,[24] but at the cost of reduced signal, 3) increase the recording bandwidth into the MHz range such that pulses less than 10 μs can be effectively analyzed, [21] or 4) incorporate a preprocessing step where proteins are first separated into groups and then interrogated independently and in parallel. An additional challenge with diverse protein mixtures is protein stability, as an inherent limitation to resistive pulse sensing is using a buffer solution with high conductivity (*i.e.* high salt concentration), which may promote protein denaturation or aggregation.[25] Instability in high-salt buffers can be overcome by: 1) decreasing the salt content to physiologic levels at the cost of reduced signal, 2) limiting the time the protein spends in the high salt buffer, either by only analyzing a brief initial period after the protein is injected, or by injecting proteins using a system analogous to an H-filter in microfluidics,[26] where a physiologic buffer flows alongside recording buffer and proteins diffuse across the interface directly before transiting a nanopore, or 3) exploring alternative

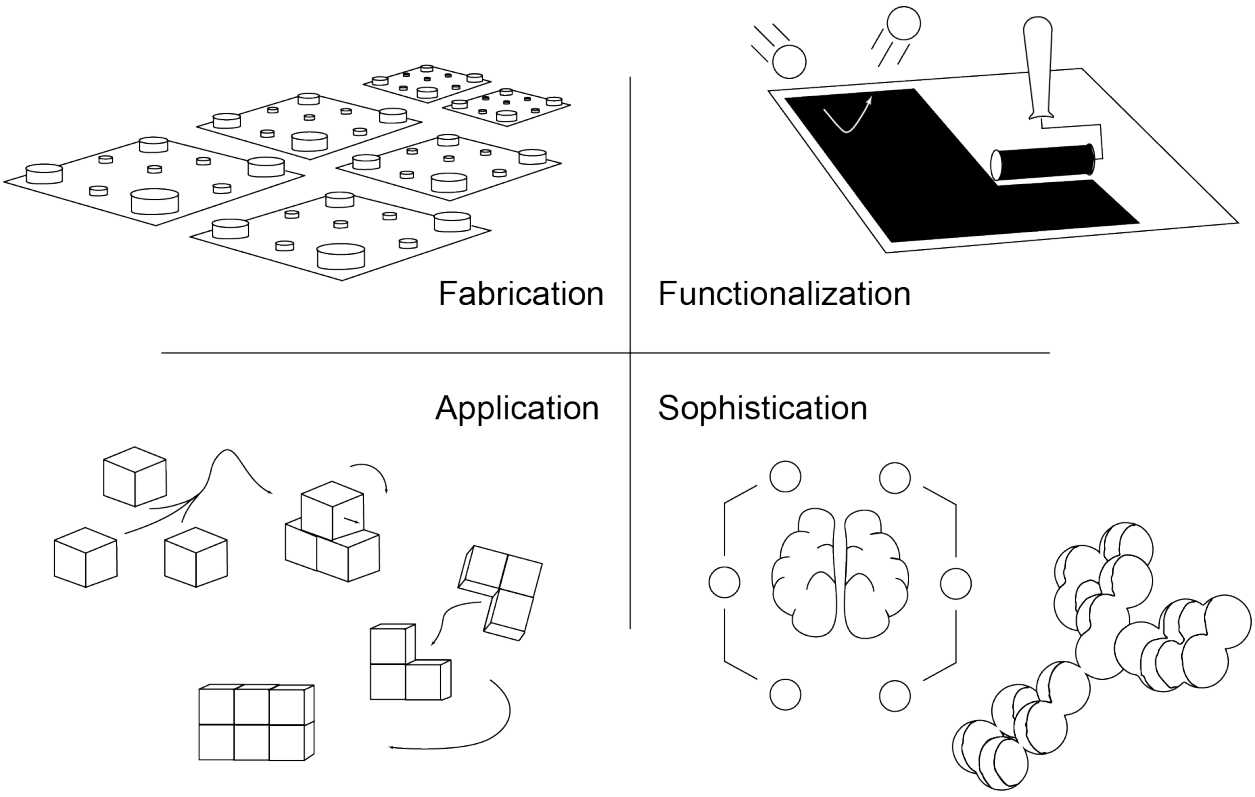


Figure 6.2. Schematic showing four areas of future development in nanopore sensing. These areas include Fabrication, as discussed in Chapter 2, Functionalization, as discussed in Chapters 3 and 4, Sophistication, as alluded to in Chapter 5, and Application, of which amyloids have been highlighted throughout this dissertation.

experimental conditions (*e.g.* different salts, temperatures, stabilizers, etc.). Many of these solutions will require designs that incorporate some combination of at least four overlapping themes of ongoing development, as shown in Figure 6.2.

6.1.1 Prospective Advances in the Fabrication of Nanopores

As discussed in Chapter 2, there are many different ways to fabricate a nanopore in a synthetic membrane. Some of these methods, like electron beam drilling,[27] require expensive equipment and technical expertise, while others, like dielectric breakdown,[28, 29] can be performed quickly on a laboratory benchtop with modest electronics. Each technique has advantages and disadvantages, as do the resulting nanopores. Several trends appear to be emerging among the nanopore community, which include nanopores being fabricated: 1) in low-noise substrates like fused silica, which also enable seamless integration with optical

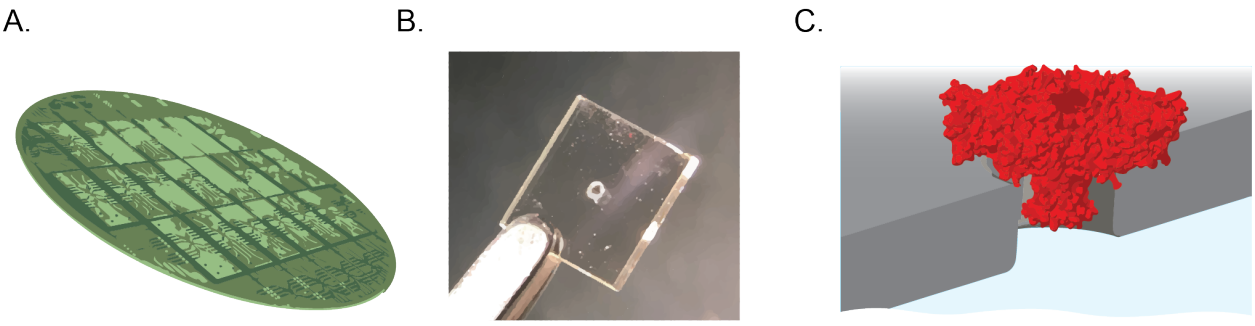


Figure 6.3. Future themes for the fabrication of synthetic nanopores. A) Cartoon showing a wafer containing multiple circuitry components embedded in nanopore chips and fabricated in batch, B) Image of a low-noise fused-silica scaffold containing a freestanding membrane with a nanopore, C) Schematic of a biological nanopore (here, α-aerolysin (5JZT)) held within a synthetic substrate as a hybrid system. Adapted with permission.[30] Copyright 2019, IOP Publishing.

techniques,[30] 2) in batch, rather than serially, to reduce cost and time burden,[31] 3) with on-board amplifiers (*e.g.* FET-based systems) to improve recording characteristics,[21] and 4) in hybrid systems, where biological components are combined with synthetic components, much like lipid bilayer coatings.[32] Figure 6.3 provides a visual overview of these themes.

We recently published work, mentioned in Chapter 2, in which we fabricated freestanding membranes on fused silica substrates on a wafer scale.[30] Not only did these chips outperform those fabricated in silicon, but they also allowed for coupled optical measurements without additional noise. Within the scope of this project, we also evaluated prior work in nanopore fabrication and found that nanopores with the best noise characteristics to date all have amplifiers integrated in close proximity to the nanopore (Table 2.1). Oxford Nanopore Technologies has popularized such a design in a commercial format with their MinION device and associated products for nucleic acid sequencing, which allow the user to monitor hundreds of nanopores independently and in parallel, although with far lower bandwidth than required for protein characterization.[1] Hybrid systems, where biological and synthetic components are combined, are also gaining popularity. For example, Wanunu *et al.* have spanned a synthetic nanopore with lipids and then incorporated a biological nanopore into the nanoscale planar lipid bilayer.[33] Other groups have recently engineered a template designed to contain a discrete number of pore-forming peptides, using differently sized templates to control

the resulting nanopore size in order to overcome traditional size limitations of biological nanopores.[34, 35]

Resistive pulses, which are typically measured across nanopores using Ag/AgCl electrodes immersed in *cis* and *trans* compartments filled with ionic solution, can also be measured using field-effect transistors (FET), which have been recently developed using graphene nanoribbons.[36–38] In this format, the FET detects and amplifies a local *change in potential* produced by the translocating biomolecule, rather than a change in current.[39] In state-of-the-art amplifiers used for taking measurements across nanopores, capacitive noise dominates high frequency regimes (*e.g.* greater than 100 kHz); these amplifiers rely on negative op-amp circuits to measure the current through the pore while keeping a consistent applied voltage, which in turn produces a current noise (stdev) of 1 nA at 10 MHz.[40] FET-based electrical measurements, on the other hand, are limited by a parasitic capacitance between the source and drain electrodes.[39] If these electrodes are sufficiently passivated, then thermal noise, rather than capacitive noise, dominates at high-bandwidths. To investigate the magnitude of thermal noise in the context of high-bandwidth protein characterization, we followed an approach similar to that presented by Parkin *et al.*,[39] estimating the thermal noise-limited using the following equation:

$$BW = \frac{(\Delta V_{en}/SNR)^2}{4kTR_{eq}(D_{eff})} \tag{6.1}$$

where BW is the bandwidth, SNR is the signal to noise ratio (set to be 5), k is the Boltzmann constant, T is the temperature of the resistor, R_{eq} is the Thevenin equivalent resistance $R_{eq} = R_a(R_a+R_p)/R_{tot}$, and V_{en} is the change in voltage at the nanopore entrance, represented by the following equation:

$$\Delta V_{en} = V_{ionic}\left(\frac{R_a(D)}{R_{tot}(D)} - \frac{R_a(D_{eff})}{R_{tot}(D_{eff})}\right) \tag{6.2}$$

where V_{ionic} is the applied ionic voltage (*e.g.* with Ag/AgCl electrodes), D is the diameter

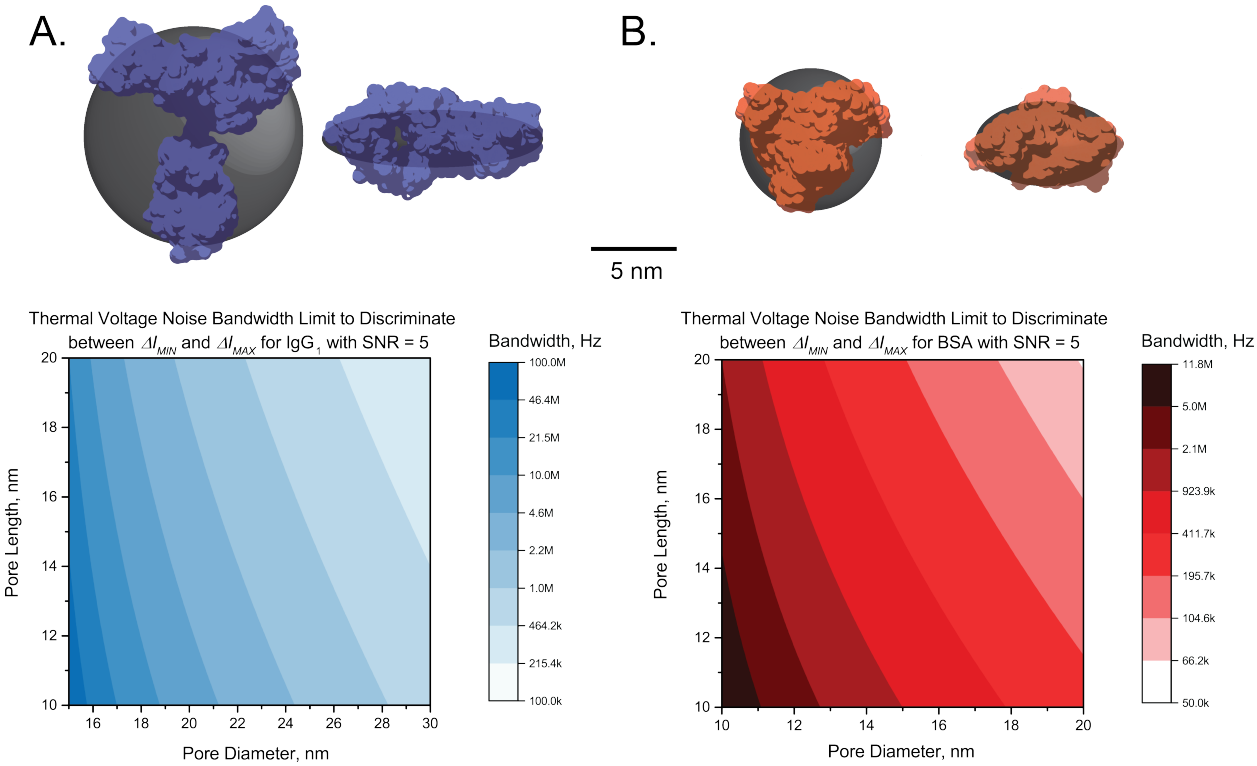


Figure 6.4. Thermal noise-limited bandwidth for characterizing proteins with an FET-based nanopore. A) Crystal structure (blue) of IgG_1 overlaid with an approximate ellipsoid estimate ($m = 0.3$) and associated heat map showing the thermal noise-limited bandwidth of resolving the difference between minimum and maximum orientations of the protein within the nanopore. B) Crystal structure (red) of BSA overlaid with an approximate ellipsoid estimate ($m = 0.5$) and associated heat map showing the thermal noise-limited bandwidth of resolving the difference between minimum and maximum orientations of the protein within the nanopore.

of the nanopore, and D_{eff} is the effective diameter of the nanopore during the translocation of a biomolecule - $D_{eff} \approxeq (D^2 - d^2)^{1/2}$. We calculated the thermal noise-limited bandwidth in order to discriminate between the minimum and maximum orientations of IgG_1 protein and Bovine Serum Albumin for a range of nanopore sizes and present the results in Figure 6.4.

From the plots in Figure 6.4, we see several trends: proteins with more extreme shapes can be analyzed at higher bandwidths as they produce larger variation between minimum and maximum orientations (100 MHz for IgG_1 *versus* 10 MHz for BSA), and pores with smaller diameters and shorter lengths are more sensitive to translocation events than larger pores for the same protein (as discussed throughout this dissertation). This estimate assumes no surface charge and complete electrode passivation passivation, and represents an upper limit

in bandwidth for an ideal recording scenario scenario. Because of the trend of increased sensitivity in smaller pores due to scaling of thermal noise, FET-based systems may be appealing for studying proteins and peptides smaller than 10 kDa, or for protein sequencing using nanopores.[22]

6.1.2 Functionalization of Nanopores with Polymer Thin Films

While bio-inspired lipid bilayer coatings have addressed many key challenges in protein characterization with synthetic nanopores by reducing nonspecific adsorption, concentrating analyte near the nanopore entrance, eliminating electroosmotic flow, and extending translocation times,[23] they have several shortcomings: 1) lipid bilayer coatings have low experimental success rates, with approximately 10 percent of attempts producing a stable coating, 2) they add an undulatory noise element to electrical recordings (see Chapter 2 for discussion on complementary conical lipids), and 3) they are single-use and their formation requires a vigorously cleaned surface which destroys the nanopore over time.[23] Thin film polymer coatings offer an intriguing alternative which could address the pitfalls of lipid bilayer coatings while retaining some of their advantages. We are currently investigating a polymer coating based on poly(N,N-dimethylacrylamide) (PDMA), and have used it successfully to determine the length-to-diameter ratio of IgG_1 protein. These results are quite encouraging considering that, until now, we have only been able to determine ellipsoidal shape accurately using nanopores coated with lipid bilayers; we have found that nanopores with other coatings like Tween-20 (see Chapter 4) lead to large errors in estimated shape. We are hoping to investigate this particular polymer coating more thoroughly in the context of protein characterization, as it may be easier to use and more suitable to different nanopore geometries and surface chemistries than lipid bilayer coatings.

We also performed a theoretical investigation of the stability of a variety of thin film coatings that have been used to prevent protein adhesion to synthetic surfaces (see Appendices). We found that many different thin films would, in theory, form stable coatings on an SiO_x

membrane, although in practice, applying them to a complex three dimensional nanopore geometry could present a challenge, as these types of films are typically applied using spin coating techniques that may not be suitable for trans-pore coatings. Additionally, there are also issues with the inability to anchor proteins to the polymeric surface in order to slow translocation; these issues could be addressed by fastening the protein to a secondary structure (*e.g.* a large engineered biomolecule unable to translocate through the nanopore) to observe them for long periods of time inside of the nanopore. However, considering the noise generated by the undulations of the lipid bilayer would be eliminated, and given that the nanopore community is shifting toward low-noise platforms with high bandwidths, there will be less of a need for "long" events in order to estimate protein parameters, depending on the desired application, of course.

6.1.3 Data-Driven Approaches for Resistive Pulse Analyses

As the nanopore community moves toward low-noise platforms, we expect the information content of each resistive pulse signal to increase, and this rich information content will make it possible to improve upon the ellipsoidal model described in detail throughout this dissertation. We have an ongoing project with this particular goal in mind, and we are in the process of developing hydrodynamic bead models to represent complex, anisotropic protein shapes more accurately. IgG_1, for instance, is a Y-shaped molecule that is not particularly suited to an ellipsoid approximation (see Figure 4.12), and thus we have developed a bead model for IgG_1 and produced a simulated distribution of resistive pulse values from that model, as shown in Figure 6.5.

While we are approaching the capability of simulating current traces from a rigid arrangement of beads in space passing through a nanopore, relating an experimentally measured current trace to the set of beads that likely produced it is a complicated inverse problem (*e.g.* multiple different bead arrangements could produce a similar looking current trace). We are in the process of combining finite element methods with machine learning techniques

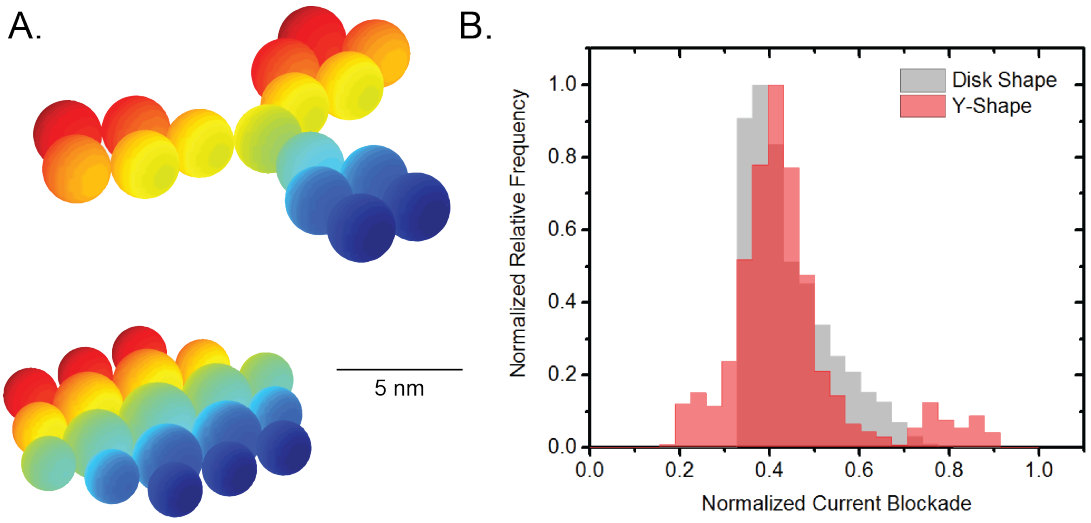


Figure 6.5. Preliminary results for hydrodynamic bead modelling of IgG_1. Beads arranged in a Y-shape produce a different distribution of current blockade values than beads arranged in a disk. A) Example of two bead models for an IgG_1 protein, both with a volume of $\sim$300 nm^3, one approximating a Y-shape (top) and the other approximating an oblate (bottom) as we are doing in this work. B) Normalized histograms of the current blockade values generated from simulated translocation events of either the disk-shaped (gray) or y-shaped (red) bead arrangement.

to tackle this problem.

With regard to machine learning techniques, we are currently in the process of extracting features from resistive pulse signals and evaluating their power in discriminating between pulses generated by particles with different physical properties. In this effort, we are investigating resistive pulses measured experimentally, as well as those generated using random-walk simulations, as described in previous chapters. In order to produce data in quantities sufficient for training machine learning algorithms, I programmed a parallel version of the random-walk simulations that is executable on a graphics processing unit (GPU). This version of the simulation reduced the computation time for simulations by several orders of magnitude relative to the MATLAB version, and is described in more detail in the Appendices.

6.1.4 Ongoing Efforts to Characterize Tau Protein and Oligomeric Aggregates

Despite the fact that low-n oligomers of microtubule associated protein tau (a.k.a. tau) have been shown to exhibit toxicity *in vivo*,[41, 42] these protein aggregates are poorly understood with regard to their structure, shape, and mechanism of action. Consequently, biophysical characterization of such oligomers on a single-particle basis may provide insight into why particular species are toxic, and could inform strategies for therapeutic intervention. Structural characterization of tau oligomers is challenging because they are highly dynamic and exist in a metastable, heterogenous population.[43, 44] We have preliminary data demonstrating that lipid bilayer-coated solid-state nanopores make it possible to characterize low-n oligomers of tau protein in aqueous solution and differentiate between them based on their shape and volume. Figure 6.6 shows a proposed aggregation pathway based on our findings from nanopore analysis, with the monomer structure determined using Nuclear Magnetic Resonance (NMR) spectroscopy.[45] We are in the process of applying supervised cluster analysis – both with k-means clustering and using a Gaussian Mixture Model – to the individual-event data in order to inform the grouping of particular events. From these clusters of events, we will determine of most probable length-to-diameter ratio, volume, and dipole moment of four distinct oligomeric sub-populations, none of which have, to the best of our knowledge, previously been reported using single-molecule, label-free measurements.

Tau aggregation is associated with a class of neurodenerative diseases known as tauopathies, and abnormal phosphorylation of tau has repeatedly been implicated as one of the driving forces of this process.[46, 47] Abnormal phosphorylation of tau destabilizes microtubules, degrades neuronal function, and has been a major target in designing drugs against tauopathies.[48] Phosphorylation of serine residues – Ser-214, Ser-258 and Ser-26 – which are otherwise important for forming stabilizing hydrogen bonds, has been shown to reduce binding affinities between tau and microtubules.[49] Hyperphosphorylation of tau, therefore, not only prevents normal function in stabilizing microtubules, but also leads to a toxic gain of function

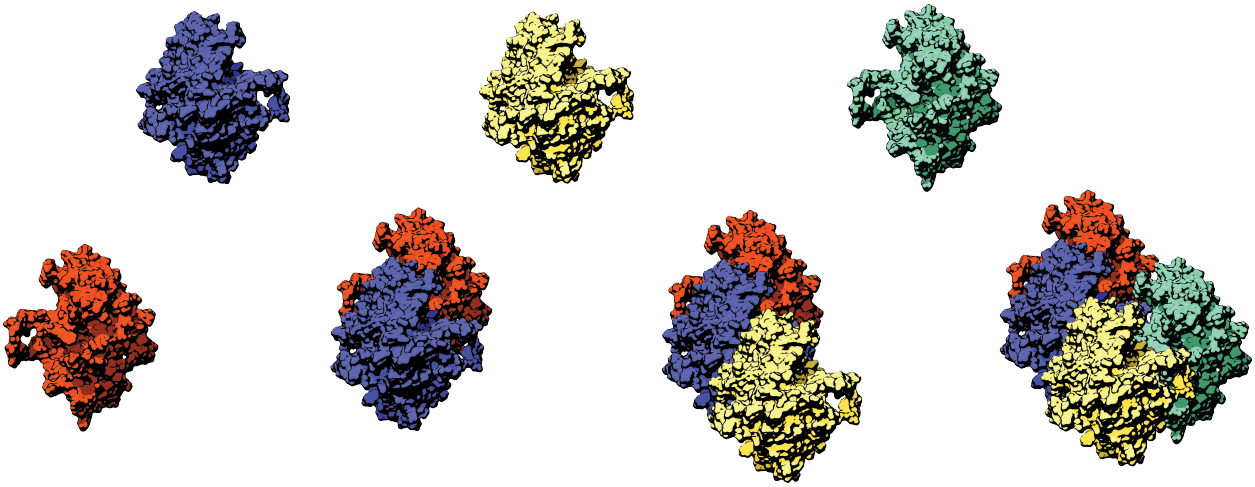


Figure 6.6. Cartoon depiction of the aggregation of Microtubule Associated Protein Tau. The monomeric structure, shown in red, was calculated from NMR analysis.[45]

that results in poor transport along axons and, ultimately, neurodegeneration.[50] In addition to distinguishing between sub-populations of tau oligomers, we also have preliminary data showing that, using lipid-coated nanopores, we can differentiate between native tau and hyperphosphorylated tau based on charge and dipole moment on an individual protein basis. Such a finding could be critical step toward using phosphorylation as a biomarker for determining the aggregation status in certain tauopathies, and may possibly provide earlier diagnoses for this class of diseases. We expect to publish these results in the coming months.

6.2 Additional Extensions

Outside of the ongoing development within our group, the work presented in this dissertation enables investigations of many unique protein targets and systems. I will outline two systems in particular: 1) observation of an individual, shape-changing protein *in situ*, and 2) manipulation of a magnetic-sensitive protein complex within a nanopore.

Adenylate Kinase, also referred to as Myokinase or AK, is an important enzyme that regulates intracellular adenine nucleotide levels by catalyzing the reversible transfer of a phosphate group between adenine nucleotides.[51] It has the primary role of maintaining the energy balance within cells, and its importance in biology is exemplified in its appearance

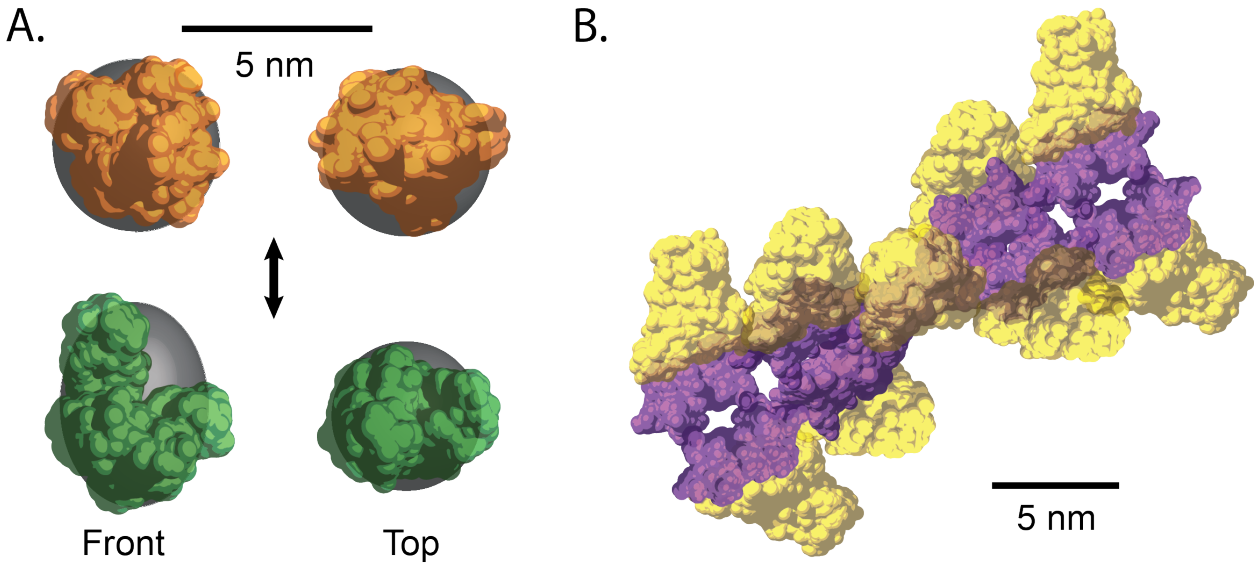


Figure 6.7. Future protein targets that are well-suited to characterization with nanopores. A) Protein crystal structures of the open Adenylate Kinase (green, PDB: 4AKE), and the closed Adenylate Kinase (orange, PDB:1AKE) overlaid over the theoretical ellipsiods of rotation. Note how the open shape resembles a prolate (m = 1.6), while the closed form is closer to a sphere (m = 0.9). B) Magnetic protein complex, from Qin *et al.* [Qin Nat Mater], showing a rod-like structure with 10 repeats of a protein from the class Cryptochromes (yellow, PDB: 4GU5), wrapping around a polymer-chain core of magnetic receptor proteins (purple, PDB: 1R94). The complex is a rod that is about 24 nm long and 18 nm in diameter (when fully bound), giving it an expected length-to-diameter ratio of 1.3. Without Cry (yellow) proteins, the polymer-chain core is a 24 $\times$ 9 nm rod with a length-to-diamer ratio of 2.7, which may be more suitable for a nanopore sensing application as it should produce a larger difference in current blockade between its minimum and maximum orientations within a nanopore.

throughout a large variety of organisms and tissue types.[52] One of the unique attributes of this particular enzyme is the relatively large conformational change it undergoes during its catalytic reaction. The protein structure has been determined for both its closed or substrate-bound state, and its open or free state, and is shown below in Figure 6.7-A. Interestingly, the flexibility of AK's structure is directly correlated with its efficiency to catalyze the adenine reaction, and its transition from closed to open following catalysis limits the rate at which the reaction progresses[53].

AK exists in an equilibrium between open and closed states in the absence of substrate, and the addition of substrate (ATP and AMP) simply shifts that equilibrium in favor of the closed state.[54] Because of its unique properties and its biological significance, AK has been the subject of an abundance of in-depth structural and functional studies, which makes it an excellent candidate for a model protein in a shape change detection experiment. We showed

previously that five-dimensional analysis makes it possible to distinguish between proteins with similar molecular weights - for example, anti-biotin IgG_1 antibody and GPI-AChE (see Chapter 3).[2] The evaluation of multiple conformational variants of a single protein, as is proposed here with AK, is analogous in many ways to this prior experiment. Three different experiments would be necessary for a preliminary investigation : 1) an experiment with the protein, either tethered or free in solution, mostly in its open state and without any substrate, 2) an experiment with AK in a mostly closed state after incubation with a high concentration of substrate, and 3) an experiment showcasing the transition from an open to a closed population of protein due to an injection of substrate *in situ*. The biggest challenge in working with AK is its size: at $\sim$ 25 kDa, it requires a small nanopore, or exceptionally low noise, in order to produce accurate shape estimates. Given the ongoing work discussed above, one or both of these conditions are likely already available.

Another unique system suitable for study with nanopores is the protein complex formed by the combination of a class of Cryptochrome proteins, and a magneto-receptive protein known as MagR.[55] The conglomerate structure has a long rod-shape (Figure 6.7-B), and is thought to be responsible for magnetic-based navigation in a variety of organisms.[56, 57] Several of its qualities are beneficial in the context of nanopore analysis; this complex 1) is not a sphere, and thus will modulate its resistive pulse according to its orientation with a nanopore, 2) has axial symmetry roughly resembling a prolate (*e.g.* rugby ball) and thus would be amenable to our convolution model-based analysis, and most importantly, 3) aligns to an external magnetic field.[55] The magnetic sensitivity of Cry/MagR means that, theoretically, it can be manipulated while inside of the nanopore simply by applying a magnetic field around the pore. This ability to control Cry/MagR could enable 1) in-depth studies of the magnetic sensitivity of different Cry/MagR varieties on a single-molecule level, and/or 2) incredibly precise calibration of a nanopore-based system for protein characterization. One can imagine that, by applying an alternating magnetic field, it may be possible to force Cry/MagR to rotate at a user-defined rate, which would then be measurable within the

resistive pulse and could help to inform and improve calculations of rotational diffusion coefficient. Alternatively, it should be possible to apply a strong enough magnetic field such that all individual Cry/MagR are forced to transit the nanopore in a single – and user-defined – orientation. Forcing a single orientation during transit using the torque from the electric field on a protein's dipole moment requires an electric field (or dipole moment) several orders of magnitude larger than what we have explored in this dissertation, and poses challenges both in material and protein stability. If successful, the combination of the Cry/MagR protein complex with nanopore sensing could produce a more thorough understanding of the complex itself, and could accelerate efforts to improve data analysis techniques for nanopore-based protein characterization.

6.3 Concluding Remarks

In summary, this dissertation has demonstrated the ability to estimate different physical parameters from the resistive pulses generated by *individual* proteins transiting through nanopores. We explored a variety of different fabrication methods for forming single nanopores in freestanding synthetic membranes and used those nanopores to determine a five-dimensional fingerprint of single tethered proteins. We then built upon our work with tethered proteins, showing that it is also possible to estimate parameters from native proteins diffusing freely in solution, and highlighted several important considerations within this label-free scenario, including effects of recording bandwidth and off-axis translocation on resulting analyses. We expect that general improvements in recording noise and bandwidth will usher in the next wave of protein characterization with nanopores, enabling applications in the detection and characterization of amyloid biomarkers for neurodegenerative disease, as well as for broader applications such as the construction of personalized proteomes.[58]

References

[1] A. S. Mikheyev and M. M. Tin. “A first look at the Oxford Nanopore MinION sequencer”. *Molecular ecology resources* 14.6 (2014), pp. 1097–1102.

[2] E. C. Yusko et al. “Real-time shape approximation and fingerprinting of single proteins using a nanopore”. *Nature Nanotechnology* 12 (2017), pp. 360–367.

[3] A. Ivankin, S. Carson, S. R. Kinney, and M. Wanunu. “Fast, label-free force spectroscopy of histoneDNA interactions in individual nucleosomes using nanopores”. *Journal of the American Chemical Society* 135.41 (2013), pp. 15350–15352.

[4] A. Ananth et al. “Reversible Immobilization of Proteins in Sensors and Solid-State Nanopores”. *Small* 14.18 (2018), p. 1703357.

[5] G. V. Soni and C. Dekker. “Detection of nucleosomal substructures using solid-state nanopores”. *Nano letters* 12.6 (2012), pp. 3180–3186.

[6] F. Rivas et al. “Label-free analysis of physiological hyaluronan size distribution with a solid-state nanopore sensor”. *Nature Communications* 9.1 (2018), p. 1037.

[7] J. J. Kasianowicz and S. M. Bezrukov. “Protonation dynamics of the alpha-toxin ion channel from spectral analysis of pH-dependent current fluctuations.” *Biophysical Journal* 69.1 (1995), pp. 94–105.

[8] J. J. Kasianowicz et al. “Analytical applications for pore-forming proteins”. *Biochimica et Biophysica Acta (BBA)-Biomembranes* 1858.3 (2016), pp. 593–606.

[9] S. M. Bezrukov and J. J. Kasianowicz. “Current noise reveals protonation kinetics and number of ionizable sites in an open protein ion channel”. *Physical Review Letters* 70.15 (1993), pp. 2352–2355.

[10] A. Oukhaled et al. “Dynamics of Completely Unfolded and Native Proteins through Solid-State Nanopores as a Function of Electric Driving Force”. *ACS Nano* 5.5 (2011), pp. 3628–3638.

[11] K. J. Freedman et al. “Chemical, thermal, and electric field induced unfolding of single protein molecules studied using nanopores”. *Analytical chemistry* 83.13 (2011), pp. 5137–5144.

[12] N. Martyushenko, N. A. Bell, R. D. Lamboll, and U. F. Keyser. “Nanopore analysis of amyloid fibrils formed by lysozyme aggregation”. *Analyst* 140.14 (2015), pp. 4882–4886.

[13] W. Li et al. “Single protein molecule detection by glass nanopores”. *ACS nano* 7.5 (2013), pp. 4129–4134.

[14] C. Wloka et al. “Label-Free and Real-Time Detection of Protein Ubiquitination with a Biological Nanopore”. *ACS nano* 11.5 (2017), pp. 4387–4394.

[15] A. E. Herr. “Disruptive by design: a perspective on engineering in analytical chemistry”. *Analytical chemistry* 85.16 (2013), pp. 7622–7628.

[16] N. Di Fiori et al. “Optoelectronic control of surface charge and translocation dynamics in solid-state nanopores”. *Nature nanotechnology* 8.12 (2013), p. 946.

[17] R. M. M. Smeets, U. F. Keyser, N. H. Dekker, and C. Dekker. “Noise in solid-state nanopores”. *Proceedings of the National Academy of Sciences* 105.2 (2008), pp. 417–421.

[18] V. Tabard-Cossa, D. Trivedi, M. Wiggin, N. N. Jetha, and A. Marziali. “Noise analysis and reduction in solid-state nanopores”. *Nanotechnology* 18.30 (2007), p. 305505.

[19] A. Balan et al. “Improving Signal-to-Noise Performance for DNA Translocation in Solid-State Nanopores at MHz Bandwidths”. *Nano Letters* 14.12 (2014), pp. 7215–7220.

[20] M.-H. Lee et al. “A Low-Noise Solid-State Nanopore Platform Based on a Highly Insulating Substrate”. *Scientific Reports* 4 (2014), p. 7448.

[21] R. L. Fraccari et al. “High-speed detection of DNA translocation in nanopipettes”. *Nanoscale* 8.14 (2016), pp. 7604–7611.

[22] L. Restrepo-Pérez, C. Joo, and C. Dekker. “Paving the way to single-molecule protein sequencing”. *Nature nanotechnology* 13.9 (2018), p. 786.

[23] E. C. Yusko et al. “Controlling protein translocation through nanopores with bio-inspired fluid walls”. *Nature Nanotechnology* 6.4 (2011), pp. 253–60.

[24] B. Luan et al. “Dynamics of DNA translocation in a solid-state nanopore immersed in aqueous glycerol”. *Nanotechnology* 23.45 (2012), p. 455102.

[25] R. Jaenicke and G. Böhm. “The stability of proteins in extreme environments”. *Current opinion in structural biology* 8.6 (1998), pp. 738–748.

[26] J. L. Osborn et al. “Microfluidics without pumps: reinventing the T-sensor and H-filter in paper networks”. *Lab on a Chip* 10.20 (2010), pp. 2659–2665.

[27] A. J. Storm, J. H. Chen, X. S. Ling, H. W. Zandbergen, and C. Dekker. “Fabrication of solid-state nanopores with single-nanometre precision”. *Nature Materials* 2.8 (2003), pp. 537–540.

[28] H. Kwok, K. Briggs, and V. Tabard-Cossa. “Nanopore fabrication by controlled dielectric breakdown”. *PloS one* 9.3 (2014), e92880.

[29] C. Ying et al. “Formation of Single Nanopores with Diameters of 2050 nm in Silicon Nitride Membranes Using Laser-Assisted Controlled Breakdown”. *ACS Nano* (2018).

[30] L. de Vreede et al. “Wafer scale fabrication of fused silica chips for low-noise recording of resistive pulses through nanopores”. *Nanotechnology* (2019).

[31] L. J. de Vreede, A. van den Berg, and J. C. Eijkel. “Nanopore fabrication by heating Au particles on ceramic substrates”. *Nano letters* 15.1 (2015), pp. 727–731.

[32] A. R. Hall et al. “Hybrid pore formation by directed insertion of α-haemolysin into solid-state nanopores”. *Nature nanotechnology* 5.12 (2010), p. 874.

[33] H. Yamazaki, Y. Qiu, X. Kang, and M. Wanunu. “Photothermally-Assisted Lipid Bilayer Coating on a Sin Nanopore for High-Throughput Protein Channel Formation”. *Biophysical Journal* 116.3 (2019), 294a.

[34] A. Henning-Knechtel, J. Knechtel, and M. Magzoub. “DNA-assisted oligomerization of pore-forming toxin monomers into precisely-controlled protein channels”. *Nucleic acids research* 45.21 (2017), pp. 12057–12068.

[35] E. Spruijt, S. E. Tusk, and H. Bayley. “DNA scaffolds support stable and uniform peptide nanopores”. *Nature nanotechnology* (2018), p. 1.

[36] F. Traversi et al. “Detecting the translocation of DNA through a nanopore using graphene nanoribbons”. *Nature nanotechnology* 8.12 (2013), p. 939.

[37] M. Puster, J. A. Rodrguez-Manzo, A. Balan, and M. Drndic. “Toward sensitive graphene nanoribbon–nanopore devices by preventing electron beam-induced damage”. *ACS nano* 7.12 (2013), pp. 11283–11289.

[38] M. Puster et al. “Cross-Talk Between Ionic and Nanoribbon Current Signals in Graphene Nanoribbon-Nanopore Sensors for Single-Molecule Detection”. *Small* 11.47 (2015), pp. 6309–6316.

[39] W. M. Parkin and M. Drndi. “Signal and noise in fet-nanopore devices”. *ACS sensors* 3.2 (2018), pp. 313–319.

[40] S. Shekar et al. “Measurement of DNA Translocation Dynamics in a Solid-State Nanopore at 100 ns Temporal Resolution”. *Nano Letters* 16.7 (2016), pp. 4483–4489.

[41] C. M. Huber, C. Yee, T. May, A. Dhanala, and C. S. Mitchell. “Cognitive decline in preclinical Alzheimers disease: amyloid-beta versus tauopathy”. *Journal of Alzheimer’s Disease* 61.1 (2018), pp. 265–281.

[42] M. Fá et al. “Extracellular tau oligomers produce an immediate impairment of LTP and memory”. *Scientific reports* 6 (2016), p. 19393.

[43] P. J. Chung et al. “Order from Disorder: The Intrinsically Disordered Protein Tau Facilitates Higher-Order Assembly of Microtubules”. *Biophysical Journal* 106.2 (2014), 354a.

[44] P. E. Wright and H. J. Dyson. “Linking folding and binding”. *Current opinion in structural biology* 19.1 (2009), pp. 31–38.

[45] M. D. Mukrasch et al. “Structural polymorphism of 441-residue tau at single residue resolution”. *PLoS biology* 7.2 (2009), e1000034.

[46] W. Noble et al. “Cdk5 is a key factor in tau aggregation and tangle formation in vivo”. *Neuron* 38.4 (2003), pp. 555–565.

[47] H.-g. Lee et al. “Tau phosphorylation in Alzheimer’s disease: pathogen or protector?” *Trends in molecular medicine* 11.4 (2005), pp. 164–169.

[48] M. P. Mazanetz and P. M. Fischer. “Untangling tau hyperphosphorylation in drug design for neurodegenerative diseases”. *Nature reviews Drug discovery* 6.6 (2007), p. 464.

[49] E. H. Kellogg et al. “Near-atomic model of microtubule-tau interactions”. *Science* 360.6394 (2018), pp. 1242–1246.

[50] E.-M. Mandelkow and E. Mandelkow. “Biochemistry and cell biology of tau protein in neurofibrillary degeneration”. *Cold Spring Harbor perspectives in medicine* 2.7 (2012), a006247.

[51] U. Olsson and M. Wolf-Watz. “Overlap between folding and functional energy landscapes for adenylate kinase conformational change”. *Nature communications* 1 (2010), p. 111.

[52] L. Noda. “Adenylate Kinase”. *The enzymes*. Vol. 8. Elsevier, 1973, pp. 279–305.

[53] M. Wolf-Watz et al. “Linkage between dynamics and catalysis in a thermophilic-mesophilic enzyme pair”. *Nature structural & molecular biology* 11.10 (2004), p. 945.

[54] J. Ådén and M. Wolf-Watz. “NMR identification of transient complexes critical to adenylate kinase catalysis”. *Journal of the American Chemical Society* 129.45 (2007), pp. 14003–14012.

[55] S. Qin et al. “A magnetic protein biocompass”. *Nature materials* 15.2 (2016), p. 217.

[56] R. Wiltschko. *Magnetic orientation in animals*. Vol. 33. Springer Science & Business Media, 2012.

[57] W. Wiltschko and R. Wiltschko. “Magnetic orientation and magnetoreception in birds and other animals”. *Journal of Comparative Physiology A* 191.8 (2005), pp. 675–693.

[58] M. J. Dunn. “Clinical Proteomics Advances”. *PROTEOMICS Clinical Applications* 8.1-2 (2014).

Appendix

A.1 Pseudocode and Explanations of Scripts for Fabrication of Nanopores, and Analysis of Resistive Pulses

During my time in the Mayer group, I wrote a variety of scripts to expedite and simplify certain aspects of the nanopore experimental process. In this section of the appendix, I will discuss the various scripts with regard to their programming design and functionality.

A.1.1 MATLAB Script for Fabrication of Nanopores *via* Dielectric Breakdown

As discussed in Chapter 2, dielectric breakdown is a process in which a high electrical potential applied across a freestanding membrane leads to defect accumulation and ultimately the formation of a nanopore.[1, 2] In order to automate this process, I designed a MATLAB script to communicate with a Kiethley picoammeter over an RS-232 serial cable, relaying precise current measurements on the order of milliseconds, and allowing a user to input a particular voltage protocol with current-based feedback in order to create a nanopore of desired size. Pseudocode for this script is shown below:

Algorithm 1: Dielectric Breakdown

```
Data: voltage protocol, desired nanopore diameter
Result: current-versus-time trace, estimated final diameter
GUI initialization;
picoammeter initialization;
while diameter estimate < desired diameter do
    read current from picoammeter;
    estimate diameter from current;
    if diameter estimate > desired diameter then
        set picoammeter potential to 0;
    else
        set picoammeter potential according to voltage protocol;
    end
    plot measured current value;
    plot applied potential value;
end
export current-versus-time trace, final diameter;
```

A.1.2 MATLAB Script for Population-Based Analysis of Resistive Pulses

As discussed in Chapters 3 and 4, populations of resistive pulses can be used to estimate the length-to-diameter ratio and volume for a pure protein in solution. Each resistive pulse is extracted and analyzed using a peak finding algorithm initially developed by Pedone *et al.*, which defines "true" resistive pulses as those that cross a threshold of 5× the standard deviation of baseline noise. The maximum value from each of these pulses, along with its associated dwell time, is arranged in a CSV file (referred to as a peak-scatter, or pscat, file), which we then analyze using our convolution model (Chapter 4). To simplify the user experience of this process, I wrote a graphical user interface (GUI) that allows a user to load these pscat files, input experimental parameters, modify the data, and fit the data using our analysis methods. Pseudocode for this program is shown below:

A.1.3 MATLAB Script for Individual-Event Analysis of Resistive Pulses

Discussed in detail in Chapters 3 and 4, single resistive pulses can be used to determine a multidimensional fingerprint of a translocating protein. As with population-based analysis,

Algorithm 2: Population-Based Analysis

```
Data: pscat files with resistive pulse maxima and dewll times
Result: length-to-diameter ratio, volume, charge
GUI initialization;
while analysis in progress do
    determine user input from GUI;
    import various parameters and variables from GUI;
    if user selects load data then
        allow user to select directory;
        parse selected directory for "pscat" files;
        import files into workspace;
    end
    if user selects crop data then
        duplicate original data;
        crop data by value OR percentage;
        plot newly cropped data;
    end
    if user selects fit shape and volume then
        generate experimental CDF & PDF from max data;
        initialize fitting values;
        perform convolution fit with lsqcurvefit;
        update GUI with fit output values;
        plot fit and data overlaid;
    end
    if user selects fit charge then
        generate experimental CDF & PDF from dwell time data;
        initialize fitting values;
        perform biased first passage fit with lsqcurvefit;
        update GUI with fit output values;
        plot fit and data overlaid;
    end
    if user selects export then
        generate CSV file for export;
        allow user to select file name and output directory;
        export file;
    end
end
close GUI;
```

each resistive pulse is extracted and analyzed using a peak finding algorithm initially developed by Pedone *et al.*, which defines "true" resistive pulses as those that cross a threshold of 5× the standard deviation of baseline noise. All values from each of these pulses are stored together in a CSV file (referred to as an intraevent file), which we then analyze using our convolution model (Chapter 4). Similar to above, I wrote a graphical user interface (GUI) that allows a user to load these intraevent files, plot and examine the raw current and current histograms of single events, input experimental parameters, modify the data, and fit the data using our analysis methods. Pseudocode for this program is shown below:

A.1.4 C++ & Cuda-C Script for Simulations of Resistive Pulses

As mentioned in Chapter 6, I wrote a random-walk simulation algorithm to simulate resistive pulses for particles with different ellipsoidal shapes on a graphics processor (GPU). FigureA.1 shows a schematic overview of this code:

By switching away from MATLAB and into C++, I was able to reduce computation time for the simulations by approximately two orders of magnitude. It is important to note, however, that single dimensional random-walk simulations are not particularly suited to parallelization, and we expect computational gains to be much greater for future simulations of hydrodynamic bead models (Chapter 6), as well as for the coupled random-walk simulations to determine the influence of off-axis effects in Chapter 5. One technique that I implemented to reduce computational time was treating the biased probability distribution – determined using equation 5.6 – as a three state decision system, as shown in Figure A.2. In this fashion, I created two probability states where a decision (*e.g.* rotate clockwise or counterclockwise) was orientation-independent and could be assigned immediately based on the associated random number, and only one probability state (approximately 1% of steps with our dipole moment and electric field parameters) that was orientation-dependent and required serial processing to compute.

Algorithm 3: Individual-Event Analysis

```
Data: intraevent files with all current values from "long events"
Result: length-to-diameter ratio, volume, dipole moment, and rotational diffusion
        coefficient
GUI initialization;
while analysis in progress do
    determine user input from GUI;
    import various parameters and variables from GUI;
    /* Data load is same as population-based, except also w workspaces */
    if user selects fit shape, volume, dipole moment then
        for all events in data stack do
            evaluate event(s) based on exclusion criteria;
            if exclusion criteria satisified then
                generate experimental CDF & PDF from single event;
                initialize fitting values;
                perform convolution fit with lsqcurvefit;
                update GUI with fit output values for single event;
                plot fit and single event histogram overlaid;
                store fit output for later analyses;
            else
                mark event as not-fit;
            end
        end
    end
    if user selects fit rotational diffusion coefficient then
        for all events in data stack with S,V,D values do
            evaluate event(s) based on additional exclusion criteria;
            if exclusion criteria satisified then
                for lowpass cutoff frequencies from 15 kHz to max BW do
                    determine MSAD from trace filtered at loop frequency;
                end
                determine and store most probable DR across all frequencies;
            else
                mark event as not-fit;
            end
        end
    end
    if user selects update population parameters then
        create ensemble datasets of fit values from each event for export and plotting;
        report median length-to-diameter ratio and median volume across all events;
        report most probable dipole moment and DR via lognormal fitting;
    end
    /* Data export is same as population-based, except more options    */
end
close GUI;
```

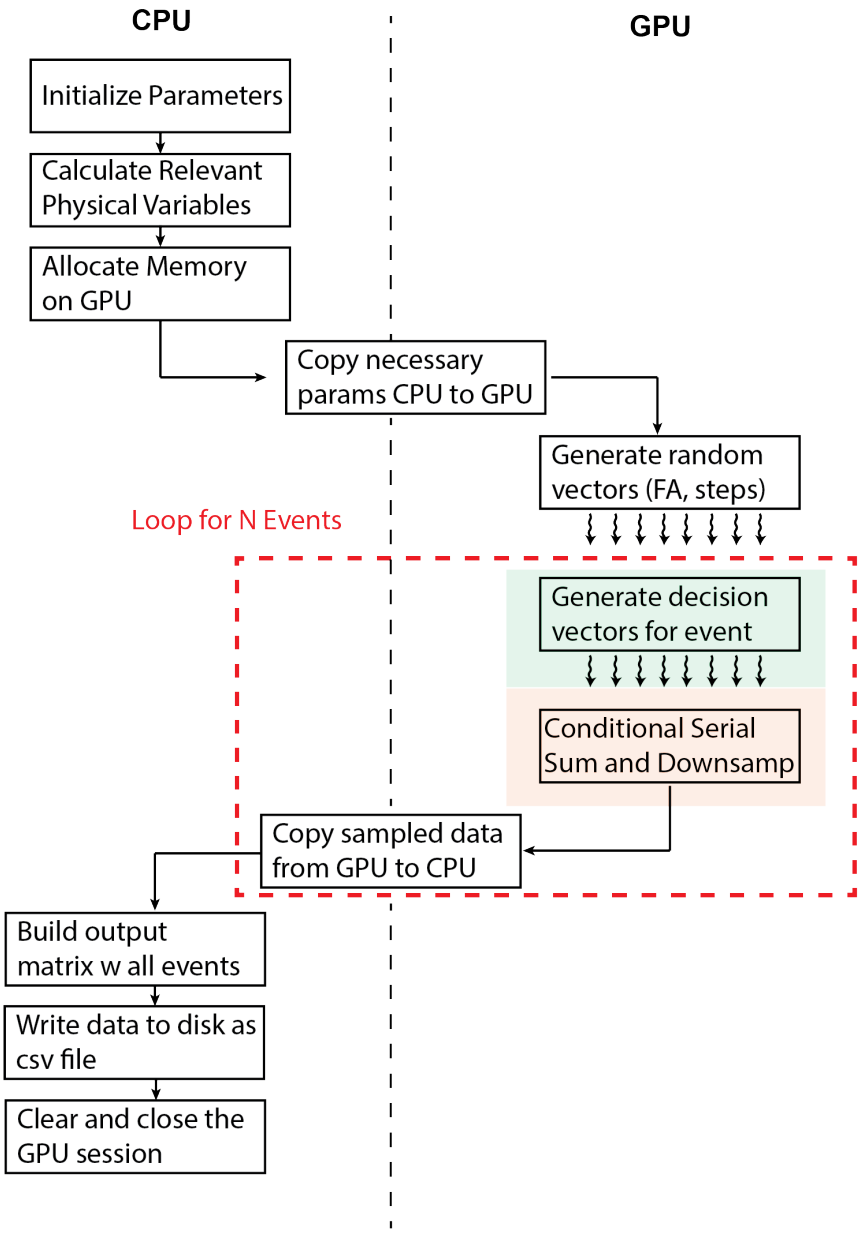


Figure A.1. Diagram showing code structure for GPU-based simulations.

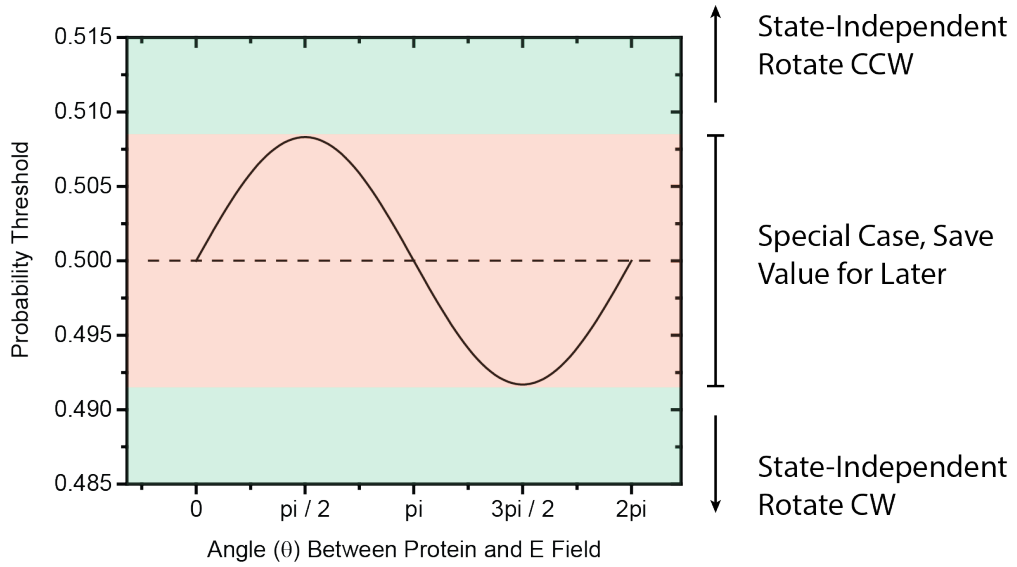


Figure A.2. Biased probability distribution used for random-walk simulations, taken from equation 5.6. Each time step was assigned one of three states in parallel before serial processing.

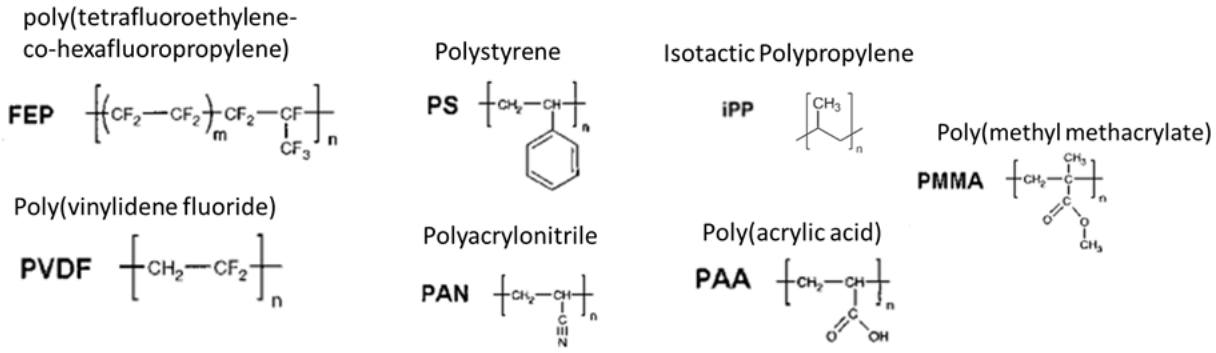


Figure A.3. Chemical structures of seven different polymers analyzed for their stability as nanopore coatings, including structures, names, and shorthand abbreviations.

A.2 Evaluation of the Theoretical Stability of Thin Polymer Films in the Context of Nanopore Coatings

Figure A.3 shows examples of the chemical structures of these adhesion-resistant films. We used similar methods to those presented by Bal *et al.* [3] for stability prediction, calculating Hamaker constants for both the SiN_x/Polymer/Air and SiO_x/Polymer/Air systems, and using the Lifshitz equation (below) to calculate the various constants based on the dielectric and refractive properties for each polymer.[4]

Polymer Shorthand	Contact Angle, H2O on Polymer	Dielectric Const.	Refractive Index	Hamaker Const SiN/Polymer/Air	Hamaker Const SiOx/Polymer/Air	State of Polymer for d_{SiOx} = 2 nm
PS	91	2.5	1.582	-6.29	1.78	Unstable
PAA	Soluble	5	1.506	-6.5	0.641	Metastable
PAN	64	6.5	1.356	-6.11	-0.92	Stable
PMMA	68	3.6	1.482	-6.54	0.277	Metastable
iPP	104	2.5	1.501	-6.53	0.516	Metastable
PVDF	89	8.4	1.407	-6.33	-0.5	Stable
FEP	109	2.1	1.348	-6.1	-1.14	Stable

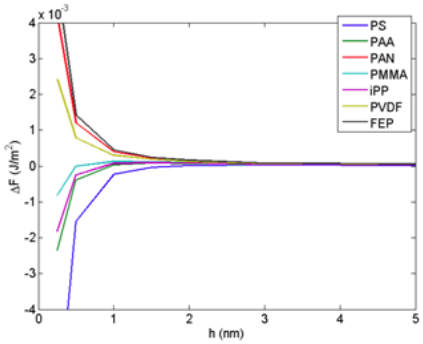


Figure A.4. Chart showing the various properties - contact angle of water, dielectric constant, and refractive index - of each of the polymers, as well as the Hamaker constants both with and without the SiOx layer incorporated. Finally, the state of the polymer was determined by calculating/plotting the free energy change using equation 2. Right) Plot of ΔF over various film thickness values, h, for each of the polymers assuming that d_{SiOx} = 2 nm.

$$A_{Total} \approx \frac{3}{4}kT\left(\frac{\epsilon_1-\epsilon_3}{\epsilon_1+\epsilon_3}\right)\left(\frac{\epsilon_2-\epsilon_3}{\epsilon_2+\epsilon_3}\right)+\frac{3h\nu_e}{8\sqrt{2}}\frac{(n_1^2-n_3^2)(n_2^2-n_3^2)}{(n_1^2+n_3^2)^{1/2}(n_2^2+n_3^2)^{1/2}\left[(n_1^2+n_3^2)^{1/2}+(n_2^2+n_3^2)^{1/2}\right]} \tag{A.3}$$

Then, using an expression (below) for free energy that incorporates these constants as well as the polymer height and SiO_x thickness, we determined the theoretical stability of each polymer:

$$\Delta F^{VDW}(h) = \left(\frac{-A_{SiN/P/Air}}{12\pi(h+d_{SiO_x})^2}\right)+\left(\frac{-A_{SiO_x/P/Air}}{12\pi h^2}\right)-\left(\frac{-A_{SiO_x/P/Air}}{12\pi(h+d_{SiO_x})^2}\right) \tag{A.4}$$

Figure A.4 shows results from these calculations, along with various polymer properties. It is important to note why the Hamaker constants were not simply the only method of stability determination here. Traditionally, the sign of a Hamaker constant has been a hallmark of the stability of a thin film.[4] However, in the case of a polymer like Polystyrene (PS) on a pure SiN substrate, the large negative Hamaker constant would indicate stability. This is not always the case, as these PS films are reported to dewet often through hole nucleation.[5] Several hypotheses may explain this discrepancy: heterogeneities may exist at the surface, there may be a silica layer at the substrate-polymer interface, spacial density variations may

exist due to film confinement, or stresses may be released during spin coating.[5] For this reason, we calculated Hamaker constants both in the presence and absence of the SiO_x layer, and then combined through a free energy relationship to determine a more accurate measure of stability.

A.3 Investigation of Insects as Inspiration for Engineering

I spent a portion of my dissertation researching and co-writing a review about the many different ways that insects solve problems, and how their solutions might be applied to some of our own engineering challenges.[6] The review is loosely related to the work presented throughout this thesis in that the inspiration for lipid bilayer coatings on nanopores comes from the silk moth, *Bombyx mori*, which uses lipid coated pores on its antennae to concentrate and guide pheromone molecules around the dendrites of its olfactory neurons. After lengthy discussion about the capabilities of nanopore sensing and its newfound prospects as a technique to characterize amyloids, it seems only fitting to conclude this thesis with a starting point for the engineers of future biomedical and material solutions. As such, I have included the abstract of the review, as well as a table-of-contents figure below:

> "Over the course of their wildly successful proliferation across the earth, the insects as a taxon have evolved enviable adaptations to their diverse habitats, which include adhesives, locomotor systems, hydrophobic surfaces, and sensors and actuators that transduce mechanical, acoustic, optical, thermal, and chemical signals. Insectinspired designs currently appear in a range of contexts, including antireflective coatings, optical displays, and computing algorithms. However, as over one million distinct and highly specialized species of insects have colonized nearly all habitable regions on the planet, they still provide a largely untapped pool of unique problem-solving strategies. With the intent of providing materials scientists and engineers with a muse for the next generation of bioinspired

materials, here, a selection of some of the most spectacular adaptations that insects have evolved is assembled and organized by function. The insects presented display dazzling optical properties as a result of natural photonic crystals, precise hierarchical patterns that span length scales from nanometers to millimeters, and formidable defense mechanisms that deploy an arsenal of chemical weaponry. Successful mimicry of these adaptations may facilitate technological solutions to as wide a range of problems as they solve in the insects that originated them."

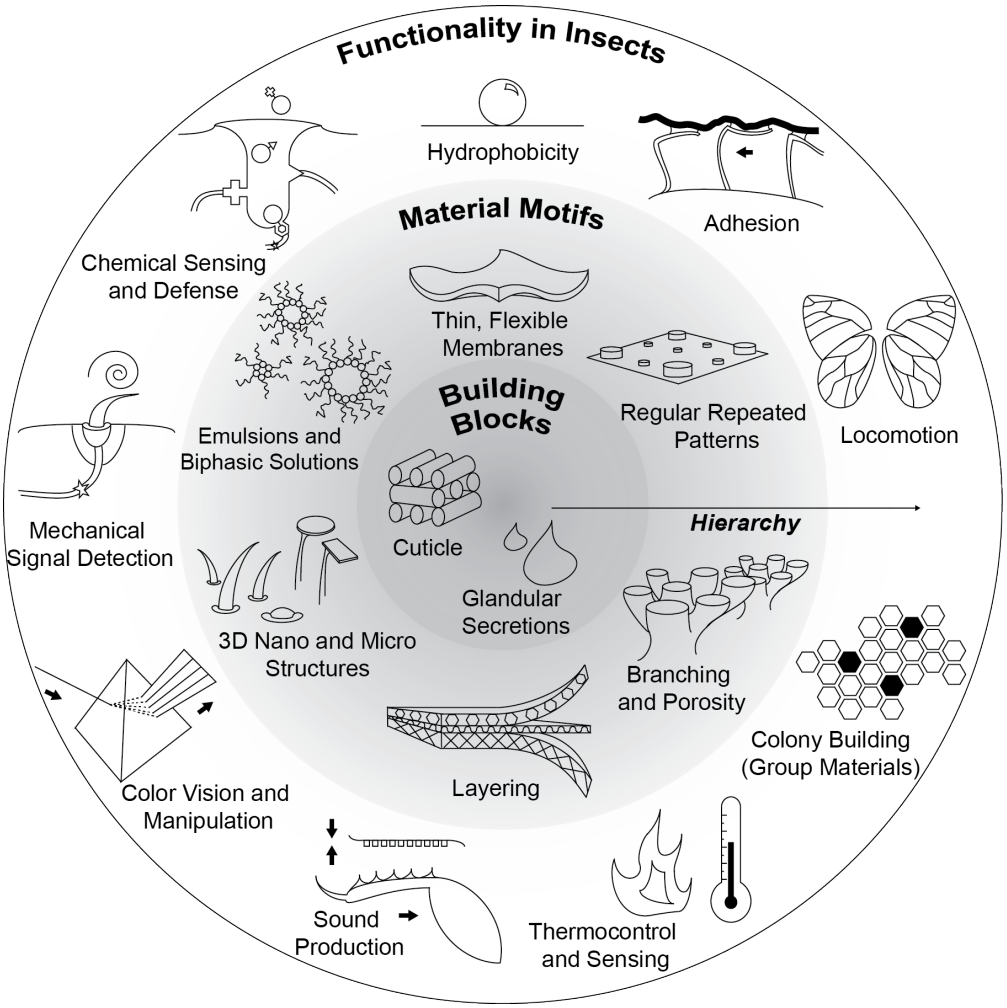


Figure A.5. The ability of insects to thrive in diverse environments is linked to two evolutionarily optimized systems: a cuticlederived exoskeleton with associated functional micro and nanostructures, and glandular complexes that secrete chemically diverse substances. Most structures rely deeply on hierarchical organization, with structural ordering on length scales ranging from nanometers to millimeters. Insects' structural adaptations function to serve a broad set of insect needs including environmental sensing and control, protection, communication, and locomotion. Adapted with permission.[6] Copyright 2018, Wiley-VCH.

References

[1] H. Kwok, K. Briggs, and V. Tabard-Cossa. “Nanopore fabrication by controlled dielectric breakdown”. *PloS one* 9.3 (2014), e92880.

[2] C. Ying et al. “Formation of Single Nanopores with Diameters of 2050 nm in Silicon Nitride Membranes Using Laser-Assisted Controlled Breakdown”. *ACS Nano* (2018).

[3] J. K. Bal et al. “Stability of polymer ultrathin films (< 7 nm) made by a top-down approach”. *ACS nano* 9.8 (2015), pp. 8184–8193.

[4] K. Feldman, T. Tervoort, P. Smith, and N. D. Spencer. “Toward a force spectroscopy of polymer surfaces”. *Langmuir* 14.2 (1998), pp. 372–378.

[5] A. Sharma, J. Mittal, and R. Verma. “Instability and dewetting of thin films induced by density variations”. *Langmuir* 18.26 (2002), pp. 10213–10220.

[6] T. B. H. Schroeder, J. Houghtaling, B. D. Wilts, and M. Mayer. “It’s Not a Bug, It’s a Feature: Functional Materials in Insects”. *Advanced Materials* (2018).