Nanopore-Based Methods for Characterizing Single Proteins

by

Brandon Robert Bruhn

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Biomedical Engineering)
in the University of Michigan
2015

Doctoral Committee:

      Associate Professor Michael Mayer, Chair
      Professor David Sept
      Associate Professor Georgios Skiniotis
      Professor Nils G. Walter

© Brandon Robert Bruhn

DEDICATION

This dissertation is dedicated *in memoriam* of the always inspiring Alan J. Hunt

and

to my family for their encouragement and support over the years.

ACKNOWLEDGMENTS

Of all the people who helped me throughout my graduate career, I would like to extend special thanks to:

My research advisor, Michael Mayer, who is first and foremost responsible for my development as a scientist. Our discussions were always entertaining, motivational, and worthwhile due to his expertise and passion for science. I will forever be grateful for his guidance and support.

Prof. David Sept for assisting me with a lot of the statistics and computational modeling presented in this dissertation. Prof. Sept was always willing to answer my questions and lend a hand.

The remainder of my dissertation committee, Prof. Georgios Skiniotis and Prof. Nils Walter, for taking the time to evaluate my work.

TABLE OF CONTENTS

xi

LIST OF TABLES

TABLE

LIST OF SYMBOLS

Symbol	Units	Description
ρ	Ω m	Resistivity of the electrolyte solution
R_A	Ω	Access resistance of a dual-pore chip
R_L	Ω	Resistance of the L-shaped channel in a dual-pore chip
R_P	Ω	Pore resistance
L_L	m	Length of the L-shaped channel in a dual-pore chip
w_L	m	Width of the L-shaped channel in a dual-pore chip
h_L	m	Height of the L-shaped channel in a dual-pore chip
h_C	m	Height of the conical segment of the recording pore in a dual-pore chip
D_1	m	Diameter of the recording pore at the top surface of the dual-pore chip
D_2	m	Diameter of the recording pore at the bottom of the conical segment in a dual-pore chip
t	m	Coverslip thickness
L_{SS}	m	Length of the single-shot aperture in a dual-pore chip
D_{SS}	m	Diameter of the single-shot aperture in a dual-pore chip
ΔI	A	Change in current in a resistive-pulse
t_d	S	Duration of a resistive pulse
z	unitless	Net charge valence of a protein
e	1.60E-19 C	Elementary charge of an electron
v	m s^{-1}	Electrophoretic drift velocity
D	m^2 s^{-1}	Translational diffusion coefficient

Symbol	Units	Description
D_L	$\mathrm{m^2\,s^{-1}}$	Lateral diffusion coefficient of a lipid or lipid-anchored protein
D_R	$\mathrm{rad^2\,s^{-1}}$	Rotational diffusion coefficient of a protein
η	Pa s	Solution viscosity
r_H	M	Hydrodynamic radius
S	unitless	Perrin shape factor
Λ	$\mathrm{m^3}$	Volume excluded by a translocating particle
m	unitless	Length-to-diameter ratio of an ellipsoid of rotation
μ	D	Dipole moment of a protein
U	J	Potential energy
f_c	Hz	Low-pass cutoff frequency
V_A	V	Applied potential
V_P	V	Voltage drop across the nanopore
d_P	m	Diameter of the nanopore
l_P	m	Length of the nanopore
d_m	m	Diameter of a translocating molecule
θ	rad	Angle of the axis of revolution of an ellipsoid of rotation relative to the electric field
γ	unitless	Electrical shape factor
$\gamma_\parallel$	unitless	Electrical shape factor when $\theta = 0$
$\gamma_\perp$	unitless	Electrical shape factor when $\theta = \pi/2$
$\Delta I_{\min}$	A	Minimum observed ΔI due to the translocation of an ellipsoid of revolution (occurs when $\theta = 0$ for a prolate and $\pi/2$ for an oblate)
$\Delta I_{\max}$	A	Maximum observed ΔI due to the translocation of an ellipsoid of revolution (occurs when $\theta = \pi/2$ for a prolate and 0 for an oblate)
k_B	1.38E-23 J/K	Boltzmann constant

T K Temperature

ABSTRACT

Proteins represent the most diverse class of biomolecules in both structure and function and are involved in nearly every physiological process; their quantification, identification, and biophysical characterization is therefore of fundamental and practical importance. This dissertation introduces two distinct techniques that use nanopores to characterize and identify single unlabeled proteins in a high-throughput manner. Whereas the most common techniques for characterizing proteins (*e.g.*, two-dimensional gel electrophoresis, mass spectroscopy, immunoassays) provide measurements from an ensemble of 10^{15} to 10^{18} molecules, the methods presented here analyze proteins one-by-one and are thus better-suited for examining heterogeneous populations, rare species, and protein dynamics.

The first technique uses femtosecond-laser-fabricated dual-pore glass chips for performing cell-attached single-ion-channel recordings. Existing planar patch-clamp platforms are generally unable to perform these types of recordings due to excess noise arising from low seal resistances and the use of substrates with poor dielectric properties (*e.g.*, silicon). While these platforms tend to use a single pore (diameter ~ 1 to 2 µm) to position a cell by suction and to establish a seal, the dual-pore glass chips employ separate pores optimized for each function, enabling the use of a relatively small patch aperture (diameter ~ 150 to 300 nm) that is more suitable for forming high-resistance seals than micropores used currently. Patch-clamp experiments with these chips consistently achieved high seal resistances (rate of gigaseal formation = 61%, mean seal

resistance = 53 GΩ), maintained gigaseals for prolonged durations (up to 6 hours), and achieved the lowest RMS noise ever reported for a planar patch-clamp platform (0.46 pA at 5 kHz). This platform enables semi-automated single-channel recordings in the cell-attached configuration that are comparable to those obtained by conventional patch-clamp, which is laborious and requires manual control of micropipette position.

The second technique uses electrolyte-filled nanopores coated with a lipid bilayer to characterize single lipid-anchored proteins *via* resistive-pulse sensing. Lipid-coated nanopores have previously been used to determine a protein's volume, charge, and ligand affinity by measuring the change in ionic current, ΔI, through the nanopore as a protein travels from one side to the other. Exploiting the dependence of ΔI on the shape and orientation of a particle in the nanopore, this work extends the capabilities of resistive-pulse sensors by enabling determination of the shape, volume, rotational diffusion coefficient, and dipole moment of individual non-spherical proteins. This research further demonstrates the utility of these additional parameters for distinguishing proteins in a mixture.

The work presented in this dissertation expands the capabilities of planar patch-clamp platforms and resistive-pulse sensors for characterizing and identifying ion channels and soluble proteins, respectively. The techniques introduced in this work may ultimately reveal insights into conformational protein dynamics, expedite biomarker and drug discovery, enable the characterization of personal proteomes, and improve our understanding of proteins and protein complexes in the context of health and disease.

CHAPTER 1

Nanopore-based, single-molecule characterization of proteins

Proteins represent the most diverse class of biomolecules in both structure and function and are involved in nearly every physiological process. Consequently, disease states generally manifest through changes in an individual's proteome, the protein counterpart of the genome that includes all actively expressed proteins[1]. Recent acknowledgment of this fact has caused rapid growth in the field of clinical proteomics (*e.g.*, roughly 1,600 relevant papers were published in 2012 compared to 1,000 in 2007), which focuses on the identification and validation of protein biomarkers[2,3]. Nevertheless, extremely few protein biomarkers have been approved by the U.S. Food and Drug Administration for clinical use (*e.g.*, seven proteins were approved between 2003 and 2008) even though more than 20,000 proteins have been recognized as potential biomarkers[3]. This deficiency is partially due to the limitations of existing proteomic technologies, particularly with regard to sensitivity, resolution, cost, and throughput[1,3]. Therefore, a significant need exists to devise improved methods to characterize, identify, and quantify proteins in a precise, cheap, and rapid manner; such methods would certainly aid in the understanding, diagnosis, and treatment of various medical conditions.

Single-molecule techniques, which first arose in the 1970s but have garnered more attention in recent years, are particularly promising in this regard[4]. The majority

of techniques currently used to analyze proteins (*e.g.*, two-dimensional gel electrophoresis[5], mass spectroscopy[6], immunoassays[7], etc.) provide measurements from an ensemble of 10^{15} to 10^{18} molecules and are thus ill-suited for characterizing heterogeneous populations, molecules present at low concentrations, and protein dynamics[1,4]. Moreover, such techniques tend to provide information regarding just one or two properties (*e.g.*, mass and charge), limiting their usefulness for distinguishing and identifying proteins. On the other hand, single-molecule techniques analyze individual molecules one-by-one, avoiding the issues posed by ensemble measurements and yielding data that is generally more comprehensive and intuitive[8]. Furthermore, the unrivaled sensitivity afforded by these techniques makes it possible to detect rare molecular states that deviate significantly from the average of the population[9]. Single-molecule techniques have already been used to answer a number of fundamental questions in different areas of protein science, such as protein folding and enzyme catalysis, and present a number of interesting possibilities that have yet to be explored[9].

This dissertation presents two different high-throughput single-molecule techniques for analyzing proteins using nanopores (*i.e.*, pores with diameters less than 1 μm). The first technique uses laser-fabricated dual-pore glass chips for performing cell-attached single-ion-channel recordings and addresses several limitations that encumber planar patch-clamp platforms. The second technique is based on the concept of resistive-pulse sensing and expands the capabilities of lipid-coated nanopores for multi-parameter characterization of single proteins in solution. The remainder of this introduction describes the evolution of both the patch-clamp technique and resistive-pulse sensing and

introduces the idea of lipid-coated nanopores, setting the stage for the results presented in Chapters 2 and 3.

1.1 The evolution of the patch-clamp technique

The patch-clamp technique is the culmination of a number of prior technical achievements in the electrophysiology field dating back to 1791 when Luigi Galvani used metal wires to stimulate frog nerve-muscle preparations and elicit contractions[10]. Nearly 150 years later, Alan Lloyd Hodgkin and Andrew Fielding Huxley (1939) obtained the first intracellular recording of an action potential by inserting a rudimentary glass electrode into the giant axon of a squid (*Loligo forbesi*); the electrode consisted of a saline-filled glass micropipette that was roughly 100 μm in diameter and contained a Ag/AgCl wire to interface with the recording amplifier[11]. In 1949, Ling and Gerard (building upon work by Graham and Gerard[12]) developed glass microelectrodes with reduced tip diameters (< 1 μm), enabling intracellular recordings in smaller cells[13]; however, the use of these sharp electrodes involves puncturing the cell membrane, which results in large leakage currents and therefore high background noise. At the same time, Cole[14] and Marmont[15] developed the voltage-clamp technique, wherein the recording electronics hold the membrane potential constant while measuring current. This technology enabled Hodgkin and Huxley (1952) to deduce the ionic basis of the action potential even though the existence of ion channel proteins had yet to be proven[16].

Erwin Neher and Bert Sakmann developed the patch-clamp technique in 1976[17], for which they received the Nobel Prize in Physiology or Medicine in 1991[18]. This groundbreaking technique, which is a refinement of an earlier approach originally

conceived by Pratt and Eisenberger in 1919[19], involves bringing the tip of a fire-polished, glass micropipette into physical contact with a cell without puncturing its membrane[17]. Due to the tight seal that spontaneously forms between the pipette and cell membrane, this procedure electrically isolates the membrane patch encompassed by the pipette rim and results in lower leakage currents and noise than when using sharp electrodes[18]. Neher and Sakmann made use of the unprecedented signal-to-noise ratio realized by their approach to record, for the first time ever, the current activity of single ion channels, definitively proving their existence[17]. Several years later, Sigworth and Neher (1980) showed that the electrical resistance of the seal between the pipette and membrane could be increased from 10^8 to 10^{10} Ω by applying gentle suction to the interior of a pristinely clean and extremely smooth pipette tip, further reducing noise and thereby enabling current recordings with unprecedented resolution[20].

In 1981, Hamill *et al.* published the quintessential paper on the patch-clamp technique (with more than 17,000 citations) that describes several commonly used recording configurations (Fig. 1.1)[21]. The cell-attached configuration (Fig. 1.1), which was employed by Neher and Sakmann in the work described above, allows the activity of single ion channels within the membrane patch to be recorded non-invasively[18]. While this configuration maintains the integrity of the cell under examination and thereby prevents the loss of its intracellular contents, it does not permit the resting membrane potential to be determined precisely as the exact electrochemical gradient of each ion is unknown[18]. In contrast, the excised patch configurations (Fig. 1.1) allow the experimenter to explicitly define the composition of both the intracellular and extracellular solutions[18]. Moreover, the inside-out and outside-out configurations (Fig.
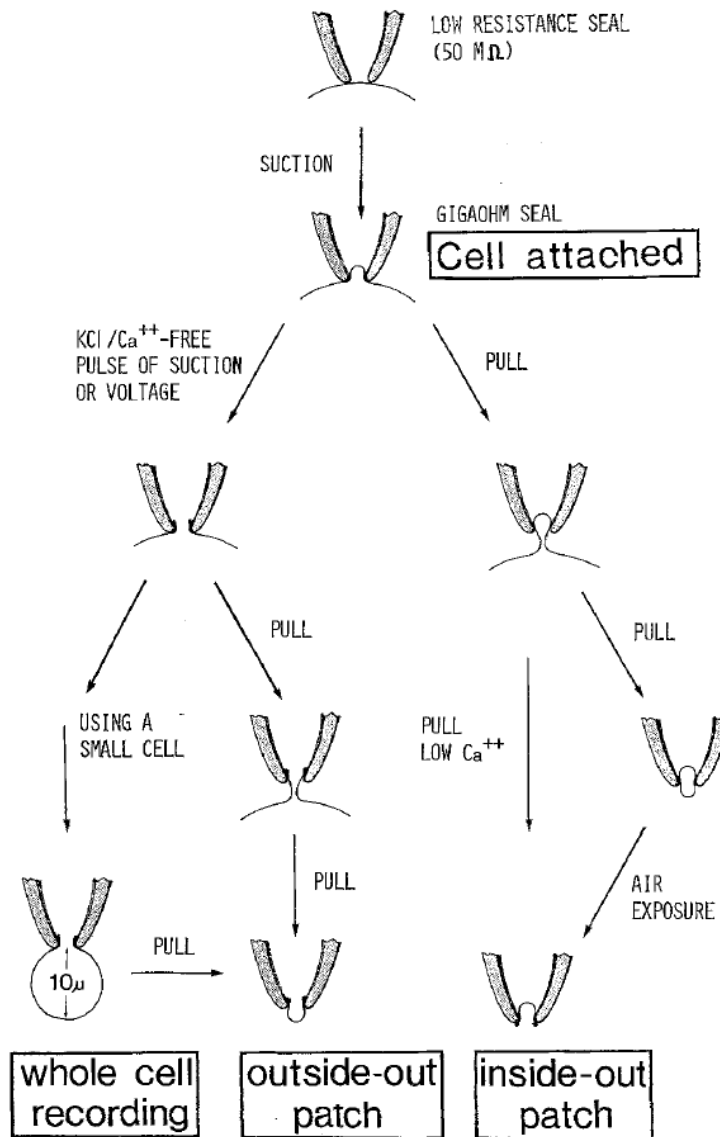
Figure 1.1 | Illustration of the four primary recording configurations of the patch-clamp technique. Initially, a micromanipulator is used to bring a fire-polished, glass micropipette into physical contact with the membrane of an adherent cell, establishing a low-resistance seal (on the order of MΩ) between the pipette and membrane. To reach the cell-attached configuration, which serves as a starting point for the three other recording configurations, gentle suction is applied to increase the seal resistance into the GΩ range; this high-resistance seal is commonly referred to as a gigaseal. The whole-cell configuration is obtained by applying a short pulse of suction or voltage that ruptures the membrane patch and provides electrical access to the cell interior. The inside-out and outside-out configurations, collectively referred to as excised or cell-free patch configurations, are reached by retracting the pipette in the cell-attached and whole-cell configurations, respectively. Adapted from (21).

1.1) enable quick exchange of the intracellular and extracellular solutions, respectively, which is useful for studying second-messenger-activated and ligand-gated channels[18].

5

Unlike the cell-attached and excised patch configurations, the whole-cell configuration (Fig. 1.1) allows the ensemble average of all ion channel activity in a cell to be recorded[18]. Consequently, whole-cell currents are typically much larger than single-channel currents (several hundred pA as opposed to a few pA)[18]. Following the work by Hamill *et al.*, a number of groups developed variants of the above configurations to address certain limitations (*e.g.*, Lindau and Fernandez (1984) developed the perforated-patch configuration to reduce the dilution of intracellular constituents that occurs in the whole-cell configuration[22]); nonetheless, the four original configurations are still frequently used. To this day, the patch-clamp technique remains the method of choice for characterizing the behavior of ion channel proteins (*e.g.*, the term "patch-clamp" was used in roughly 7,500 articles in 2013 according to Web of Science).

While patch-clamping yields electrophysiological data that is unrivaled in quality, the conventional version of the technique is low-throughput (tens of data points per day) and requires highly trained personnel; thus, conventional patch-clamping is impractical for certain applications, particularly drug discovery[23,24]. To address these limitations, companies and academics have developed a variety of automated patch-clamp platforms over the last two decades, starting with the invention of the NeuroPatch at NeuroSearch in the late 1990s[25]. The NeuroPatch, which later evolved into the Apatchi-1, and several other platforms fully automate the conventional, pipette-based technique; however, these systems only provide modest improvements in throughput (hundreds of data points per day)[23]. In 2002, Fertig *et al.*[26] and Klemic *et al.*[27] established the first planar patch-clamp platforms, wherein a micropore in a planar substrate replaces the patch pipette, a concept first touched on by Kostyuk *et al.* in the 1970s[28]. The concept of planar patch-clamp has

since been widely adopted by the electrophysiology field due to its simplicity and amenability to parallelization[23]. While planar platforms enable extremely high throughput (up to 18,000 data points per day), these systems are generally limited to whole-cell recordings, leaving the need for an automated platform that can perform single-channel recordings in a high-throughput manner[23]. Chapter 2 discusses this topic in further detail.

1.2 Resistive-pulse sensing: from detecting cells to single molecules

In 1949, Wallace H. Coulter invented the concept of resistive-pulse sensing (*i.e.*, Coulter counting) for counting particles suspended in a fluid, facilitating the development of the first automated cell counters[29,30]. This elegantly simple concept is illustrated in Fig. 1.2. A voltage source applies a potential across an electrolyte-filled pore that separates two fluidic chambers, resulting in a constant baseline current (Fig. 1.2a). The total resistance of the circuit is dominated by the contribution of the pore due to its relatively small size. When a particle passes through the pore, it excludes electrolyte and causes a transient increase in resistance (*i.e.*, resistive-pulse), thereby decreasing the measured
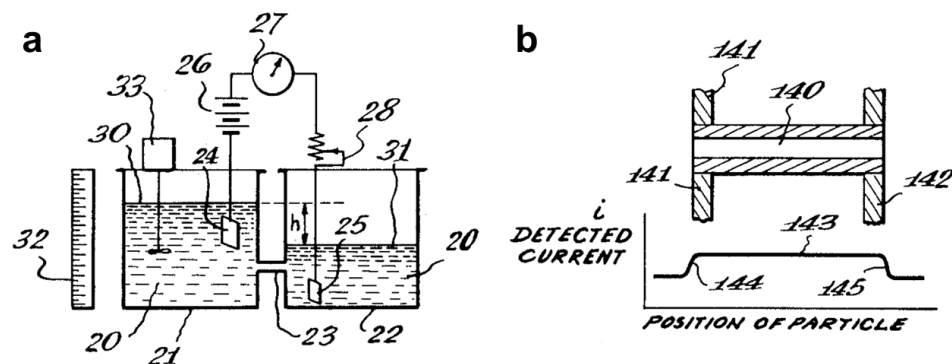
Figure 1.2 | Resistive-pulse sensing as envisioned by Wallace H. Coulter. a) A voltage source applies a constant potential across an electrolyte-filled pore as an ammeter measures current. A difference in fluid height between the two chambers generates a pressure gradient and drives particles (*e.g.*, cells) from one chamber to the other through the pore. b) A particle passing through the pore reduces the measured current. The magnitude of the current reduction depends on the volume of the particle, resistivity of the electrolyte solution, and pore dimensions. Adapted from (29).

current (Fig. 1.2b). Thus, one can count particles by merely tallying resistive-pulses. Moreover, the magnitude of a resistive-pulse is proportional to the excluded volume of electrolyte, providing information regarding particle size[31]. Coulter published his sole technical paper on resistive-pulse sensing in 1956, unveiling the concept to the scientific community[31]. Many years later, DeBlois *et al.* (1970, 1976) demonstrated that the duration of a resistive-pulse is inversely proportional to the electrophoretic mobility of the particle and the frequency of resistive-pulses is directly proportional to particle concentration[32,33].

Expanding upon Coulter's work, Herbert E. Kubitschek (1958) formulated a theoretical expression relating the magnitude of a resistive-pulse to the volume of the translocating particle (in addition to the resistivity of the electrolyte solution and pore geometry) and used the expression to size individual bacteria[34]. This expression, however, did not account for the distortion of the electric field due to the presence of a particle in the pore, resulting in large error ($\geq 50\%$) for spherical and ellipsoidal particles[35]. Long before, Maxwell and Lord Rayleigh considered the deformation of an electric field due to the presence of spherical particles in bulk solution in order to calculate the resistivity of such a solution[36,37]. Fricke (1924, 1953) as well as Velick and Gorin (1940) expanded this theory to estimate the resistivity of solutions containing ellipsoidal particles[38–41]. Following Kubitschek's study, various groups began accounting for the effect of a translocating particle's shape and orientation in resistive-pulse sensing experiments to estimate particle volume more accurately[32,35,42–48].

In 1973, Golibersuch clearly demonstrated this effect by recording resistive-pulses resulting from the passage of oblate-shaped erythrocytes through relatively long

pores[47]. Due to the shear field inside the pore, the erythrocytes rotated in a fixed orbit and caused periodic fluctuations in resistance (Fig. 1.3)[47]. By accounting for this orbital motion and the distortion of the electric field in the pore, Golibersuch developed a sophisticated model that accurately described the empirical distribution of resistance values resulting from the translocation of red blood cells[47]. Berge *et al.* (1989, 1990) and Carbonaro *et al.* have reported similar resistive-pulse signatures to those observed by Golibersuch resulting from the translocation of fused polystyrene spheres and murine erythroleukemia cells, respectively[49–51].
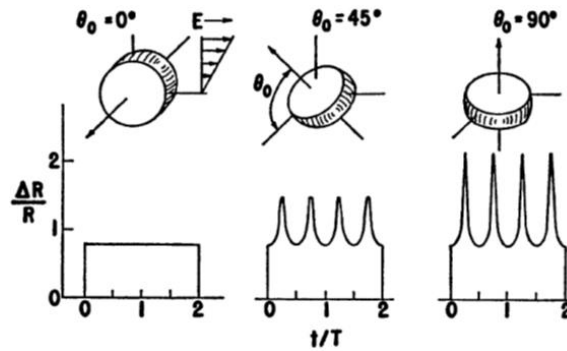
Figure 1.3 | Time-varying resistive-pulses resulting from the rotation of oblate-shaped erythrocytes as predicted by Golibersuch. The resistive-pulses undergo periodic fluctuations due to the effect of the particle's shape and orientation on the deformation of the electric field. The minimum and maximum values of each pulse depend on the axis about which the particle rotates. Adapted from (47).

The minimum detectable particle size in resistive-pulse sensing is primarily governed by the dimensions of the pore; thus, advancements in pore technology have been essential to the evolution of the technique. Early on, the commercially-available Coulter counter had a detection limit of approximately 500 nm and therefore the majority of studies focused on characterizing cells[32]. In 1970, DeBlois and Bean developed a method to fabricate pores with diameters around 450 nm in 3-μm-thick track-etched polycarbonate sheets[32], enabling the detection of particles as small as 60 nm (*e.g.*,

nanoparticles[46,52], virus particles[33,53,54], clusters of antigen crosslinked by antibody[55,56]). The next major breakthrough occurred over two decades later in 1994 when Bezrukov *et al.* established the concept of using a biological pore (*i.e.*, alamethicin) in a planar lipid bilayer to detect single molecules (*i.e.*, PEGs with molecular masses between 200 and 1000 Da) *via* resistive-pulse sensing (Fig 1.4a)[57]. In 1996, Kasianowicz *et al.* published the hallmark paper on this concept in the *Proceedings of the National Academy of Sciences*. Kasianowicz *et al.* used the bacterial protein α-hemolysin to detect single-stranded polynucleotides, raising the possibility of using resistive-pulse sensing to sequence DNA or RNA in a rapid and inexpensive manner (Fig. 1.4b)[58].
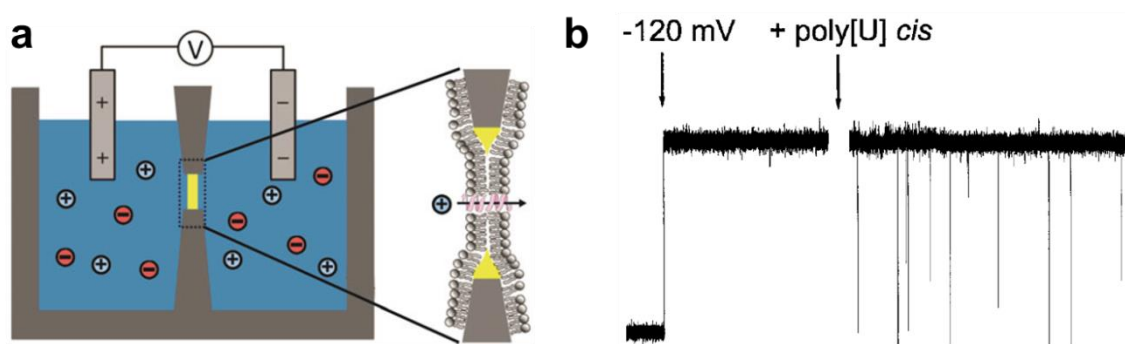
Figure 1.4 | Resistive-pulse sensing of single molecules using a biological pore in a planar lipid bilayer. a) Illustration of a resistive-pulse sensor wherein single molecules pass through a biological pore that spans across a free-standing lipid bilayer. b) Current recording obtained by Kasianowicz *et al*. Upon the introduction of poly[U] strands on one side of the pore, resistive-pulses occur due to the translocation of these molecules through the pore. Adapted from (59) and (58).

Following the initial work by Kasianowicz *et al.*, a number of groups have continued using different biological pores (*i.e.*, porins and ion channels) as resistive-pulse sensors in order to detect and characterize polynucleotides[60] as well as proteins and small molecules[61–69]. Additionally, researchers have used these sensors to study enzymatic activity[70], binding affinities and kinetics[71–74], protein folding[75–78], chemical reactions[79–81], and polymerization[82]. Notwithstanding the success of this work, the use of biological

pores has two key shortcomings. First, the free-standing lipid bilayer that contains the pore is fragile and sensitive to mechanical perturbations, which generally limits its lifetime to less than an hour. Second, biological pores are too small to permit the vast majority of proteins to pass through in their native folded conformation. Researchers addressed both of these issues with the advent of synthetic (*i.e.*, solid-state) nanopores.

In 2001, Li *et al*. introduced a feedback-controlled method known as ion-beam sculpting for reliably fabricating nanopores in silicon nitride with diameters and lengths ranging from 1 to 50 nm and 10 to 20 nm, respectively[83]. Li *et al.* used one such pore with a diameter of 5 nm to record resistive-pulses due to the translocation of double-stranded DNA, a feat that was impossible with the biological pores in use at the time[83]. Five years later, Han *et al*. (2006) finally used a synthetic nanopore to detect and characterize individual folded proteins[84]. Using the theory developed by DeBlois and Bean many years earlier, Han *et al*. measured the diameter of bovine serum albumin with reasonable accuracy[84]. This report prompted a wave of research that examined single folded proteins using synthetic nanopores as resistive-pulse sensors.

At the onset of this research, several groups reported observing two types of events due to the translocation of the same protein: 1) events with small magnitudes and durations on the order of a few μs and 2) events with large magnitudes and durations ranging from tens of μs to a few ms[85–88]. The second type of event occurred less frequently and varied widely in duration compared to the first type of event [85–88]. Researchers later showed that the second type of event results from non-specific adsorption of the translocating protein to the pore wall, confirming earlier hypotheses[89,90]. Such interactions are highly unfavorable since they often cause the pore to clog[91,92] and

11

result in translocation times that are difficult to predict and not strictly governed by electrophoresis[93]. In the absence of interactions with the pore wall, Talaga and Li showed that translocation times are inversely proportional to the diffusion coefficient and charge of the analyte[87,94]. The biased first-passage-time model used by Talaga and Li, which was later appended by Ling and Ling[95], predicts that a folded protein should normally translocate through a nanopore in a few μs, indicating that the first type of event described above occurs in the absence of non-specific interactions. These events, however, are generally too short to resolve in time such that their magnitudes are attenuated, as indicated above[88,91,96]. Plesa *et al.* (2013) clearly illustrated this point and further demonstrated that the majority of protein translocation events are not detected under normal conditions due to their short durations[97]. Di Fiori *et al.* (2013) developed one approach for increasing event duration by focusing a laser on a synthetic nanopore to induce electroosmotic flow, thereby retarding the translocation of proteins and DNA[98].

Regardless of the limitations described above, researchers have still successfully used solid-state nanopores as resistive-pulse sensors to study protein size[85–88,98–103] and conformation[75,87,103–106], protein-ligand interactions[107–109], protein-DNA interactions[110–117], and antibody-antigen interactions[86,99,101,102,108,118].

1.3 Lipid-coated nanopores for resistive-pulse sensing of proteins

Taking inspiration from the olfactory sensilla of insects, Yusko *et al.* introduced the concept of lipid-coated nanopores in 2011 to address many of the shortcomings associated with the use of synthetic nanopores for resistive-pulse sensing of proteins (Fig. 1.5)[91]. To form such a coating, Yusko *et al.* simply added a solution containing small
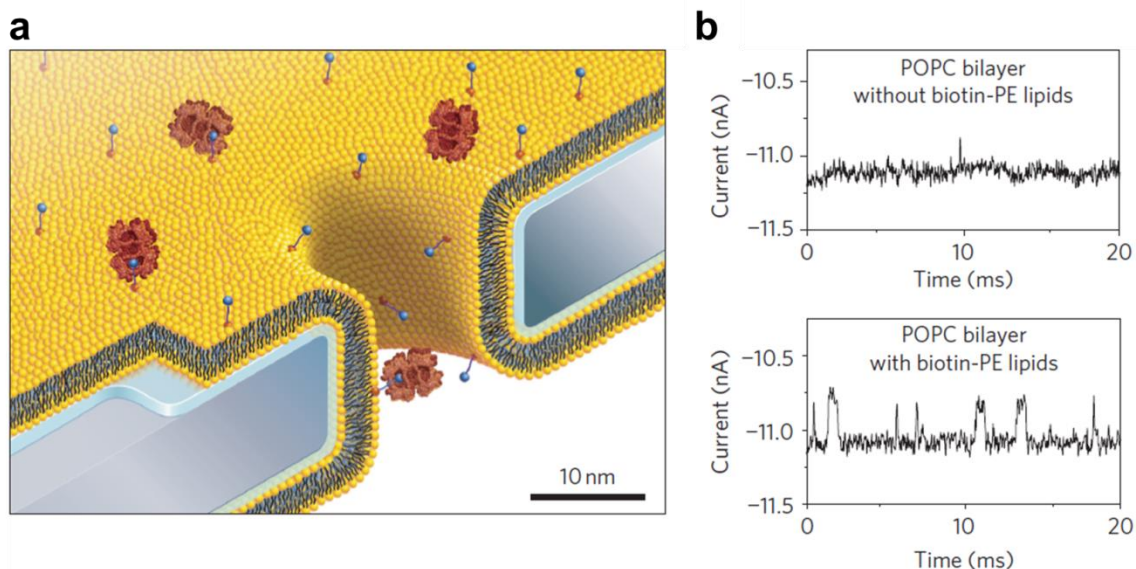
Figure 1.5 | Lipid-coated nanopores for resistive-pulse sensing of proteins as envisioned by Yusko *et al.* a) Illustration of a lipid-coated (yellow) nanopore in a silicon nitride membrane (grey). A water layer (blue) exists between the bilayer and substrate. Lipid-anchored ligands (dark blue) bind a protein of interest (red). **b)** Current traces showing resistive-pulses due to the translocation of streptavidin in the absence and presence of lipid-anchored biotin. Streptavidin bound to a lipid anchor passes through the pore much more slowly than free streptavidin. Furthermore, pre-concentration of the protein on the surface of the bilayer in the presence of lipid-anchored ligand results in a markedly higher event frequency. Adapted from (91).

unilamellar vesicles (SUVs) to one side of the chip, resulting in the spontaneous formation of a lipid bilayer that spread through the pore to the other side of the chip[91]. Whereas previous coatings for synthetic nanopores were immobile[90,119–121], the lipid coatings developed by Yusko *et al.* are fluid and therefore present a number of unique capabilities[91]. First, these coatings minimize non-specific interactions between proteins and the pore wall, preventing pore clogging and enabling accurate determination of protein charge *via* the first-passage-time model initially developed by Talaga and Li[91]. Second, anchoring proteins to mobile lipids in the coating significantly increases translocation times (Fig 1.5) such that the majority of events are fully time-resolved after filtering, enabling accurate determination of protein volume[91]. Third, incorporating lipid-anchored ligands in the coating pre-concentrates particular analytes and therefore

provides specificity (Fig 1.5)[91]. Fourth, the bilayer thickness determines the pore diameter, allowing precise control of the diameter *in situ*[91]. Finally, the surface chemistry of the pore is defined by the bilayer composition and is therefore amenable to different applications[91].

Following the initial work by Yusko *et al*., a few additional studies have used lipid-coated nanopores as resistive-pulse sensors in order to study the aggregation of amyloid-beta proteins[92], improve the detection of λ-DNA[122], and measure the threading force acting on double-stranded DNA in the absence of electroosmotic flow[123].

1.4 Summary of dissertation

In this introduction, I first established the need for high-throughput single-molecule techniques to characterize and identify proteins. Next, I described the history and current state of the patch-clamp technique, the so-called gold standard for characterizing single ion channel proteins, and briefly touched on the need for a high-throughput platform capable of assessing single-channel activity. Subsequently, I presented the fundamentals behind resistive-pulse sensing and detailed the development of the technique from its invention to its use for characterizing single molecules. Furthermore, I described the problems associated with the use of both biological and synthetic nanopores for resistive-pulse sensing of proteins and introduced the concept of lipid-coated nanopores as a potential solution to many of these problems. In the next two chapters, I present advancements regarding both patch-clamp and resistive-pulse sensing that enhance the ability of these methods to characterize single proteins.

Chapter 2 describes the femtosecond-laser-based fabrication and use of dual-pore glass chips for cell-attached single-ion-channel recordings. The dual-pore design allows one pore (with an outer diameter of roughly 8 μm) to position a cell and a second smaller pore (approximately 150 to 300 nm in diameter) to establish a seal and record single-channel activity. Patch-clamp experiments with these chips consistently achieved high seal resistances, maintained gigaseals for up to 6 hours, achieved the lowest RMS noise ever reported for a planar patch-clamp platform, and enabled single-channel recordings in the cell-attached configuration with good fidelity.

Chapter 3 presents methods that use lipid-coated nanopores to determine the shape, volume, charge, rotational diffusion coefficient, and dipole moment of single non-spherical proteins simultaneously; previously, researchers had only been able to measure the volume and charge of spherical proteins. This work draws upon theory previously developed by Golibersuch in the 1970s describing variation in the magnitude of resistive-pulses that results from the rotation of a translocating particle. Chapter 3 further demonstrates the ability to distinguish proteins in a binary mixture based on the multiparametric information yielded by the methods presented here. Finally, this chapter quantitatively assesses the advantage of five-dimensional fingerprinting over standard two-dimensional characterization for protein identification.

In Chapter 4, I summarize the key results of this work and possible avenues for future research.

References

1. Boja, E. *et al.* Evolution of Clinical Proteomics and its Role in Medicine. *J. Proteome Res.* **10,** 66–84 (2011).

2. Dunn, M. J. Clinical Proteomics Advances. *Proteom Clin Appl* **8,** 1–2 (2014).
3. Chiou, S.-H. & Wu, C.-Y. Clinical proteomics: Current status, challenges, and future perspectives. *Kaohsiung J. Med. Sci.* **27,** 1–14 (2011).
4. Serdyuk, I. N., Zaccai, N. R. & Zaccai, J. *Methods in Molecular Biophysics: Structure, Dynamics, Function*. (Cambridge University Press, 2007).
5. O'Farrell, P. H. High resolution two-dimensional electrophoresis of proteins. *J. Biol. Chem.* **250,** 4007–4021 (1975).
6. Andersson, C.-O. Mass spectrometric studies on amino acid and peptide derivatives. in *Acta Chemica Scandinavica* **12,** 1353 (1958).
7. Lequin, R. M. Enzyme Immunoassay (EIA)/Enzyme-Linked Immunosorbent Assay (ELISA). *Clin. Chem.* **51,** 2415–2418 (2005).
8. *Single Molecules and Nanotechnology*. (Springer, 2008).
9. Oberhauser, A. F. *Single-molecule Studies of Proteins*. (Springer Science & Business Media, 2012).
10. Galvani, L. De viribus electricitatis in motu musculari commentarius. *Bon. Sci. Art. Inst. Acad. Comm.* **7,** 363–418 (1791).
11. Hodgkin, A. L. & Huxley, A. F. Action Potentials Recorded from Inside a Nerve Fibre. *Nature* **144,** 710–711
12. Graham, J. & Gerard, R. W. Membrane potentials and excitation of impaled single muscle fibers. *J. Cell. Comp. Physiol.* **28,** 99–117 (1946).
13. Ling, G. & Gerard, R. W. The normal membrane potential of frog sartorius fibers. *J. Cell. Comp. Physiol.* **34,** 383–396 (1949).
14. Cole, K. S. Dynamic electrical characteristics of the squid axon membrane. *Arch. Sci. Physiol. (Paris)* **3,** 253–258 (1949).
15. Marmont, G. Studies on the axon membrane. I. A new method. *J. Cell. Comp. Physiol.* **34,** 351–382 (1949).
16. Hodgkin, A. L., Huxley, A. F. & Katz, B. Measurement of current-voltage relations in the membrane of the giant axon of Loligo. *J. Physiol.* **116,** 424–448 (1952).
17. Neher, E. & Sakmann, B. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature* **260,** 799–802 (1976).
18. *Single-Channel Recording*. (Plenum Press, 1995).
19. Pratt, F. H. & Eisenberger, J. P. The Quantal Phenomena in Muscle: Methods, with Further Evidence of the All-or-None Principle for the Skeletal Fiber. *Am. J. Physiol.* **49,** 1–54 (1919).
20. Sigworth, F. J. & Neher, E. Single Na+ channel currents observed in cultured rat muscle cells. *Nature* **287,** 447–449 (1980).
21. Hamill, O. P., Marty, A., Neher, E., Sakmann, B. & Sigworth, F. J. Improved patch-clamp techniques for high-resolution current recording from cells and cell-free membrane patches. *Pflüg. Arch. Eur. J. Physiol.* **391,** 85–100 (1981).
22. Lindau, M. & Fernandez, J. M. IgE-mediated degranulation of mast cells does not require opening of ion channels. *Nature* **319,** 150–153 (1986).
23. Remillard, C. V. & Yuan, J. X. J. *Conventional Patch Clamp Techniques and High-Throughput Patch Clamp Recordings on a Chip for Measuring Ion Channel Activity*. (Springer, 2011).
24. Zheng, W. & Kiss, L. Screening Technologies for Ion Channel Targets in Drug Discovery. *Am. Pharm. Rev.* **6,** 85–92 (2003).

25. Asmild, M. *et al.* Upscaling and automation of electrophysiology: toward high throughput screening in ion channel drug discovery. *Receptors Channels* **9,** 49–58 (2003).
26. Fertig, N., Klau, M., George, M., Blick, R. H. & Behrends, J. C. Activity of single ion channel proteins detected with a planar microstructure. *Appl. Phys. Lett.* **81,** 4865–4867 (2002).
27. Klemic, K. G., Klemic, J. F., Reed, M. A. & Sigworth, F. J. Micromolded PDMS planar electrode allows patch clamp electrical recordings from cells. *Biosens. Bioelectron.* **17,** 597–604 (2002).
28. Kostyuk, P. G., Krishtal, O. A. & Pidoplichko, V. I. Effect of internal fluoride and phosphate on membrane currents during intracellular dialysis of nerve cells. *Nature* **257,** 691–693 (1975).
29. Coulter, W. H. Means for counting particles suspended in a fluid. (1953).
30. Graham, M. The Coulter principle: foundation of an industry. *J. Assoc. Lab. Autom.* **8,** 72–81 (2003).
31. Coulter, W. H. High speed automatic blood cell counter and cell size analyzer. in *Proc Natl Electron Conf* **12,** 1034–1040 (1956).
32. DeBlois, R. W. & Bean, C. P. Counting and Sizing of Submicron Particles by the Resistive Pulse Technique. *Rev. Sci. Instrum.* **41,** 909–916 (1970).
33. DeBlois, R. & Wesley, R. Viral sizes, concentrations, and electrophoretic mobilities by nanopar analyzer. **16,** A178–A178 (1976).
34. Kubitschek, H. E. Electronic Counting and Sizing of Bacteria. *Nature* **182,** 234–235 (1958).
35. Gregg, E. C. & Steidley, K. D. Electrical Counting and Sizing of Mammalian Cells in Suspension. *Biophys. J.* **5,** 393–405 (1965).
36. Maxwell, J. C. *Treatise on Electricity and Magnetism, Vol. 1*. (Dover Publications, 1954).
37. Rayleigh, Lord. The influence of obstacles arranged in rectangular order upon the properties of a medium. *Phil Mag* **34,** 481 (1892).
38. Fricke, H. A mathematical treatment of the electrical conductivity of colloids and cell suspensions. *J. Gen. Physiol.* **6,** 375–384 (1924).
39. Fricke, H. Relation of the Permittivity of Biological Cell Suspensions to Fractional Cell Volume. *Nature* **172,** 731–732 (1953).
40. Fricke, H. The Electric Permittivity of a Dilute Suspension of Membrane-Covered Ellipsoids. *J. Appl. Phys.* **24,** 644–646 (1953).
41. Velick, S. & Gorin, M. The electrical conductance of suspensions of ellipsoids and its relation to the study of Avian erythrocytes. *J. Gen. Physiol.* **23,** 753–771 (1940).
42. Grover, N. B., Naaman, J., Ben-Sasson, S. & Doljanski, F. Electrical sizing of particles in suspensions. I. Theory. *Biophys. J.* **9,** 1398–1414 (1969).
43. Grover, N. B., Naaman, J., Ben-Sasson, S., Doljanski, F. & Nadav, E. Electrical sizing of particles in suspensions. II. Experiments with rigid spheres. *Biophys. J.* **9,** 1415–1425 (1969).
44. Grover, N. B., Naaman, J., Ben-Sasson, S. & Doljanski, F. Electrical Sizing of Particles in Suspensions. III. Rigid Spheroids and Red Blood Cells. *Biophys. J.* **12,** 1099–1116 (1972).
45. Hurley, J. Sizing Particles with a Coulter Counter. *Biophys. J.* **10,** 74–79 (1970).

46. Anderson, J. L. & Quinn, J. A. The Relationship between Particle Size and Signal in Coulter-Type Counters. *Rev. Sci. Instrum.* **42,** 1257–1258 (1971).

47. Golibersuch, D. C. Observation of aspherical particle rotation in Poiseuille flow via the resistance pulse technique. I. Application to human erythrocytes. *Biophys. J.* **13,** 265–280 (1973).

48. Golibersuch, D. C. Observation of aspherical particle rotation in Poiseuille flow via the resistance pulse technique. II. Application to fused sphere 'dumbbells'. *J. Appl. Phys.* **44,** 2580–2584 (1973).

49. Berge, L. I., Feder, J. & Jo/ssang, T. A novel method to study single-particle dynamics by the resistive pulse technique. *Rev. Sci. Instrum.* **60,** 2756–2763 (1989).

50. Berge, L. I., Feder, J. & Jøssang, T. Single particle flow dynamics in small pores by the resistive pulse and the pressure reversal technique. *J. Colloid Interface Sci.* **138,** 480–488 (1990).

51. Carbonaro, A., Mohanty, S. K., Huang, H., Godley, L. A. & Sohn, L. L. Cell characterization using a protein-functionalized pore. *Lab. Chip* **8,** 1478–1485 (2008).

52. DeBlois, R. W., Bean, C. P. & Wesley, R. K. A. Electrokinetic measurements with submicron particles and pores by the resistive pulse technique. *J. Colloid Interface Sci.* **61,** 323–335 (1977).

53. DeBlois, R. W. & Wesley, R. K. Sizes and concentrations of several type C oncornaviruses and bacteriophage T2 by the resistive-pulse technique. *J. Virol.* **23,** 227–233 (1977).

54. Feuer, B. I., Uzgiris, E. E., Deblois, R. W., Cluxton, D. H. & Lenard, J. Length of glycoprotein spikes of vesicular stomatitis virus and sindbis virus, measured in Situ using quasi elastic light scattering and a resistive-pulse technique. *Virology* **90,** 156–161 (1978).

55. Von Schulthess, G., Benedek, G. & De Blois, R. Measurement of the Cluster Size Distributions for High Functionality Antigens Cross-Linked by Antibody. *Macromolecules* **13,** 939–945 (1980).

56. Von Schulthess, G. K., Benedek, G. B. & De Blois, R. W. Experimental measurements of the temporal evolution of cluster size distributions for high-functionality antigens cross-linked by antibody. *Macromolecules* **16,** 434–440 (1983).

57. Bezrukov, S. M., Vodyanoy, I. & Parsegian, V. A. Counting polymers moving through a single ion channel. *Nature* **370,** 279–281 (1994).

58. Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci.* **93,** 13770–13773 (1996).

59. *Nanopores: Sensing and Fundamental Biological Interactions*. (Springer, 2011).

60. Wanunu, M. Nanopores: A journey towards DNA sequencing. *Phys. Life Rev.* **9,** 125–158 (2012).

61. Movileanu, L., Howorka, S., Braha, O. & Bayley, H. Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nat. Biotechnol.* **18,** 1091–1095 (2000).

62. Kasianowicz, J. J., Henrickson, S. E., Weetall, H. H. & Robertson, B. Simultaneous Multianalyte Detection with a Nanometer-Scale Pore. *Anal. Chem.* **73,** 2268–2272 (2001).

63. Stefureac, R., Long, Y., Kraatz, H.-B., Howard, P. & Lee, J. S. Transport of α-Helical Peptides through α-Hemolysin and Aerolysin Pores. *Biochemistry (Mosc.)* **45,** 9172–9179 (2006).

64. Oukhaled, G., Bacri, L., Mathé, J., Pelta, J. & Auvray, L. Effect of screening on the transport of polyelectrolytes through nanopores. *EPL Europhys. Lett.* **82,** 48003 (2008).

65. Howorka, S. *et al.* A Protein Pore with a Single Polymer Chain Tethered within the Lumen. *J. Am. Chem. Soc.* **122,** 2411–2416 (2000).

66. Mereuta, L. *et al.* Slowing down single-molecule trafficking through a protein nanopore reveals intermediates for peptide translocation. *Sci. Rep.* **4,** (2014).

67. Danelon, C., Nestorovich, E. M., Winterhalter, M., Ceccarelli, M. & Bezrukov, S. M. Interaction of zwitterionic penicillins with the OmpF channel facilitates their translocation. *Biophys. J.* **90,** 1617–1627 (2006).

68. Kullman, L., Winterhalter, M. & Bezrukov, S. M. Transport of maltodextrins through maltoporin: a single-channel study. *Biophys. J.* **82,** 803–812 (2002).

69. Berkane, E. *et al.* Nanopores: maltoporin channel as a sensor for maltodextrin and lambda-phage. *J. Nanobiotechnology* **3,** 3 (2005).

70. Zhao, Q., de Zoysa, R. S. S., Wang, D., Jayawardhana, D. A. & Guan, X. Real-Time Monitoring of Peptide Cleavage Using a Nanopore Probe. *J. Am. Chem. Soc.* **131,** 6324–6325 (2009).

71. Howorka, S., Nam, J., Bayley, H. & Kahne, D. Stochastic Detection of Monovalent and Bivalent Protein–Ligand Interactions. *Angew. Chem. Int. Ed.* **43,** 842–846 (2004).

72. Movileanu, L., Schmittschmitt, J. P., Scholtz, J. M. & Bayley, H. Interactions of peptides with a protein pore. *Biophys. J.* **89,** 1030–1045 (2005).

73. Rotem, D., Jayasinghe, L., Salichou, M. & Bayley, H. Protein Detection by Nanopores Equipped with Aptamers. *J. Am. Chem. Soc.* **134,** 2781–2787 (2012).

74. Gu, L.-Q., Braha, O., Conlan, S., Cheley, S. & Bayley, H. Stochastic sensing of organic analytes by a pore-forming protein containing a molecular adapter. *Nature* **398,** 686–690 (1999).

75. Oukhaled, G. *et al.* Unfolding of Proteins and Long Transient Conformations Detected by Single Nanopore Recording. *Phys. Rev. Lett.* **98,** 158101 (2007).

76. Payet, L. *et al.* Thermal Unfolding of Proteins Probed at the Single Molecule Level Using Nanopores. *Anal. Chem.* **84,** 4071–4076 (2012).

77. Cressiot, B., Oukhaled, A., Bacri, L. & Pelta, J. Focus on Protein Unfolding Through Nanopores. *BioNanoScience* **4,** 111–118 (2014).

78. Mereuta, L., Asandei, A., Seo, C. H., Park, Y. & Luchian, T. Quantitative Understanding of pH- and Salt-Mediated Conformational Folding of Histidine-Containing, β-Hairpin-like Peptides, through Single-Molecule Probing with Protein Nanopores. *ACS Appl. Mater. Interfaces* **6,** 13242–13256 (2014).

79. Shin, S.-H., Luchian, T., Cheley, S., Braha, O. & Bayley, H. Kinetics of a Reversible Covalent-Bond-Forming Reaction Observed at the Single-Molecule Level. *Angew. Chem. Int. Ed.* **41,** 3707–3709 (2002).

80. Luchian, T., Shin, S.-H. & Bayley, H. Kinetics of a Three-Step Reaction Observed at the Single-Molecule Level. *Angew. Chem. Int. Ed.* **42,** 1926–1929 (2003).
81. Luchian, T., Shin, S.-H. & Bayley, H. Single-Molecule Covalent Chemistry with Spatially Separated Reactants. *Angew. Chem. Int. Ed.* **42,** 3766–3771 (2003).
82. Shin, S.-H. & Bayley, H. Stepwise growth of a single polymer chain. *J. Am. Chem. Soc.* **127,** 10462–10463 (2005).
83. Li, J. *et al.* Ion-beam sculpting at nanometre length scales. *Nature* **412,** 166–169 (2001).
84. Han, A. *et al.* Sensing protein molecules using nanofabricated pores. *Appl. Phys. Lett.* **88,** 093901 (2006).
85. Fologea, D., Ledden, B., McNabb, D. S. & Li, J. Electrical characterization of protein molecules by a solid-state nanopore. *Appl. Phys. Lett.* **91,** 053901 (2007).
86. Han, A. *et al.* Label-Free Detection of Single Protein Molecules and Protein−Protein Interactions Using Synthetic Nanopores. *Anal. Chem.* **80,** 4651–4658 (2008).
87. Talaga, D. S. & Li, J. Single-Molecule Protein Unfolding in Solid State Nanopores. *J. Am. Chem. Soc.* **131,** 9287–9297 (2009).
88. Pedone, D., Firnkes, M. & Rant, U. Data Analysis of Translocation Events in Nanopore Experiments. *Anal. Chem.* **81,** 9689–9694 (2009).
89. Niedzwiecki, D. J., Grazul, J. & Movileanu, L. Single-Molecule Observation of Protein Adsorption onto an Inorganic Surface. *J. Am. Chem. Soc.* **132,** 10816–10822 (2010).
90. Sexton, L. T. *et al.* An Adsorption-Based Model for Pulse Duration in Resistive-Pulse Protein Sensing. *J. Am. Chem. Soc.* **132,** 6755–6763 (2010).
91. Yusko, E. C. *et al.* Controlling protein translocation through nanopores with bio-inspired fluid walls. *Nat. Nanotechnol.* **6,** 253–260 (2011).
92. Yusko, E. C. *et al.* Single-Particle Characterization of Aβ Oligomers in Solution. *ACS Nano* **6,** 5909–5919 (2012).
93. Howorka, S. & Siwy, Z. Nanopore analytics: sensing of single molecules. *Chem. Soc. Rev.* **38,** 2360–2384 (2009).
94. Li, J. & Talaga, D. S. The distribution of DNA translocation times in solid-state nanopores. *J. Phys. Condens. Matter* **22,** 454129 (2010).
95. Ling, D. Y. & Ling, X. S. On the distribution of DNA translocation times in solid-state nanopores: an analysis using Schrodinger's first-passage-time theory. *J Phys Condens Matter* **25,** 375102 (2013).
96. Uram, J. D., Ke, K. & Mayer, M. Noise and Bandwidth of Current Recordings from Submicrometer Pores and Nanopores. *ACS Nano* **2,** 857–872 (2008).
97. Plesa, C. *et al.* Fast Translocation of Proteins through Solid State Nanopores. *Nano Lett.* **13,** 658–663 (2013).
98. Di Fiori, N. *et al.* Optoelectronic control of surface charge and translocation dynamics in solid-state nanopores. *Nat. Nanotechnol.* **8,** 946–951 (2013).
99. Sexton, L. T. *et al.* Resistive-Pulse Studies of Proteins and Protein/Antibody Complexes Using a Conical Nanotube Sensor. *J. Am. Chem. Soc.* **129,** 13144–13152 (2007).
100. Steinbock, L. J. *et al.* Probing the size of proteins with glass nanopores. *Nanoscale* **6,** 14380–14387 (2014).

101. Uram, J. D., Ke, K., Hunt, A. J. & Mayer, M. Submicrometer Pore-Based Characterization and Quantification of Antibody–Virus Interactions. *Small* **2,** 967–972 (2006).

102. Saleh, O. A. & Sohn, L. L. Direct detection of antibody-antigen binding using an on-chip artificial pore. *Proc. Natl. Acad. Sci. U. S. A.* **100,** 820–824 (2003).

103. Rodriguez-Larrea, D. & Bayley, H. Multistep protein unfolding during nanopore translocation. *Nat. Nanotechnol.* **8,** 288–295 (2013).

104. Freedman, K. J., Haq, S. R., Edel, J. B., Jemth, P. & Kim, M. J. Single molecule unfolding and stretching of protein domains inside a solid-state nanopore by electric field. *Sci. Rep.* **3,** (2013).

105. Freedman, K. J. *et al.* Chemical, thermal, and electric field induced unfolding of single protein molecules studied using nanopores. *Anal. Chem.* **83,** 5137–5144 (2011).

106. Wu, L. *et al.* Electrically facilitated translocation of protein through solid nanopore. *Nanoscale Res. Lett.* **9,** 1–10 (2014).

107. Wei, R., Gatterdam, V., Wieneke, R., Tampé, R. & Rant, U. Stochastic sensing of proteins with receptor-modified solid-state nanopores. *Nat. Nanotechnol.* **7,** 257–263 (2012).

108. Vlassiouk, I., Kozel, T. R. & Siwy, Z. S. Biosensing with Nanofluidic Diodes. *J. Am. Chem. Soc.* **131,** 8211–8220 (2009).

109. Tiwari, P. B. *et al.* Quantitative study of protein–protein interactions by quartz nanopipettes. *Nanoscale* **6,** 10255 (2014).

110. Ding, S., Gao, C. & Gu, L.-Q. Capturing Single Molecules of Immunoglobulin and Ricin with an Aptamer-Encoded Glass Nanopore. *Anal. Chem.* **81,** 6649–6655 (2009).

111. Hall, A. R., van Dorp, S., Lemay, S. G. & Dekker, C. Electrophoretic Force on a Protein-Coated DNA Molecule in a Solid-State Nanopore. *Nano Lett.* **9,** 4441–4445 (2009).

112. Kowalczyk, S. W., Hall, A. R. & Dekker, C. Detection of Local Protein Structures along DNA Using Solid-State Nanopores. *Nano Lett.* **10,** 324–328 (2010).

113. Smeets, R. M. M., Kowalczyk, S. W., Hall, A. R., Dekker, N. H. & Dekker, C. Translocation of RecA-Coated Double-Stranded DNA through Solid-State Nanopores. *Nano Lett.* **9,** 3089–3095 (2009).

114. Soni, G. V. & Dekker, C. Detection of Nucleosomal Substructures using Solid-State Nanopores. *Nano Lett.* **12,** 3180–3186 (2012).

115. Hall, A. R. & Dekker, C. in *Nanopores* (eds. Iqbal, S. M. & Bashir, R.) 35–49 (Springer US, 2011).

116. Hornblower, B. *et al.* Single-molecule analysis of DNA-protein complexes using nanopores. *Nat. Methods* **4,** 315–317 (2007).

117. Spiering, A., Getfert, S., Sischka, A., Reimann, P. & Anselmetti, D. Nanopore translocation dynamics of a single DNA-bound protein. *Nano Lett.* **11,** 2978–2982 (2011).

118. Uram, J. D. & Mayer, M. Estimation of solid phase affinity constants using resistive-pulses from functionalized nanoparticles. *Biosens. Bioelectron.* **22,** 1556–1560 (2007).

119. Nilsson, J., Lee, J. R. I., Ratto, T. V. & Létant, S. E. Localized Functionalization of Single Nanopores. *Adv. Mater.* **18,** 427–431 (2006).
120. Wang, G., Zhang, B., Wayment, J. R., Harris, J. M. & White, H. S. Electrostatic-gated transport in chemically modified glass nanopore electrodes. *J. Am. Chem. Soc.* **128,** 7679–7686 (2006).
121. Wanunu, M. & Meller, A. Chemically Modified Solid-State Nanopores. *Nano Lett.* **7,** 1580–1585 (2007).
122. Hernández-Ainsa, S. *et al.* Lipid-coated nanocapillaries for DNA sensing. *The Analyst* **138,** 104 (2013).
123. Galla, L. *et al.* Hydrodynamic Slip on DNA Observed by Optical Tweezers-Controlled Translocation Experiments with Solid-State and Lipid-Coated Nanopores. *Nano Lett.* **14,** 4176–4182 (2014).

CHAPTER 2

Dual-pore glass chips for cell-attached single-channel recordings

While high-throughput planar patch-clamp instruments are now established to perform whole-cell recordings for drug screening, the conventional micropipette-based approach remains the gold standard for performing cell-attached single-channel recordings. Generally, planar platforms are not well-suited for such studies due to excess noise resulting from low seal resistances and the use of substrates with poor dielectric properties. Since these platforms tend to use the same pore to position a cell by suction and establish a seal, biological debris from the cell suspension can contaminate the pore surface prior to seal formation, reducing the seal resistance. Here, femtosecond laser ablation was used to fabricate dual-pore glass chips optimized for use in cell-attached single-channel recordings that circumvent this problem by using different pores to position a cell and to establish a seal. This dual-pore design also permitted the use of a relatively small patch aperture (diameter ~ 150 to 300 nm) that is better-suited for establishing high-resistance seals than the micropores used typically in planar patch-clamp setups (diameter ~ 1 to 2 μm) without compromising the ability of the device to position a cell. Taking advantage of the high seal resistances and low capacitive and dielectric noise realized using glass substrates, patch-clamp experiments with these dual-pore chips consistently achieved high seal resistances (rate of gigaseal formation = 61%,

mean seal resistance = 53 GΩ), maintained gigaseals for prolonged durations (up to 6 hours), achieved RMS noise values as low as 0.46 pA at 5 kHz bandwidth, and enabled single-channel recordings in the cell-attached configuration that are comparable to those obtained by conventional patch-clamp.

2.1 Introduction

Ion channels comprise a diverse family of tightly regulated, pore-forming membrane proteins that permit the passive transport of ions across biological membranes and play a vital role in signal transduction and gene transcription, among other functions[1]. More than 50 different disorders (*i.e.*, channelopathies) such as cystic fibrosis and epilepsy are known to result from mutations in genes encoding for ion channels. Moreover, ion channel dysfunction is involved in many other conditions such as hypertension and chronic pain[2,3]. Consequently, drugs that target ion channels account for more than 13 percent of the pharmaceutical market, making ion channels the second most targeted family of proteins behind G protein-coupled receptors[4]. Nevertheless, ion channels are still underutilized as drug targets in part due to inadequate target validation and a dependence on *indirect* screening technologies (*e.g.*, fluorescence-based assays)[5].

The patch-clamp technique has remained the gold standard for *directly* screening ion channel activity since its invention by Erwin Neher and Bert Sakmann in the 1970s[6,7]. In a conventional patch-clamp experiment, the tip of a fire-polished, glass micropipette (diameter ~ 1 to 2 μm) is carefully positioned in contact with an adherent cell and gentle suction is applied to establish a high-resistance seal between the micropipette and cell

membrane (Fig. 2.1a). The electrical resistance of the seal is inversely proportional to background noise (*i.e.*, Johnson and shot noise) and therefore must be sufficiently large (typically ≥ 1 GΩ) to distinguish ion channel activity from noise[8]. In a whole-cell recording, the membrane patch encompassed by the micropipette tip is ruptured to provide electrical access to the interior of the cell and a high-gain amplifier records the ensemble average of all ion channel activity *via* electrodes in the electrolyte-filled pipette and bath solution. Conversely, in a cell-attached single-channel recording, the membrane patch is kept intact to record the activity of only those ion channels in the electrically isolated region of membrane. Unlike whole-cell recordings, single-channel recordings permit detailed kinetic analyses of individual ion channels, allow related or mutated channels to be distinguished based on their unitary conductances or their open-state and closed-state probabilities, and enable the investigation of drug-ion channel interactions at the single-molecule level[2,9]. Whole-cell and single-channel recordings each provide valuable yet complementary information that is critical for understanding ion channel behavior and selecting viable drug targets.
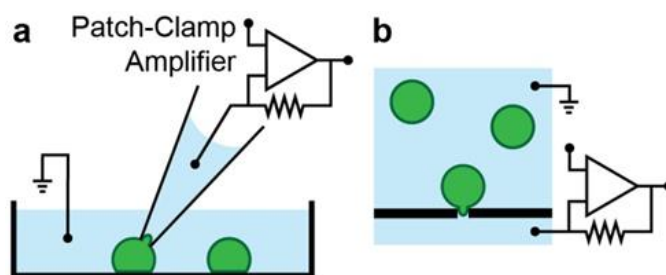
Figure 2.1 | Illustration of (a) conventional and (b) planar patch-clamp recordings. In a planar patch-clamp recording, a microfabricated pore replaces the micropipette used in a conventional recording.

Despite its widespread use, conventional patch-clamp is a low-throughput technique (tens of data points per day) that requires highly trained personnel and

expensive equipment such as a micromanipulator and optical microscope[6]. As a result, automated patch-clamp platforms have been developed to enable multiple recordings in parallel. Planar patch-clamp platforms are the most common variety; here, a micropore in a planar substrate is used to patch a cell from suspension (Fig. 2.1b). Borosilicate glass and quartz substrates typically provide the highest rates of $G\Omega$ seal (*i.e.*, gigaseal) formation[10–14], though various other materials have been used, including silicon coated with SiO_2 or phosphosilicate glass (PSG)[15–25], PEG/SU-8[26], PDMS[27–30], polyimide[31,32], and a cyclic olefin copolymer (COC)[33]. While planar patch-clamp platforms offer low-cost, high-throughput electrophysiological data with up to 18,000 data points per day, these devices seldom obtain seal resistances that are comparable to the conventional micropipette-based technique and are largely limited to whole-cell recordings in which the signal-to-noise ratio is high[10,34,35]. The inability of these devices to perform cell-attached single-channel recordings reliably is exemplified by the scarcity of published papers that demonstrate such recordings on an automated platform[10,11,28,36].

Here, we used femtosecond laser ablation to fabricate dual-pore glass chips optimized for use in cell-attached single-channel recordings. The dual-pore design is similar to that of the CytoPatch chip by Cytocentrics, wherein one pore (*i.e.*, the positioning pore) positions a cell by suction while another nearby pore (*i.e.*, the recording pore) avoids contamination by maintaining positive pressure until a cell is positioned and then establishes a seal[15]. Other planar patch-clamp platforms typically use the same pore to position a cell and establish a seal, hence increasing the chance of contaminating the pore surface with biological debris from the cell suspension prior to seal formation, which reduces the seal resistance. In contrast to the CytoPatch chip, however, the chips

developed here are made of borosilicate glass; this material generally yields higher seal resistances and has lower dielectric and capacitive current noise than SiO_2 on silicon[10]. Furthermore, the recording pore is smaller in diameter (150 to 300 nm instead of ~1.5 µm), which is advantageous for forming high-resistance seals[8]. Such a small pore would not be practical in a standard one-pore design as it would make it difficult to position a cell due to its high resistance to fluid flow. Our design takes inspiration from conventional patch-clamp, wherein smaller patch pipette openings are generally used for cell-attached recordings in comparison to those used for whole-cell recordings[8]. Patch-clamp experiments with these dual-pore chips consistently achieved high seal resistances ($\geq$ 10 GΩ), maintained gigaseals for prolonged durations (up to 6 hrs), and enabled cell-attached single-channel recordings that are comparable to those obtained by conventional patch-clamp.

2.2 Experimental

2.2.1 Machining Setups

To machine the recording and positioning pores (Fig. 2.2a,b), we used a diode-pumped Nd:glass chirped-pulse amplification laser system (Intralase) to generate 600-fs-long pulses at a wavelength of 1053 nm that were later frequency doubled by a KTP crystal to clean their temporal and spatial profile. We used a photodiode to measure the average power of the laser, which we adjusted with a reflective variable-density filter. We directed the laser into the epifluorescence path of an Axiovert 200M inverted microscope (Zeiss) and used a 40x, 0.65 NA Achroplan air objective (Zeiss) to focus the laser into a 150-µm-thick, borosilicate glass coverslip (72228, Electron Microscopy
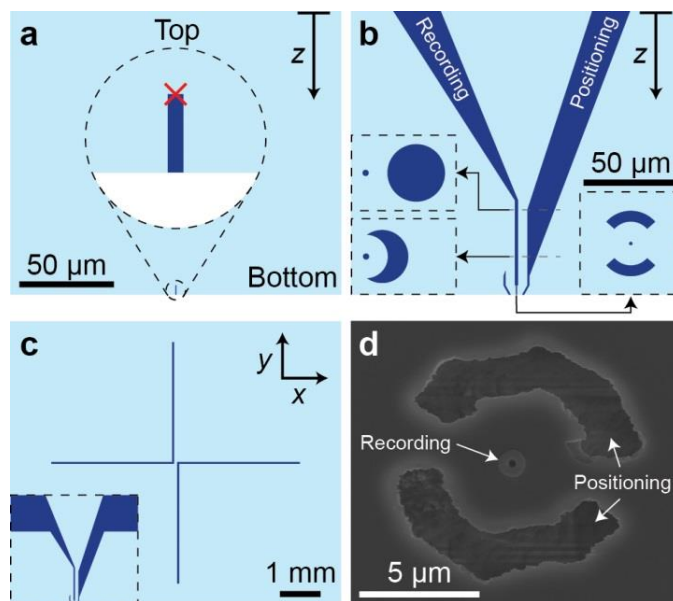
Figure 2.2 | Major stages in the fabrication of a dual-pore glass chip. a) Illustration showing a vertical cross-section after using a single femtosecond pulse to machine the aperture used for establishing seals. The red 'x' indicates the approximate focal depth of the laser. The diameter of the aperture varies between 150 to 300 nm. b) Same cross-section shown in pane (a) after using high-repetition-rate femtosecond laser ablation to machine the remainder of the recording pore (left) and entirety of the positioning pore (right). At the top surface, the pores are 20 μm in diameter and spaced by 100 μm from center to center. Moving downward, the pores steadily decrease in diameter and approach one another until the recording pore is centered above the single-shot aperture on the bottom surface ($z \approx 100$ μm). As z further increases, the positioning pore continues to decrease in diameter as it gradually encircles the recording pore, which remains fixed at 1.6 μm in diameter (see left insets). A 5-μm-thick layer of glass separates the pores in this region. Once the positioning pore is centered about the recording pore ($z \approx 140$ μm), the positioning channel splits into two segments (see right inset) to avoid fracturing the wall that separates the pores. Each subsequent layer is identical until the upper portion of the recording pore intersects the single-shot aperture, whereupon the distance between the two pores gradually decreases. At the bottom surface (right inset), the inner and outer diameters of the positioning pore are typically 6 and 8 μm, respectively. c) Top view after using high-repetition-rate ablation to machine low-resistance channels on the surface of the glass for interfacing electrically and fluidically with each pore. Each segment of the L-shaped channels is 3 mm long, 45 μm wide, and 50 μm deep. The pores are located at the vertices of the L-shaped channels. The inset shows a vertical cross-section. d) SEM image of the bottom surface. All panes are drawn to scale.

Sciences). We used a piezoelectric positioner (P-725.1CL, Physik Instrumente) to control the focal plane of the objective and an xy-nanostage (PI-629.2CL, Physik Instrumente) to move the coverslip laterally with respect to the focal spot. We put Milli-Q water on top of the coverslip and a solution containing 2 M KCl underneath to enhance debris removal[37] from the pores and to monitor the electrical connectivity across the chip with a picoammeter/voltage source (6487, Keithley). Last, we used custom-written MATLAB

(MathWorks) software to interface with the various electronics described here *via* a data acquisition board (PCI-6259, National Instruments). Section 2-App.S1 shows a simplified schematic of this machining setup.

We used a similar setup to machine channels on the surface of the coverslip for interfacing with each pore (Fig. 2.2c). Here, we used a fiber laser (Satsuma, Amplitude Systems) to generate 400-fs-long pulses at a wavelength of 1030 nm. To generate a relatively large focal volume, we focused the laser into the substrate with a 20x, 0.50 NA Achroplan air objective (Zeiss). We used the motorized focusing drive of an Axiovert 200M inverted microscope (Zeiss) to control the focal plane of the objective and an xy-microstage (BioPrecision2, Ludl) to move the coverslip laterally. In addition, we controlled the lateral position of the focal spot by adjusting the angle of two scanning galvo mirrors (GVS012, Thorlabs). As with the other setup, we kept Milli-Q water above the chip to enhance debris removal and used MATLAB to interface with the electronics.

2.2.2 Fabrication procedure

To machine the recording and positioning pores (Fig. 2.2a-b), we first ablated the opening of the recording pore on the bottom side of the coverslip by focusing a single femtosecond pulse as deep into the glass as possible while still yielding visible damage (Fig. 2.2a). To yield a relatively long aperture, we used a laser power that was two times larger than the single-shot ablation threshold as measured on the top surface (see Section 2-App.S2 for details). Next, we used high-repetition-rate (2 kHz) femtosecond laser ablation to machine the remainder of the recording pore and the entirety of the positioning pore layer by layer from top to bottom (Fig. 2.2b). We machined each layer

by scanning the laser focus no faster than 150 µm s^{-1} (corresponding to 1 pulse per 75 nm) along circular or semi-circular trajectories that are evenly spaced in the radial direction and centered about each pore. While machining the upper 100 µm of the coverslip, we used a laser power that was 25 percent above the high-repetition-rate threshold as measured on the top surface (see Section 2-App.S2) and a radial step size of 1 µm. We limited the vertical step size (*i.e.*, spacing between layers) to a maximum of 1 µm and gradually decreased its value such that the ablated regions in adjacent layers overlapped by at least 90 percent in terms of area. Once the positioning pore started to encircle the recording pore (Fig. 2.2b, left insets), we began ablating each layer of the positioning pore twice and reduced the laser power to 10 percent above threshold as measured on the bottom surface (see Section 2-App.S2) and reduced the radial and vertical step sizes to 0.2 and 0.4 µm, respectively. Once the positioning pore was centered about the recording pore, we resumed ablating each layer only once. We halted machining of the recording pore once the current increased by 0.2 nA at an applied potential of 500 mV. Machining of the positioning pore continued until the focal plane was positioned below the bottom surface of the chip. This procedure took approximately 45 minutes to complete. Immediately following ablation, we placed the coverslip in Milli-Q water with its upper surface (Fig. 2.2b) facing downward to allow additional debris to settle out of the pores.

We used high-repetition-rate (200 kHz) ablation to machine low-resistance L-shaped microchannels for interfacing with each pore (Fig. 2.2c). As before, machining proceeded layer by layer from top to bottom. We used a laser power that was roughly twice that of the high-repetition-rate threshold as measured on the top surface (see

Section 2-App.S2) and a vertical step size of roughly 2 μm. We machined each layer by scanning the laser focus about the length of each channel at a rate of 200 μm s^{-1}. Simultaneously, we oscillated a galvo mirror at 100 Hz to quickly move the laser focus back-and-forth along the channel width. To account for tilt in the coverslip, we discretely adjusted the position of the objective along the length of the channel each time the position of the surface changed by more than 0.5 μm (see Section 2-App.S3 for details). This procedure took approximately 20 minutes to complete.

We optimized all parameters to minimize the failure rate of the machining process and to prevent cracking of the glass substrate; for instance, excessive laser power results in the formation of microcracks in the vicinity of the ablated structures. Following laser fabrication, we often etched the chips for a short duration ($\leq$ 30 s) in a buffered hydrofluoric acid solution (Buffer HF Improved, Transene) to remove residual debris from the pores.

2.2.2 Cell culture

HEK-293 cells transfected with large-conductance Ca^{2+}-activated K$^+$ (BK) channels were obtained from Dr. Heike Wulff (University of California, Davis) and were cultured in Dulbecco's modified Eagle medium (DMEM, ATCC) with 10% (v/v) fetal bovine serum (Gibco), 0.5 mg/mL G418 (Sigma-Aldrich), 100 units/mL penicillin (Gibco), and 100 μg/mL streptomycin (Gibco). To prepare these cells for patching, we first rinsed the cell culture flask in PBS to remove extracellular proteins. Next, we treated the cells with 0.05% (w/v) trypsin-EDTA (Gibco) for 3 minutes at 37 ℃, adding fresh medium to stop the trypsinization process. We then aspirated and centrifuged the

suspension. We resuspended the cells in cell culture medium and waited 15 minutes as we measured the cell density. Subsequently, we centrifuged the suspension and resuspended the cells in an electrolyte solution (see Section 2.2.4) to achieve a density of roughly 10^6 cells mL^{-1}. We passed the suspension through a 40 μm cell strainer (Becton Dickinson) to remove cell aggregates. Last, we placed the suspension in a 1 mL Eppendorf tube, which was constantly mixed at 800 rpm and 37 ℃ *via* a Thermomixer R (Eppendorf). This procedure has been shown to reduce the number of cell aggregates and to maintain cell viability at 90% for up to 4.5 hrs; however, it is widely accepted that the cells form higher quality seals if used within 45 minutes of passaging[21].

2.2.3 Device set-up

Prior to use, we cleaned the chips overnight with a piranha solution consisting of 3:1 (v/v) concentrated sulphuric acid and 30% (v/v) aqueous hydrogen peroxide solution at 90 °C. Immediately preceding each experiment, we rinsed a chip with Milli-Q water, dried it with argon gas, and mounted it in the scaffold shown in Fig. 2.3a-b. We placed the scaffold on a vibration isolation table (BM-4, Minus K Technology) inside of a Faraday cage to minimize noise. A constant-pressure pump (Suction Control Pro, Nanion) and shutoff valve were located upstream and downstream of each pore, respectively. To fill the pores with solution, we applied positive pressure while the shutoff valves were open. Once the L-shaped channels (Fig. 2.2c) were filled, as indicated by a significant decrease in the flow rate, we closed both shutoff valves to stop perfusion and placed solution in the well above the chip. The fluidic resistance of each pore is relatively high in comparison to that of the L-shaped channels; hence, the amount
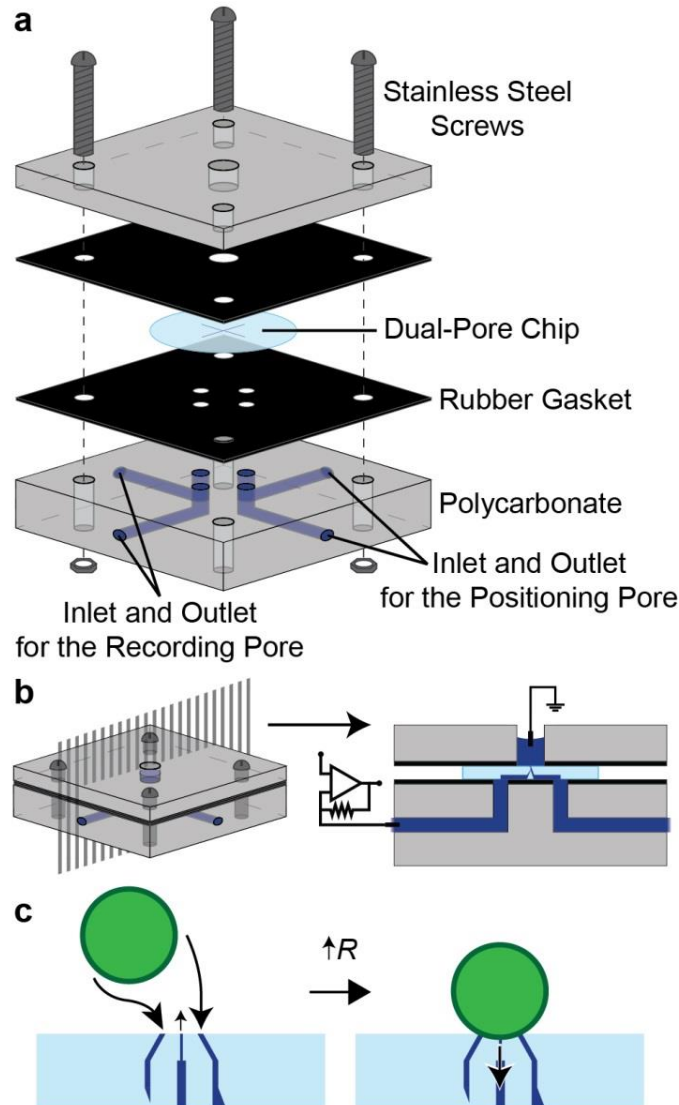
Figure 2.3 | Using a dual-pore glass chip in a planar patch-clamp experiment. a) Schematic showing the assembly of the scaffold used to interface with a dual-pore chip. We oriented the chip such that the low-resistance L-shaped channels (Fig. 2.2c) were facing downward. The ends of these channels were continuous with channels in the polycarbonate, providing two separate flow paths for perfusing to each pore. b) Assembled scaffold; the cross-section shown on the right corresponds to the plane indicated on the left. c) Simplified representation of the strategy used to establish a high-resistance seal with a cell. Briefly, we applied suction to the positioning pore and positive pressure to the recording pore until we detected an increase in resistance resulting from a cell being positioned in the vicinity of the recording pore. At this point, we ceased to apply suction to the positioning pore and applied suction to the recording pore to establish a seal.

of time needed to fill the scaffold can be dramatically reduced by allowing air to flow out

the downstream end of each L-shaped channel before solution reaches the chip. An

33

Ag/AgCl electrode located upstream of the recording pore served as the command electrode, which we attached to the headstage of a patch-clamp amplifier (EPC 10 Plus, HEKA), while a second electrode in the bath solution served as ground. We used commercially available software for data acquisition (PatchMaster, HEKA).

2.2.4 Patch-clamp experiments

We filled the recording and positioning pores with an electrolyte solution consisting of 140 mM KCl, 5 mM NaCl, 10 mM HEPES, 5 mM glucose, 2 mM $CaCl_2$, and 2 mM $MgCl_2$ (pH = 7.4). After filling each pore, we applied a positive pressure of 3 kPa to the recording pore until we attempted to establish a seal. After adding cells to the well above the chip, we applied a negative pressure of -2.5 kPa to the positioning pore to aspirate a cell (Fig. 3.3c). During this process, we regularly applied a 10 mV voltage pulse to monitor the resistance across the recording pore. Once a significant increase in resistance was observed, we immediately applied a negative pressure of -3 kPa to the recording pore and stopped applying pressure to the positioning pore (Fig. 3.3c). If the seal resistance stabilized for a prolonged duration (> 15 s) prior to reaching the gigaohm range, we gradually increased the magnitude of the negative pressure up to a maximum of -30 kPa. Once we established a gigaseal, we ceased to apply pressure to either pore. We obtained all recordings in the cell-attached configuration and voltage-clamp mode (*i.e.*, at constant applied potential). We set the low-pass 4-pole Bessel filter of the amplifier to a cut-off frequency of 5 kHz and used a sampling rate of 25 kHz.

After use in a patch clamp experiment, each chip was placed in a solution of 1% Micro-90 (International Products Corporation), an alkaline cleaning solution that is used

to remove biological debris, and desiccated for roughly 30 minutes to ensure the pores were filled with the solution. The chips were kept in this solution until they were cleaned in piranha for additional experiments.

2.3 Results and discussion

2.3.1 Device characterization

As dual-pore glass chips (Fig. 2.2b,c) cannot be machined using traditional microfabrication techniques, we used femtosecond laser ablation to fabricate our design in borosilicate glass coverslips. Herbstman and Hunt had previously shown that high-aspect ratio nanochannels can be machined by focusing a single femtosecond laser pulse just below the surface of a glass coverslip[38]. We used this technique to fabricate the opening of the recording pore (Fig. 2.2a), then used high-repetition-rate ablation to machine the rest of the geometry. Fig. 2.2d shows a scanning electron microscopy (SEM) image of the surface of a completed dual-pore chip. Due to differences between single-shot and high-repetition-rate ablation, the recording pore is smoother than the positioning pore; in fact, the recording pore appears devoid of discernable roughness, similar to the tip of a patch pipette. Smoother pores tend to yield higher seal resistances, highlighting the importance of using single-shot ablation to machine the recording pore[39]. The recording pore is also smaller in diameter than a typical planar patch-clamp pore and hence is better suited for forming high-resistance seals due to geometric factors[8].

We characterized the resistance of the recording pore in series with the L-shaped channel (*i.e.*, the access resistance) using a standard extracellular solution ($\rho \sim 0.5\ \Omega$ m). The resistance varied between 24 and 116 MΩ with an average value of 60 ± 25 MΩ ($N =$

18 chips). While these resistances are generally too large for whole-cell recordings without using series resistance compensation, they are suitable for cell-attached single-channel recordings. Assuming the diameter of the single-shot aperture (Fig. 2.2a) is constant at 250 nm, these resistances suggest that the aperture typically accounts for the majority of the access resistance (78%) and that its length varies between 1 to 10 μm and is 5 ± 2 μm on average (see Section 2-App.S4). Based on the dimensions shown in Fig. 2.2b,c, we expect the resistance of the positioning pore in series with the L-shaped channel to be less than 2 MΩ, which is similar to the resistance of other planar patch-clamp pores that use suction to position a cell[10–12]. Hence, we expect the positioning pore to be able to position a cell as effectively as these other devices.

Low capacitance is critical for minimizing dielectric noise and the distributed *RC*-noise of the recording pore. Therefore, we determined the total capacitance of the recording setup by cancelling the fast capacitive transients that occur upon the application of a voltage pulse. On average, we measured a capacitance of 1.8 ± 0.4 pF (*N* = 39; see Section 2-App.S5 for a boxplot of this data). The amplifier and headstage contribute 1 to 1.5 pF to this capacitance, which indicates that the polycarbonate scaffold, electrode leads, and dual-pore chip contribute approximately 0.3 to 0.8 pF in total. The combined capacitance of the scaffold, leads, and chip is comparable to the lowest values reported in the literature for other planar patch-clamp devices and is smaller than the capacitance of a typical patch pipette[10,11]. In order to achieve such a low capacitance, we found that it was crucial to minimize the length of the electrode leads and to interface with the backside of each pore *via* a microchannel (Fig. 2.2c).

2.3.2 Seal quality

High seal resistances are critical for providing high fidelity single-channel recordings with minimal noise and leakage currents. To assess the ability of the dual-pore chips to establish high-resistance seals, we conducted planar patch-clamp experiments with HEK-293 cells using the seal formation strategy shown in Fig. 2.3c. In over 90 percent of the experiments, we observed a significant increase in resistance (*i.e.*, greater than ~5% of the access resistance), indicating that a cell was positioned in the vicinity of the recording pore. Typically, this increase in resistance occurred within 1 to 2 minutes after adding cells. Fig. 2.4a shows the maximum seal resistance we obtained in each experiment with respect to the access resistance and estimated length of the single-shot aperture (Fig. 2.2a). While Nagarah *et al.* previously observed a positive correlation between seal resistance and pore length with quartz pores, we observed no such correlation in this work (Pearson's r = -0.10)[12]. This discrepancy may result from the relatively small diameter of the recording pore, which could reduce the protrusion of the membrane into the pore upon the application of suction. Nonetheless, the rate of gigaseal formation was 61 percent, the mean seal resistance was 53 GΩ, the median seal resistance was 3 GΩ, and the maximum seal resistance was 650 GΩ. Excluding experiments in which we did not form a gigaseal, the mean and median seal resistances were 87 and 15 GΩ, respectively. Whereas the rate of gigaseal formation is comparable to that of other planar patch-clamp devices, the magnitude of the seal resistance is relatively large when a gigaseal is formed[10]. For instance, van Stiphout *et al.* obtained a median gigaseal resistance of approximately 2 GΩ using the CytoPatch device[15]. In certain cases, a gigaseal may have failed to form as a result of off-center positioning of

37

the cell; this is supported by the observation that the initial increase in resistance upon positioning a cell tended to be lower when a gigaseal did not form.
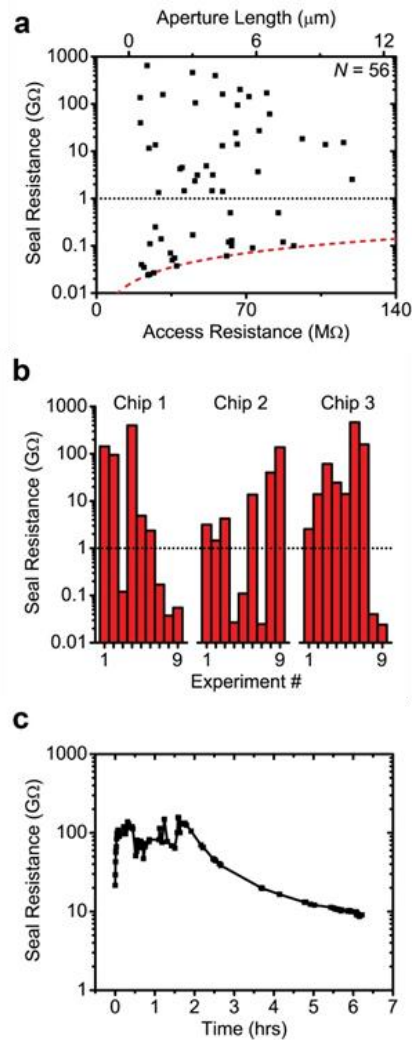
Figure 2.4 | Dual-pore chips repeatedly form and maintain seals in patch-clamp experiments with HEK-293 cells. a) Maximum seal resistance versus access resistance and estimated length of the single-shot aperture (Fig. 2.2a). We excluded experiments in which the access resistance was significantly larger (> 10%) than what we previously measured. The red dashed line corresponds to no change in the seal resistance during an experiment and the black dotted line indicates a seal resistance of 1 GΩ. b) Ability of the same chip to repeatedly form gigaseals. We used each chip in nine patch-clamp experiments. Between experiments, we cleaned the chips in 1% Micro-90 and piranha solution. Chips 1, 2, and 3 formed gigaseals in 56, 67, and 78 percent of the experiments, respectively. The black dotted line indicates a seal resistance of 1 GΩ. c) Maintaining a gigaseal for a prolonged duration. We were able to maintain a gigaseal for over 6 hrs using a dual-pore chip. The seal resistance reached its maximum value of 154 GΩ at approximately 1.6 hrs into the experiment. We measured all seal resistances while in the cell-attached configuration.

Fig. 2.4b shows that the same dual-pore chip can be used to form a gigaseal on multiple occasions by cleaning the chip between experiments (see Section 2.2 for details). Seal resistance was not clearly correlated with the number of times we used each chip (Pearson's r = -0.43, 0.67, and 0.17), suggesting that the same chip can be re-used indefinitely without degrading the average seal quality. Kao *et al.* obtained a similar result using a slightly different cleaning procedure with the Nanion Port-a-Patch system[40]. Consequently, it might be possible to improve the rate of gigaseal formation and average seal resistance in future experiments by re-using only the best-performing dual-pore chips (*e.g.*, Chip 3 in Fig. 2.4b).

Finally, Fig. 2.4c shows the seal resistance from a single patch-clamp experiment as a function of time. We were able to maintain a gigaseal for over 6 hrs, which is long in comparison to the typical duration of a conventional patch-clamp experiment performed in the cell-attached configuration. Prolonging the duration of a gigaseal permits longer recordings that ultimately increase data throughput *via* the exploration of a wider parameter space; accordingly, a platform such as the one presented here that is capable of maintaining a seal for an extended duration is ideal for maximizing the utility of each experiment.

2.3.3 Noise Characterization

From experiment to experiment, the root-mean-square (RMS) current (*i.e.*, noise) we measured after forming a gigaseal with a dual-pore chip varied between 0.46 to 1.3 pA and was 0.92 pA on average at a bandwidth of 5 kHz and an applied potential of ±50 mV (N = 13; see Section 2-App.S6 for a boxplot of this data). To our knowledge, the

lowest RMS current ever reported for a planar patch-clamp platform in the cell-attached configuration is 0.27 pA at a bandwidth of 1 kHz[11]. Here, we achieved RMS currents as low as 0.12 pA at 1 kHz bandwidth. Using a conventional patch-clamp setup, the RMS current varied between 0.48 to 0.83 pA and was 0.58 pA on average under the same conditions as used with the dual-pore chip ($N = 6$; see Section 2-App.S6 for a boxplot of this data). While the average RMS current of the dual-pore platform is approximately 60 percent higher than that of the conventional setup, the minimum noise values achieved by both platforms are nearly identical. Furthermore, nearly 40 percent of the experiments conducted with the dual-pore platform achieved levels of noise that fell within the range observed for the conventional setup. Hence, the platform developed here was often able to perform as well as a conventional patch-clamp setup after establishing a gigaseal.

The RMS current generated by the headstage, analog low-pass Bessel filter, amplifier, and digitizer was approximately 0.15 pA at a bandwidth of 5 kHz. The additional noise observed during a patch-clamp experiment was likely dominated by the distributed RC-noise of the recording pore since other sources of noise (*i.e.*, dielectric and Johnson noise) should be relatively small according to theory[41]. To reduce this noise, the access resistance of the recording pore could be reduced by, for example, increasing the depth of the L-shaped channel leading to the pore (Fig. 2.2c). This was attempted but increasing the depth made the chips too fragile to handle easily.

2.3.4 Cell-attached single-channel recordings

To demonstrate the capability of the dual-pore platform to perform cell-attached single-channel recordings with similar fidelity to the conventional method, we performed

patch-clamp experiments with HEK-293 cells that were transfected with large-conductance Ca^{2+}-activated K^+ (BK) channels. We used an extracellular solution with a high concentration of potassium (140 mM) for these experiments to establish a resting membrane potential that was near zero. Fig. 2.5 shows the resulting single-channel recordings obtained using the dual-pore platform as well as a conventional patch-clamp setup (see Section 2-App.S7 for additional current traces). In each experiment, we could clearly distinguish two levels of current. We measured a single-channel conductance of 203 and 253 pS using the dual-pore and conventional setups, respectively. These two values are both within the expected range for BK channels (100 to 270 pS) and are in fairly good agreement with each other (22% difference)[42]. Furthermore, the RMS noise varied by less than 5 percent between the two experiments. These results show that the dual-pore platform is capable of performing low-noise single-channel recordings in the cell-attached configuration that are comparable to those obtained using the conventional technique.
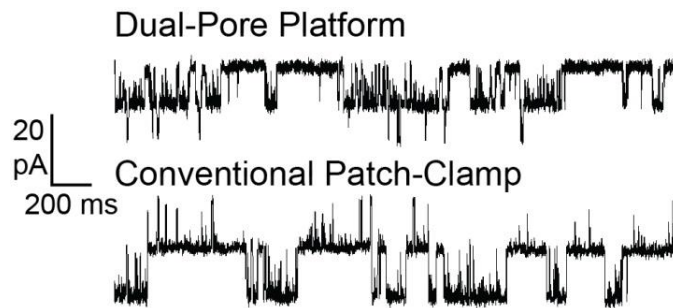
Figure 2.5 | Comparison of cell-attached single-channel recordings obtained using the dual-pore platform and a conventional patch-clamp setup. The activity of single BK channels was monitored at a bandwidth of 5 kHz (as plotted) and an applied potential of -50 mV. The seal resistance was in excess of 10 GΩ in both experiments. The RMS current in the absence of single-channel activity was 0.95 and 0.90 pA at 5 kHz bandwidth for the dual-pore and conventional setup, respectively.

2.4 Conclusions

Even with the advent of automated planar patch-clamp platforms, the conventional micropipette-based technique is still the method of choice for performing cell-attached single-channel recordings largely due to the high signal-to-noise ratios that are required. Here, we developed dual-pore glass chips designed explicitly for performing such recordings in an automated manner. By employing a dual-pore design, we were able to use a patch aperture that is much smaller than a typical planar patch-clamp pore and hence better-suited for forming high-resistance seals without sacrificing the ability of the device to position a cell *via* suction. Glass is an ideal substrate for planar patch-clamp due to its excellent dielectric and seal forming properties; however, it is difficult to fabricate complex three-dimensional structures, such as the design described here, in glass *via* conventional microfabrication techniques. Only laser-based machining as used here makes it possible to fabricate such structures in glass. Using the dual-pore platform, we achieved exceptionally high seal resistances and the lowest noise ever reported in the cell-attached configuration for a planar platform. Furthermore, we obtained single-channel recordings with similar fidelity to the gold standard (*i.e.* conventional) technique. Ideally, the dual-pore platform developed here will inspire future designs for performing high-throughput screening of single ion channels and help to expedite the drug discovery process by providing information that is complementary to whole-cell recordings. In future work, the design described here could be modified to accommodate smaller cells and organelles by using an array of single-shot pores for positioning instead of one or two large pores. In addition, two sets of dual-pores could be fabricated in close proximity to one another for performing automated on-chip gap junction recordings.

2-App.S1 Schematic of machining setup
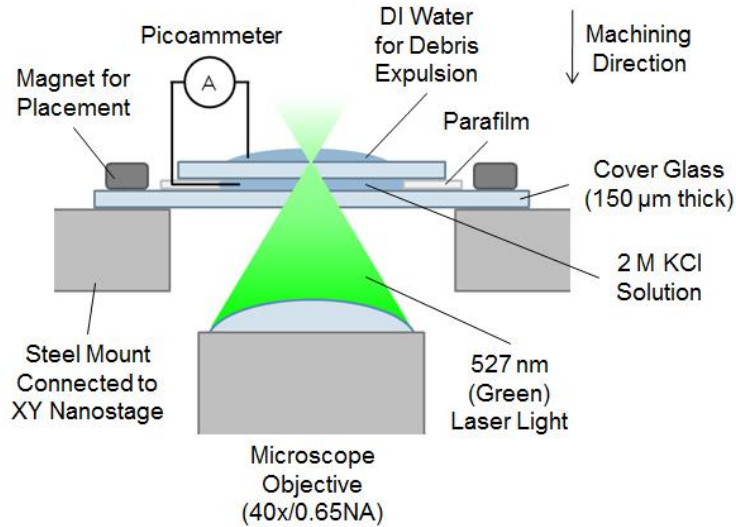
Figure 2-App.1 | Schematic representation of the machining setup used for femtosecond laser ablation of the recording and positioning pores (Fig. 2.2a,b). Since the laser passes through media with different indices of refraction, changes in the vertical position of the objective do not directly equal the corresponding changes in the position of the focal spot. For instance, we only need to move the objective by 92 μm in order for the focal spot to traverse the entire 150 μm thickness of a glass coverslip. As a result, we multiplied all vertical step sizes by a factor of 0.61 to determine how far to move the objective during machining.

2-App.S2 Measuring the ablation threshold

We measured single-shot and high-repetition-rate ablation thresholds at the beginning of each day of machining to account for drift in the laser profile and pulse width. Here, we define an ablation threshold as the lowest laser power that produces visible damage on the surface of the glass coverslip as viewed under brightfield microscopy. We used the same objective that focuses the laser into the substrate to observe the laser-induced damage. The procedure for measuring the ablation threshold is simple. First, we adjusted the position of the objective to bring the image of the glass surface into focus. Next, we attempted to ablate the surface as we varied the focal plane

about the starting position; this was necessary to account for a slight difference between the position of the laser focus and the imaging plane. When measuring the single-shot threshold, we adjusted the lateral position of the laser focus between shots to avoid effects from sub-threshold modification to the glass. When measuring the high-repetition-rate threshold, we scanned the laser focus about a line at each focal plane. After attempting to ablate the surface at a particular power, we adjusted the power either up or down to find the threshold at which damage started to occur or no longer occurred. The ablation threshold was generally higher at the top surface of the glass in comparison to the bottom surface (see Fig. 2.2b), which is likely due to spherical aberration.

2-App.S3 Measuring tilt in the coverslip surface

When machining the L-shaped channels shown in Fig. 2.2c, we accounted for tilt in the coverslip to avoid cracking or incomplete ablation resulting from focusing the laser too far below the surface. At three separate non-collinear *x-y* coordinates, we determined the position of the uppermost focal plane that yielded visible single-shot damage in order to calculate the tilt (assuming a flat surface). Tilt was negligible when machining the positioning and recording pores (Fig. 2.2b) as they span a relatively small lateral distance.

2-App.S4 Estimating the resistance contribution and length of the single-shot aperture

The access resistance of a dual-pore chip can be described by the following equation:

$$R_A = R_L + R_P \tag{2-App.1}$$

where R_A is the access resistance, R_L is the resistance of the segment of the L-shaped channel that is upstream of the recording pore, and R_P is the resistance of the recording pore. Using Pouillet's law, we can derive an expression for R_L:

$$R_L = \frac{\rho L_L}{w_L h_L} \tag{2-App.2}$$

where ρ is the resistivity of the electrolyte solution, L_L is the channel length, w_L is the channel width, and h_L is the channel height. Assuming the single-shot aperture located at the entrance of the recording pore (Fig. 2.2a) is a perfect cylinder and neglecting access resistance, R_P can be described as follows:

$$R_P = \frac{4\rho h_C}{\pi D_1 D_2} + \frac{4\rho(t - h_C - L_{SS})}{\pi D_2^2} + \frac{4\rho L_{SS}}{\pi D_{SS}^2} \tag{2-App.3}$$

where h_C is the height of the conical segment of the pore (see Fig. 2.2b), D_1 is the pore diameter at the top surface of the coverslip, D_2 is the diameter at the bottom of the conical segment, t is the thickness of the coverslip, L_{SS} is the length of the single-shot aperture, and D_{SS} is the diameter of the single-shot aperture. The first term of this expression is the resistance of the conical segment of the pore, the second term is the resistance of the cylindrical segment of the pore, and the third term is the resistance of the single-shot aperture. Combining the above equations and solving for L_{SS} yields:

$$L_{SS} = \left[\frac{\pi D_2^2}{4} \left(\frac{R_A}{\rho} - \frac{L_L}{w_L h_L} \right) + h_C \left(1 - \frac{D_2}{D_1} \right) - t \right] \left(\frac{D_2^2}{D_{SS}^2} - 1 \right)^{-1} \tag{2-App.4}$$

Assuming D_{SS} equals 250 nm, L_{SS} varies between 1 to 10 μm and is 5 ± 2 μm on average ($N = 18$ chips). Based on these estimates of length, the resistance of the aperture varies

between 9 to 103 MΩ (37 to 89% of the access resistance) and is 49 ± 25 MΩ (78% of

the access resistance) on average. Furthermore, the resistance of the remainder of the

recording pore is 13 ± 1 MΩ and accounts for 10 to 60% of the access resistance. Finally,

R_L is equal to 0.7 MΩ and accounts for less than 3% of the access resistance.

2-App.S5 Characterizing the capacitance of the recording setup
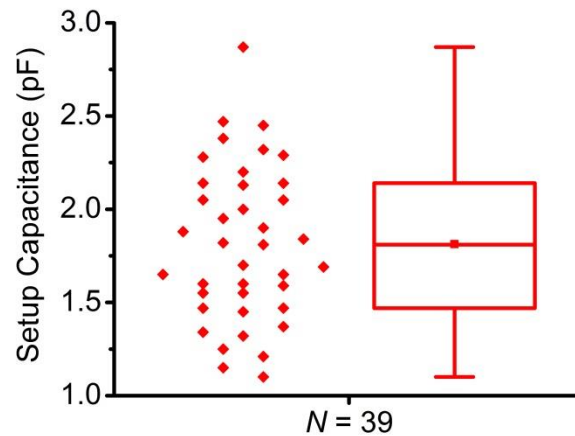
Figure 2-App.2 | Total capacitance of the recording setup. Each data point is from a different experiment. The raw data is shown on the left and a corresponding boxplot is shown on the right. The box encompasses the middle 50 percent of the data, the horizontal line and point inside the box show the median and mean values, respectively, and the whiskers extend to data points that are within 1.5 * IQR from the 25th and 75th percentiles. We obtained capacitance values by using the "Auto C-fast" function in the PatchMaster software to cancel the fast capacitive transients that occur upon the application of a voltage pulse.

2-App.S6 Characterizing noise in the presence of a gigaseal
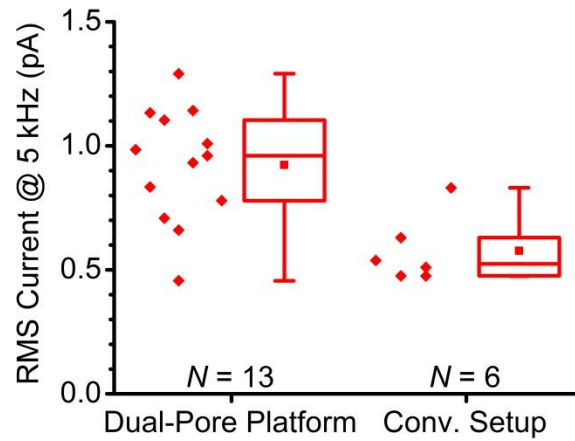
Figure 2-App.3 | RMS current measured after forming a gigaseal with both the dual-pore platform (left) and a conventional patch-clamp setup (right). Each data point shows the minimum RMS current measured in each experiment. The raw data is shown on the left and a corresponding boxplot is shown on the right. The box encompasses the middle 50 percent of the data, the horizontal line and point inside the box show the median and mean values, respectively, and the whiskers extend to data points that are within 1.5 * IQR from the 25th and 75th percentiles. We obtained all current traces at a bandwidth of 5 kHz and an applied potential of ±50 mV. We calculated each RMS current value using a region of the current trace in which no single-channel activity was present (*i.e.*, when all ion channels were in the closed state or no channels were present in the patch).

2-App.S7 Cell-attached single-channel recordings of BK channels
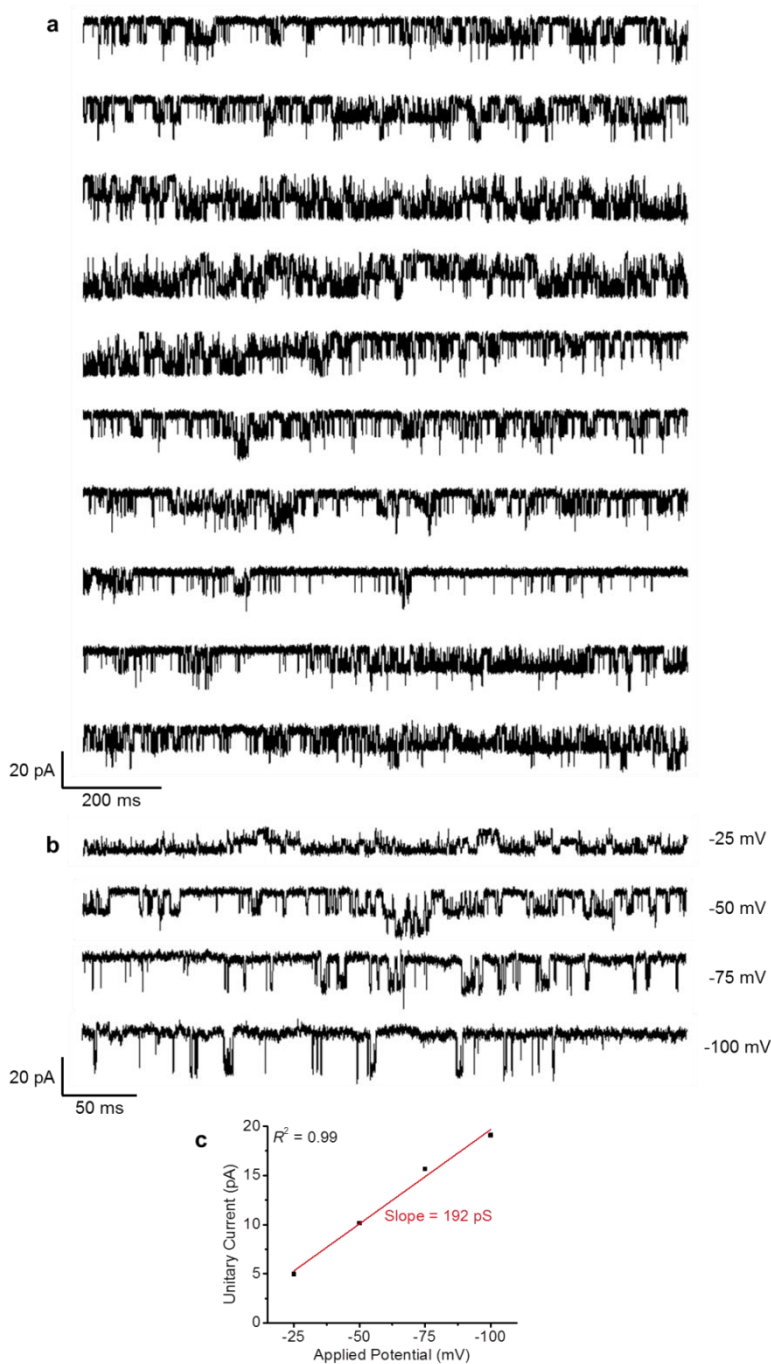
Figure 2-App.4 | Measuring the activity of single BK channels in the cell-attached configuration. a) 12-s-long current trace obtained at a bandwidth of 5 kHz (as plotted) and an applied potential of -50 mV. b) Current traces obtained at -25, -50, -75, and -100 mV. With increasing depolarization, the open-state probability of the BK channels increases, as expected. c) Plot of the unitary (*i.e.*, single-channel) current versus the applied potential. The red line is a best-fit. The slope of this line provides an estimate of the single-channel conductance, which agrees well with the value measured at -50 mV of 203 pS (6% difference) and falls within the range of previously published values (100 to 270 pS).

References

1. Zheng, W. & Kiss, L. Screening Technologies for Ion Channel Targets in Drug Discovery. *Am. Pharm. Rev.* **6,** 85–92 (2003).
2. Remillard, C. V. & Yuan, J. X. J. *Conventional Patch Clamp Techniques and High-Throughput Patch Clamp Recordings on a Chip for Measuring Ion Channel Activity*. (Springer, 2011).
3. Terstappen, G. C., Roncarati, R., Dunlop, J. & Peri, R. Screening technologies for ion channel drug discovery. *Future Med. Chem.* **2,** 715–730 (2010).
4. Mathie, A. Ion channels as novel therapeutic targets in the treatment of pain. *J. Pharm. Pharmacol.* **62,** 1089–1095 (2010).
5. Wickenden, A., Priest, B. & Erdemli, G. Ion channel drug discovery: challenges and future directions. *Future Med. Chem.* **4,** 661–679 (2012).
6. Martina, M. Ion channel screening: advances in technologies and analysis. *Front. Pharmacol.* **3,** 86 (2012).
7. Neher, E. & Sakmann, B. Single-channel currents recorded from membrane of denervated frog muscle fibres. *Nature* **260,** 799–802 (1976).
8. *Single-Channel Recording*. (Plenum Press, 1995).
9. Mortensen, M. & Smart, T. G. Single-channel recording of ligand-gated ion channels. *Nat. Protoc.* **2,** 2826–2841 (2007).
10. Chen, C. Y., Tu, T. Y., Chen, C. H., Jong, D. S. & Wo, A. M. Patch clamping on plane glass--fabrication of hourglass aperture and high-yield ion channel recording. *Lab. Chip* **9,** 2370–2380 (2009).
11. Fertig, N., Klau, M., George, M., Blick, R. H. & Behrends, J. C. Activity of single ion channel proteins detected with a planar microstructure. *Appl. Phys. Lett.* **81,** 4865–4867 (2002).
12. Nagarah, J. M. *et al.* Batch Fabrication of High-Performance Planar Patch-Clamp Devices in Quartz. *Adv. Mater.* **22,** 4622–4627 (2010).
13. Stava, E., Yu, M. R., Shin, H. C. & Blick, R. H. Single-Ion Channel Recordings on Quartz Substrates. *Ieee Trans. Nanobioscience* **9,** 307–309 (2010).
14. Stava, E. *et al.* Mechanical actuation of ion channels using a piezoelectric planar patch clamp system. *Lab. Chip* **12,** 80–87 (2012).
15. Van Stiphout, P., Knott, T., Danker, T. & Stett, A. 3D Microfluidic Chip for Automated Patch-Clamping. *Mikrosystemtechnik Kongr.* 435–438 (2005).
16. Boussaoud, A. *et al.* A miniaturized planar patch-clamp system for transportable use. *Biosens. Bioelectron.* **32,** 96–103 (2012).
17. Py, C. *et al.* A Novel Silicon Patch-Clamp Chip Permits High-Fidelity Recording of Ion Channel Activity From Functionally Defined Neurons. *Biotechnol. Bioeng.* **107,** 593–600 (2010).
18. Pandey, S., Mehrotra, R., Wykosky, S. & White, M. H. Characterization of a MEMS BioChip for planar patch-clamp recording. *Solid-State Electron.* **48,** 2061–2066 (2004).
19. Matthews, B. & Judy, J. W. Design and fabrication of a micromachined planar patch-clamp substrate with integrated microfluidics for single-cell measurements. *J. Microelectromechanical Syst.* **15,** 214–222 (2006).

20. Zhang, Z. L. *et al.* Fabrication of Si-based planar type patch clamp biosensor using silicon on insulator substrate. *Thin Solid Films* **516,** 2813–2815 (2008).
21. Sordel, T. *et al.* The development of high quality seals for silicon patch-clamp chips. *Biomaterials* **31,** 7398–7410 (2010).
22. Lehnert, T., Gijs, M. A. M., Netzer, R. & Bischoff, U. Realization of hollow SiO2 micronozzles for electrical measurements on living cells. *Appl. Phys. Lett.* **81,** 5063–5065 (2002).
23. Nagarah, J. M. *et al.* Silicon chip-based patch-clamp electrodes integrated with PDMS microfluidics. *Biophys. J.* **88,** 522A–522A (2005).
24. Weerakoon, P. *et al.* Patch-clamp amplifiers on a chip. *J. Neurosci. Methods* **192,** 187–192 (2010).
25. Schmidt, C., Mayer, M. & Vogel, H. A Chip-Based Biosensor for the Functional Analysis of Single Ion Channels. *Angew. Chem. Int. Ed Engl.* **39,** 3137–3140 (2000).
26. Xu, B. J., Liu, Z. B., Lee, Y. K., Mak, A. & Yang, M. A PDMS microfluidic system with poly(ethylene glycol)/SU-8 based apertures for planar whole cell-patch-clamp recordings. *Sens. Actuators -Phys.* **166,** 219–225 (2011).
27. Klemic, K. G., Klemic, J. F., Reed, M. A. & Sigworth, F. J. Micromolded PDMS planar electrode allows patch clamp electrical recordings from cells. *Biosens. Bioelectron.* **17,** 597–604 (2002).
28. Klemic, K. G., Klemic, J. F. & Sigworth, F. J. An air-molding technique for fabricating PDMS planar patch-clamp electrodes. *Pflugers Arch.-Eur. J. Physiol.* **449,** 564–572 (2005).
29. Lau, A. Y., Hung, P. J., Wu, A. R. & Lee, L. P. Open-access microfluidic patch-clamp array with raised lateral cell trapping sites. *Lab. Chip* **6,** 1510–1515 (2006).
30. Li, X. H., Klemic, K. G., Reed, M. A. & Sigworth, F. J. Microfluidic system for planar patch clamp electrode arrays. *Nano Lett.* **6,** 815–819 (2006).
31. Stett, A., Bucher, V., Burkhardt, C., Weber, U. & Nisch, W. Patch-clamping of primary cardiac cells with micro-openings in polyimide films. *Med. Biol. Eng. Comput.* **41,** 233–240 (2003).
32. Martinez, D. *et al.* High-fidelity patch-clamp recordings from neurons cultured on a polymer microchip. *Biomed. Microdevices* **12,** 977–985 (2010).
33. Tanzi, S. *et al.* Ion channel recordings on an injection-molded polymer chip. *Lab. Chip* **13,** 4784–4793 (2013).
34. Estacion, M. *et al.* Can robots patch-clamp as well as humans? Characterization of a novel sodium channel mutation. *J. Physiol.-Lond.* **588,** 1915–1927 (2010).
35. Wood, C., Williams, C. & Waldron, G. J. Patch clamping by numbers. *Drug Discov. Today* **9,** 434–441 (2004).
36. Lepple-Wienhues, A. *et al.* The Chiptip: A Novel Tool for Automated Patch Clamp. *Comb. Chem. High Throughput Screen.* **12,** 73–77 (2009).
37. Ke, K., Hasselbrink, E. & Hunt, A. J. Nanofabrication with ultrafast lasers at critical intensity. in **5714,** 53–62 (2005).
38. Herbstman, J. F. & Hunt, A. J. High-aspect ratio nanochannel formation by single femtosecond laser pulses. *Opt. Express* **18,** 16840–16848 (2010).

39. Malboubi, M., Gu, Y. & Jiang, K. Experimental and simulation study of the effect of pipette roughness on giga-seal formation in patch clamping. *Microelectron. Eng.* **87,** 778–781 (2010).

40. Kao, L., Abuladze, N., Shao, X. M., McKeegan, K. & Kurtz, I. A new technique for multiple re-use of planar patch clamp chips. *J. Neurosci. Methods* **208,** 205–210 (2012).

41. Uram, J. D., Ke, K. & Mayer, M. Noise and Bandwidth of Current Recordings from Submicrometer Pores and Nanopores. *ACS Nano* **2,** 857–872 (2008).

42. Salkoff, L., Butler, A., Ferreira, G., Santi, C. & Wei, A. High-conductance potassium channels of the SLO family. *Nat. Rev. Neurosci.* **7,** 921–931 (2006).

CHAPTER 3

Real-time shape determination and 5-D fingerprinting of single proteins

This work exploits the zeptoliter sensing volume of electrolyte-filled nanopores to determine, simultaneously and in real time, the shape, volume, charge, rotational diffusion coefficient, and dipole moment of individual proteins. We have developed the theory for a quantitative understanding and analysis of modulations in ionic current that arise from rotational dynamics of single proteins as they move through the electric field inside a nanopore. The resulting multi-parametric information raises the possibility to characterize, identify, and quantify individual proteins and protein complexes in a mixture. This approach interrogates single proteins in solution and determines parameters such as the shape and dipole moment, which are excellent protein descriptors and cannot be obtained otherwise from single protein molecules in solution. Therefore, this five-dimensional characterization of proteins at the single particle level has the potential for instantaneous protein identification, quantification, and possibly sorting with implications for structural biology, proteomics, biomarker detection, and routine protein analysis.

3.1 Introduction

Methods to characterize, identify, and quantify unlabeled, folded proteins in solution on a single molecule level do not currently exist[1]. If feasible, such methods would likely have disruptive impact on the life sciences and clinical assays by simplifying routine protein analysis, enabling rapid biomarker detection[2], and allowing the analysis of personal proteomes[3]. Furthermore, if such methods could instantly provide low-resolution shape information in solution that is complementary to high-resolution methods, they may help to reveal the shape of transient protein complexes or large assemblies that are not accessible by analysis with electron microscopy, NMR spectroscopy, X-ray crystallography, or small-angle X-ray scattering[4]. Here, we demonstrate that interrogation of single proteins or protein-protein complexes during their passage through the electric field inside a nanopore can characterize these particles based on their shape, volume, charge, rotational diffusion coefficient, and dipole moment.

Dipole moment has mostly been neglected as a protein descriptor following the pioneering work by Debye[5] and others[6] because first, neither its usefulness for protein identification nor its importance for concentrated protein solutions has hitherto been widely appreciated and second, measuring protein dipole moments is tedious and limited to ensemble measurements of concentrated, purified protein solutions carried out by specialized laboratories. We propose, however, that dipole moment provides a powerful new dimension in label-free protein analysis, since its magnitude is widely distributed among different proteins (absolute values typically range from 0 to 4,000 Debye)[7]. Dipole moment may therefore approach the usefulness of protein size for identification and would likely exceed the usefulness of protein charge (whose values typically range

from -40 to +40 times e)[7]. Moreover, the pharmaceutical industry is increasingly recognizing the importance of dipole moments for antibody formulations[8], in part because subcutaneous injection of highly concentrated solutions of monoclonal antibodies (the fastest growing class of therapeutics) can be impractical due to high viscosity and aggregation resulting from dipole alignment[8]. Measurements of antibody dipole moments could therefore provide a criterion to select early candidates in the drug discovery process and reduce development costs[9]. Additionally, certain protein families, such as DNA-binding proteins and acetycholinesterases from various species[10], share conserved dipole moments that are important for their function. Therefore, measuring dipole moments may allow grouping of proteins by functional criteria in contrast to physical descriptors such as size and isoelectric point, which typically have no correlation with function. Finally, real-time measurements of dipole moments (and shapes) of single proteins may be useful for detecting phosphorylation and ligand binding[11] as well as conformational changes during protein activation or folding[7].

Interrogating single protein particles during their passage through a pore is simple in principle[12-16]. It requires a single electrolyte-filled pore that connects two solutions across a thin insulating membrane and serves as a conduit for ions and proteins (Fig. 3.1a). Electrodes connect the solutions on either side of the membrane to a high-gain amplifier that applies a constant electric potential difference while measuring the ionic current through the nanopore. This arrangement ensures that virtually the entire voltage drop occurs within the constriction of the pore, rendering this zone supremely sensitive to transient changes in its ionic conductivity. Consequently, each protein that is driven electrophoretically through the pore displaces conductive electrolyte, distorts the electric
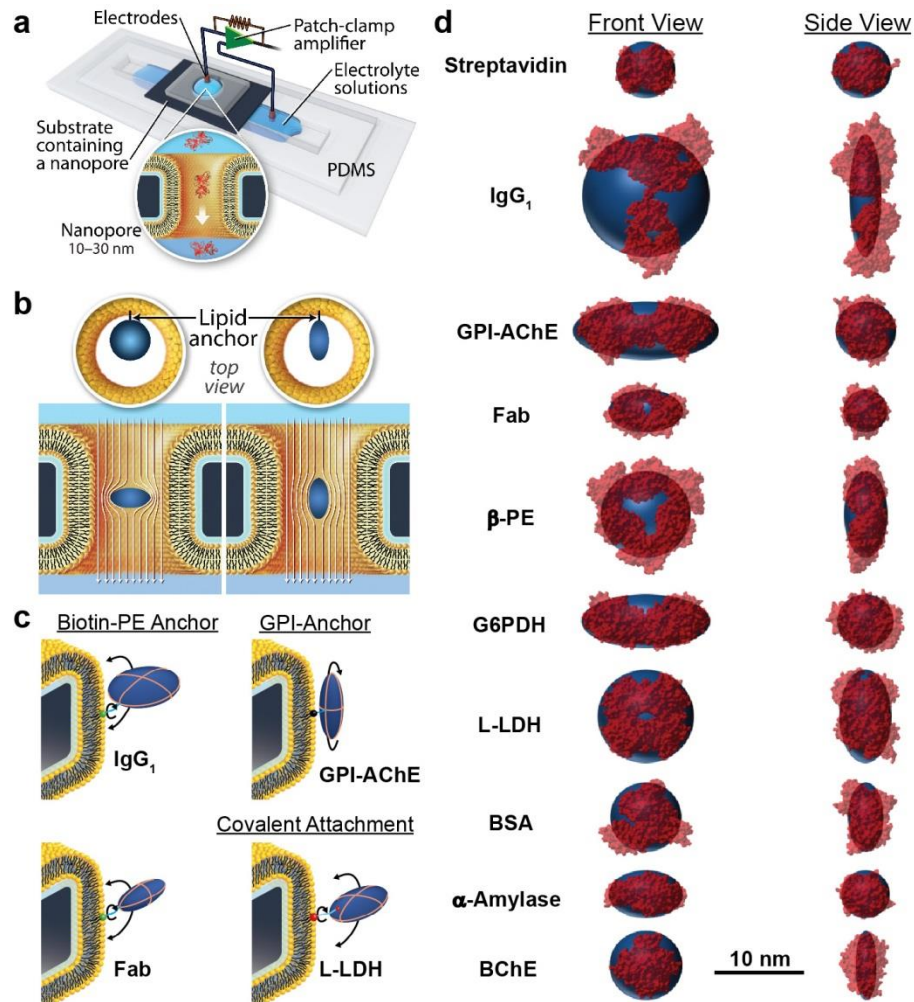
Figure 3.1 | Current recordings through nanopores reveal the shape of single proteins as they translocate through the pore. a) Experimental setup to measure resistive-pulses from the translocation of individual proteins. b) Top and side views of a nanopore illustrating the two extreme orientations of an ellipsoidal protein that is anchored to a fluid lipid coating on the pore wall. A crosswise orientation disturbs the field lines inside the pore more than a lengthwise orientation. c) Three different strategies of anchoring proteins to the lipid coating were used to slow down translocation such that rotational diffusion of the proteins could be resolved in time. A lipid anchor with a biotin group selectively captured anti-biotin antibodies and Fab fragments, an intrinsic GPI anchor captured acetylcholinesterase, and a bi-functional, amine-reactive crosslinker provided a general strategy to attach proteins of interest covalently to ethanolamine lipids in the bilayer coating. All proteins analyzed in this work were tethered with a phospholipid anchor to the bilayer by one of these three strategies. These tethers were sufficiently long ($\geq$ 1.5 nm in their extended conformation) and flexible ($\geq$ 12 σ-bonds) and nanopore diameters were at least twice the volume-equivalent spherical diameter of the examined proteins, such that the proteins were able to rotate and sample all possible orientations. d) Comparison of the shape of ten proteins as determined by analysis of resistive-pulses (blue ellipsoids) with crystal structures from the Protein Data Bank in red (streptavidin: 3RY1, anti-biotin immunoglobulin G_1: 1HZH, GPI-anchored acetylcholinesterase: 3LII, anti-biotin Fab fragment: 1F8T, β-phycoerythrin: 3V57, glucose-6-phosphate dehydrogenase: 4EM5, L-lactate dehydrogenase: 2ZQY, bovine serum albumin: 3V03, α-amylase: 1BLI, and butyrylcholinesterase: 1P0I).

field, and reduces the ionic current through the pore. If the volume of the electrolyte-filled pore is sufficiently small compared to the volume of the particle, then the change in ionic current due to the translocating particle is measurable and characterized by its magnitude, ΔI, and duration, t_d[17-20]. In addition to its exquisite sensitivity to conductivity changes, this small volume transiently separates single proteins from other macromolecules in solution making it possible to interrogate the rotational dynamics of one protein without artifacts from other macromolecules. For this reason, time-dependent modulations of ionic current as a single protein passes through a nanopore can, under appropriate conditions, relate uniquely to the time-dependent molecular orientation of that protein as well as its shape, volume, charge, rotational diffusion coefficient, and dipole moment (Fig. 3.2, Section 3-App.S1-3, and Fig. 3-App.1-9). Several groups have recently considered in qualitative terms the effect of a protein's shape when analyzing distributed ΔI signals[18,20-24] as well as the effect of a protein's dipole moment on its translocation through an alpha-hemolysin pore in the presence of an AC field[25]. The work presented here develops the theory for a *quantitative understanding* of the dependence of measured ΔI values on the shape and rotational dynamics of proteins. We show that the introduction of this theoretical model makes it possible to exploit rotational dynamics for determining the volume, shape, rotational diffusion coefficient, and dipole moment of non-spherical proteins.

Measuring five parameters of a single, unlabeled protein simultaneously has previously not been feasible, and the possibility to analyze single molecules and hence dilute samples circumvents artifacts encountered in concentrated protein solutions. For instance, established techniques to determine the dipole moment of proteins suffer from
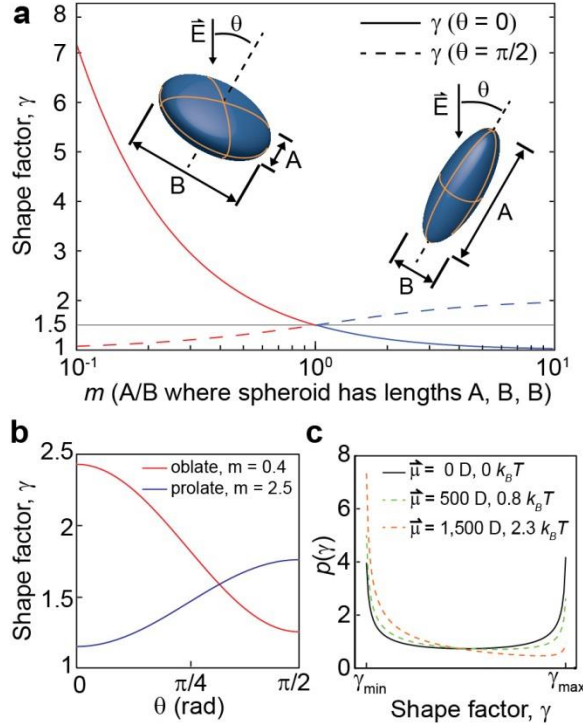
Figure 3.2 | Possible values of electrical shape factors, γ, for ellipsoids of rotation and their probability distribution. a) Shape factor of ellipsoids (prolates in blue curves and oblates in red curves) as a function of their aspect ratio, m, for two extreme orientations: when the angle, θ, between the axis of rotation of the ellipsoid relative to the electric field is 0, i.e. $\theta = 0$ (solid curves), and when $\theta = \pi/2$ (dashed curves). For reference, a sphere has a m value equal to 1, and a shape factor of 1.5 that is independent of its angle θ (grey line). b) Shape factor as a function of θ for prolates with a defined m value of 2.5 and oblates with an m value of 0.4. c) Bimodal probability distribution of shape factors, $p(\gamma)$, for ellipsoids without a dipole moment as predicted by Golibersuch (black curve) and for ellipsoidal proteins with a dipole moment of 500 and 1500 Debye) pointed parallel to the longest axis of the protein (dashed curves). For the different magnitudes of the dipole moment, the energy difference between $\theta = 0$ and $\theta = \pi/2$ is listed in units of k_bT for a typical electric field of 2×10^6 V m^{-1}. See Section 3-App.S2 and Section 3-App.S.9 for details.

the need for concentrated and pure protein samples and from the associated challenges of high viscosity and aggregation[8,26,27]. Techniques for determining the charge of native proteins suffer from possible artifacts due to interactions with capillary walls in the case of capillary electrophoresis or with chemical groups on hydrogels in the case of gel electrophoresis (the latter is also limited by semi-quantitative analysis and extended analysis times)[2]. Techniques for determining the shape of proteins (e.g. cryo-electron microscopy, atomic force microscopy, or small-angle X-ray scattering) are slow, carried

out off-line[4], and may distort the shape of proteins during crystallization or surface adsorption[28]. Techniques for determining the rotational diffusion coefficient of single proteins typically monitor the motion of a fluorescent tag as opposed to that of the protein itself[29]. Finally, most methods for determining the volume of proteins cannot be performed on a single molecule; rather, these methods report ensemble averages and, in the case of solution-based techniques such as dynamic light scattering, are not well suited for characterizing non-spherical shapes[30]. Hence, the ability to measure five parameters simultaneously on single proteins in real time (see Fig. 3-App.10) based on their dynamics in an electric field has fundamental implications. For instance, analyzing individual proteins one-by-one may mean inherently that these proteins do not have to be purified for determining their shape or the other four parameters. 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.

In order to measure several protein characteristics on a single-molecule level with nanopores as demonstrated here, a set of enabling breakthroughs had to occur. These include the possibility to: (i) engineer single nanopores with molecular scale interrogation volumes[31]; (ii) measure low-noise currents with microsecond time resolution[32,33]; (iii) eliminate non-specific interactions of proteins with the pore walls[19,20]; (iv) slow down protein translocation to enable adequate data collection from a single translocation through a pore that is only five times longer than the protein[20,34]; and (v) understand the effect of an object's volume, shape, and orientation on the current through a nanopore based on fundamental effects of ellipsoids in an electric field[35-38] combined with the

development of the model and algorithm introduced here to interpret the noise-convoluted current signal[35,36,39].

3.2 Results

3.2.1 Ionic currents through a nanopore reveal the size and shape of proteins

Fig. 1 shows the low-resolution shape of ten different proteins determined from nanopore recordings compared to the respective crystal structure for each protein. We obtained these shapes from analysis of distributions of maximum ΔI values from hundreds of individual protein translocations through a nanopore (Fig. 3.1; see Section 3-App.S2 and Fig. 3-App.5-8 for details). For spherical proteins, these ΔI distributions are Normal (Fig 3.3c); however, for non-spherical proteins, these distributions deviate from a Normal distribution because the rotation of non-spherical proteins modulates the ionic current in the nanopore as a function of the protein's orientation with respect to the long axis of the pore (Fig. 3.1b and Fig. 3.2). For instance, an oblate (*i.e.*, a lentil-shaped) protein perturbs field lines and hence reduces the ionic current more strongly when its flat side is oriented perpendicular to the luminal axis of the pore than when it is oriented parallel to it (Fig. 3.1b and Fig. 3.2). Similarly, a prolate (*i.e.*, a rugby ball-shaped) protein perturbs field lines more strongly when its long side is oriented crosswise to the luminal axis of the pore than when it is oriented parallel to it (Fig. 3.1b). Fricke and later Velick and Gorin described these shape- and orientation-dependent effects of ellipsoids on the electrical field lines analytically with a so-called electrical shape factor, γ,[40-42] and Golibersuch described the probability distribution of shape factors assuming all orientations are equally probable (black curve in Fig. 3.2c; see Section 3-App.S2 for
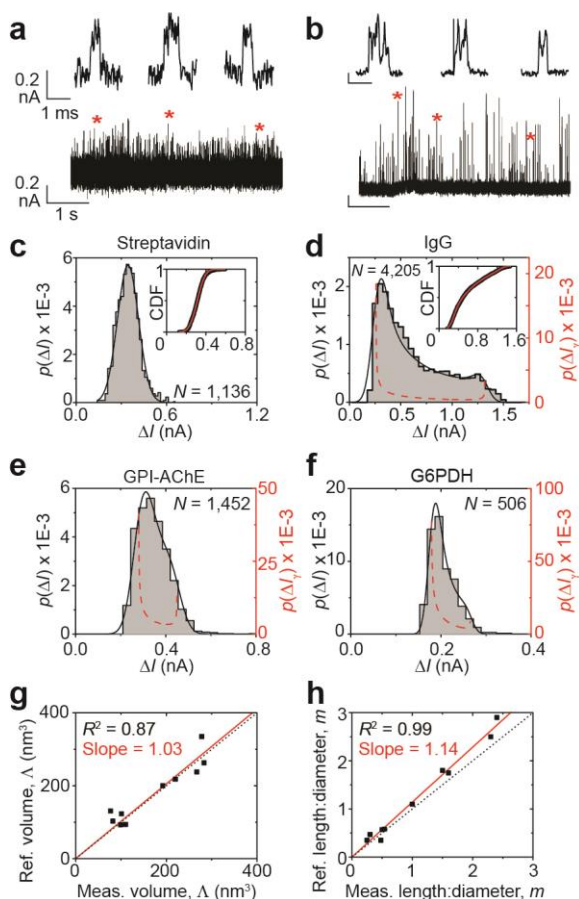
Figure 3.3 | Determination of protein shape and volume from histograms of maximum ΔI values from resistive-pulse recordings. a,b) Examples of original current traces as a function of time: upward spikes indicate individual resistive current pulses towards zero current due to the translocation of single streptavidin (a) or IgG (b) proteins. Resistive-pulses marked by an asterisk are shown in detail above. c-f) Histograms of maximum ΔI values from resistive-pulse recordings with streptavidin (c), IgG$_1$ (d), GPI-AChE (e), and G6PDH (f) proteins. Black curves show the solution of the convolution model, $p(\Delta I)$, after a non-linear least squares fitting procedure, and red dashed curves show the estimated distribution of ΔI values due to the distribution of shape factors, $p(\Delta I_\gamma)$. Table 3-App.1 lists the values of all fitting parameters and the electric field strength used in each experiment. Section 3-App.S2 and Fig. 3-App.5-7 explain the convolution model and fitting procedure in detail and extend the analyses to all proteins characterized in this work. (**g**) Comparison of the measured volume by nanopore-based analysis with the expected reference volume. (**h**) Comparison of the measured length-to-diameter ratios m of all proteins with the expected reference values of m.

details). We use this dependence of γ on shape and orientation to determine the shape and

volume of single proteins. Two proteins with the same volume but different shapes result

in different minimal and maximal ΔI values (see Section 3-App.S2) although their

average ΔI value may be the same (compare, for instance, the ellipsoids determined for

the IgG antibody and the acetylcholinesterase proteins in Fig. 3.1d). This shape and orientation dependence of ΔI values has general consequences for nanopore-based protein characterization: Since volume determination is typically based on maximum ΔI values, it can only be accurate if the shape of non-spherical proteins is considered in the analysis.

In order to obtain time-resolved values of ΔI from the translocation of single proteins, we slowed down their translocation by tethering them to a lipid anchor that was embedded in the fluid lipid bilayer coating inside the nanopores (Fig. 3.1b; see Section 3-App.S4 for a detailed discussion on the effects of lipid tethering)[20,24,34]. 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. Importantly, however, rotation of the tethered proteins occurred within the low-viscosity aqueous environment, and we maximized the possibility that the proteins could sample all orientations in the nanopore by employing long and flexible tethers (Fig. 3.1c). The lipid coating also minimized non-specific interactions of proteins with the pore wall[20], thus enabling to extract quantitative data on Brownian rotational and translational dynamics of proteins while they are in the pore[43]. 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-App.11).

Determining the shape and volume of ellipsoids of rotation from protein translocations through nanopores proceeds in three steps; Fig. 3.3 shows the results from each step (see Section 3-App.S2 and Fig. 3-App.5-7). First, an algorithm detects resistive-pulses from the translocations of hundreds to thousands of proteins and

determines the amplitude of the current modulation from the baseline current, ΔI (Fig. 3.3a,b). Second, the resulting distribution of ΔI values is converted to an empirical cumulative density function, CDF, and fit iteratively with a theoretical distribution developed in this work and referred to as the "convolution model"[39] (Fig. 3.3c,d; see Section 3-App.S2 and Fig. 3-App.5-6). This approach derives the theoretical distribution by convolving the expected bimodal distribution of ΔI values that results from the variation in the electrical shape factor (Fig. 3.2) with a Gaussian distribution that describes the electrical current noise (Fig. 3-App.5). The minimal and maximal ΔI values returned from this fitting procedure reflect the two extreme orientations of proteins: fully cross-wise or fully lengthwise (red dashed curves in Fig. 3.3d-f). Third, these extreme ΔI values are converted to a maximal and minimal shape factor γ. The combination of these extreme ΔI values and shape factor values yields the shapes and volumes of ellipsoids of rotation with length-to-diameter ratios, m, that agree best with the experimental distribution of ΔI values from each protein. Fig. 3.1d shows the results from this analysis for ten proteins with regard to their shape, and Fig. 3.3g,h shows that the volume and m values agree well with the expected reference values; the average deviation of both the volumes and m values is less than 20% (Tables 3-App.1-4 list the results of this analysis and provide values from independent experiments for comparison). Independent from these experimental results, we confirmed the accuracy of this algorithm for shape and volume determination on simulated data that were generated from the theory of biased one-dimensional Brownian diffusion and convoluted with current noise. Analysis of these simulated data returned values of shape and volume that were in excellent agreement with the input parameters (Section 3-App.S5 and Fig. 3-App.12).

3.2.2 Rotation of a single protein modulates the ionic current through a nanopore

In addition to analyzing the distribution of maximum ΔI values from hundreds of resistive-pulses, Fig. 3.4 shows that it is possible to determine the shape of proteins from individual translocation events and therefore from a single protein in real time. We restricted the analysis to resistive-pulses with durations of at least 400 μs to ensure that, on average, each protein samples the full range of electrical shape factors. Based on the mean-square-angular displacement equation, we expect that a protein will sample all possible orientations in less than 400 μs as long as it has a rotational diffusion coefficient of at least 3,000 rad^2 s^{-1}. This threshold value of D_R is about 30% smaller than the minimum rotational diffusion coefficient determined in this work. In this case, all recorded values of the electrical current within each individual resistive-pulse are analyzed; for example, Fig. 3.4b shows an empirical probability density distribution of ΔI values from all of the sampled values of the electrical current during a single resistive-pulse (see Section 3-App.S6 and Fig. 3-App.15-19). These distributions of single event (or intra-event) ΔI values are analyzed in the same way as the distributions of maximum ΔI values from hundreds of resistive-pulses. We find that the intra-event ΔI distributions retain their key features (*e.g.*, minimal and maximal ΔI values) even though the current recordings are smoothed due to filtering (see Fig. 3-App.21). This intra-event analysis has the additional benefit that it can determine the dipole moment and rotational diffusion coefficient of single proteins by relating time-dependent changes in current to time-dependent changes in the shape factor, γ, which originate from rotations of single proteins during their translocation through the nanopore. To estimate the dipole moment, μ, we characterized the bias in each protein's orientation under the influence of the

63
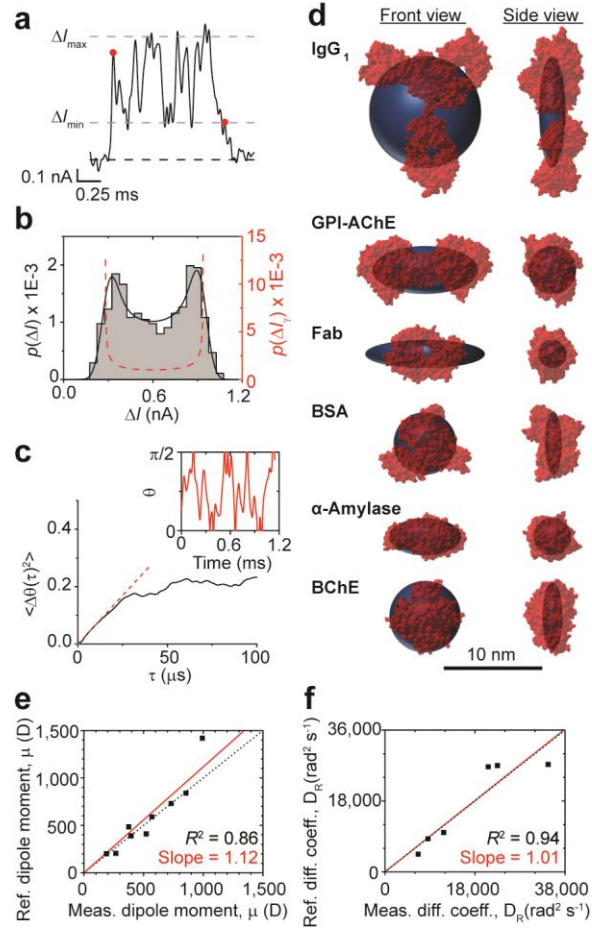
Figure 3.4 | Shape, dipole moment, and rotational diffusion coefficient obtained from current modulations *within individual* resistive-pulses from the translocation of a single protein. a) Resistive-pulse from the translocation of a single IgG$_1$ molecule. Red dots mark the beginning and end of the resistive-pulse as identified by an automated algorithm. b) Distribution of all current values within this one resistive-pulse. The black curve shows the solution of the convolution model, $p(\Delta I)$, after a non-linear least squares fitting procedure, and the red dashed curve shows the estimated distribution of ΔI values due to the distribution of shape factors, $p(\Delta I_\gamma)$. c) Mean-squared-angular displacement curve (black trace) and the initial slope (dashed red line). The inset shows the transformation of intra-event $\Delta I(t)$ to $\theta(t)$. d) Comparison of the shape of proteins as determined by analysis of individual resistive-pulses (blue) with crystal structures in red (blue ellipsoids show the median values of m and volume from single event analyses of each protein; see Fig. 3-App.15 for complete distributions from the single event analyses). e) The most frequently observed dipole moments of G6PDH, L-LDH, α-amylase, β-phycoerythrin, BSA, Fab, GPI-AChE, IgG$_1$, and BChE agree well with expected reference values of their dipole moments. f) The most frequently observed rotational diffusion coefficients of IgG$_1$, β-phycoerythrin, GPI-AChE, BChE, Fab, and α-amylase agree well with the expected reference values. The signal-to-noise ratio for G6PDH, L-LDH, and BSA was too small to determine accurate values of D_R.

electric field[5,6] within the nanopore by fitting the cumulative ΔI distribution to a model that considers the energy difference of a dipole rotating in an electric field (Fig. 3.4b; see Section 3-App.S6 and Fig. 3-App.16-17). To estimate the rotational diffusion coefficient,

D_R, we transformed the time-dependent intra-event current signal into a time-dependent change in the angle of the protein over short time scales (Sections 3-App.S6 and 7 and Fig. 3-App.18-20), plotted the mean-square-angular displacement during a time interval, τ, and fit with a 1-D model for rotational diffusion (Fig. 3.4c). Figure 3.4d shows that the median protein shapes determined from this intra-event analysis are in reasonable agreement with their crystal structure, and Fig. 3.4e,f shows that the most probable values of dipole moment and D_R from this nanopore-based analysis agree well with expected reference values; the average deviation was less than 25% for both parameters. The analysis on simulated intra-event data again returned values of the determined shape, volume, dipole moment, and rotational diffusion coefficient that were in excellent agreement with the input parameters for the simulation (Fig. 3-App.13 and 3-App.14).

3.2.3. Multiparameter-characterization of individual proteins improves protein classification

To assess the potential of nanopore-based identification and characterization of different proteins in a mixture, we repeated the characterization of glucose-6-phosphate dehydrogenase described in Fig. 3.3f and added an anti-G6PDH IgG antibody. Thus, in the same experiment, single proteins of G6PDH and protein-protein complexes of G6PDH-IgG were passing through the nanopore. Analysis of current modulations within each translocation event returned an estimate of the volume, shape, charge, rotational diffusion coefficient, and dipole moment for each particle passing through the pore. Fig. 3.5 shows that this multiparameter-fingerprinting approach made it possible to distinguish G6PDH from the G6PDH-IgG complex by using a clustering algorithm to
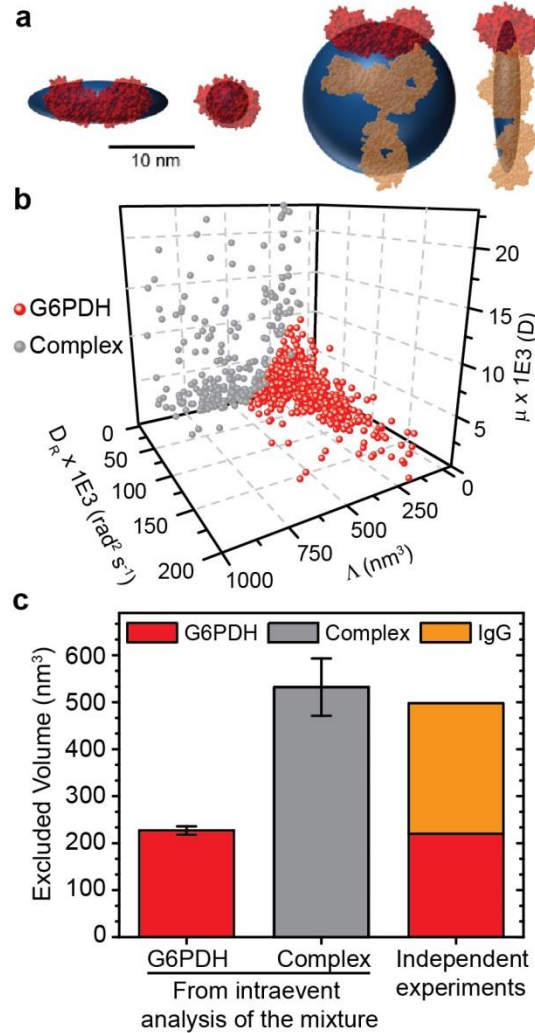
Figure 3.5 | Fingerprinting of individual translocation events permits identification and characterization of G6PDH and a G6PDH-IgG complex from a mixture. a) The volume, Λ, and shape of G6PDH (left side) and G6PDH-IgG complex (right side) as determined by analysis of individual resistive-pulses is similar to the crystal structures in red. Blue ellipsoids show the median values of *m* and Λ determined from single-event analyses and classification of each event. b) Values for the volume, rotational diffusion coefficient, and dipole moment determined for each event. The *kmeans* clustering algorithm in MATLAB classified each event as corresponding to a single G6PDH (red points) or to the G6PDH-IgG complex (grey points) (see Section 3-App.S8 and Fig. 3-App.22). This single event classification estimated that 28% of events were due to the complex, which is nearly the same proportion of events estimated to be in the complex based on analysis of maximum ΔI values from distributions of hundreds of resistive-pulses (Fig. 3-App.22). c) The volume of G6PDH and the G6PDH-IgG complex determined by single-event analysis and classification of events from the mixture are nearly identical (< 10% deviation) to the volumes obtained for G6PDH in an experiment without anti-G6PDH IgG and the volume of an individual IgG. Error bars represent the standard error of the median volume value (Section 3-App.S8).

66

classify each translocation event (Fig. 3.5b; see Section 3-App.S8 and Fig. 3-App.22 for details)[24,44]. This analysis returned excellent estimates of the size and shape of G6PDH and the G6PDH-IgG complex (Fig. 3.5a and 3.5c). In contrast, employing the current standard practice of distinguishing proteins by the ΔI values and translocation times of each resistive-pulse[45,46] underestimated the amount of the G6PDH-IgG complex formed by 90% and overestimated its volume by 70% (Section 3-App.S8). Figure 3.5b also confirms several expectations with regard to the difference between G6PDH and its complex with IgG. For instance, individual resistive-pulses assigned to the complex correspond to significantly larger molecular volumes and smaller rotational diffusion coefficients than resistive-pulses assigned to G6PDH by itself. In addition, the dipole moment of G6PDH is relatively clustered as expected for a protein with well-defined shape and position of amino acids. In contrast, the dipole moment of the complex between G6PDH and the polyclonal anti-G6PDH IgG antibody varies widely as expected for a complex that involves a protein antigen with multiple binding sites and binding of a relatively floppy IgG molecule. This analysis provides proof-of-principle for nanopore-based characterization, identification, and quantification at the single protein level and demonstrates the advantage of simultaneous multiparameter characterization for identifying individual proteins or protein-protein complexes over single-variate or bi-variate characterization.

These first results also raise the question, what benefit is gained by determining additional descriptors for distinguishing individual molecules in a mixture of hundreds of different proteins. Fig. 3.6 takes a bioinformatics-based approach to address this question. Every pixel in this plot represents the normalized distance between one protein-protein
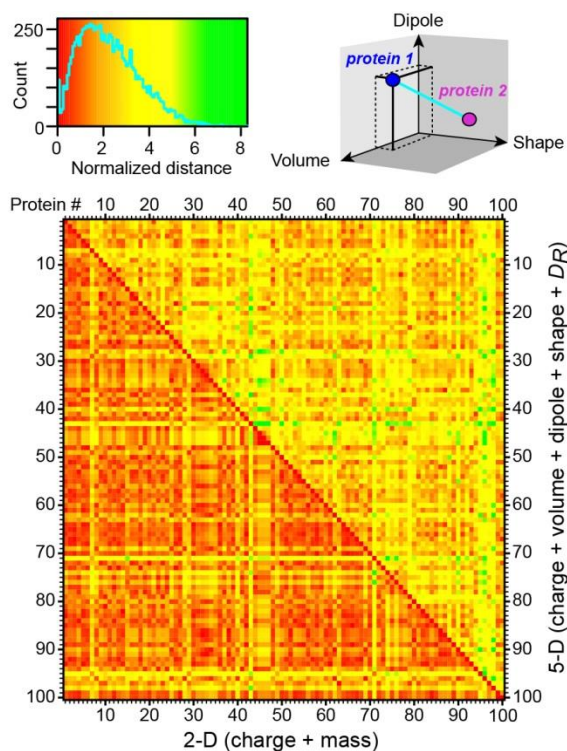
Figure 3.6 | Advantage of 5-D fingerprinting over the standard 2-D characterization for protein identification. Using structural and sequence data from the Protein Data Bank, we randomly selected a group of proteins and determined their mass, volume, rotational diffusion constant, shape factor, dipole moment, and charge. Each parameter can be thought of as a dimension, and the heat map shows the separation between each pair of 100 randomly sampled proteins for two dimensions (lower left corner) or five dimensions (upper right corner) calculated using standard normal distributions for each descriptor. This separation is calculated as $\sqrt{\sum_{i=1}^{n} d_i^2}$ where n is the number of dimensions and d_i is the difference between the values of two different proteins in one parameter. Red squares mark protein-protein pairs that are similar in all descriptors (*i.e.*, closely spaced), while yellow and green squares indicate increasing separation. Physical descriptors beyond protein charge and mass such as shape and dipole moment create additional dimensions and facilitate protein identification by increasing the separation between each protein-protein pair.

pair in either two or five dimensions. The normalized distances between most protein pairs shift from less than one standard deviation in the two dimensional analysis (lower left corner of the plot) to more than three standard deviations in the five dimensional analysis (upper right corner). The graph therefore illustrates that additional descriptors of proteins beyond the oft-employed protein size and charge make it significantly easier to distinguish proteins from each other. Another question is which protein descriptors are

most useful for distinguishing proteins from each other. Ideal descriptors are not correlated with each other and therefore provide orthogonal distinguishing power. Analysis of 780 randomly sampled proteins from the Protein Data Bank revealed that mass, volume, and rotational diffusion constant of proteins are strongly correlated with each other (see Fig. 3-App.23 and 3-App.24), while protein size (*i.e.*, mass or volume) did not correlate strongly with protein charge, shape factor *m*, or dipole moment. Protein charge spanned a range from -40 to +40 $\times$ *e* with a majority between -10 and +10 $\times$ *e* and is therefore a somewhat degenerate descriptor. In contrast, dipole moment and the length-to-diameter ratio, *m* – the descriptors made accessible on a single-molecule level by the work introduced here – are both widely distributed. Hence, dipole moment and protein shape are compelling candidates for protein identification by multidimensional fingerprinting.

3.3 Discussion

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 isolates 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, which minimizes non-specific interactions and slows

down translocations of lipid-anchored proteins, enables studying the rotational and translational dynamics of single proteins long enough in time to determine their shape, volume, charge, rotational diffusion coefficient, and dipole moment.

Based on the spectacular progress in nanopore-based DNA sequencing in the last 17 years[3,47-49], we predict that improvements of the approach introduced here will increase the potential of nanopore-based protein characterization[50]. For instance, the approach to single event (intra-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. Novel fabrication methods such as He-ion beam fabrication produce pores that are almost perfectly cylindrical and should therefore minimize possible artifacts from this source of error[51]. In addition, the recent development of integrated CMOS current amplifiers[33], 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[33], 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[34] while their improved signal to noise ratio will likely reduce the errors of each determined parameter[32,33]. Furthermore, computational approaches that can model proteins with shapes more complex than simple ellipsoids may increase the resolution of shape determination, while the capability to monitor current modulations with MHz bandwidths[33,34] opens up the

possibility to follow transient changes in protein conformation and folding as well as to determine the shape of short-lived protein complexes whose structure and dynamics are not 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[2]. Despite its tremendous capabilities, mass spectrometry has currently limited throughput and is not broadly applicable to meet demand for routine protein analysis[1,2]. 2-D 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[52]. We propose that multi-dimensional analysis and fingerprinting of single proteins in nanoscale volumes may be one alternative. The work presented here is a first step in this direction, and 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 while providing additional protein descriptors, improved quantification, increased sensitivity, reduced analysis time (*i.e.*, in real time) and therefore lower cost. Such a capability may ultimately make it feasible to characterize and monitor an individual's proteome with significant implications for personalized medicine[1]. Protein characterization on a single-molecule level may also reveal biochemically- or clinically-relevant heterogeneities, such as small fractions of phosphorylated proteins, that are often

hidden in ensemble measurements. Moreover, real-time identification of single proteins might ultimately enable single-molecule sorting in a fashion analogous to cell sorting.

Finally, this report focused on one of the most relevant and challenging applications of nanoscale shape determination, namely the characterization of single proteins. The same approach may, however, apply to particles such as DNA origami[53], synthetic nanoparticles[54,55], and nanoparticle assemblies[56], whose characterization is important since they are typically more heterogeneous than proteins and since their charge, shape, volume, and dipole moment affects their assembly characteristics and function[57-59].

3.4 Methods

3.4.1 Materials

All phospholipids were obtained from Avanti Polar Lipids. Bis(succinimidyl) penta(ethylene glycol) (21581) was purchased from Thermo Scientific. Monoclonal anti-biotin IgG$_1$ (B7653), GPI-anchored acetylcholinesterase (C0663), glucose-6-phosphate dehydrogenase (G5885), L-lactate dehydrogenase (59747), bovine serum albumin (A7638), α-amylase (A4551), and streptavidin were purchased from Sigma Aldrich, Inc. Polyclonal anti-biotin IgG-Fab fragments (800-101-098) were purchased from Rockland and β-phycoerythrin (P-800) was purchased from Life Technologies.

3.4.2 Methods of Nanopore-Based Sensing Experiments

To sense proteins, we first formed a supported lipid bilayer containing either 0.15 mol% 1,2-dipalmitoyl-*sn*-glycero-3-phosphoethanolamine-N-capbiotinyl (biotin-PE)

lipids or 1 mol% 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphoethanolamine (POPE) lipid in a background of 1-palmitoyl-2-oleoyl-*sn*-glycero-3-phosphocholine (POPC) lipids (Avanti Polar Lipids, Inc.). We described details of the bilayer formation in Yusko et al.[20] The dimensions of all nanopores are shown in Fig. 3-App.25. When biotin-PE lipids were present in the bilayer, we added a solution containing anti-biotin IgG_1, Fab, or GPI-anchored acetylcholinesterase to the top solution compartment of the fluidic setup such that the final concentration of protein ranged from 5 pM to 10 nM. When sensing GPI-anchored acetylcholinesterase, we started recording resistive-pulses after incubating the bilayer-coated nanopore for 1 h with GPI-anchored acetylcholinesterase (where the solution was 150 mM KCl, 10 mM HEPES, pH = 7.4) to allow time for the GPI-lipid anchor of the protein to insert into the fluid lipid bilayer coating. When POPE lipids were present in the bilayer, we first dissolved bis(succinimidyl) penta(ethylene glycol), a bifunctional crosslinker, in a buffer containing 2 M KCl and 100 mM $KHCO_3$ (pH = 8.4) and immediately added this solution to the top compartment of the fluidic setup such that the final concentration of crosslinker was 10 mg/mL. After 10 min, we rinsed away excess crosslinker and subsequently added β-phycoerythrin, glucose-6-phosphate dehydrogenase, L-lactate dehydrogenase, bovine serum albumin, α-amylase, or butyrylcholinesterase dissolved in the same buffer as the preceding step to the top compartment such that final protein concentration ranged from 1 to 3 μM. After at least 30 minutes, we rinsed away excess protein and began recording. We recorded resistive-pulses at an applied potential difference of -0.04 to -0.115 V with the polarity referring to the top fluid compartment relative to the bottom fluid compartment, which was connected to ground. The electrolyte contained 2 M KCl with either 10 mM HEPES at pH 6.5 for

experiments with GPI-anchored acetylcholinesterase; 10 mM HEPES at pH 7.4 for experiments with IgG, Fab, α-amylase, butyrylcholinesterase, and streptavidin; 10 mM $C_6H_7KO_7$ at pH 5.1 for experiments with β-phycoerythrin; 10 mM $C_6H_7KO_7$ at pH 5.2 for experiments with bovine serum albumin; or 10 mM $C_6H_7KO_7$ at pH 6.1 for experiments with glucose-6-phosphate dehydrogenase and L-lactate dehydrogenase. We used Ag/AgCl pellet electrodes (Warner Instruments) to monitor ionic currents through electrolyte-filled nanopores with a patch-clamp amplifier (Axopatch 200B, Molecular Devices Inc.) in voltage-clamp mode (*i.e.*, at constant applied voltage). We set the analog low-pass filter of the amplifier to a cutoff frequency of 100 kHz. We used a digitizer (Digidata 1322) with a sampling frequency of 500 kHz in combination with a program written in LabView to acquire and store data[32]. To distinguish resistive-pulses reliably from the electrical noise, we first filtered the data digitally with a Gaussian low-pass filter (f_c =15 kHz) in MATLAB and then used a modified form of the custom written MATLAB routine described in Pedone *et al.*[60,61]. We calculated the translocation time, t_d, as the width of individual resistive-pulse at half of their peak amplitude, also known as the full-width-half-maximum value[20,61]. From this analysis we obtained the ΔI and t_d values for each resistive-pulse, and we only analyzed ΔI values for resistive-pulses with t_d values greater than 50 µs, since resistive-pulses with translocation times faster than 50 µs have attenuated ΔI values due to the low-pass filter[20,60].

3-App.S1 Control experiments indicate that broad distributions of ΔI values were not due to impurities or simultaneous translocations

To confirm that the distributions of ΔI values during experiments with monoclonal anti-biotin IgG_1 antibodies were not affected by potential impurities in the solution, we performed three control experiments. In one control experiment, we competitively inhibited the binding of IgG_1 antibodies to the biotin-PE lipids on the surface by adding an excess concentration of soluble biotin to the aqueous solution of an ongoing experiment (Fig. 3-App.2a and 3-App.2b). Fifteen minutes after the addition of the soluble biotin we observed the frequency of resistive pulses decrease from 34 s^{-1} to 1.3 s^{-1}. In the second control experiment, we generated a lipid bilayer coated nanopore that did not contain biotin-PE lipids in the coating and therefore was not specific for the translocation of IgG_1 antibodies (Fig. 3-App.2c). In this experiment, the concentration of the IgG_1 antibody was even higher (25 nM compared to 20 nM) than in the original experiment (Fig. 3-App.2a), and the frequency of translocation events was 2 s^{-1}. Since the frequency of events is proportional to concentration, we estimated that if the concentration of IgG_1 in this control experiment was 20 nM, we would expect to observe an event frequency of approximately 1.6 s^{-1}. From these two control experiments, we estimated that during experiments with biotin-PE lipids in the bilayer coating only 3.8 to 4.7% of translocation events were due to proteins that were not bound to biotin-PE lipids. Furthermore, almost all (~90%) of the translocation times calculated from resistive-pulses observed in control experiments (where binding to biotin-PE was not possible) were less than 50 μs, and we did not include resistive-pulses with translocation times less

than 50 µs in the analysis of ΔI distributions because the amplitude would be attenuated due to electronic filtering[20,32]. Consequently, we concluded that the protein we detected in the solution of anti-biotin IgG_1 antibodies was bound to biotin-PE lipids specifically. In the final control experiment, we removed any fragments of IgG proteins (*i.e.*, Fab fragments) or other proteins in the IgG stock solution by purifying the solution with a Protein A spin column (Thermo Scientific 89952). Using this purified solution in a nanopore-based sensing experiment, we observed distributions of ΔI values similar to those seen with the stock solution, and we determined the same volume and shape of anti-biotin IgG_1, within error (Table 3-App.1). Results from these three control experiments indicate that the resistive pulses in experiments with IgG_1 were due to anti-biotin IgG_1 proteins and not due to the presence of other proteins or Fab fragments.

Since IgG antibodies can occasionally form dimers[62], we performed dynamic light scattering (DLS) experiments to characterize the hydrodynamic diameter of the IgG_1 antibodies. If dimers of IgG_1 antibodies were present in solution and contributing to the bimodal distribution of ΔI values in Fig. 3.3 of the main text, we would expect them to be reflected in DLS experiments in a significant fraction because approximately ½ of the resistive pulses had ΔI values that can be attributed to either of the bimodal peaks in the distribution of ΔI values. Consequently, if dimers were present, we would expect to observe two peaks in the distributions of estimated hydrodynamic diameters of the particles (proteins in this case) in DLS experiments[62]. Fig. 3-App.2d shows that we only observed one peak corresponding to a hydrodynamic diameter of 10.5 ± 2.0 nm. This value is in good agreement with previously published hydrodynamic diameters of IgG antibodies of 10.9 to 11.0 nm, which were determined in physiologic buffers[62,63]. As

additional evidence, we added urea to a concentration of 8 M to denature the IgG protein and disassociate potential aggregates. Again we only observed one peak corresponding to a hydrodynamic diameter of 12.9 ± 2.7 nm (Fig. 3-App.2d). This hydrodynamic diameter is larger because of the random-coil and ball-like structure of denatured IgG$_1$ antibodies compared to their native, oblate-shaped structures[62]. Thus, the results presented in Fig. 3-App.2 confirm that dimers of IgG$_1$ antibodies were not responsible for the bimodal distribution of ΔI values and that the IgG$_1$ antibodies were stable and functional in the buffered solutions used here.

GPI-anchored acetylcholinesterase purified from human erythrocyte membranes naturally occurs in a dimeric, prolate-shaped form that is held together by disulfide bonds near the C-terminal tail of the protein[10,64-67]. To confirm that the GPI-AChE used in this work remained in dimeric form and to detect impurities in solution, we performed a SDS-PAGE experiment (Fig. 3-App.3). We ran three lanes on the SDS-PAGE gel corresponding to three different treatments of the protein: incubation with SDS, incubation with SDS and β-mercaptoethanol to dissociate the disulfide bond, and incubation with only β-mercaptoethanol to assess whether the disulfide bond in the folded protein was accessible to β-mercaptoethanol as reported in literature[67]. After staining, we observed only one protein band in each lane. When the protein was denatured with only SDS, we observed the dimeric form that ran with an apparent molecular weight of ~140 kDa. In both lanes where the protein was treated with β-mercaptoethanol, we observed only one protein band, running at an apparent molecular weight of ~60 kDa. These apparent molecular weights are lower than the values in reported literature of 160 kDa for the dimer and 80 kDa for the monomer because GPI-

AChE is an amphiphilic protein and likely has a higher binding capacity for SDS than more commonly run soluble proteins[10,64-69]. This increased binding of SDS yields a greater charge to mass ratio and therefore greater migration speed of the protein compared to most soluble proteins, causing GPI-AChE to migrate in the gel as if it had a lower molecular weight. This phenomenon is well known for amphiphilic proteins[70].

The fact that we only observe one band in each lane of the gel indicates that our samples contained high concentrations of GPI-AChE relative to other contaminants. In the lanes treated with β-mercaptoethanol the absence of a band at ~140 kDa coincident with the appearance of a band at ~60 kDa is consistent with breakage of the disulfide bond holding the dimer together. Moreover, as reported in literature, the disulfide bond was accessible in the native structure of the protein, as indicated by the appearance of a single band at the monomer molecular weight when the protein was treated only with β-mercaptoethanol (and no SDS) prior to running the gel[67]. Consequently, this gel confirms that the GPI-AChE protein in our sample was in its natural dimeric, prolate-shaped form. Moreover, the control experiments in Fig. 3-App.2a-c indicate that if there were soluble (*i.e.*, not lipid-anchored) contaminants in the solution, they would not be detected, since soluble proteins would not be concentrated on the lipid surface or slowed during translocation through the nanopore.

To rule out the possibility that the widely distributed ΔI values were due to two proteins passing through the nanopore simultaneously, we compared the frequency of translocation events with the translocation times for each protein[71]. In the case of streptavidin translocations, we observed approximately 45 translocation events per second and a most-probable translocation time of about 115 µs. Consequently, on average

there was a 0.52% probability of a molecule occupying the nanopore at any time, and the probability of two streptavidin proteins occupying the nanopore at the same time would be 0.003%. In the case of the IgG_1 translocation events, the maximum frequency we observed was approximately 30 events per second and a most probable translocation time of about 55 μs. Consequently, on average there was a 0.16% probability of an IgG_1 protein occupying the nanopore at any time, and the probability of two IgG_1 proteins occupying the nanopore at the same time would then be 0.0027%. Even if the first translocation event of an IgG antibody would be exceptionally long lived (*e.g.*, 1000 μs), the probability of a second antibody to enter the pore during that time would still only be around 3% at an average translocation frequency of 30 Hz. This analysis neglects steric effects, which we expect would be significant given the size of an IgG_1 antibody and the dimensions of the nanopores. For GPI-anchored acetylcholinesterase the estimated probability of a two proteins being in the nanopore at the same time was 0.000036%. For Fab fragments, β-phycoerythrin, glucose-6-phosphate dehydrogenase, L-lactate dehydrogenase, BSA, α-amylase, and butyrylcholinesterase, the estimated probability was less than 0.00001%.

Even during the resistive-pulse sensing experiments with streptavidin in which we estimated the highest probability of observing a protein in the nanopore, we did not observe resistive-pulses with multiple current levels that might suggest the translocation of two proteins simultaneously. Consequently, we conclude that the resistive pulses due to each protein detected in this work resulted from the translocation of one protein at a time.

3-App.S2 Determining the volume and shape of proteins from fitting distributions of maximum ΔI values

Equation relating the amplitude of resistive pulses to the volume and electrical shape factor of particles

The relationship between the magnitude of ΔI and the volume of a particle stems from Maxwell's derivation[72], and it is shown in equation (3-App.1)[73-76].

$$\frac{\Delta I}{I} = -\frac{4 \Lambda \gamma}{\pi d_P^2 (l_P + 0.8d_P)} S\left(\frac{d_M}{d_P}\right) \Rightarrow \Delta I = -\frac{\Lambda V_A \gamma}{\rho (l_P + 0.8d_P)^2} S\left(\frac{d_M}{d_P}\right), \qquad \textbf{(3-App.1)}$$

where γ is the electrical shape factor[21,22,35,74,77,78], Λ (m^3) is the excluded volume of the particle, l_P (m) is the length of the pore, d_P (m) is the diameter of the pore, ΔI (A) is the magnitude of the change in the current during translocation of a particle, I (A) is the baseline current, V_A (V) is the applied voltage, and ρ (Ω m) is the resistivity of the electrolyte. $S\left(\frac{d_M}{d_P}\right)$ is a correction factor applied when the diameter of the particle, d_M, approaches the diameter of the pore, d_P, (i.e. $d_M > 0.5\, d_P$)[73,74]. Under these conditions the electric field in the pore is additionally distorted between the particle and the pore walls resulting in a non-linear increase in the resistance with increasing particle volume[73,74]. Qin *et al.* recently reviewed these correction factors and showed that the most accurate correction factor for all d_M/d_P ratios was developed by Smythe[38] and Deblois *et al.*[73], equation (3-App.2)[17]:

$$S\left(\frac{d_M}{d_P}\right) = \frac{1}{1 - 0.8\left(\frac{d_M}{d_P}\right)^3}. \qquad \textbf{(3-App.2)}$$

Note that in the majority of resistive-pulse sensing literature, particles and proteins have been considered spherical and consequently γ was set to a value of 1.5 and Λ was constrained to equal $\frac{1}{6} \pi d_M{}^3$. Substituting these values into equation (3-App.1) simplifies it to the more commonly seen form in equation (3-App.3)[17,72,73,75,76,78]:

$$\frac{\Delta I}{I} = - \frac{d_M{}^3}{d_P{}^2 \left(l_P + 0.8d_P\right)} S\left(\frac{d_M}{d_P}\right) \Rightarrow \Delta I = - \frac{\pi\, V_A\, d_M{}^3}{4\, \rho \left(l_P + 0.8d_P\right)^2} S\left(\frac{d_M}{d_P}\right), \qquad \textbf{(3-App.3)}$$

Since in this work we analyzed resistive-pulses due to the translocation of non-spherical proteins and we expected d_M to be less than ½ d_P, we set the correction factor to a value of 1[20,75,76]. We used equation (3-App.1) and expressed the impeded flow of ions through the nanopore during protein translocation events as reductions in current, ΔI.

The volume exclusion model shown in equation (3-App.1) has yielded accurate estimates of volume in a number of prior publications[14,16,18,19,21,75,79]; however, it has also been inadequate under a variety of different experimental conditions[80-84]. The model fails to describe certain current pulses because it does not account for heterogeneity in the distribution of ions, and thus the conductivity of the solution, in the nanopore. Heterogeneity in the distribution of ions results from electrostatic interactions with the surface of the pore and translocating particle. For instance, Lan *et al.* observed biphasic current pulses resulting in part from the accumulation of chloride ions on one side of the particle[80]. In this case, the flow of chloride ions around the particle was inhibited as the particle and pore were both negatively charged. To determine whether such effects are likely to occur under the experimental conditions used here, we performed finite-element simulations nearly identical to those described by Lan *et al.* Fig. 3-App.9a shows similar

local variations in the conductivity of the solution to those reported by Lan *et al.* at a low ionic strength of 10 mM KCl due to the accumulation and depletion of chloride ions on opposite sides of the protein. In contrast, the conductivity of the solution is nearly constant at the high ionic strength of 2 M KCl that we used in the experiments presented here (Fig. 3-App.9b). In this case, the ΔI signature (Fig. 3-App.9c) is well described by the volume exclusion model shown in equation (3-App.1). Consequently, the volume exclusion model is appropriate under the experimental conditions used in this work.

Electrical shape factor and distributions of shape factors

The electrical shape factor has been reported in literature since Maxwell derived equations to describe the conductance of solutions that contain insulating (*i.e.*, non-conducting) spheres[72]. Maxwell considered both the volume fraction of the spheres in solution and the deformation of the electric field around these spheres. To account for the distortion of the electric field, Maxwell derived a scaling factor that is dependent on the shape of the insulating particles (*i.e.*, electrical shape factor) and equal to 3/2 or 1.5 for spheres. Several years later, Fricke derived the electrical shape factor for spheroids, and Velick and Gorin developed analytical equations to describe the shape factor for ellipsoids of a general shape[40-42]. In 1954, Smythe numerically tested Maxwell's theory for the specific case of a particle residing in a pore; this work verified the electrical shape factor of 1.5 for spheres as well as the methods described by Fricke, Velick, and Gorin[38]. Around the same time, many groups experimentally proved these theories during resistive-pulse sensing experiments with holes that were micrometers in diameter while sensing various micrometer-sized particles[35-37,74,78]. In 1973, Golibersuch observed the

rotation of red blood cells within the pore of a resistive-pulse sensor and derived the distribution for electrical shape factors to explain the periodic variations in ΔI that occurred during the rotation of the blood cell.

The mathematical descriptions for shape factors are analogous among many systems and can be used to describe how electric and magnetic fields deform around insulating particles as well as how ideal fluids flow around obstacles in wind tunnels or in aqueous solutions with laminar flow[35,36,38]. Spheres alter flow and electric fields to the same extent regardless of their orientation; however, spheroid particles alter these fields to a different extent depending on their orientation relative to the direction of the field. Thus, the electrical shape factor is a function of a particle's shape and orientation.

To relate the value of ΔI to the volume and shape of non-spherical proteins, we considered the possible values of the electrical shape factor, γ, with the condition that a protein may have an oblate, prolate, or spherical shape. Oblates and prolates have an axis of revolution (shown as the dashed blue line in Fig. 3.2 of the main text) with length A and secondary axes with length B. Golibersuch elegantly pointed out that equation (3-App.4) describes the electrical shape factor, γ, for these ellipsoids as a function of the angle between the axis of symmetry and the electric field, θ, (Fig. 3.2)[35,36]:

$$\gamma(\theta) = \gamma_{\perp} + (\gamma_{\parallel} - \gamma_{\perp})\cos^2(\theta) \qquad \textbf{(3-App.4)}$$

where $\gamma_{\parallel}$ and $\gamma_{\perp}$ are the electrical shape factors when the axis of symmetry is parallel to the electric field (i.e. $\theta = 0, \pi, ...$) and perpendicular to the electric field (i.e. $\theta = \pi/2, 3\pi/2, ...$), respectively. Equation (3-App.4) implies that the shape factor for any orientation will range between the values of $\gamma_{\parallel}$ and $\gamma_{\perp}$. These factors, $\gamma_{\parallel}$ and $\gamma_{\perp}$, are related to the

well-described depolarization factors for ellipsoids, $n_{\parallel}$ and $n_{\perp}$, by equation (3-App.5) and are a function of the length to diameter ratio, $m = A/B$, of an ellipsoid[35,37,77,85].

$$\gamma_{\parallel} = \frac{1}{1 - n_{\parallel}} \text{ and } \gamma_{\perp} = \frac{1}{1 - n_{\perp}}$$ (3-App.5)

where $n_{\parallel}$ for a prolate spheroid with $m = A/B > 1$ is described by equation (3-App.6):

$$n_{\parallel} = \frac{1}{m^2 - 1}\left[\frac{m}{\sqrt{m^2 - 1}}\ln\left(m + \sqrt{m^2 - 1}\right) - 1\right]$$ (3-App.6)

and $n_{\parallel}$ for an oblate spheroid with $m = A/B < 1$ is described by equation (3-App.7):

$$n_{\parallel} = \frac{1}{1 - m^2}\left[1 - \frac{m}{\sqrt{1 - m^2}}\cos^{-1}(m)\right]$$ (3-App.7)

and $n_{\perp} = (1 - n_{\parallel})/2$[35,78,85].

To derive the distribution of shape factors, we assume that ellipsoidal proteins rotate freely such that all angles of θ are equally likely when ΔI is measured. By symmetry, values of θ range between 0 and $\pi/2$. According to Golibersuch, these assumptions enable using substitution of variables to write a probability distribution function for electrical shape factors $P(\gamma)$ based on the probability of observing a certain orientation $P(\theta(\gamma))$, where θ is a function of γ (equation (3-App.8))[35]:

$$P(\gamma)d\gamma = P[\theta(\gamma)]\frac{d\theta}{d\gamma}d\gamma$$ (3-App.8)

Since, by symmetry, values of θ range between 0 and $\pi/2$ and we assumed that all angles of θ are equally likely, we solved for $P(\theta)$ by noting that the integral of a probability distribution function equals 1:

$$\int_0^{\pi/2} P(\theta)d\theta = 1 = \int_0^{\pi/2} \frac{2}{\pi}d\theta \Rightarrow P(\theta)d\theta = \frac{2}{\pi}d\theta \qquad \text{(3-App.9)}$$

Combining equation (3-App.8) with (3-App.9), we obtained:

$$P(\gamma)d\gamma = \frac{2}{\pi}\left(\frac{d\gamma}{d\theta}\right)^{-1}d\gamma . \qquad \text{(3-App.10)}$$

Differentiating equation (3-App.4) with respect to θ (i.e., $\frac{d\gamma}{d\theta}$) and combining the result with equation (3-App.10), we obtained a probability density function for the possible shape factors[35].

$$P(\gamma)d\gamma = \frac{1}{\pi\left[(\gamma - \gamma_\perp)(\gamma_\parallel - \gamma)\right]^{1/2}}d\gamma \qquad \text{(3-App.11)}$$

Fig. 3.2c of the main text (black line) shows that this probability density function (equation (3-App.11)) is bimodal and symmetric with peaks at $\gamma_\parallel$ and $\gamma_\perp$. The bimodal character of this distribution reflects the fact that for small deviations in θ near 0 and near $\pi/2$, there is little change in the value of the shape factor compared to deviations in θ around $\pi/4$ (Fig. 3.2b).

Before attempting to describe the non-Normal distributions of ΔI values as a consequence of $p(\gamma)$, we considered whether the non-spherical proteins could sample

various orientations, and therefore shape factors, in these experiments as well as whether the time-scale of rotation would bias the measurement of maximum ΔI values. We first considered potential steric limitations on the orientations of the proteins in the nanopore. Figure 3.1c in the main text shows the expected lipid anchoring locations for the anti-biotin IgG antibody, anti-biotin Fab fragment, GPI-AChE[64]. Since the chemical linker between the lipid head group and the ligand for the IgG_1 and Fab fragments was approximately 1.5 nm in length, we expect the anchoring positions shown in Fig. 3.1c to permit rotation of the proteins in orientations that could generate the minimum and maximum shape factors. We attached the remaining non-spherical proteins characterized in this work to the bilayer *via* a homobifunctional crosslinker with a flexible, 2.2-nm-long polyethylene glycol spacer arm. Since the crosslinker reacted with primary amines (*e.g.*, lysines and glutamines), the anchoring locations on these proteins were randomly distributed across their surface. Consequently, we also expect these proteins to sample the full range of electrical shape factors while passing through the nanopore.

We next examined whether the dipole moment of a protein may align completely in the large electric field in the nanopore ($\sim 10^6$ V m^{-1}). Combining the potential energy, ΔU, of a dipole moment in an electric field and the Boltzmann distribution of energies while assuming that the dipole moment was pointed parallel to one of the principal axes of the protein, we expanded on Golibersuch's probability distribution of shape factors to develop a $p(\gamma)$ for a protein with a dipole moment (Fig. 3.2c in the main text and Section 3-App.S9). To expand on the theories developed by Golibersuch, we considered the possible probability distribution of shape factors if the orientation of the protein were biased by the electric field in the nanopore. The electric field in the nanopore is on the

86

order of 10^6 V m^{-1}, and consequently, we expect the orientation of a protein to be biased by alignment of its dipole moment, $\bar{\mu}$ (Debye $\approx 3.33564 \times 10^{-30}$ C m), in the electric field, $\bar{E}$ (V m^{-1}). Taking into account the potential energy of a dipole in an electric field, $\Delta U = \bar{E} \cdot \bar{\mu} = -E\mu \cos(\phi)$, using the Boltzmann distribution of energies, and assuming the dipole was aligned along the symmetry or equatorial axis, we derived equations (3-App.12a) and (3-App.12b), respectively (Section 3-App.S9). Equations (3-App.12a) and (3-App.12b) describe probability distribution functions of shape factors for spheroid proteins when their orientation is biased by the dipole energy in an electric field.

$$P(\gamma)d\gamma = \frac{1}{A}\cosh\left[\frac{E\mu\left(\dfrac{\gamma - \gamma_\perp}{\gamma_\parallel - \gamma_\perp}\right)^{1/2}}{k_B T}\right]\left[\frac{1}{\pi\left[(\gamma - \gamma_\perp)(\gamma_\parallel - \gamma)\right]^{1/2}}\right]d\gamma \qquad \textbf{(3-App.12a)}$$

$$P(\gamma)d\gamma = \frac{1}{A}\cosh\left[\frac{E\mu\left(\dfrac{\gamma - \gamma_\parallel}{\gamma_\perp - \gamma_\parallel}\right)^{1/2}}{k_B T}\right]\left[\frac{1}{\pi\left[(\gamma - \gamma_\perp)(\gamma_\parallel - \gamma)\right]^{1/2}}\right]d\gamma \qquad \textbf{(3-App.12b)}$$

In equations (3-App.12a) and (3-App.12b), A is a normalization constant described in Section 3-App.S9. Fig. 3.2c of the main text demonstrates that for spheroid proteins with dipoles of several thousand Debyes, it is theoretically possible to observe a bimodal distribution of shape factors. The average dipole moments of proteins is approximately 550 Debye (http://bioinfo.weizmann.ac.il/dipol/indexj.html), suggesting that many non-spherical proteins should generate a skewed bimodal distribution of shape factors. Additional factors may bias the orientation of proteins in the nanopore including steric effects, interactions with the pore wall, and alignment of slender proteins prior to entering

the nanopore. All of these factors could affect the estimated value of ΔU or μ in this model. Therefore, an alternative interpretation of these parameters is that they describe the overall bias of the protein's orientation toward $\theta = 0$ or $\pi/2$. Equations (3-App.12a) and (3-App.12b) cannot describe distributions of ΔI accurately for proteins that are significantly biased (*i.e.*, $\Delta U > \sim 4$ $k_B T$ or $\mu > \sim 3000$ D for a typical pore at 100 mV applied potential) toward intermediate orientations relative to the electric field (*i.e.*, $\theta = \pi/4$). Under these circumstances, the model would not resolve ΔI_{min} and ΔI_{max} accurately, underestimating the shape of the protein (*i.e.*, m would approach 1) and overestimating the volume of the protein. Consequently, equations (3-App.12a) and (3-App.12b) are an approximation of how the orientation, and therefore distribution of shape factors, of a protein with a dipole moment may be biased, and they allow the theoretical distribution of shape factors to become asymmetric.

We also considered whether the orientation of the protein would be significantly biased due to the hydrodynamic drag force, which is orientation dependent for non-spherical particles. To this end, we calculated the drag for an oblate with a relatively extreme shape (4 x 16 x 16 nm) when its axis of symmetry is aligned perpendicular and parallel to the direction of fluid flow (*i.e.*, the direction of translational motion). Assuming the pore is 25 nm long and the protein transits this distance in 100 μs, the average speed of the proteins is 2.5 x 10^{-4} m s^{-1}, and the corresponding orientation-dependent drag force would range between 26 and 33 fN[86]. Based on these forces, the difference in energy required to move the protein through the entire length of the pore varies by a maximum of roughly 0.04 $k_B T$. As a result, we do not expect hydrodynamic drag to significantly bias the orientation of the protein.

Finally, we considered whether the proteins would rotate in the pore too quickly to be time resolved or whether their rotation would bias the measurement of ΔI values such that we would only observe ΔI values corresponding to γ_{max}, and therefore, not resolve ΔI values corresponding to γ_{min}. Axelrod observed that GPI-AChE has rotational diffusion coefficients, D_R, of $10{,}000 \pm 4{,}000$ rad^2 s^{-1} and Timbs *et al.* have observed dramatically reduced mobility (*i.e.*, $D_R \approx 0.003$ rad^2 s^{-1}) of IgG antibodies binding to lipids in a substrate-supported monolayer[87-89]. Consequently, we estimate that the average time for a protein to rotate $\pi/2$ radians to be at least 125 μs. Since the majority of the translocation times in these experiments were between 50 and 100 μs (Fig. 3-App.2), we expect the majority of ΔI values to reflect a single orientation or a very limited range of orientations of the protein in the nanopore. Consequently, we expect the bimodal distributions of ΔI values observed here to reflect accurately the underlying distribution of shape factors with modes at γ_{min} and γ_{max}[35]. This prediction is supported by our recent discovery of bimodal distributions of ΔI values from translocation of a single, pure protein[20] and subsequent observations made by Raillon *et al.*[21].

Since the value of ΔI is directly proportional to the electrical shape factor, γ, according to equation (3-App.1), we expressed equations (3-App.12a) and (3-App.12b) in terms of ΔI. For an oblate this procedure results in equations (3-App.13a) and (3-App.13b), where the parameters ΔI_{min} and ΔI_{max} correspond to γ_{min} and γ_{max}.

$$P(\Delta I_\gamma)d\,\Delta I_\gamma = \frac{1}{A}\cosh\left[\frac{E\,\mu\left(\dfrac{\Delta I - \Delta I_{max}}{\Delta I_{min} - \Delta I_{max}}\right)^{1/2}}{k_B T}\right]\left[\frac{1}{\pi\left[\left(\Delta I - \Delta I_{max}\right)\left(\Delta I_{min} - \Delta I\right)\right]^{1/2}}\right]d\,\Delta I_\gamma \textbf{(3-App.13a)}$$

and

$$P(\Delta I_\gamma) d \Delta I_\gamma = \frac{1}{A} \cosh \left[\frac{E \mu \left(\dfrac{\Delta I - \Delta I_{min}}{\Delta I_{max} - \Delta I_{min}} \right)^{1/2}}{k_B T} \right] \left[\frac{1}{\pi \left[\left(\Delta I - \Delta I_{max} \right) \left(\Delta I_{min} - \Delta I \right) \right]^{1/2}} \right] d \Delta I_\gamma \textbf{ (3-App.13b)}$$

For a prolate, equations (3-App.13a) and (3-App.13b) are interchanged. These probability distributions are the expected distributions of ΔI values due only to the possible values of the shape factor – they do not include effects such as experimental or analytical errors in determining ΔI values.

Fitting the convolution model to distributions of ΔI values

To account for experimental and analytical errors in determining ΔI values, we convolved the expected distribution of ΔI values due to variation in the electrical shape factor, $p(\Delta I_\gamma)$ (equations (3-App.13a) and (3-App.13b)), with a Normal distribution, $p(\Delta I_\sigma)$, to generate the a distribution of ΔI values that one expects to observe experimentally, $p(\Delta I)$. We used this theoretical distribution (herein referred to as the "convolution model") to fit all of the empirical distributions of ΔI values presented in this work. Figure 3-App.5 illustrates this method.

When constructing empirical distributions of ΔI values from many translocation events, we represented each event by its maximum ΔI value as opposed to its average (*e.g.*, Fig. 3.3 in the main text). We followed this strategy because representing events by their average value causes bias towards intermediate ΔI values and may introduce an additional mode besides the two expected modes at ΔI_{min} and ΔI_{max}, which would result in an improper fit with the convolution model (Fig. 3-App.5). On the other hand, representing events by their maximum value likely biases the distribution of ΔI values

toward ΔI_{max} such that the amplitude of the peak corresponding to ΔI_{max} increases. In this instance, however, the location of ΔI_{max} and ΔI_{min} should be preserved such that the shape of the protein can still be determined accurately.

Since the distribution of ΔI values resulting from the distribution of shape factors, $p(\Delta I_\gamma)$, is different depending whether the dipole moment is assumed to be parallel to the symmetry or equatorial axis of the protein (equations (3-App.13a) and (3-App.13b), respectively), we fit each empirical distribution of ΔI values, $P(\Delta I)$, with both of the resulting solutions to the convolution model. Subsequently, we selected the fit that yielded the larger adjusted R^2 value as the correct solution. Since the orientation of the dipole moment dictates the preferred orientation of the protein, this procedure effectively determined whether the distribution of ΔI values was skewed towards ΔI_{min} or ΔI_{max}.

When fitting the distributions of ΔI values for Fab, α-amylase, and BChE, we excluded outliers from the upper end of the distributions to determine their shape correctly. For each distribution, we excluded ΔI values that were greater than a threshold value, which we chose such that the R^2 value of the fit with the convolution model was maximized. Conversely, ΔI values were not excluded for any of the other proteins detected in this work. Finally, we low-pass filtered the data for BChE at 10 kHz as opposed to 15 kHz in order to improve the signal-to-noise ratio.

Using ΔI_{min} and ΔI_{max} to solve for the volume and shape of proteins

Given that the probability distribution of shape factors has modes at $\gamma_{||}$ and $\gamma_{\perp}$ corresponding to either ΔI_{min} or ΔI_{max} values according to equation (3-App.1), we expected that if the value of ΔI_{min} and ΔI_{max} could be determined quantitatively from the

empirical distribution of ΔI values then the volume and shape of a protein could also be determined. For example, the minimum shape factor for an oblate spheroid occurs at $\theta = \pi/2$ and has a value of $\gamma_\perp(m)$ (equation (3-App.4)). Thus, according to equation (3-App.1), the minimum mode in the bimodal ΔI distribution, ΔI_{min}, is a function of Λ and $\gamma_\perp(m)$, and the maximum mode in the bimodal ΔI distribution, ΔI_{max}, is a function of Λ and $\gamma_{\parallel}(m)$. Since both $\gamma_{\parallel}$ and $\gamma_\perp$ are solely a function of m, we developed the system of equations (3-App.14) and (3-App.15) in which the values of m and Λ are the only two unknowns and the values of ΔI_{min} and ΔI_{max} are determined from fitting the empirical distributions of ΔI with the convolution model. By rearranging equation (3-App.3), we can write for oblate spheroids with $m < 1$:

$$\Lambda(m) = \begin{cases} \Lambda\left(\gamma_\perp(m), \Delta I_{min}\right) \\ \Lambda\left(\gamma_{\parallel}(m), \Delta I_{max}\right) \end{cases} \quad \text{if } m < 1, \qquad \textbf{(3-App.14)}$$

and for prolate spheroids with $m > 1$:

$$\Lambda(m) = \begin{cases} \Lambda\left(\gamma_{\parallel}(m), \Delta I_{min}\right) \\ \Lambda\left(\gamma_\perp(m), \Delta I_{max}\right) \end{cases} \quad \text{if } m > 1. \qquad \textbf{(3-App.15)}$$

Since this system of equations has a piecewise dependence on the value of m, we substituted the determined values of ΔI_{min} and ΔI_{max} into equations (3-App.14) and (3-App.15) and used MATLAB to solve the system for the excluded volume of the protein, Λ, and the value of m. For all prolates and relatively spherical oblates, two solutions to this system of equations exist as shown in Fig. 3-App.7. The solutions for all experiments are summarized in Table 3-App.1.

For many of the fits, the value of σ is reasonable given the standard deviation of the baseline noise, which was typically between 20 and 60 pA. On the other hand, several of the fits returned relatively low estimates of σ (*e.g.*, α-Amylase using Pore 10), which may be a result of using maximum ΔI values to represent long events or due to partial truncation of the ΔI distributions since only values larger than a certain threshold were detected. Nevertheless, the excellent agreement between the estimated volume of the proteins and their respective shapes (Table 3-App.4) provide strong evidence that this procedure enables one to approximate the shape and determine the volume of non-spherical proteins by analyzing the distributions of maximum ΔI values. This method does not assume any information about the protein to extract the parameters shown in Table 3-App.1.

While results for *m* and Λ from different pores are in good agreement for G6PDH and BSA (<10% difference in *m* and <20% difference in Λ), we observed significant pore-to-pore variability for GPI-AChE and IgG$_1$ (Table 3-App.4). Using all of the 9 possible pore-to-pore comparisons from the results presented in Table 3-App.1, we found that pore-to-pore variability in *m* and Λ is weakly correlated with differences in pore diameter and length (*i.e.*, $-0.3 \leq$ Pearson's $r \leq 0.3$). In fact, we observed the lowest pore-to-pore variability in *m* for G6PDH despite the fact that the pores used to characterize this protein have the largest difference in radii of any of the possible pore-to-pore comparisons. Based on these results, it appears that pore-to-pore variability of determined *m*- and Λ-values does not depend on pore diameter or length. This variability is likely due to variations in the pore geometry that are not accounted for in the model. The model assumes that the pore is perfectly cylindrical (*i.e.*, constant diameter); however,

nanopores prepared by ion-beam sculpting generally have an hourglass shape[90]. Even if the maximum ΔI value for each event is obtained when the protein is centered about the narrowest constriction of the pore, ΔI_{min} and ΔI_{max} will vary with the degree of pore tapering. Moreover, sterics may also introduce pore-to-pore variability by restricting certain protein orientations or conformations, particularly for IgG_1 since it is a relatively large protein and composed of three domains that move relative to one another. Based on these arguments, it is perhaps unsurprising that we find the largest pore-to-pore variability for the determination of the shape factor, m, for IgG. However, even in this most challenging case with a large protein whose shape deviates significantly from an ellipsoid of rotation, the standard deviation of m-values is smaller than $\pm 50\%$, while it is smaller than $\pm 40\%$ for AChE and $\pm 6\%$ for G6PDH and BSA.

Estimating the volume of spheroidal proteins via dynamic light scattering

For comparison to the nanopore-based method, we used the technique of dynamic light scattering (DLS) to estimate the volume of each protein detected in this work. We assumed that the proteins were either spherical or spheroidal in shape in order to calculate their volume from the hydrodynamic radius, r_H, returned from the DLS measurements (Table 3-App.3). For spheroidal proteins, we used the length-to-diameter ratio, $m = A/B$, of the particle (listed in Table 3-App.1) with the corresponding Perrin shape factor, S, to calculate the volume based on the following equation[91]:

$$D = \frac{k_B T}{6\pi\eta r_H} = \frac{k_B T}{f_{sphere}\left. 2m^{2/3} \middle/ S \right.} = \frac{k_B T}{\left[6\pi\eta\left(ab^2\right)^{1/3}\right]\left. 2m^{2/3} \middle/ S \right.}$$

where f_{sphere} is the friction coefficient of a sphere with the same volume as a spheroid with semi-axes a, b, and b. Furthermore, S for an oblate spheroid is equal to:

$$S_{oblate} = \frac{2 \tan^{-1} \varepsilon}{\varepsilon}$$

and S for a prolate spheroid is equal to:

$$S_{prolate} = \frac{2 \tanh^{-1} \varepsilon}{\varepsilon}$$

where $\varepsilon = \sqrt{|a^2 - b^2|} / a$. Combining the above equations yields:

$$r_{H,oblate} = \frac{\sqrt{|a^2 - b^2|}}{\tan^{-1}\left(\frac{\sqrt{|a^2 - b^2|}}{a}\right)}$$

and

$$r_{H,prolate} = \frac{\sqrt{|a^2 - b^2|}}{\tanh^{-1}\left(\frac{\sqrt{|a^2 - b^2|}}{a}\right)} .$$

We solved the preceding two equations numerically in MATLAB to determine the dimensions of each spheroidal protein and calculated the corresponding volume. The resulting spheroidal volumes were in excellent agreement with the volumes that we determined by fitting the convolution model to distributions of ΔI values. For reference, we used the crystal structures of these proteins to determine their length-to-diameter ratio, m, and subsequently determine their spheroidal volume; these volumes were also in

95

excellent agreement with volumes obtained from analysis of DLS and resistive-pulse sensing experiments (Table 3-App.3). In contrast, if we assumed the particles were a perfect sphere, the volumes that we determined from the hydrodynamic radius were overestimated for every non-spherical protein. These experiments provide additional evidence that the methods we present in this paper accurately describe the distribution of ΔI values for determining the shape and volume of spheroidal proteins.

Low applied potentials yield consistent estimates of protein shape

The value of the shape parameter, m, determined from fitting distributions of maximum ΔI values for IgG_1 and GPI-AChE is consistent at relatively low applied potentials but decreases or increases, respectively, as the applied potential is increased (Fig. 3-App.8). This deviation might result from deformation of the protein due to the electrophoretic force acting on it while in the nanopore as was observed by Freedman *et al.*[92]; however, the amount of deformation that is expected based on theory (see proceeding subsection) is not large enough to account for the change in m observed here. Furthermore, Pelta *et al.* have previously shown that proteins do not change shape under similar electric field intensities[93]. Alternatively, this deviation could be due to changes in the size and shape of the hydration shell surrounding the protein or increasing alignment of the protein in the electric field gradient prior to entering the pore with increasing field intensity. In response to this observation, we limited our analyses to distributions of ΔI values that were obtained at relatively low potentials where the distributions appeared to be resolved fully. This approach consistently returned accurate estimates of the shape and volume of non-spherical proteins (Table 3-App.1).

Forces acting on proteins in a nanopore

Since the magnitude of the electric field is on the order of 10^6 V m^{-1} in the nanopore, we considered theoretically whether it was possible for the shape of proteins to be affected by forces in the nanopore. In this work, we expect proteins in the nanopore to be subjected to the following forces:

1) Instantaneous forces due to collisions with water will be on the order of ~500 pN with a net force equal to 0 on time scales of roughly 1 ps[94].

2) Net torque due to the dipole moment in the electric field will have magnitudes similar to thermal energy. Fig. 3.2c in the main text shows that we expect the torque on a protein to be on the order of 0 to 4 $k_B T$ in this work.

3) Average force due to the net charge of the protein in the electric field, F_q, is in the range of 0.1 to 4 pN for the electric field strengths and net charges of proteins used in this work.

4) Average force on the protein due to viscous drag in the aqueous solution, F_w, which opposes the electrophoretic force. We approximated this force to range from about 0.026 to 0.033 pN.

5) Average force on the lipid anchor, F_L, which also opposes the electrophoretic force, is thus on the order of 0.1 to 4 pN, since the force due to drag in the aqueous solution is negligible (*i.e.*, $F = 0 = F_q - F_w - F_L$).

Since we expect these five forces to be nearly constant through the length of the nanopore, the shape of proteins in the nanopore should also be constant. This expectation is based on the fact that the internal stiffness of a protein and the viscosity of the solution

97

result in highly over-damped motion of the protein. Any external force that affects the global conformation of a protein results in a gradual deformation of the protein toward its equilibrium conformation over a period of nanoseconds and without oscillations. In other words: "the global motions of proteins, especially less rigid ones, are highly overdamped: They creep rather oscillate when subject to applied forces".[94]

The largest constant force listed here is the possible tension within a protein due to the electrophoretic force acting on the net charge of the protein and the opposing drag force exerted by the lipid anchor. To estimate the deformation of the protein acted upon by a net force of 4 pN, we note that the Young's modulus (E) of most rigid proteins is on the order of 1 GPa[94]. Considering a protein similar in size and shape to GPI-AChE (*e.g.*, a cross sectional area, A, of 5 nm x 5 nm and a length, L, of 13 nm), the total deformation (*i.e.*, change in length) of the protein in response to a force of 4 pN is:

$$\Delta L = F * L / (E * A) = 4 \text{ pN} * 13 \text{ nm} / (1 \text{ GPa} * 5 \text{ nm} * 5 \text{ nm}) = 2 \text{ pm} = 0.002 \text{ nm}$$

This estimate for the deformation of a protein due to 4 pN of force illustrates that forces due to the electric field are unlikely to deform the proteins used in this work.

Proteins in the nanopore may also experience transient collisions with the pore wall in which the average force acting on the protein during the collision is equal to the rate of change in momentum of the protein. To estimate this force conservatively, we consider that a 100 kDa protein has an instantaneous velocity of 8.6 m s^{-1} (this velocity is indeterminable on short time scales due to collisions with water molecules, corresponding to about 2 ps or 0.024 nm of distance traveled)[94] and that it collides directly with the nanopore wall, bouncing straight back with the same speed. If this protein collides with

the pore wall over a period of 1 ps, the average force acting on the protein during that collision would be roughly 1 nN. This approximation estimates that the protein would deform between 0.076 nm to 0.52 nm depending on which face of the protein struck the wall. Again these deformations are small compared to the sizes of the protein we used in this work.

What kind of deformations might take place if the forces were far larger than we estimate? Suppose the forces acting on the protein in the nanopore did work equal to ~30 k_BT (1.23 E-19 J); this energy is ten times larger than the energy we estimate for the protein's dipole moments within the electric field of the nanopore. The deformation of the protein can be estimated by considering the stiffness of the protein, $k = E * L$.[94] Using the dimensions of the hypothetical protein that we described in the previous paragraphs, the stiffness of the protein to be $k = 1$ GPa $* 13$ nm $= 13$ N m^{-1}. Since the energy of a spring is $\frac{1}{2}k\Delta x^2$, we can estimate the deformation, Δx, of the protein to be on the order of 0.14 nm. Consequently, we do not expect rigid proteins to deform significantly due to forces in the nanopore.

Proteins with multiple domains and flexible connecting regions may change shape in the nanopore, however their motion will be overdamped and not subject to oscillatory changes while in the nanopore. As an example, consider IgG$_1$ which has three separate domains that move relative to one another. Similarly, myosin head-groups are linked to the rest of the protein through a flexible domain known to have a stiffness of 4 pN nm^{-1} (0.004 N m^{-1})[94-96]. Using this stiffness, we estimated the maximum distance that the domains of IgG$_1$ might stretch relative to one another by considering the maximum applied force acting on the molecule of 4 pN. Under this force, IgG$_1$ may stretch on

average roughly 1 nm. Because we expect these forces to remain constant through the nanopore and since the global motions of proteins are highly overdamped, especially for flexible proteins, this deformation would be nearly constant through the length of the nanopore.

Based on the magnitude of the forces discussed above, we do not expect the proteins used in this work to change shape significantly while in the nanopore. This expectation is supported by the accurate measurements of the size and shape of the ten different proteins detected in this work compared to the size and shape of these proteins as determined from crystal structures (Fig. 3.1 in the main text and Table 3-App.4). In further support of this expectation, proteins that bound non-covalently to biotinylated lipids (IgG$_1$, Fab, and streptavidin) translocated through pores in the bound, lipid-anchored state as confirmed by their distributions of translocation times and measured charges (see Fig. 3-App.11); if the binding pockets were denatured, antigen-binding would likely not occur.

Description of the assumptions underlying the convolution model

The following section describes the primary assumptions underlying the convolution model in particular with regard to their validity. To derive this model (*i.e.*, equation), we made four key assumptions:

1) The protein is a spheroid.
2) The dipole moment of the spheroidal protein is aligned with one of the principal axes.

3) While residing in the nanopore, the orientation of the protein is only biased due to its dipole moment.

4) The pore is perfectly cylindrical.

The first assumption states that the protein is a spheroid with principle axes having lengths *A, B,* and *B* (see Fig 3.2a). We examined approximately 1,000 randomly sampled proteins from the Protein Data Bank and found that the lengths of two of the three principal axes are less than 20% different on average, indicating that most proteins can be approximated as spheroids. Based on our results for IgG_1, we have also shown that our approach can be used to characterize proteins with highly irregular shapes. Although the complexity of the shape of IgG_1 is not captured in full, our approach still provides low-resolution shape information and yields accurate values for the dipole moment and rotational diffusion coefficient of the protein.

The second assumption is based on the expectation that the dipole moment is most often aligned with a principle axis of a spheroidal protein. For an asymmetrical protein, we expect the dipole moment to be aligned along the longest axis of the protein because the residues that are furthest from the center of the protein contribute most to the magnitude of the dipole moment. For a multimeric protein with rotational symmetry, such as GPI-AChE or β-PE, we expect the dipole moment to lie along the axis of symmetry since the off-axis components of the dipole moment from each subunit cancel each other out. Thus, in both cases it seems reasonable that the dipole moment will be in near alignment with one of the principal axes of the protein. In support of this expectation, we found that the dipole moment was aligned close to one of the principal

axes for each of the nine non-spherical proteins examined here by using the Weizmann server to analyze the protein crystal structures.

The third assumption states that the orientation of a protein in the nanopore is only biased by its dipole moment. We expect this to be true since we coated the nanopores with a lipid bilayer to eliminate non-specific interactions[20], anchored the proteins to the coating *via* long (≥ 1.5 nm) and flexible (≥ 12 σ-bonds) tethers so they could sample most orientations, and used nanopore diameters that were at least twice the volume-equivalent spherical diameter of the proteins to minimize steric effects. Under these conditions, we obtained dipole moment measurements for nine different proteins that were in excellent agreement with reference values (see Fig. 3.4e), supporting our assumption. We expect this assumption to be valid as long as the protein being characterized does not interact with lipids in the nanopore coating. In such cases, however, interactions with the coating could likely be avoided by modifying the bilayer composition (*e.g.*, including a small fraction of PEG-conjugated lipids).

The fourth assumption is that the pore is cylindrical. This assumption does not depend on the protein under investigation and, consequently, does not limit the general applicability of our approach toward other proteins. The good agreement between the measured and expected values of volume for the ten different proteins examined here (see Fig. 3.3g) supports this assumption. Additionally, the change in the baseline current observed upon coating a nanopore is generally close to the value predicted by theory in which the pore geometry is assumed to be cylindrical[20], further supporting this assumption. Regardless, we discuss how pores that are not perfectly cylindrical may affect the analysis of intra-event ΔI values in Section 3-App.S6.

3-App.S3 Interpretation of the observed bimodal distributions of Δ*I* values from the translocation of non-spherical proteins

To determine whether any explanation might exist for the bimodal distributions of Δ*I* values observed here besides the theory presented in Section 3-App.S2, we closely examined the literature to ascertain whether other groups had observed similar signals in nanopore-based sensing experiments resulting from alternative mechanisms. To the best of our knowledge, there is only one such report. In this study, Spiering *et al.* used optical tweezers to characterize the force response of threading a protein bound to a negatively-charged DNA molecule through a nanopore. For a finite range of optical trap positions, the authors found that the potential landscape "exhibits two minima (potential wells), corresponding to two metastable 'states'… with the charged protein on either side of the membrane," resulting in bimodal force *versus* time signals. Since the protein is located *outside* of the pore in both of these states (*i.e.*, where the electric field is negligible), the resulting Δ*I* values should be close to zero in resistive-pulse sensing experiments. Hence, we do not think that these two states do not correspond to the two modes in the Δ*I* distributions that we observe. The potential landscape in our experiments is different than that described by Spiering *et al.* due to the following reasons: (1) the lipid tethers are shorter (length ~1.5 nm) than the pore length (~30 nm) and hence cannot contract and extend to allow a protein to transition from one side of the pore to the other, (2) the charge of the protein-lipid complex is dominated by the charge of the protein rather than the tether, whereas in the case of a DNA tether, the opposite is true, (3) the lipid tethers only extend on one side of the protein instead of both sides as with the DNA tethers, and (4) there is no optical trap potential present in our experiments. Hence, it is extremely

unlikely that the two metastable states described by Spiering *et al.* exist under the experimental conditions used in our work.

Skewed and bimodal distributions of ΔI values have been observed with increasing frequency in the last few years and in each instance the authors suggested that the shape of the ΔI distributions may have been influenced by the shape and orientation of the macromolecule. For example, early indications that the shape and orientation of a macromolecule can affect the ΔI signal were reported by Mathé *et al.*, who observed orientation-dependent translocation signals of DNA through α-hemolysin pores[97], and Fologea *et al.* who observed a unimodal but skewed distribution of ΔI values due to the translocation of nodular fibrinogen proteins through nanopores[18]. More recently, Raillon *et al.* observed distributions of ΔI values that appeared to be bimodal due to the translocation of an untethered, non-spherical RNA polymerase through a nanopore; without additional quantification, the authors attributed this result to different orientations of the RNA polymerase[21]. Finally, Fiori *et al.* observed a bimodal distribution of ΔI values due to the translocation of untethered, prolate-shaped protein ubiquitin[23]. Together, these reports indicated that the bimodal distributions presented in our work do not result from the effect of the lipid tether on the potential landscape but rather the shape and orientation of the translocating proteins. Until the work presented here, however, the origin of these biomodal distributions was not understood and it was unknown whether useful information could be obtained from the shape of these distributions of ΔI values.

While we considered a number of other possible explanations for the current signatures that we observe (see Section 3-App.S1 and the subsection titled "Forces acting

on proteins in a nanopore" in Section 3-App.S2), eight observations indicate that ΔI reflects the rotational dynamics of proteins passing through the nanopore:

1) Streptavidin, which is spherical with a shape factor, m, of 1.1, yielded a Normal distribution of ΔI values (Fig. 3.3c in the main text).

2) The values of $\Delta I_{\min}$ and $\Delta I_{\max}$ that we determined for each protein are consistent with the values predicted by established theory for large particles; Golibersuch originally developed this theory to describe the periodic variations in ΔI that occurred during the rotation of a red blood cell within a resistive-pulse sensor.

3) Simulations based on a spheroidal particle undergoing biased random rotation in one dimension yield ΔI signals that are comparable to those that we obtained experimentally (Section 3-App.S5).

4) The values of volume (Λ), length-to-diameter ratio (m), rotational diffusion coefficient (D_R), and dipole moment (μ) that we determined for 9 different proteins are in good agreement with expected values (Fig. 3.3g-h and Fig. 3.4e-f); the methods used to determine these parameters critically depend on the assumption that ΔI reflects the orientation of non-spherical proteins, as described by the theory in Section 3-App.S2.

5) D_R of bivalently-bound IgG$_1$ is significantly less than D_R of monovalently-bound IgG$_1$ (see Section 3-App.S7), indicating that ΔI reflects the rotational dynamics of the protein in the nanopore.

6) Translocation of IgG_1 and GPI-AChE through the same nanopore result in markedly different distributions of ΔI values despite their similar molecular weights, indicating that ΔI is related to the shape of the protein.

7) Kolmogorov Smirnov (KS) tests indicate that the convolution model, which incorporates the effect of protein shape and orientation combined with noise to predict distributions of ΔI values, is not significantly different from the empirical distribution in 11 out of 13 cases (Fig. 3-App.6), indicating that the model that underlies our analysis and employs the effect of protein orientation and shape on ΔI describes the data very well.

8) From a fundamental physical chemistry perspective it is also reasonable to assume that proteins rotate while moving through the pore. In the case of non-spherical proteins this rotation will change the electric field lines and hence modulate the current based on Maxwell's and Golibersuch's equations. This expectation is supported by simulations (see Point 3 of this list). We think it is extremely unlikely that proteins translocate through the pores in one constant orientation over several hundreds of microseconds given that we demonstrated before that the fluid bilayer coating circumvents non-specific protein adsorption to the pore walls.

3-App.S4 Effect of lipid anchoring on the measurement of protein properties

Since we anchored each protein to a lipid in the bilayer coating of the nanopore to slow down translocation, we considered whether anchoring may have any other effects on the five parameters measured in this work. First, we do not expect protein shape or

volume to be affected by anchoring. As discussed in Section 3-App.S2, the force exerted by the lipid anchor that opposes the electrophoretic force is unlikely to deform the protein. In addition, the chemical modifications involved in the crosslinking procedure are unlikely to cause denaturation as such modifications are standard practice in various biochemical assays that rely on retention of protein function. Our expectation that protein shape and volume are unaffected by anchoring is supported by the excellent agreement between the measured size and shape of the ten different proteins detected in this work with reference values (Fig. 3.3g-h in the main text).

We do expect the distribution of translocation times to reflect the net charge, z, of the protein-lipid complex as a whole. Hence, we subtracted 1 from the expected value of z (Fig. 3-App.11j and Table 3-App.4) for each protein (except GPI-AChE) to correct for the net charge of the lipid anchor. For each protein that was crosslinked to the bilayer, we also subtracted 0.93 from the expected value of z based on the "charge regulation" model by Menon and Zydney[132] to account for the consumption of a positively charged amine group.

Tethering a protein to a lipid anchor is known to slow rotation significantly, which we exploited in order to resolve in time the rotational dynamics of proteins residing in the nanopore. Proteins in free solution generally have rotational diffusion coefficients, D_R, that are on the order of 10^6 to 10^7 rad^2 s^{-1},[98,99] while lipid anchored proteins have been shown to rotate over 2 orders of magnitude more slowly[87-89]. To account for this reduction in D_R when determining the theoretically expected value for each protein, we first estimated the value of D_R for each protein in free solution from its crystal structure using the software HydroPRO. Next, we multiplied each value by the

known value of D_R for GPI-AChE and divided by the theoretical estimate for GPI-AChE in free solution, thereby assuming that the rotation of each protein was slowed by tethering to the same degree. This information is described in the footnotes for Table 3-App.4. The resulting theoretical estimates of D_R are in good agreement with values measured in nanopore experiments (Fig. 3.4f), showing that D_R values of tethered proteins are indicative of their values in free solution.

We do not expect the tether itself to bias protein orientation and thereby affect the measurements of dipole moment, μ; however, the crosslinking reaction consumes a positively charged amine and thus will affect μ. To determine the extent by which crosslinking and removal of the positively charged amine affect μ, we modified the crystal structure for BSA (PDB ID: 3V03) by replacing a single, randomly-chosen lysine residue on the protein surface with a glycine residue and calculated μ for the modified protein using the Weizmann server (http://bioinfo.weizmann.ac.il/dipol/). We found that the median percent difference between μ for the native protein and 10 modified versions of the protein was roughly 12 percent and ranged from 1 to 38 percent. In line with this relatively small change, we observed good agreement between the values of μ determined in nanopore experiments and those measured with impedance spectroscopy (Fig. 3.4e).

3-App.S5 Simulating translocation events due to spheroidal particles

We numerically simulated translocation events due to spheroidal particles in MATLAB in order to provide support for the analysis methods developed in this work. Input parameters for the simulations included ΔI_{min}, ΔI_{max}, the dipole moment or μ, the rotational diffusion coefficient or D_R, pore geometry (*i.e.*, length and diameter), the

resistivity of the solution, the standard deviation of the noise, and the duration of each event or t_d. To generate an intra-event ΔI signal, we first simulated a spheroidal particle undergoing a biased random walk in one dimension by adapting the model developed by Gauthier and Slater for translational motion[100]. In our model, bias is introduced solely due to the electric field acting on the dipole moment of the particle, which was assumed to be pointed parallel to one of the principal axes. We simulated discrete 1-ns-long time steps in which the angle of the particle relative to the electric field, θ, changed by a fixed step size, $\Delta\theta = $ sqrt($2D_R\Delta t$). For each time step, the following equation gives the probability that the particle will move in the positive or negative direction:

$$p_{\pm} = \frac{1}{1 + e^{\Delta U / k_B T}} = \frac{1}{1 + e^{\pm E \mu \left[\cos(\theta - \Delta\theta) - \cos(\theta + \Delta\theta) \right] / (2 k_B T)}} \tag{3-App.16}$$

which was implemented in the simulations *via* the random number generator in MATLAB. Note that the change in potential energy, ΔU, is divided by a factor of 2 since the particle is initially located halfway in between the two possible final orientations. After simulating the entire event, we converted $\theta(t)$ to $\Delta I(t)$ based on equation (3-App.4) and sampled the signal at a rate of 500 kHz to mimic the sampling conditions of the real electronic recordings. Finally, we added Gaussian noise to the signal (unless indicated otherwise) and proceeded with analyzing these simulated signals in the same manner as the resistive-pulse signals obtained during an experiment.

Fig. 3-App.12 shows results from fitting the convolution model to a cumulative distribution of maximum ΔI values from simulated translocation events. The convolution model described the experimental data extremely well ($R^2 = 0.999$) and yielded estimates of the length-to-diameter ratio, m, and excluded volume, Λ, that were within 10% of their

expected values. As hypothesized in Section 3-App.S2, the distribution was biased toward ΔI_{max} more than expected (*i.e.*, based on the dipole moment only), which is likely a result of representing each event by its maximum value. These results suggest that fitting distributions of maximum ΔI values yields accurate estimates of shape and volume but not dipole moment.

Fig. 3-App.13 shows distributions of the length-to-diameter ratio, *m*, and excluded volume, Λ, determined from fitting the convolution model to simulated intra-event ΔI signals (analysis of intra-event ΔI values is presented in Section 3-App.6). The median values of *m* and Λ exactly match the expected values despite the relatively low signal-to-noise ratio of the data (SNR = $[I_{RMS, \, Signal} / I_{RMS, \, Noise}]^2$), which is lower than that observed in any of the experiments summarized in Fig. 3-App.15 wherein the signal-to-noise ratio was at least 1.4 and the noise was also Gaussian. These results suggest that the error in determining *m* and Λ from fitting experimental intra-event ΔI signals, as described in Section 3-App.S6, is not due to low signal-to-noise ratios. Furthermore, these results highlight the ability of the convolution model to account for the presence of noise.

Fig. 3-App.14 shows the distribution of μ and D_R that we obtained from analyzing simulated intra-event ΔI signals. These distributions were described well by a lognormal distribution ($R^2 > 0.98$) similar to our experimental results. The most probable value of μ determined from fitting each intra-event ΔI signal with the convolution model was in excellent agreement with the expected (*i.e.*, input) value over the range of values measured in this work (Fig. 3-App.14c). Similarly, the most probable value of D_R determined from analyzing each intra-event ΔI signal similarly was in agreement with the input value; however, our analysis methods systematically underestimated D_R by about

10 percent (Fig. 3-App.14d). This underestimation is likely due to a slight leveling off of the MSAD curve between the first two points (an example MSAD curve is shown in Fig. 3.4 in the main text), which might be rectified by increasing the sampling frequency of the signal. Regardless, these results suggest that the analysis methods developed in this work yield accurate estimates of the dipole moment and rotational diffusion coefficient of a spheroidal particle as long as its orientation is biased purely by its dipole moment.

We want to emphasize that these simulation results were not acquired by performing a simple backwards calculation. The data used here was simulated based on the probability of the particle rotating in one direction or another (equation (3-App.16)), and thus it is accomplished in a manner that is independent from the analysis methods described in equations (3-App.8) through (3-App.13).

3-App.S6 Analysis of intra-event ΔI values

Distributions of m and Λ determined from fitting intra-event ΔI values

Fig. 3-App.15 shows distributions of the length-to-diameter ratio, m, and excluded volume, Λ, determined from fitting the convolution model to all intra-event ΔI signals longer than 0.4 ms from experiments with IgG_1, GPI-AChE, Fab, BSA, α-amylase, and BChE. In general, the median value of m from each experiment corresponds to a shape that is more elongated than we expect (*i.e.*, the median values were less than expected for oblates and greater than expected for prolates), and the median value of Λ is lower than we expect based on the crystal structure of each protein and the results we obtained by analyzing distributions of maximum ΔI values (Table 3-App.1). The discrepancy between the values of these parameters may result from the shape of the nanopore, which the

111

model assumes is perfectly cylindrical (*i.e.*, constant diameter); however, the pore may have a varying diameter. The intra-event ΔI signal would be expected to reflect changes in pore diameter[101] and will include ΔI values from when the protein is in the widest regions of the pore. In contrast, the maximum ΔI value from each event most likely occurs when the protein is near the tightest constriction of the pore. One might also expect low ΔI values as the protein enters and exits the nanopore; however, the electric field is highly non-uniform and dense at the edges of the pore, which is thought to offset this effect or even result in larger than expected ΔI values[102,103]. In the current model, the effect of pore shape and the non-uniformity of the electric field near the pore entrance and exit are not considered and could result in lower than expected ΔI values. These low ΔI values would yield more elongated shapes and lower volumes than expected. If these hypotheses are true, this analysis could be improved by using pores that more closely match a perfect cylinder, by excluding ΔI values from the beginning and end of the signals, and by knowing the exact geometry of the nanopore in combination with an improved description of the electric-field in and around the pore. Regardless, the values m and Λ determined from fitting intra-event ΔI signals still can be used to identify and characterize proteins as evidenced by the repeatability between different experiments for IgG$_1$, Fab, BSA, and α-amylase.

Determining the dipole moment of a protein from fitting intra-event ΔI values

In the main text, we plotted the most probable value of the biasing parameter or dipole moment, μ, determined from fitting the convolution model to all intra-event signals longer than 0.4 ms for IgG$_1$, GPI-AChE, Fab, β-PE, G6PDH, L-LDH, BSA, α-

amylase, and BChE (Fig. 3.4e in the main text). Fig. 3-App.16 shows histograms of the values of μ that were returned from fitting each event in all experiments. In every case, the distribution of μ was described well by a lognormal distribution ($R^2 > 0.94$); we expected distributions of this shape based on simulations (see Section 3-App.S5). Moreover, the most probable value of μ in each distribution was indicative of the dipole moment of the protein. Only the permanent dipole moment of a protein biases its orientation inside the nanopore as the dipole moment induced by the electric field is roughly parallel to the field and hence does not affect the torque exerted on the protein[104]. The dipole moment estimates were in good agreement with measurements from dielectric impedance spectroscopy and calculations from crystal structures returned by the software HydroPro and the Weizmann server (Table 3-App.4). Dielectric impedance spectroscopy was performed as described previously[105] using a buffer of 1 mM KCl and 1 mM HEPES (pH = 7.4) for IgG_1 and Fab or 1 mM phosphate (pH = 5.2) for BSA. Moreover, these results were repeatable between different nanopores; the difference in the estimated dipole moment (*i.e.*, most probable values of μ) from experiments with different nanopores was always less than 20 percent, indicating that pore-dependent effects did not significantly bias the orientation of the protein.

Determining the rotational diffusion coefficient of a protein in a nanopore

To determine the rotational diffusion coefficient, D_R, of a protein during a translocation event, we first fit the convolution model to the intra-event ΔI signal at a bandwidth of 15 kHz to estimate ΔI_{min} and ΔI_{max}. Using these values, we determined the volume and shape of the protein as described in Section 3-App.S2; this procedure also

reveals the maximum and minimum shape factors of the protein based on equations (3-App.5) through (3-App.7). Using these values we calculated $\theta(t)$ based on equation (3-App.4). From this trajectory, we calculated the mean-squared-angular displacement (MSAD) of the protein using overlapping time intervals (*i.e.*, 0 to 4 μs, 2 to 6 μs, 4 to 8 μs, etc.). Since $\theta(t)$ can be "clipped" (*i.e.*, equation (3-App.4) yields imaginary values of $\theta(t)$ for ΔI values that are not between ΔI_{min} and ΔI_{max}), we only calculated angular displacement between two non-clipped values when computing the MSAD. By symmetry of the spheroid, multiple orientations of the particle are equivalent to θ in the range of 0 to $\pi/2$ (for example, the orientation of $3\pi/2$ is equivalent in this equation to the orientation of $\pi/2$). This degeneracy in the estimate of θ means that the trajectory of the MSAD will fail to describe the rotation of the protein accurately for long time scales; rather, the trajectory of $\theta(t)$ should be used only to estimate changes in θ over short time scales. This degeneracy, combined with the periodicity of rotation, causes the MSAD curve to level off asymptotically (see Fig. 3.4c in the main text for an example). Hence, we only fit the MSAD curve with a tangent line that passes through the origin to estimate the initial slope of the MSAD curve and reveal the rotational diffusion coefficient, D_R. According to the Langevin torque equation, D_R is equal to the initial slope of the MSAD curve divided by 2 for one-dimensional rotation[106]. Since filtering attenuates frequency components of the ΔI signal at which rotation occurs, we calculated D_R at various cut-off frequencies and fit this data with the logistic equation to estimate the value of D_R at infinite bandwidth, which corresponds to the upper horizontal asymptote of the fit (Fig. 3-App.18a shows an example). On average, these fits described the experimental data

extremely well ($R^2 > 0.96$). We calculated the overall bandwidth of the signal according to the following equation[107]:

$$f_c = \sqrt{1 / \left[1 / f_{c1}^2 + 1 / f_{c2}^2 \right]}$$

where f_{c1} is the cutoff frequency of the recording electronics (57 kHz)[32] and f_{c2} is the cutoff frequency of the digital Gaussian filter (ranges from 15 to 57 kHz).

Fig. 3-App.S18 shows histograms of the values of D_R that were returned from fitting all events longer than 0.4 ms for experiments with IgG$_1$, GPI-AChE, Fab, β-PE, α-amylase, and BChE. We excluded all other experiments from this analysis due to their relatively low signal-to-noise ratios, which yielded values of D_R that were erroneously high and similar to values obtained from analyzing signals consisting of only Gaussian noise (~50,000 rad^2 s^{-1}). As with the distributions of μ, each distribution of D_R was described well by a lognormal distribution ($R^2 > 0.96$), wherein the most probable value was in reasonable agreement with the expected rotational diffusion coefficient for each protein (Table 3-App.4). The rotational diffusion coefficient of the relatively flexible IgG$_1$ antibody was similar in two of the three nanopores; this result suggests that additional pore-dependent effects (*e.g.*, steric effects) not taken into account by this model might impact the rotation of proteins in a nanopore. Section 3-App.S5 shows results from simulated data that support the methods described in this section.

3-App.S7 Bivalently-bound IgG$_1$ rotates slower than monovalently-bound IgG$_1$

To provide additional evidence that ΔI values reflect the orientation of a non-spherical protein residing in the nanopore, we measured resistive-pulses resulting from

the translocation of anti-biotin IgG_1 bound to one or two biotin-PE lipids in the nanopore coating. Bivalently-bound IgG_1 should have reduced translational and rotational diffusion coefficients compared to monovalently-bound IgG_1 due to the additional drag associated with the second lipid anchor. To test this hypothesis, we performed an experiment in which the conditions initially favored bivalent binding of IgG_1 to the lipid coating, and gradually throughout the experiment, we changed the conditions to favor monovalent binding of IgG_1. To favor bivalent binding of IgG_1, we used a ratio of lipid-anchored biotin to IgG_1 that was 33-fold greater than that used in other experiments involving the same protein (*i.e.*, 2 nM IgG_1 and 1 mol% biotin-PE versus 10 nM IgG_1 and 0.15 mol% biotin-PE). To shift toward conditions favoring monovalent binding, we introduced soluble biotin at sequentially higher concentrations (1, 10, and 100 nM for 30 min each) to out-compete the lipid-anchored biotin in binding IgG_1, thereby reducing the fraction of bivalently-bound IgG_1 and increasing the fraction of monovalently-bound IgG_1 throughout the course of the experiment. IgG_1 proteins that were not bound to a lipid-anchored ligand were not detected[20].

To determine the *translational* diffusion coefficient of lipid-anchored IgG_1 in the presence of 0 and 100 nM of soluble biotin, we fit each distribution of translocation times with Schrödinger's first-passage probability density function. In the absence of soluble biotin wherein bivalent binding is favored, the translational diffusion coefficient was 1.05 nm^2 μs^{-1}, whereas in the presence of 100 nM of soluble biotin wherein monovalent binding is favored, the diffusion coefficient increased to 1.37 nm^2 μs^{-1}. This increase by a factor of 1.3 is in agreement with work by van Lengerich *et al.*, who previously estimated that a particle with a single lipid anchor should diffuse laterally about 1.5 times faster

116

than a particle with two lipid anchors[108]. This result supports our expectation that the ratio of bivalently-bound to monovalently-bound IgG_1 decreases with the concentration of soluble biotin. We also found that the charge of the protein-lipid complex changed from -3.25 in the absence of soluble biotin (*i.e.*, conditions favoring bivalent binding) to -1.53 in the presence of 100 nM of soluble biotin (*i.e.*, conditions favoring monovalent binding); this change in the value of the charge by -1.7 is slightly larger in magnitude than the theoretically expected value of -1 (the expected charge of one biotin-PE lipid) but this deviation is likely within the error of the measurement. The main aspect for the discussion here is that the negative charge *decreased* in magnitude as expected when fewer IgG molecules are bound bivalently.

We next obtained distributions of *rotational* diffusion coefficients (D_R) by analyzing intra-event ΔI values (see Section 3-App.S6) for IgG_1 in the presence of increasing concentrations of soluble biotin (Fig. 3-App.19). The most probable value of D_R increases with the concentration of soluble biotin as expected for conditions that favor monovalent over bivalent binding (Fig. 3-App.19c; Pearson's correlation coefficient $r = 1.00$). In the absence of soluble biotin, D_R was approximately 100 rad^2 s^{-1}. In the presence of 100 nM of soluble biotin, D_R increased more than an order of magnitude to 1,744 rad^2 s^{-1}, approaching the expected value for monovalently-bound IgG_1 of 4,500 rad^2 s^{-1} (see Table 3-App.4). As with the results for the translational diffusion coefficient, this trend indicates that the ratio of bivalently-bound to monovalently-bound IgG_1 decreases with the concentration of soluble biotin. Together, these results provide strong evidence that ΔI values reflect the rotational dynamics of the protein since we observed more than 17-

times faster rotational diffusion in the same nanopore as we change the experimental

conditions from favoring bivalent binding to favoring monovalent binding.

The distribution of maximum ΔI values is also affected by the ratio of

monovalently-bound to bivalently-bound IgG_1, as shown in Fig. 3-App.20. The

distribution becomes more biased toward low ΔI values as the fraction of bivalently-

bound IgG_1 increases, suggesting that bivalently-bound IgG_1 is less likely to sample

cross-wise orientations during a translocation event than monovalently-bound IgG_1. One

likely explanation for this result is that non-spherical proteins orient length-wise prior to

entering the pore due to the electric field gradient that they experience once they enter the

area surrounding the pore[102,109], and as a consequence of the reduced rotational diffusion

coefficient of bivalently-bound IgG_1 compared to monovalently-bound IgG_1, bivalently-

bound IgG_1 is less likely to reorient during an event of a given duration. Alternatively,

the bias toward low ΔI values might result from steric effects that limit crosswise

orientations since the second lipid anchor of bivalently-bound IgG_1 may restrict the

possible range of configurations the protein can assume for a given position of the first

lipid anchor. Regardless, these results strongly support the conclusion that ΔI reflects the

orientation and shape of non-spherical protein residing in the nanopore.

3-App.S8 Distinguishing an antigen and antibody-antigen complex in a single nanopore experiment

Fig. 3-App.22 illustrates the ability of the methods developed in this work to

characterize and identify a single protein, G6PDH, and a protein-protein complex,

G6PDH-IgG, in the same solution. Fig. 3-App.22a-i shows results from analysis of

maximum ΔI values (the procedures for this analysis are described in the figure caption). Fig. 3-App.22j-l shows results from analysis of all intra-event ΔI values.

To classify each translocation event as either G6PDH or G6PDH-IgG, we analyzed intra-event ΔI values as described in Section 3-App.S6 to determine the volume, shape, charge-related t_d value, rotational diffusion coefficient, and dipole moment from each protein or protein complex moving through the nanopore. This procedure identified 787 translocation events that were longer than 400 μs. We normalized the values for each parameter by their standard deviations and classified each event using the clustering algorithm *kmeans* in MATLAB[24,44]. Briefly, the *kmeans* clustering algorithm minimizes, across all clusters, the sum of the distance between all points in the cluster to the centroid of the cluster. To assess the quality of all cluster analyses and provide an error for the values assigned to parameters, we ran a bootstrap method in which 1,000 datasets were created by random resampling with replacement of the original dataset[110]. We then ran the cluster analysis on these 1,000 datasets. The clustering procedure was always robust with approximately 90% of the data (727 events) consistently being classified as either G6PDH or G6PDH-IgG (at least 95% of the time).

We performed the cluster analysis on several combinations of these five parameters and found that a 3-D cluster analysis based on the volume, dipole moment, and rotational diffusion coefficient provided the best separation between clusters as well as the most accurate characterization of the volumes for G6PDH (3% difference) and the G6PDH-IgG complex (7% difference). For instance, Fig. 3.5c in the main text shows that this technique determined a volume for G6PDH of 227 ± 9 nm^3 compared to the volume of 220 nm^3 determined from distributions of maximum ΔI values in an independent

experiment; similarly, this analysis determined the volume of the complex to be 530 ± 64 nm^3, and we expected a volume for the complex of 497 nm^3 (the volume of G6PDH plus the volume of an IgG protein). The volume of the complex determined from this intra-event analysis was also in excellent agreement with that determined from analysis of distributions of maximum ΔI values, which is shown in Fig. 3-App.22i. Furthermore, both the analysis of maximum ΔI values (Fig. 3-App.22f) and analysis of intra-event ΔI values followed by cluster analysis revealed that after the addition of anti-G6PDH IgG, the proportion of events due to the G6PDH-IgG complex was between 27 to 28 percent. The agreement between these two values provides additional evidence that the classification of events from single-event analysis was accurate. For reference, two-dimensional projections of the 3-D scatter plot in Fig. 3.5b of the main text are shown in Fig. 3-App.22j-l.

Prior to this work, the standard practice for distinguishing between proteins in a mixture would have been to analyze scatter plots of t_d values $vs.$ ΔI values. To illustrate the benefits of the multi-parameter characterization based on methods developed in this work, we performed a two-dimensional cluster analysis on the same data set used above, using only t_d values and average ΔI values. This analysis found that the protein complex represented only 2.5 ± 0.5 % of events, which is ~90% lower than the values determined by single-event analysis or analysis of distributions of maximum ΔI values (Fig. 3-App.22). Moreover, this analysis failed to determine the volume of the complex accurately as it returned a value of 833 ± 50 nm^3, which is 68% greater than the estimated volume of the complex of 497 nm^3 determined from independent experiments.

3-App.S9 Derivation of probability distribution of shape factors for proteins with a dipole moment

To derive a probability distribution of shape factors that takes into account a bias for a specific orientation based on the dipole moment of a protein and the electric field, we used the Boltzmann distribution of energies:

$$\frac{N_i}{N} = \frac{g_i \exp\left[-U_i \Big/ k_B T\right]}{\sum g_j \exp\left[-U_j \Big/ k_B T\right]} \tag{3-App.17}$$

where g_i is the number of states that have the same energy level, U_i is the energy level of state i, N_i is the number of molecules with energy level i, N is the total number of molecules in the system, and $k_B T$ is the thermal energy. The denominator of equation (3-App.17) is the partition function, and we will label it Z. Assuming that all of the energy affecting the orientation of the protein is in the form of the potential energy of a dipole in an electric field, then g_i is constant for all energy states and cancels out of equation (3-App.17). The potential energy of a dipole in an electric field is:

$$\Delta U = \vec{E} \cdot \vec{\mu} = -E\mu\cos(\phi) \tag{3-App.18}$$

where E is the electric field, μ is the dipole moment, and ϕ is the angle between the moment and the electric field. Combining equations (3-App.17) and (3-App.18), the proportion of molecules at an angle, ϕ, is:

$$\frac{N_\phi}{N} = \frac{1}{Z}\exp\left[\frac{E\mu\cos(\phi)}{k_B T}\right] \tag{3-App.19}$$

and therefore the probability of observing an angle ϕ is:

$$P_\phi = \frac{c}{Z} \exp\left[\frac{E\,\mu\cos(\phi)}{k_B T}\right] \tag{3-App.20}$$

where c is a normalization constant.

Considering a simple scenario in which the dipole moment is parallel with the symmetry or equatorial axis and accounting for the two possible orientations of the dipole moment relative to the electric field for a given orientation (*i.e.*, θ) due to symmetry, we obtained equations (3-App.21a) and (3-App.21b) for $\phi = \theta$ and $\theta + \pi/2$ from equation (3-App.20):

$$P_\theta = \frac{c}{Z} \exp\left[\frac{E\,\mu\cos(\theta)}{k_B T}\right] + \frac{c}{Z}\exp\left[\frac{E\,\mu\cos(\pi-\theta)}{k_B T}\right] = \frac{c}{Z}\cosh\left[\frac{E\,\mu\cos(\theta)}{k_B T}\right] \tag{3-App.21a}$$

and

$$P_\theta = \frac{c}{Z}\exp\left[\frac{E\,\mu\cos\left(\theta+\dfrac{\pi}{2}\right)}{k_B T}\right] + \frac{c}{Z}\exp\left[\frac{E\,\mu\cos\left(\pi-\left(\theta+\dfrac{\pi}{2}\right)\right)}{k_B T}\right] = \frac{c}{Z}\cosh\left[\frac{E\,\mu\sin(\theta)}{k_B T}\right] \tag{3-App.21b}$$

To express $\cos(\phi)$ in terms of the electrical shape factor we first rearranged equation (3-App.4), which describes γ as a function of θ, to obtain:

$$\cos(\theta) = \left(\frac{\gamma - \gamma_\perp}{\gamma_\parallel - \gamma_\perp}\right)^{1/2} \tag{3-App.22}$$

Substituting equation (3-App.22) into equations (3-App.21a) and (3-App.21b), we obtain:

$$P_{\theta} = \frac{c}{Z} \cosh \left[\frac{E \mu \left(\frac{\gamma - \gamma_{\perp}}{\gamma_{\parallel} - \gamma_{\perp}} \right)^{1/2}}{k_B T} \right]$$

(3-App.23a)

and

$$P_{\theta} = \frac{c}{Z} \cosh \left[\frac{E \mu \sin \left(\cos^{-1} \left[\left(\frac{\gamma - \gamma_{\perp}}{\gamma_{\parallel} - \gamma_{\perp}} \right)^{1/2} \right] \right)}{k_B T} \right] = \frac{c}{Z} \cosh \left[\frac{E \mu \left(\frac{\gamma - \gamma_{\parallel}}{\gamma_{\perp} - \gamma_{\parallel}} \right)^{1/2}}{k_B T} \right]$$

(3-App.23b)

Equations (3-App.23a) and (3-App.23b) express the probability of observing an angle θ as a function of the shape factor, $P(\theta(\gamma))$. As in the derivation by Golibersuch, we used substitution of variables to transform $P(\theta(\gamma))$ into $P(\gamma)$:

$$P(\gamma) d\gamma = P_{\theta} \frac{d\theta}{d\gamma} d\gamma = P_{\theta} \left(\frac{d\gamma}{d\theta} \right)^{-1} d\gamma$$

(3-App.24)

and differentiated equation (3-App.4) with respect to θ, $\frac{d\gamma}{d\theta}$. Substituting this result into equation (3-App.24), we obtained equations (3-App.25a) and (3-App.25b):

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

(3-App.25a)

and

123

$$P(\gamma)d\gamma = \frac{c}{Z}\cosh\left[\frac{E\,\mu\left(\dfrac{\gamma - \gamma_{\parallel}}{\gamma_{\perp} - \gamma_{\parallel}}\right)^{1/2}}{k_{B}T}\right]\left[\frac{1}{\pi\left[\left(\gamma - \gamma_{\perp}\right)\left(\gamma_{\parallel} - \gamma\right)\right]^{1/2}}\right]d\gamma \qquad \textbf{(3-App.25b)}$$

To solve for the normalization constants, we integrated equations (3-App.25a) and (3-App.25b) and set each equation equal to 1 (*i.e.*, $\int_{\gamma_{\parallel}}^{\gamma_{\perp}} P(\gamma)d\gamma = 1$). This procedure cancels

out the partition function *Z* and yields:

$$P(\gamma)d\gamma = \frac{1}{A}\cosh\left[\frac{E\,\mu\left(\dfrac{\gamma - \gamma_{\perp}}{\gamma_{\parallel} - \gamma_{\perp}}\right)^{1/2}}{k_{B}T}\right]\left[\frac{1}{\pi\left[\left(\gamma - \gamma_{\perp}\right)\left(\gamma_{\parallel} - \gamma\right)\right]^{1/2}}\right]d\gamma \qquad \textbf{(3-App.26a)}$$

and

$$P(\gamma)d\gamma = \frac{1}{A}\cosh\left[\frac{E\,\mu\left(\dfrac{\gamma - \gamma_{\parallel}}{\gamma_{\perp} - \gamma_{\parallel}}\right)^{1/2}}{k_{B}T}\right]\left[\frac{1}{\pi\left[\left(\gamma - \gamma_{\perp}\right)\left(\gamma_{\parallel} - \gamma\right)\right]^{1/2}}\right]d\gamma \qquad \textbf{(3-App.26b)}$$

where *A* is described by:

$$A = \int_{\gamma_{\parallel}}^{\gamma_{\perp}}\cosh\left[\frac{E\,\mu\left(\dfrac{\gamma - \gamma_{\perp}}{\gamma_{\parallel} - \gamma_{\perp}}\right)^{1/2}}{k_{B}T}\right]\left[\frac{1}{\pi\left[\left(\gamma - \gamma_{\perp}\right)\left(\gamma_{\parallel} - \gamma\right)\right]^{1/2}}\right]d\gamma \qquad \textbf{(3-App.27a)}$$

And

$$A = \int_{\gamma_{\parallel}}^{\gamma_{\perp}} \cosh\left[\frac{E\mu\left(\dfrac{\gamma - \gamma_{\parallel}}{\gamma_{\perp} - \gamma_{\parallel}}\right)^{1/2}}{k_B T}\right]\left[\frac{1}{\pi\left[(\gamma - \gamma_{\perp})(\gamma_{\parallel} - \gamma)\right]^{1/2}}\right]d\gamma \qquad \textbf{(3-App.27b)}$$

Equations (3-App.26a) and (3-App.26b) are identical to equations (3-App.12a) and (3-App.12b) in Section 3-App.S2.

Appendix Figures
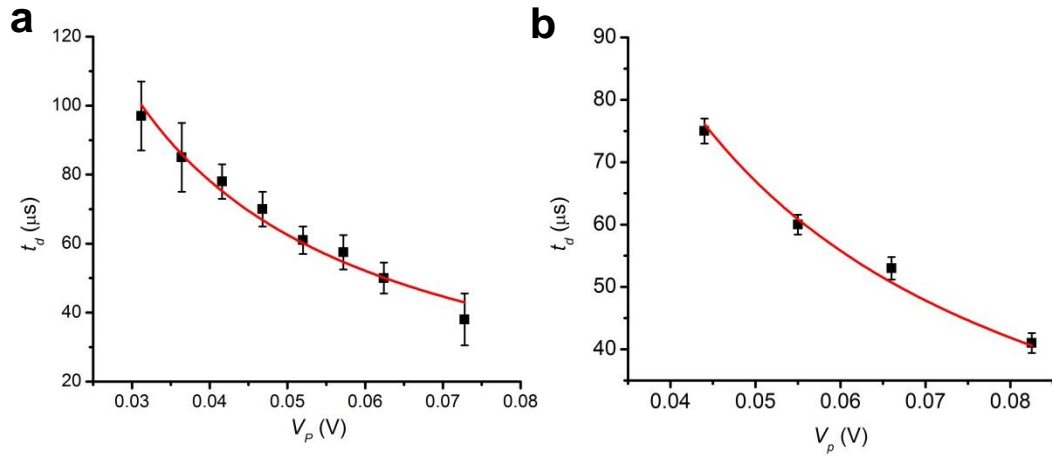
Figure 3-App.1 | Most probable t_d values for the monoclonal anti-biotin IgG$_1$ antibody (a) and GPI-AChE (b) as a function of the voltage drop, V_P, across a bilayer-coated nanopore containing biotin-PE. The inverse relationship between translocation time and applied voltage as well as the excellent agreement between theory (red curve) and experiment indicate that the lipid-anchored proteins completely passed through the nanopore. The red curve was obtained by a best-fit of equation $t_d = l_P^2 k_B T / (|z| e V_p D_L)$ as described in Yusko *et al.*[20], where the only fitting parameter is the net charge of the protein, *z*. l_P is the length of the nanopore with the bilayer coating, $k_B T$ is the thermal energy (1.38E-23 J K^{-1} × 295 K), V_p is the voltage drop across the nanopore, and D_L is the diffusion coefficient of the lipids in the bilayer as determined from FRAP experiments. For the IgG$_1$ antibody (a), the fit returned a value for *z* of -3.5 ± 0.1 (in 2 M KCl with pH = 7.4 in 10 mM HEPES) with $R^2 = 0.98$, *p-value* < 0.001 (*N* = 8), which is the expected value for the charge of this monoclonal antibody based on capillary electrophoresis experiments[20]. The value used for D_L was 1.35E-12 m^2 s^{-1} determined from FRAP experiments[20], and the value of l_P was 24 nm. For the GPI-AChE (b), the fit returned a value for *z* of -2.7 ± 0.1 (in 2 M KCl with pH = 6.1 in 10 mM HEPES) with $R^2 = 0.99$, *p-value* < 0.001 (*N* = 4). For comparison, the theoretical charge of GPI-AChE at zero ionic strength and pH 7.4 is -12 to -16[66,68]. D_L was 1.6 E-12 m^2 s^{-1} and l_P = 24 nm. The bilayer coating in (a) contained 0.15% biotin-PE, 0.8% Rh-PE, and ~99% POPC, and the bilayer coating in (b) contained only 0.8% Rh-PE, and ~99.2% POPC.
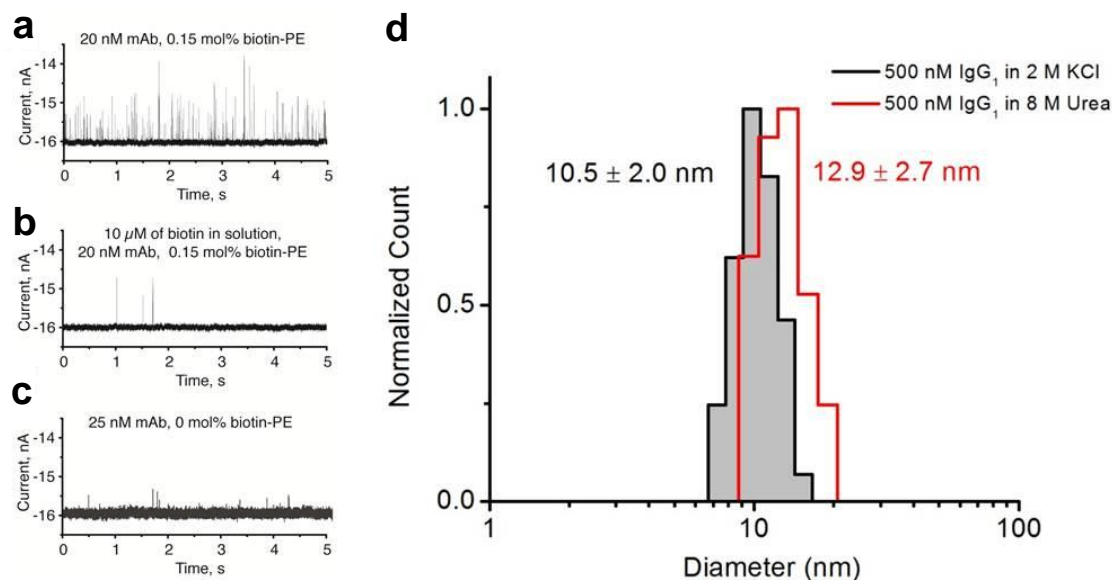
Figure 3-App.2 | Detection of monoclonal anti-biotin IgG$_1$ antibody with a bilayer-coated nanopore and dynamic light scattering experiments. a) Current *versus* time trace showing resistive pulses due to translocation of IgG$_1$ antibodies that were bound to biotin-PE lipids in the bilayer coating. Resistive pulses occurred at a frequency of 34 s^{-1}. b) Current *versus* time trace recorded after the addition of excess biotin (10 μM) to the solution and containing a reduced frequency of resistive pulses (1.3 s^{-1}). c) Current *versus* time trace recorded using the same nanopore as (a) and (b) but with a bilayer coating that did not contain biotin-PE lipids. Resistive-pulses occurred at a frequency of 2 s^{-1}. The experiments were performed using pore 2 (Fig. 3-App.25). d) Hydrodynamic diameter of IgG$_1$ antibodies determined from dynamic light scattering experiments. IgG$_1$ antibodies were at a concentration of 500 nM in aqueous solutions identical to the recording electrolyte (2 M KCl and 10 mM HEPES at pH = 7.4) during the dynamic light scattering experiment. Where indicated, 8 M of urea was added to the solution in order to denature all proteins. The dynamic light scattering results are the combination of 5 runs, each 60 s in duration. Results show the intensity-weighted calculation for the hydrodynamic diameter. The instrument was a Brookhaven 90Plus Particle Sizer and used a 658 nm laser at an angle of 90º to the detector. The absence of a second peak indicates that IgG$_1$ antibodies were not fragmented or present in dimers in 2 M KCl even at concentrations 500 fold greater than in the resistive-pulse sensing experiments.
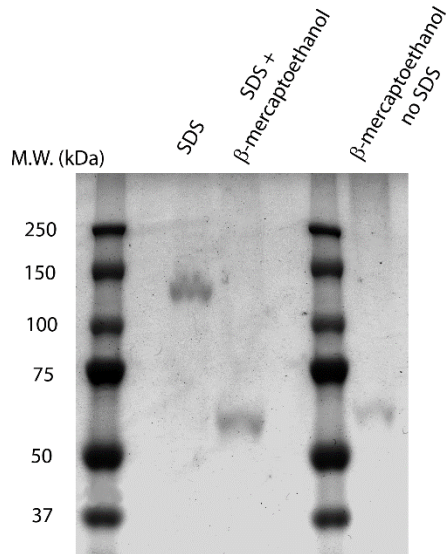
Figure 3-App.3 | Solutions containing GPI-AChE contained the dimeric, prolate shaped form of GPI-AChE. 2 μg of protein in Tris-Tricine sample buffer was added to each lane after treatment with 5% w/v SDS, 5% w/v SDS and 7.5 % v/v β-mercaptoethanol, or 7.5% v/v β-mercaptoethanol only. In the samples that contained SDS, the solution was heated to 95°C for 5 min to denature the protein. The gel was a 7.5% Tris-HCl TGX gel from BioRad, and the running buffer was Tris-Glysine buffer (25 mM Tris, 192 mM Glycine, 0.1% SDS). After running the gel, the gel was placed in 100 mL of deionized water and placed in the microwave for 30 s (careful not to boil the solution). The gel was rinsed twice for 3 to 5 min each time. The gel was then immersed in Coomassie staining solution (70 mg of Coomassie brilliant blue in 1 L of water; after 4 h, 3 mL of concentrated HCl was added) and heated in the microwave for 10s (again careful not to boil). The gel was left to stain overnight and destained with pure water[111].
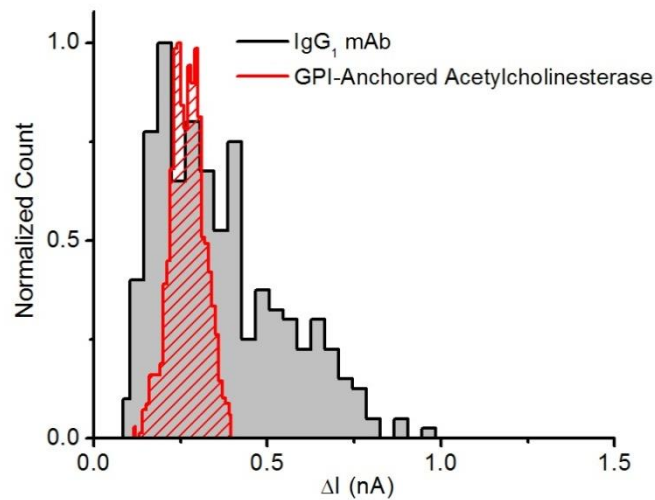
Figure 3-App.4 | Histograms of the Δ*I* values due to the translocation of the IgG$_1$ antibody (150 kDa) and GPI-anchored acetylcholinesterase (160 kDa) through the same nanopore. The experiments were performed using pore 3 (Fig. 3-App.25). Though both distributions are bimodal, the relatively narrow distribution of Δ*I* values due to GPI-anchored acetylcholinesterase compared to that of the IgG$_1$ antibody confirms that the large molecular weight of the IgG$_1$ antibody was not the reason for broadly distributed Δ*I* values. Currents were recorded at an applied potential difference of -100 mV.
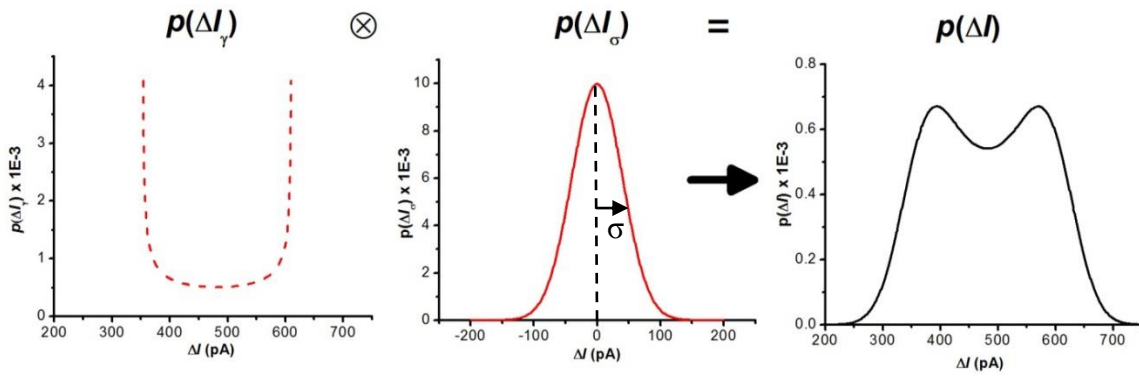
Figure 3-App.5 | Example convolution of the probability distribution of ΔI values one expects due to the distribution of shape factors, $p(\Delta I_\gamma)$ (equations (3-App.13a) and (3-App.13b)), and the error in determining individual ΔI values, $p(\Delta I_\sigma)$ (a Normal distribution function). The solution to the convolution is the probability distribution of ΔI values one expects to observe, $p(\Delta I)$. During the fitting procedure, the theoretical cumulative distribution, $p(\Delta I)$, is compared to the empirical cumulative distribution of ΔI values, $P(\Delta I)$, and the Levenberg-Marquardt non-linear least squares fitting algorithm in MATLAB generates new values for the fitting parameters ΔI_{min}, ΔI_{max}, μ, and σ, thereby creating new iterations of $p(\Delta I_\gamma)$ and $p(\Delta I_\sigma)$. This process repeats until the fit converges, which typically takes around 20 iterations.
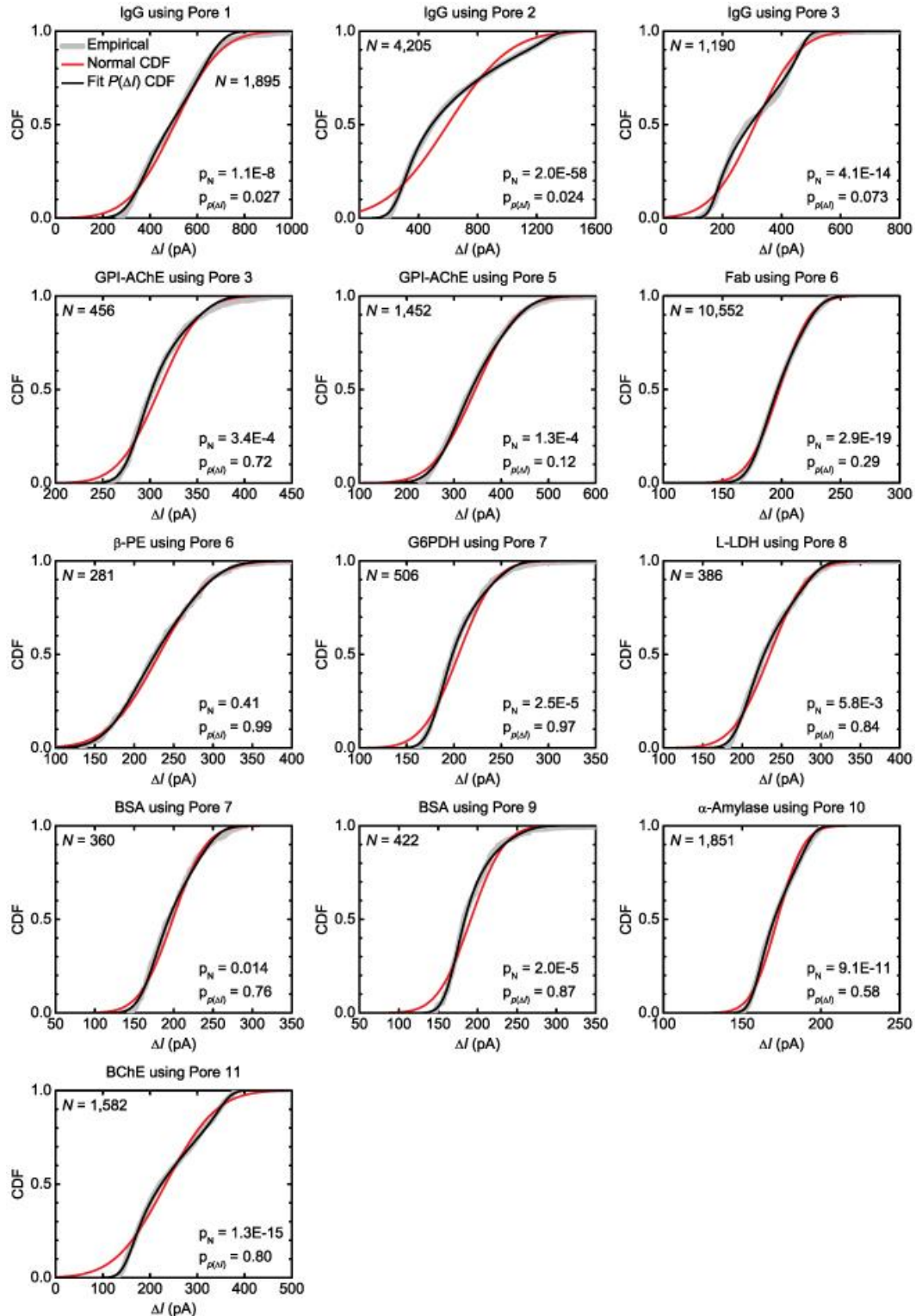
Figure 3-App.6 | Empirical cumulative distributions (grey curves) of ΔI values due to the translocation of non-spherical proteins compared to a best-fit Normal distribution (red curves) and the solution to the convolution model, *p*(Δ*I*) (black curves). In each case, Kolmogorov Smirnov (KS) tests were used to determine if the empirical distribution was different from the best-fit Normal distribution or $p(\Delta I)$. Resulting *p*-values are shown in the figure panels. In KS-tests, the null hypothesis is that the two

distributions are the same, and therefore, a p-value ≤ 0.05 indicates that the difference between two distributions is statistically significant at the $\alpha = 0.05$ level. For all of the non-spherical proteins except β-phycoerythrin, the distribution of ΔI values was different from a Normal distribution ($p_N < 0.05$). In contrast, the difference between the empirical distribution and convolution model, $p(\Delta I)$, was not statistically significant in 11 out of 13 cases (p-value ≥ 0.05).
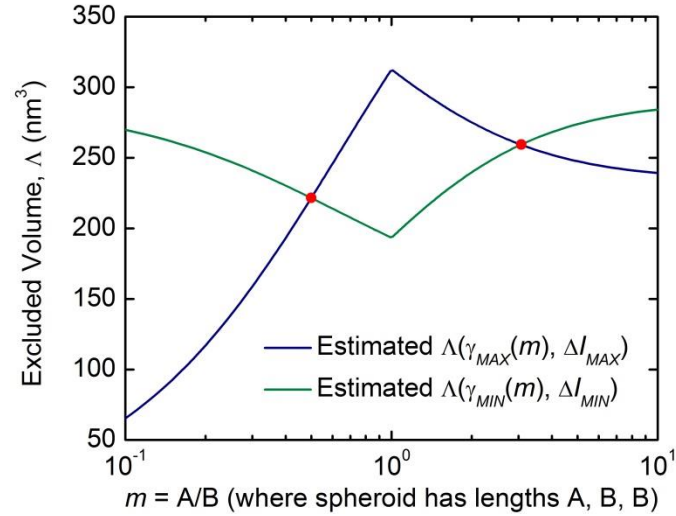
Figure 3-App.7 | Estimating the excluded volume as a function of *m* using $\Delta I_{\min}$ and $\Delta I_{\max}$ values illustrates that there are two solutions to equations (3-App.14) and (3-App.15) for prolate shaped proteins. This figure shows this result graphically by plotting the estimated volume of GPI-anchored acetylcholinesterase as a function of *m* for Pore 5. The two red dots indicate the two solutions to the system of equations ($m = 0.50$, $\Lambda = 222$ nm^3 and $m = 3.1$, $\Lambda = 259$ nm^3). In order to simplify the graph, we described the electrical shape factor with the notation γ_{MAX} or γ_{MIN}. We used this notation because for prolates ($m > 1$) $\gamma_{MAX} = \gamma_{\perp}$ and for oblates ($m < 1$) $\gamma_{MAX} = \gamma_{\parallel}$ (see equations (3-App.14) and (3-App.15)). The opposite is true for γ_{MIN}.
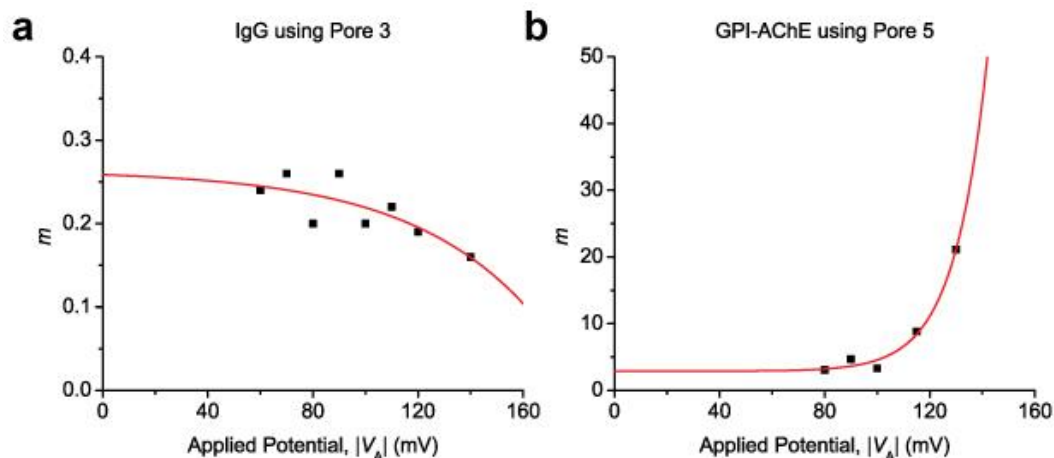
Figure 3-App.8 | The dependence of a protein's length-to-diameter ratio, *m*, on the applied potential, V_A, for IgG$_1$ (a) and GPI-AChE (b). We determined the value of *m* at different applied potentials by fitting the convolution model to distributions of maximum ΔI values. Interestingly, *m* is consistent at low potentials, while its value changes to indicate an increasingly elongated protein (*i.e.*, *m* approaches 0 for oblates or approaches ∞ for prolates) with increasing potential. To clearly illustrate this trend, we fit the results with an exponential growth function, $m = m_0 + A \cdot \exp\left(|V_A|/\tau\right)$ where A may be positive or negative. Considering that the fits asymptotically approached $m \approx 0.26$ for IgG$_1$ and $m \approx 2.9$ for GPI-AChE and the expected value of *m* is between 0.2 and 0.5 for IgG$_1$ and 2.9 for GPI-AChE (Table 3-App.4), this result suggests that low potentials yield accurate estimates of the protein shape.
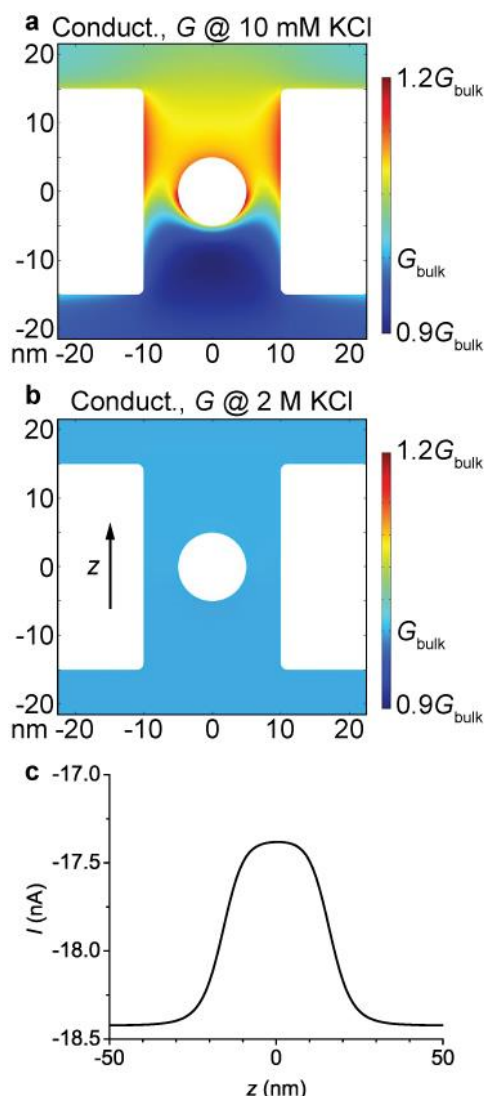
Figure 3-App.9 | Finite-element simulations indicate that local variations in the conductivity of the solution are negligible under the experimental conditions used in this work. We performed the simulations in COMSOL Multiphysics 4.4 (COMSOL Inc.). The electric field intensity inside the pore was set to 3 MV m^{-1}, the protein charge was set to -10, the charge density of the pore walls was set to 2 mC m^{-2} to account for the non-zwitterionic lipids in the nanopore coating, the protein diameter was set to10 nm, and the pore diameter and length were set to 20 and 30 nm, respectively. All boundary conditions were identical to those used by Lan *et al*.[80]. The upper semi-infinite boundary at z = 20 μm had a fixed negative potential relative to the lower boundary. a-b) 2-D heat maps showing the conductivity of the electrolyte solution throughout a vertical cross-section of the nanopore in the presence of (a) 10 mM KCl and (b) 2 M KCl. The color scale of each map was normalized to the conductivity in bulk solution, G_{bulk}. At low ionic strength, the conductivity varies significantly due to the accumulation and depletion of chloride ions on opposite sides of the protein; at high ionic strength as used in our experiments, this effect is essentially absent. c) A position-current (*I-z*) curve obtained by varying the position of the protein in the presence of 2 M KCl. The Δ*I* value of this curve is roughly 1.04 nA, which is in excellent agreement with the expected value of 1.00 nA obtained by using the volume exclusion model shown in equation (3-App.1). The near perfect symmetry of this curve further indicates that variations in conductivity are negligible at the high ionic strength used in this work.
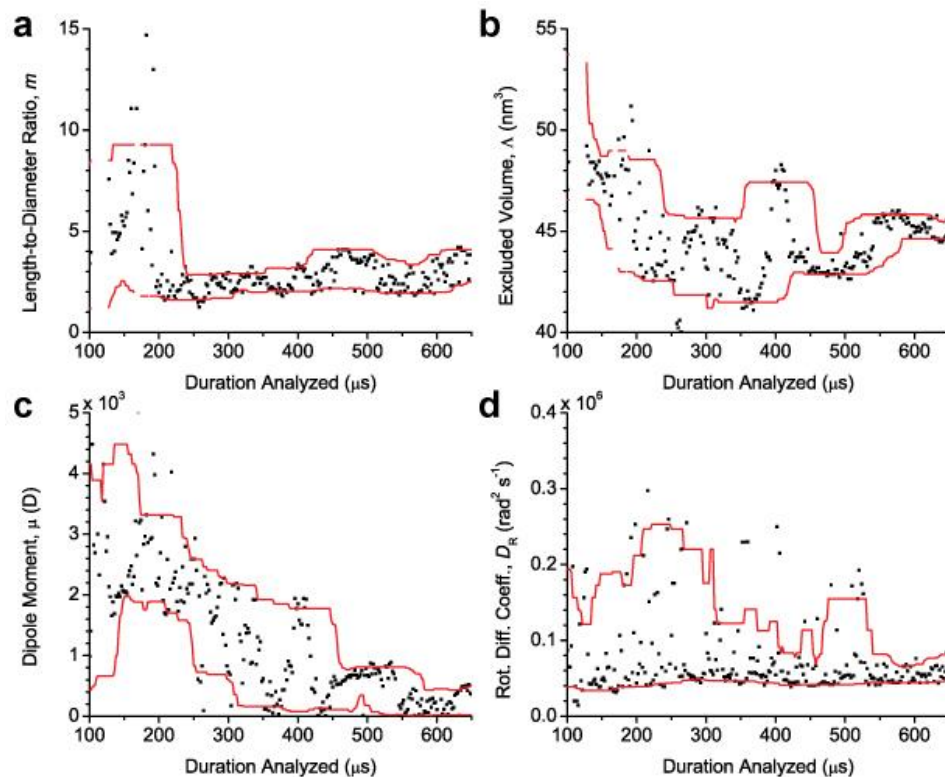
Figure 3-App.10 | Analysis of intra-event Δ*I* signals can yield parameter estimates in real-time. Measurements of the (a) length-to-diameter ratio, (b) excluded volume, (c) dipole moment, and (d) rotational diffusion coefficient obtained by progressively analyzing the current modulations (*i.e.*, intra-event Δ*I* values) of a single resistive-pulse resulting from the translocation of an individual anti-biotin Fab fragment. The red lines are moving 10th and 90th percentiles (smoothing window = 50 points). As the protein spends additional time in the pore, more data is acquired and analyzed; consequently, the spread in the determined parameter values narrows and the determined magnitudes of each parameter converge to their final values. The figure also shows that, for this particular event due to the translocation of a single anti-biotin Fab fragment, the variation in each parameter had narrowed to about 20% of its initial spread after approximately 550 μs before the end of the resistive-pulse. These results show that by analyzing a translocation event as it is occurring, it is possible to obtain parameter estimates while the protein still resides in the pore.

Figure 3-App.11 | Determining the charge of proteins by fitting translocation time distributions with a first-passage-time model. a-i) Histograms of translocation times from the nanopore experiments summarized in Table 3-App.1 (bin width = 15 μs). We fit each distribution with Schrödinger's first-passage probability density function

$$P(t_d) = \frac{l_P}{\sqrt{4\pi D_L t_d^3}} e^{-(l_P - v t_d)^2 / 4 D_L t_d} \quad \text{as}$$

described by Ling and Ling[112], where the electrophoretic drift velocity $v = |z| e V_P D_L / l_P k_B T$ as described by Yusko et al.[20] and the fitting parameters are the protein charge, z, and the diffusion coefficient of the lipids in the bilayer coating, D_L. We used a bin width of 2 μs when fitting the data, which corresponds to the sampling period of the current recordings. The most probable value of the translocation time is indicated by the dotted black line and corresponds to the maximum of the fit. The error in z is shown in parentheses next to its best-fit value, which we estimated by fitting the data with D_L fixed at its best-fit value ± standard error of the mean. j) Measured versus expected charges. Measured and expected values for anti-biotin IgG$_1$, anti-biotin Fab, and streptavidin were previously determined by Yusko et al.[20] via nanopore and capillary electrophoresis experiments, respectively (black squares). The expected value for BSA was acquired from literature[113] (green circle). The expected values for the remaining proteins were estimated from protein crystal structures via the PROPKA web interface (http://propka.ki.ku.dk/)[114-117] (blue triangles). GPI-AChE and BChE were excluded from this plot due to a lack of a reference value. For each protein that was covalently attached to the bilayer, we subtracted 0.93 from the expected value of charge to account for the reaction of a primary amine on the protein surface with an NHS ester on the crosslinker molecule to form an amide bond[118]. There is a strong positive correlation ($r = 0.95$) between the measured and expected values; however, the measured values are systematically lower in magnitude than the expected values. This underestimation may be due to inaccuracies in the PROPKA method or the high ionic strength of the recording solution used in nanopore experiments.


137
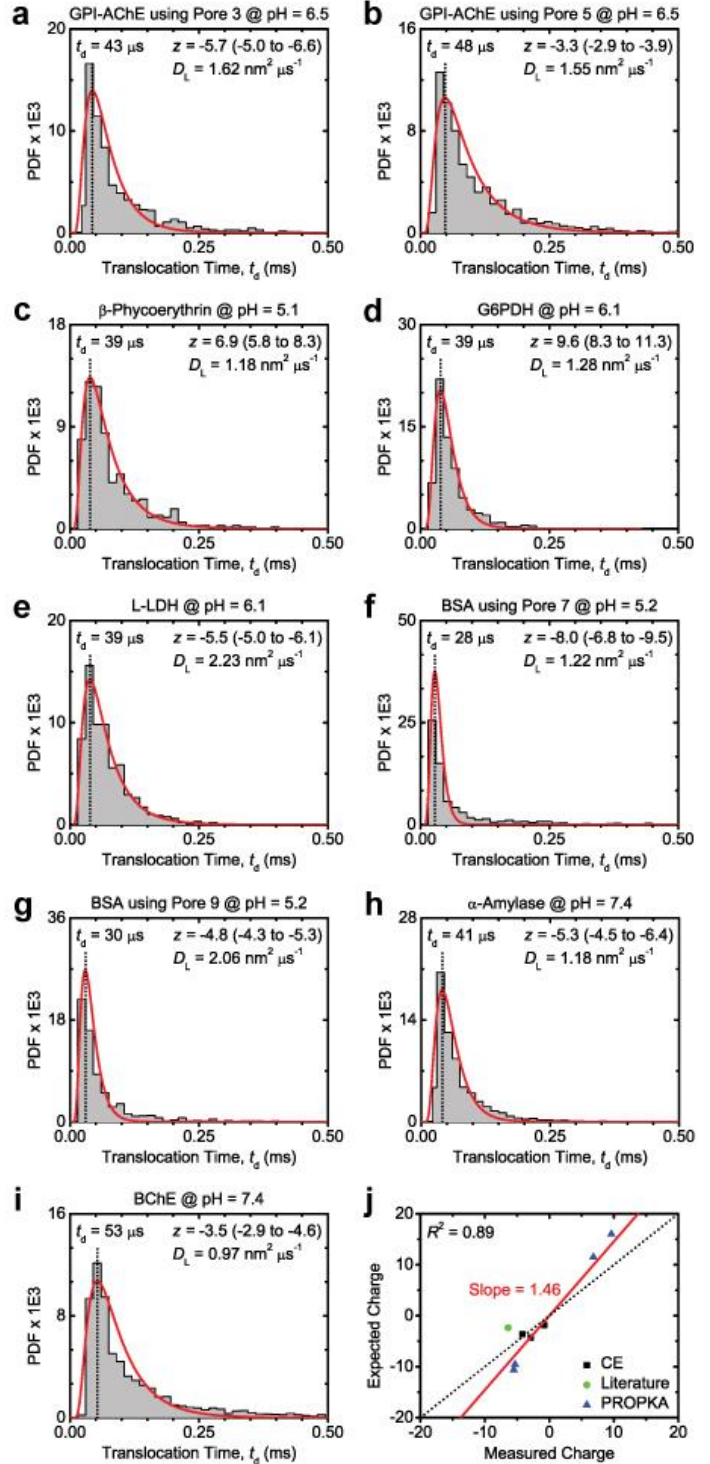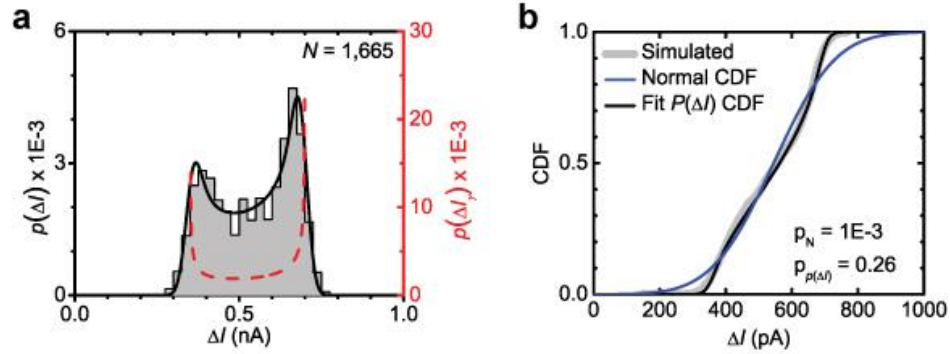
Figure 3-App.12 | Distributions of maximum ΔI values from simulated translocation events. a) A histogram of maximum ΔI values from simulated translocation events. The black curve shows the solution of the convolution model, $p(\Delta I)$, after a non-linear least squares fitting procedure, and the red dashed curve shows the estimated distribution of ΔI values due to the distribution of shape factors, $p(\Delta I_\gamma)$. b) The cumulative distribution of the same data shown in (a) (grey curve) compared to a best-fit Normal distribution (blue curve) and the solution to the convolution model (black curve). p-values shown in the figure resulted from Kolmogorov Smirnov (KS) tests that compared the simulated, empirical cumulative distribution to the model Normal distribution, p_N, or the convolution model, $p_{p(\Delta I)}$. Since, the value p_N was less than 0.05, the KS-test indicated that the distribution was not Normal at the $\alpha = 0.05$ level. In contrast, the value of $p_{p(\Delta I)}$ was greater than 0.05 and therefore not significantly different from the convolution model; this result indicates that the model describes the empirical distribution well. For the simulations, we used input parameters that were based on the experiment done with IgG_1 in pore 1 (*e.g.*, ΔI_{min} and ΔI_{max} were 329 and 678 pA, corresponding to values of m and Λ of 0.37 and 292 nm^3). We simulated 2,000 events with translocation times that were sampled from Schrödinger's first-passage probability density function[112]. The signal processing algorithm detected 1,922 events wherein 1,665 of these events were fully time resolved (*i.e.*, $t_d > 50$ μs). From the fit, we calculated values for m and Λ of 0.38 (2.7% greater than the expected value) and 310 nm^3 (6.2% greater than the expected value).
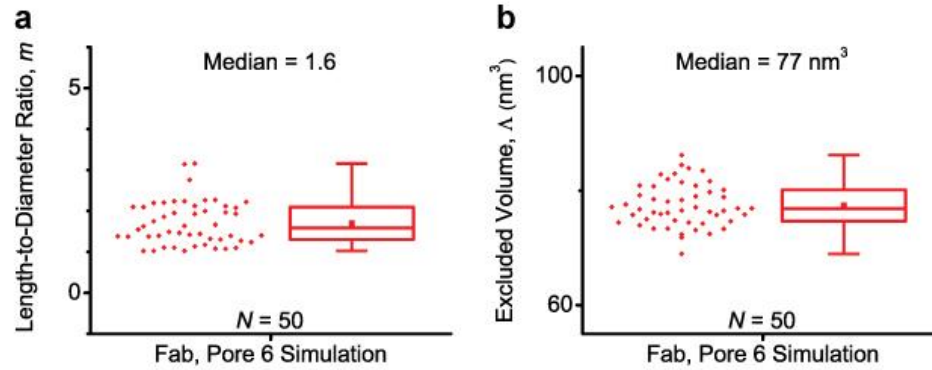
Figure 3-App.13 | Distributions of the length-to-diameter ratio, *m* (a), and excluded volume, Λ (b), determined from fitting the convolution model to simulated intra-event Δ*I* signals. The box represents the 1st and 3rd quartiles of the data, the horizontal line is the median value, the point inside the box shows the mean value, and the whiskers extend to data points that are within $1.5 \times$ IQR. For the simulations, we used input parameters that were based on the experiments done with Fab in pore 6 (*e.g.*, ΔI_{min} and ΔI_{max} were 178 and 231 pA, corresponding to values of *m* and Λ of 1.6 and 77 nm^3). The data was low-pass filtered at 15 kHz. The standard deviation of the noise added to each signal was 26.5 pA, while the standard deviation of the intra-event Δ*I* signals was typically around 26.8 pA, corresponding to a signal-to-noise ratio of roughly 1.02. The duration of each event was 1 ms.
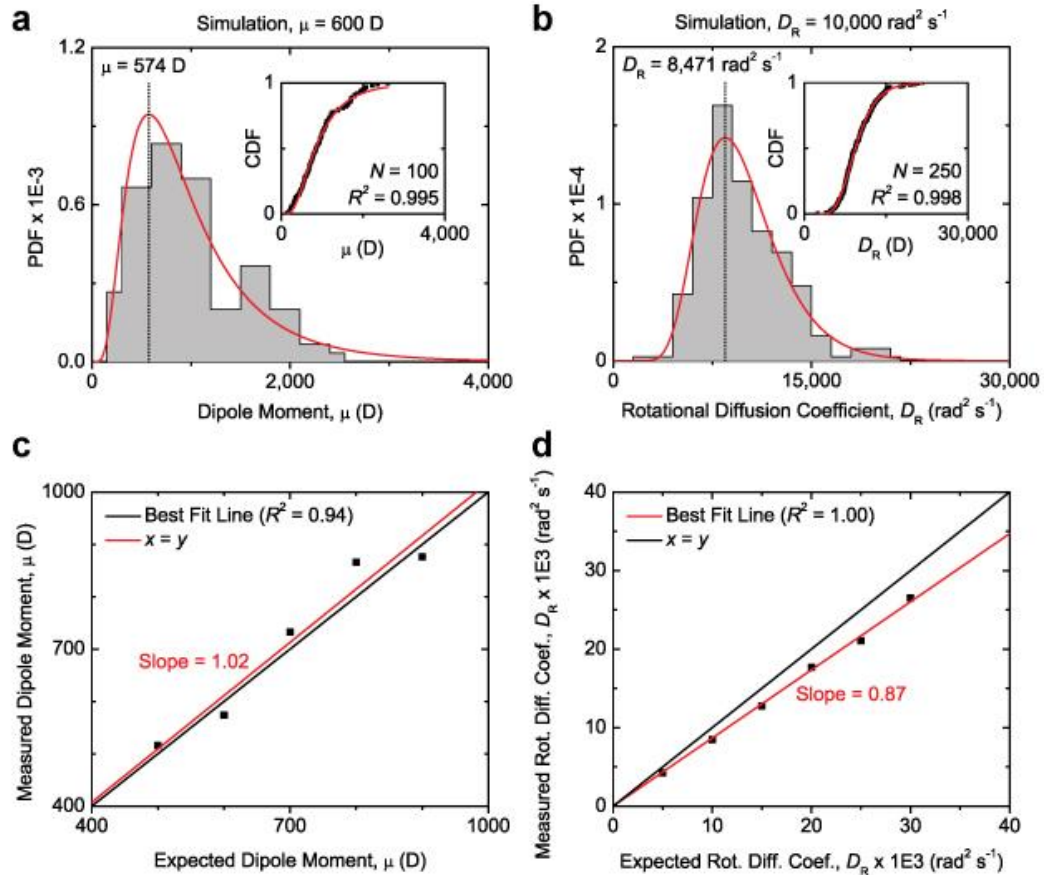
Figure 3-App.14 | Dipole moments, μ, and rotational diffusion coefficients, D_R, determined from analyzing simulated translocation events due to spheroidal particles. a-b) Distributions of dipole moments and rotational diffusion coefficients determined from analyzing simulated, 1-ms-long translocation events. The inset in each plot shows the empirical cumulative distribution (black squares) fit with a lognormal cumulative distribution function (CDF) (red line). KS-tests indicated the difference between the empirical distribution and best-fit curve was not significant in all cases at a confidence level of α = 0.10. The derivative of the CDF is the probability density function (PDF), which is plotted in red with the histograms of dipole moments and rotational diffusion coefficients. The most probable value is indicated by the dotted black line and corresponds to the maximum of the lognormal fit. c-d) Measured *versus* expected (*i.e.*, input) dipole moments and rotational diffusion coefficients. The ideal outcome wherein the measured values are equal to the input values is shown in black and the best fit line is shown in red. For the simulations where we varied μ, we used input parameters that were based on the experiment done with IgG_1 in pore 1. For the simulations where we varied D_R, we kept μ fixed at 500 Debyes and did not add noise to the signal; lower signal-to-noise ratios resulted in additional error as expected. Furthermore, we calculated D_R at bandwidths ranging up to 60 kHz for each event to determine the value of D_R at infinite bandwidth, as illustrated in Fig. 3-App.18a.

Figure 3-App.15 | Distributions of the length-to-diameter ratio, *m*, and excluded volume, Λ, determined from fitting the convolution model to all intra-event Δ*I* signals longer than 0.4 ms for IgG$_1$ (a-b), GPI-AChE (c-d), Fab (e-f), BSA (g-h), α-Amylase (i-j), and BChE (k-l). The box represents the 1st and 3rd quartiles of the data, the horizontal line is the median value, the point inside the box shows the mean value, and the whiskers extend to data points that are within 1.5 × IQR. Only prolate solutions are shown for GPI-AChE, Fab, and α-amylase. The applied potential was -60 mV for IgG$_1$ using pore 3, -115 mV for GPI-AChE, and -100 mV for all other experiments.
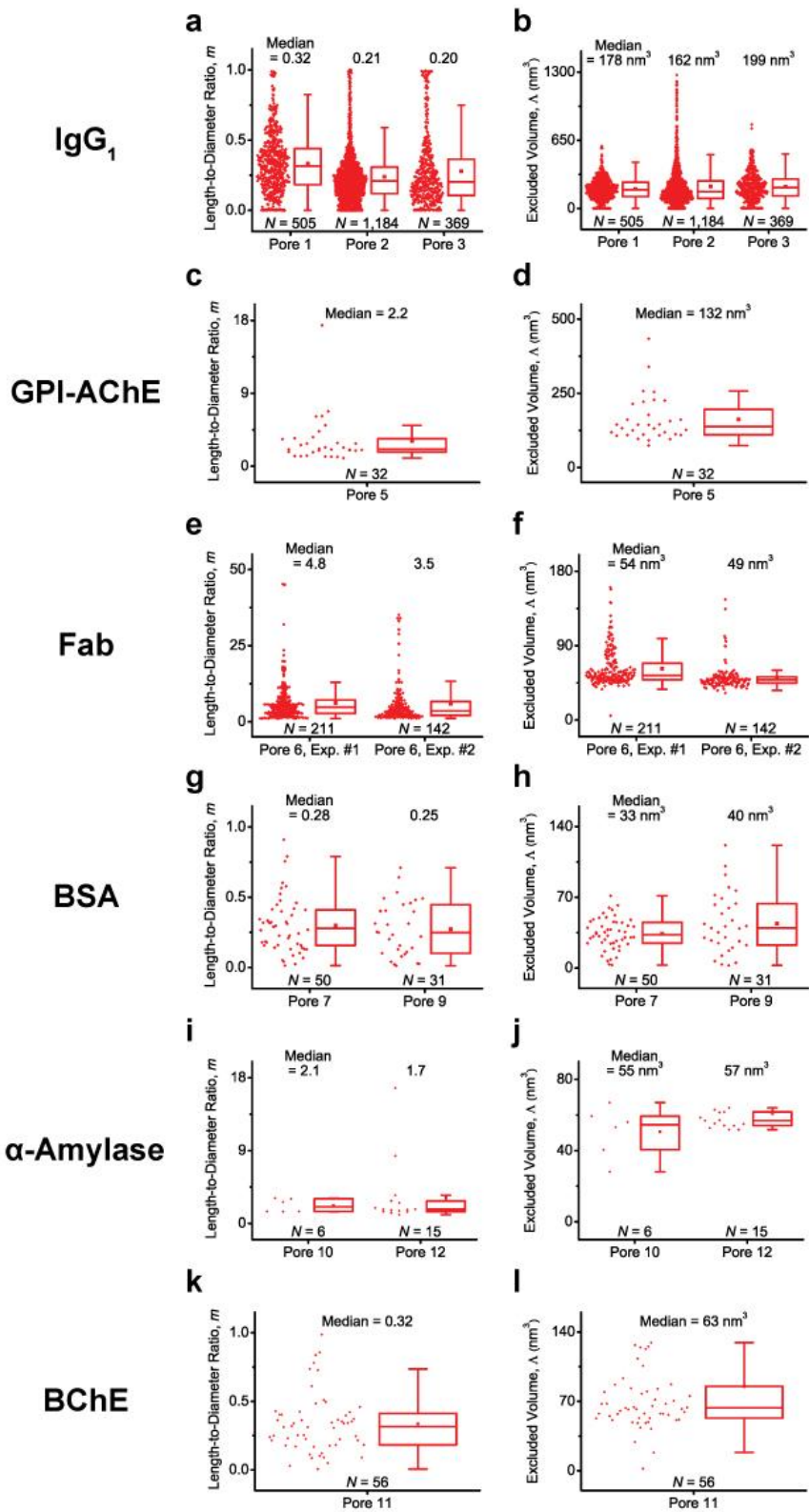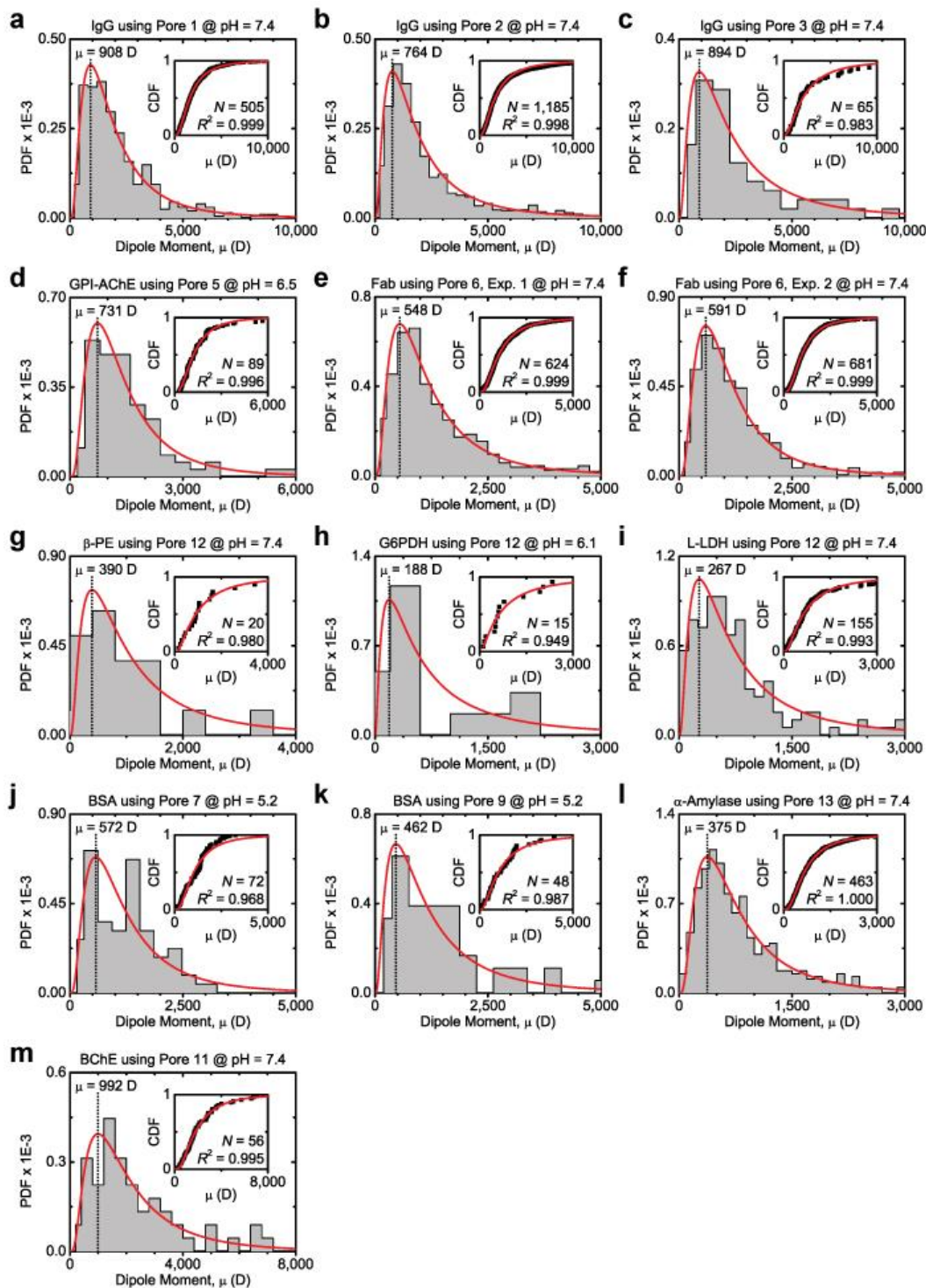
Figure 3-App.16 | Dipole moments, μ, of IgG₁ (a-c), GPI-AChE (d), Fab (e-f), β-PE (g), G6PDH (h), L-LDH (i), BSA (j-k), α-amylase (l), and BChE (m) determined from fitting intra-event ΔI values with the convolution model.. The inset in each plot shows the empirical cumulative distribution (black squares) and corresponding fit with a lognormal cumulative distribution function (CDF) (red line). KS-tests indicated the difference between the empirical distribution and best-fit curve was not significant in all cases at a confidence level of α = 0.10. The derivative of the CDF is the probability density function (PDF), which is plotted in red with the histogram of dipole moments. The most probable value of the dipole

moment is indicated by the dotted black line and corresponds to the maximum of the lognormal fit. During the fitting procedure, only events with durations greater than 0.4 ms were analyzed. The applied potential was -100 mV for all experiments with IgG$_1$, Fab, G6PDH, BSA, and BChE; -115 mV for the experiment with GPI-AChE; -140 mV for the experiment with β-PE; and -200 mV for the experiments with L-LDH and α-amylase.
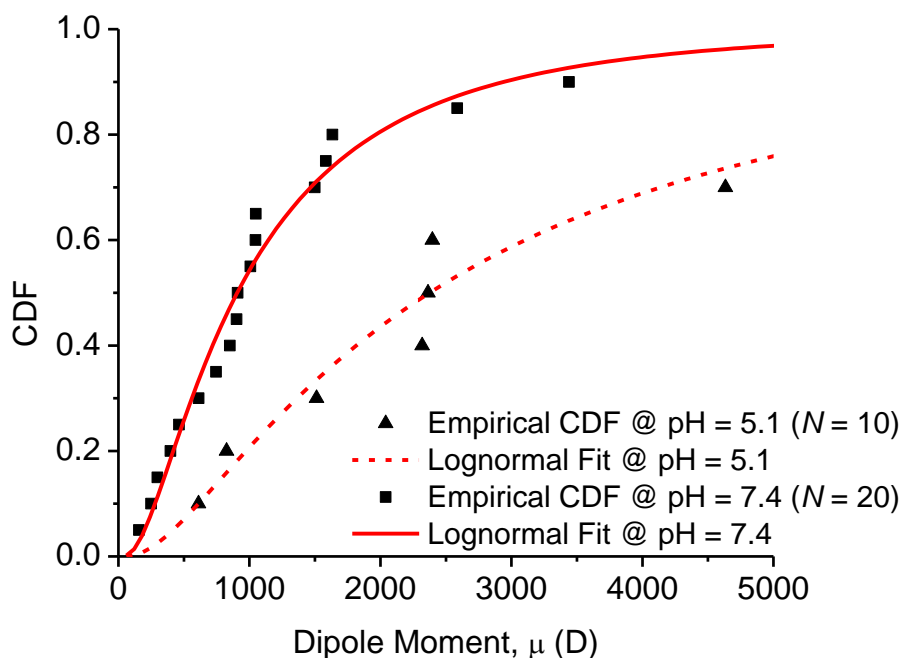
Figure 3-App.17 | Variation of the dipole moment, μ, of β-phycoerythrin (β-PE) as a function of pH.. Empirical cumulative distributions of experimentally-determined dipole moments of β-PE (black points) and corresponding lognormal fits (red curves) shown for pH 5.1 and 7.4. Based on theory, β-PE's expected values of μ are 395 D at pH 7.4 and 489 D at pH 5.1. The most probable values of the lognormal fits determined by fitting data from nanopore experiments are 390 D at pH 7.4 and 774 D at pH 5.1 and hence show the same trend. The theoretical estimates and experimentally-determined values are in reasonable agreement considering the uncertainty in theoretically estimating dipole moments at pH values different from pH 7. KS-tests indicated that the two data sets were lognormal and statistically different from each other. The applied potential was -100 mV.
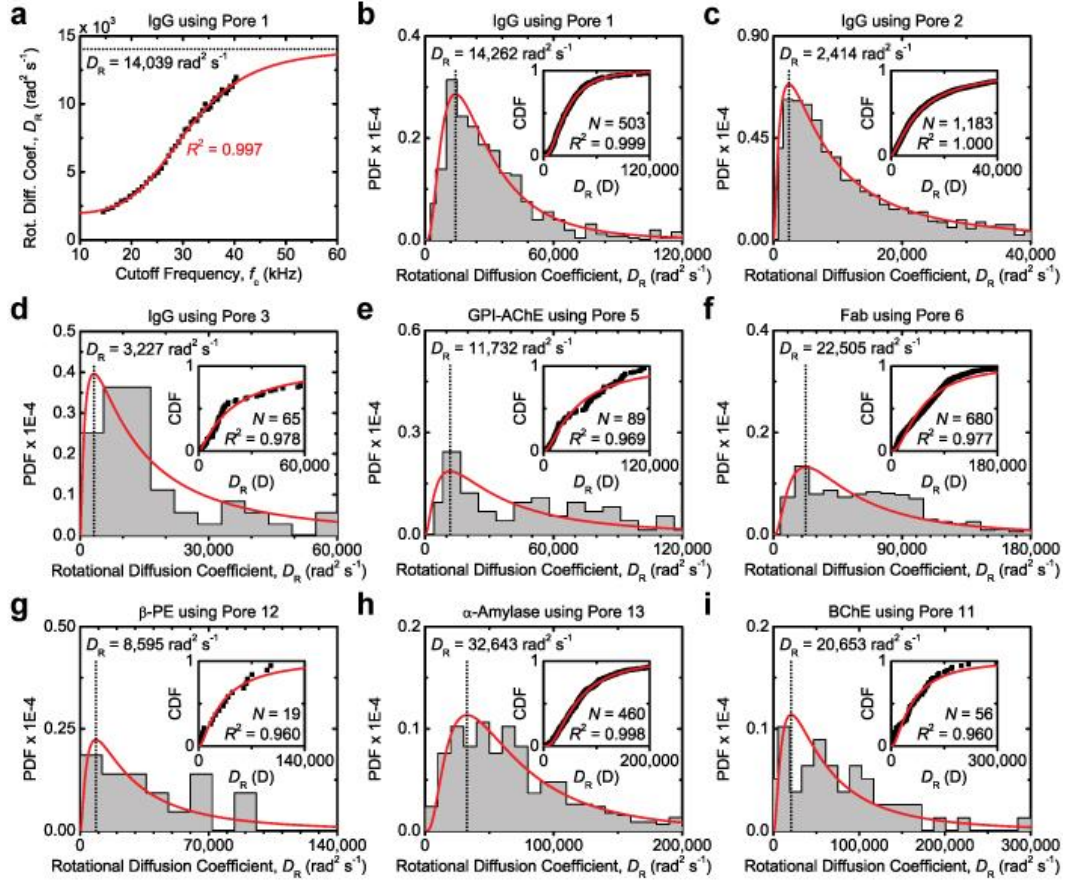
Figure 3-App.18 | Rotational diffusion coefficients, D_R, of IgG$_1$ (a-d), GPI-AChE (e), Fab (f), β-PE (g), α-Amylase (h), and BChE (i) determined from analysis of intra-event Δ*I* values. a) Rotational diffusion coefficient *versus* the low-pass cutoff frequency for a single event due to IgG$_1$ in pore 1. The curve was fit with the logistic equation to determine D_R at infinite bandwidth, which is denoted by the dotted black line. We used this procedure to determine the values of D_R for all proteins and subsequently generate the histograms in panes (b) through (i). b-i) The inset in each plot shows the empirical cumulative distribution (black squares) fit with a lognormal cumulative distribution function (CDF) (red line). KS-tests indicated the difference between the empirical distribution and best-fit curve was not significant in every case except for the experiment with Fab using Pore 6 (panel (f)) at a confidence level of α = 0.10. The derivative of the CDF is the probability density function (PDF), which is plotted in red with the histogram of rotational diffusion coefficients. The most probable value of the rotational diffusion coefficient is indicated by the dotted black line and corresponds to the maximum of the lognormal fit. Only events with durations greater than 0.4 ms were analyzed. The applied potential was -100 mV for all experiments with the IgG$_1$ antibody, Fab, and BChE; -115 mV for the experiment with GPI-AChE; -140 mV for the experiment with β-PE; and -200 mV for the experiment with α-amylase.
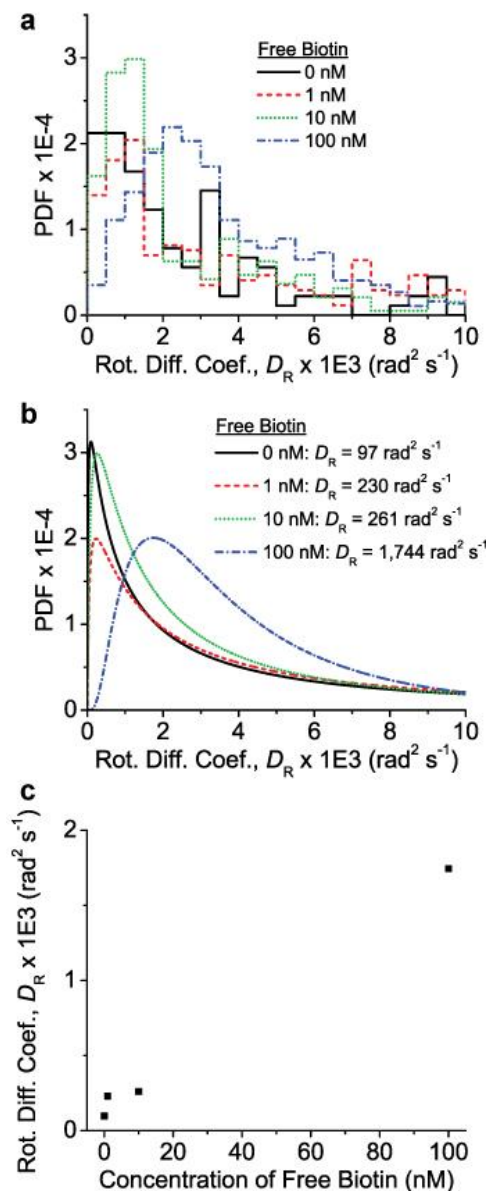
Figure 3-App.19 | The measured rotational diffusion coefficient, D_R, of lipid-anchored IgG$_1$ decreases with the ratio of bivalently-bound to monovalently-bound IgG$_1$. a) Histograms of D_R values determined from analysis of intra-event ΔI values. We analyzed 179, 343, 382, and 739 events obtained in the presence of 0, 1, 10, and 100 nM free biotin, respectively. b) Lognormal fits of the D_R distributions. We fit each empirical cumulative distribution with a lognormal cumulative distribution function (not shown). The R^2 values of the fits are 0.989, 0.998, 0.987, and 0.996 for 0, 1, 10, and 100 nM free biotin, respectively. The legend displays the most probable value of each fit. c) The most probable value of D_R as a function of the concentration of free biotin. We attribute the increase in D_R with the concentration of free biotin to a decrease in the ratio of bivalently-bound to monovalently-bound IgG$_1$.
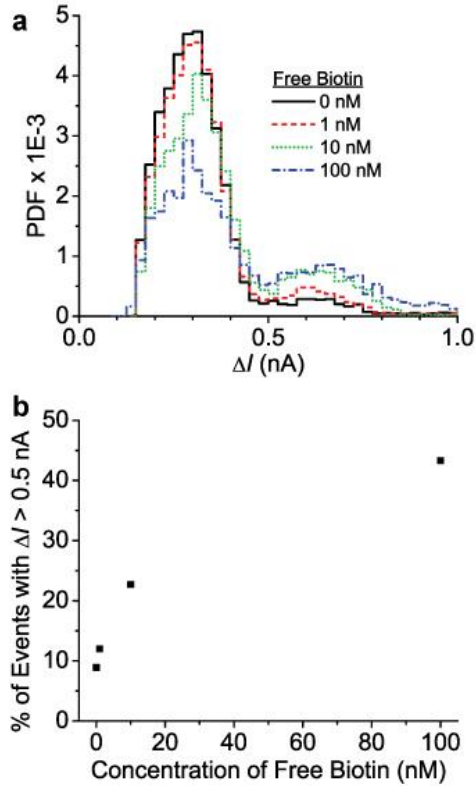
Figure 3-App.20 | The distribution of maximum ΔI values for IgG$_1$ is more biased toward low values when the fraction of bivalently-bound IgG$_1$ is relatively high. a) Histograms of maximum ΔI values from resistive-pulse recordings obtained in the presence of 0, 1, 10, and 100 nM free biotin. b) The percentage of events with a maximum ΔI value greater than 0.5 nA as a function of the concentration of free biotin. The distribution of maximum ΔI values becomes less biased toward low values as the concentration of free biotin increases (*i.e.*, the ratio of monovalently-bound to bivalently-bound IgG$_1$ increases).
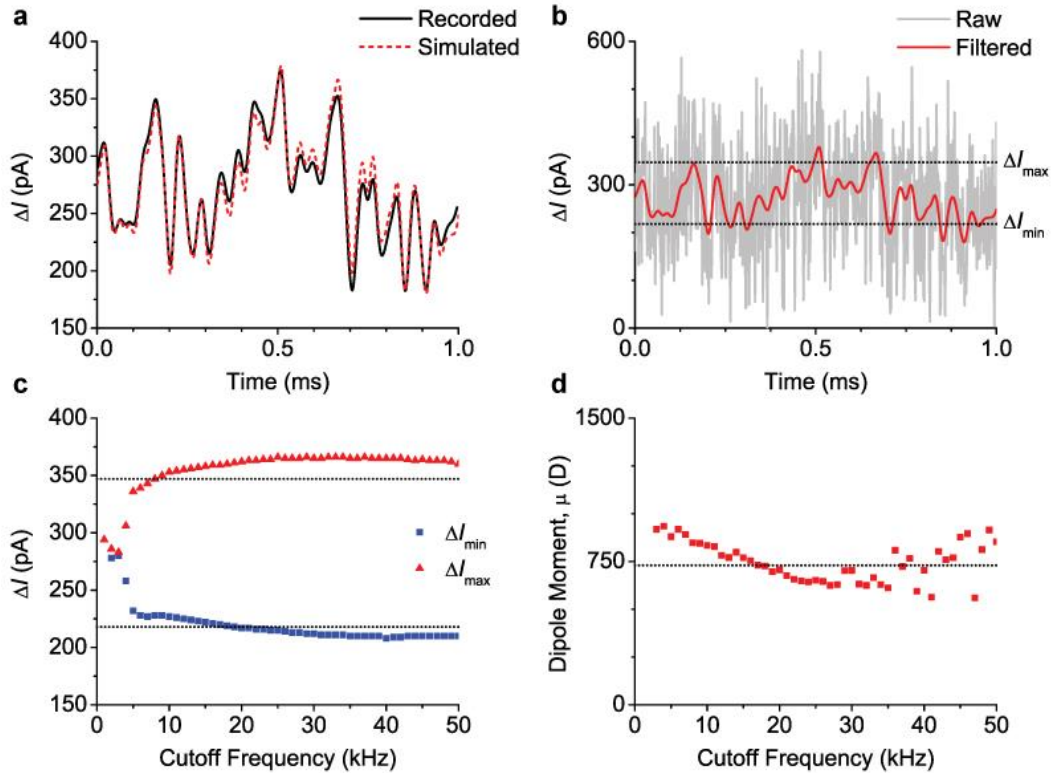
Figure 3-App.21 | Effect of the recording electronics and low-pass filtering on intra-event ΔI values.
a) A comparison between a simulated intra-event ΔI signal that was filtered digitally at 15 kHz (dashed red curve) and a waveform obtained by inputting the unfiltered simulated signal into the experimental setup using a function generator, recording at 500 kHz, and filtering digitally at 15 kHz (black curve). The two curves are nearly identical with an average difference of less than 0.3 pA and a Pearson correlation coefficient of 0.98, indicating the recording electronics do not significantly distort the signal at a bandwidth of 15 kHz. Any deviation between the two curves likely results from additional noise introduced by the recording setup. b) The same intra-event ΔI signal from (a) before and after filtering digitally at 15 kHz (gray and red curves, respectively). The dotted black lines show the known values for ΔI_{min} and ΔI_{max}. Although filtering smoothes the signal (and dramatically reduces the noise), the filtered signal still samples ΔI_{min} and ΔI_{max} and maintains its bias toward ΔI_{min}. Consequently, fitting the filtered signal with the convolution model still yields accurate values for ΔI_{min}, ΔI_{max}, and μ as shown in (c) and (d). c) Values of ΔI_{min} and ΔI_{max} determined from analyzing the same simulated intra-event ΔI signal at different cutoff frequencies. The dotted black lines show the known values for ΔI_{min} and ΔI_{max}. The values do not vary considerably with cut-off frequency except at low frequencies (<5 kHz). For instance, there is a 5.6% difference between the values of ΔI_{min} at 15 and 50 kHz and a 0.6% difference between the values of ΔI_{max} at 15 and 50 kHz. d) Values of the dipole moment, μ, determined from analyzing the same simulated intra-event ΔI signal at different cutoff frequencies. The dotted black line shows the known value for μ. Dipole moment has little dependence on cut-off frequency (*e.g.*, there is a 10.2% difference between the values at 15 and 50 kHz), although the results are scattered to a greater degree at high frequencies likely due to a decrease in the signal-to-noise ratio with cut-off frequency. For the simulations, we used input parameters that were based on the expected values for GPI-AChE (see Table 3-App.4).

Figure 3-App.22 | Determining the volume and shape of an antibody-antigen complex from individual resistive-pulses. a) Current trace showing resistive pulses due to the translocation of G6PDH in the absence of antibody. b) Current trace recorded after incubation with 15 μM polyclonal anti-G6PDH IgG for 1 hr. After incubation, we rinsed the chip with recording buffer to remove unbound IgG. c-d) Histograms of maximum ΔI values recorded before and after incubation with anti-G6PDH IgG. Insets show the same data over a reduced y-axis scale. We observed a significant increase in the number of events with large ΔI values after incubation with IgG (*e.g.*, the percentage of events with values larger than 500 pA increased from 0.01 to 9 percent). e) Empirical cumulative distribution (CDF) of ΔI values due to the translocation of G6PDH (grey curve) and the fit of this data to the convolution model (black curve). f) Empirical CDF of ΔI values due to the translocation of *both* G6PDH and the antibody-antigen complex (red dotted curve). To generate a CDF due to the translocation of the complex *only* (*i.e.*, remove ΔI values due to the translocation of unbound G6PDH), we subtracted the CDF due to the translocation of G6PDH *only* (e) after scaling this distribution such that the difference between the two empirical CDFs was minimized at low ΔI values (250 to 350 pA). We expect the majority of ΔI values in this range to result from the translocation of unbound G6PDH. The optimal scaling factor was 0.73, suggesting that roughly 27 percent of translocation events were due to the antibody-antigen complex. g-h) Blue spheroids show the volume and shape of G6PDH and the antibody-antigen complex determined by fitting the empirical CDFs shown in panes (e) and (f). The crystal structure of G6PDH and IgG are shown in red and orange, respectively. i) Bar plot showing excellent agreement

between the volume of the antibody-antigen complex determined from analyzing maximum ΔI values from this experiment and the sum of the volumes of G6PDH and IgG determined that were determined individually in other nanopore experiments (see Table 3-App.1). j-i) Scatter plots showing the 2-D projections of the 3-D plot in Fig 3.5c of the main text. These plots show that resistive pulses assigned to the complex correspond to larger molecular volumes and smaller rotational diffusion coefficients than resistive pulses assigned to G6PDH. The dipole moment of G6PDH is relatively clustered as expected for a protein with well-defined shape and position of amino acids. In contrast, the dipole moment of the complex between G6PDH and the polyclonal anti-G6PDH IgG antibody varies widely as expected since IgG may bind at multiple locations and is a relatively floppy molecule. All recordings were obtained with pore 14 at an applied potential of -100 mV and pH of 6.1. We purchased polyclonal anti-G6PDH IgG (A9521) from Sigma Aldrich, Inc.
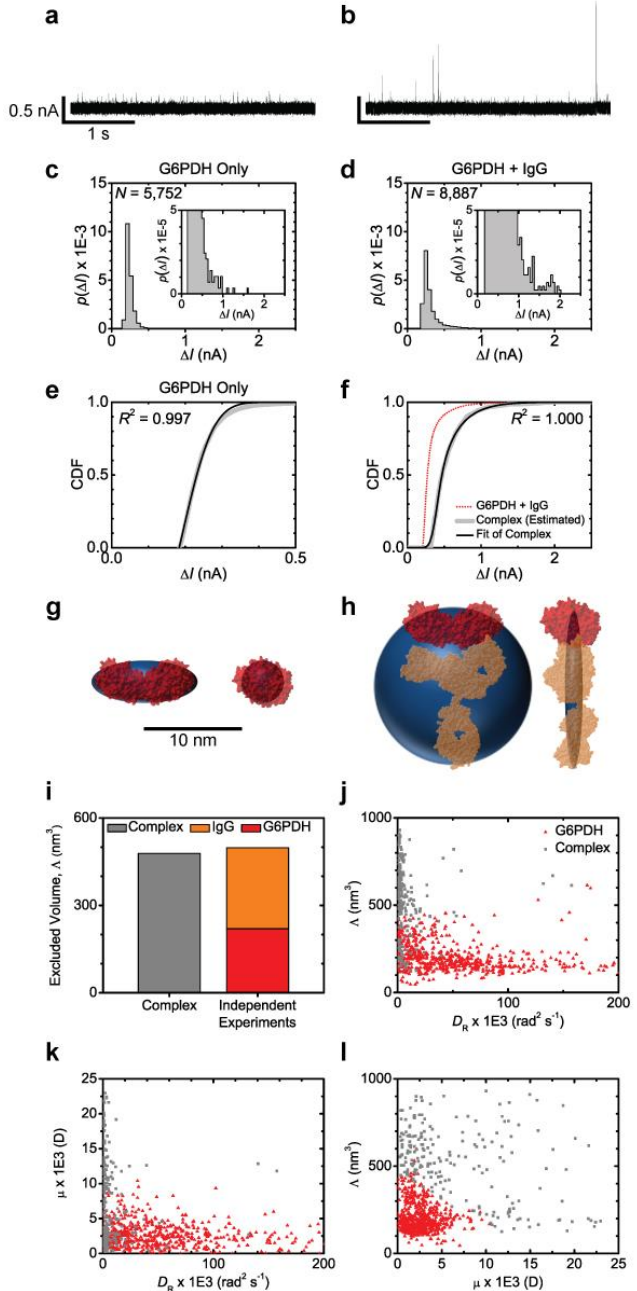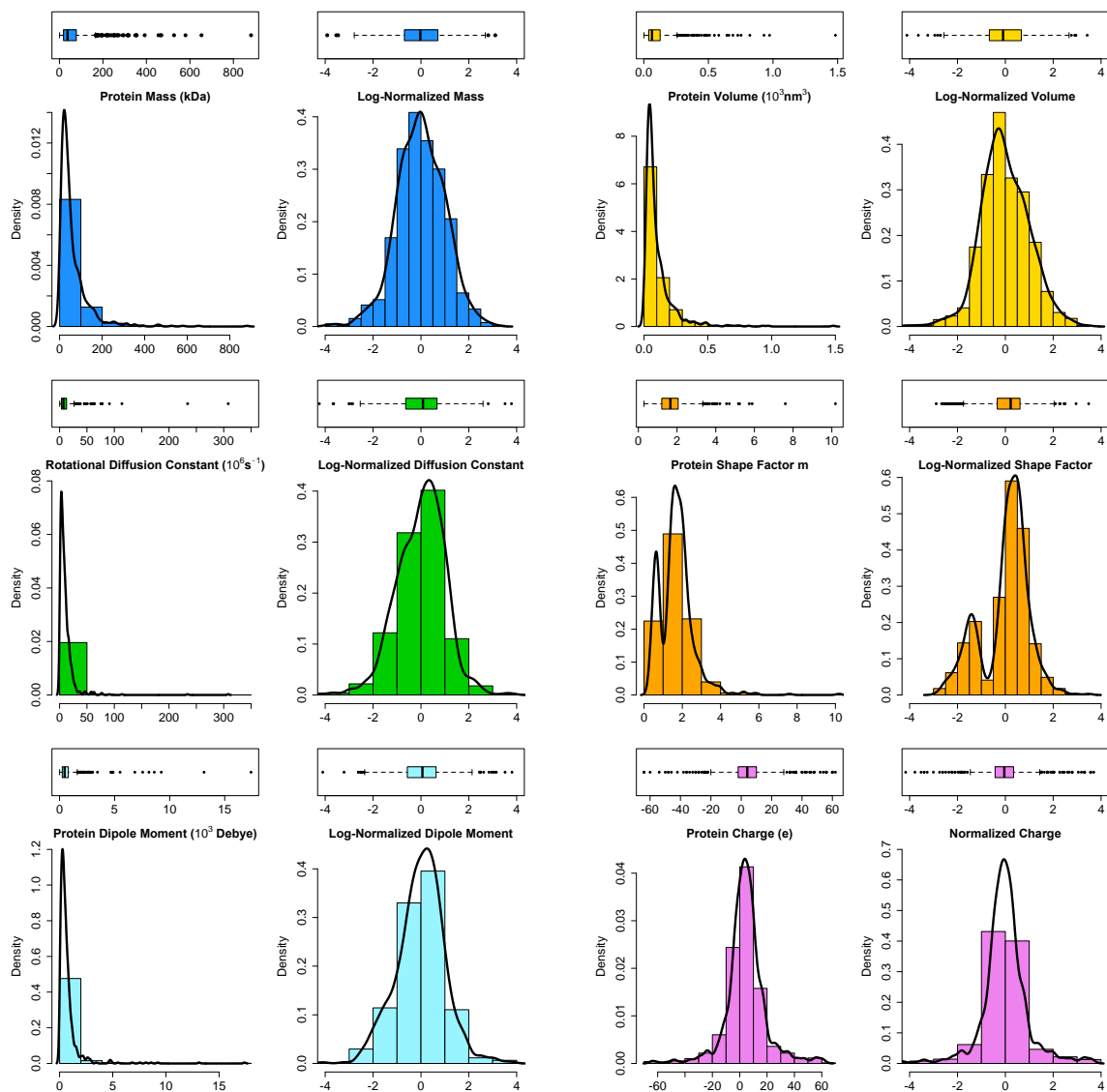
Figure 3-App.23 | Histograms, boxplots, and density distributions of calculated physical descriptors for 780 proteins. Using structural and sequence data, we randomly selected a group of proteins and determined their mass, volume, rotational diffusion constant, shape factor, dipole moment, and charge. The distributions on the left show the raw data for each quantity. To properly normalize the data, we first did log-transforms of all quantities, except charge, and then calculated standard normal distributions (shown on the right). As dimensionless, standard normal distributions, we can define a meaningful protein-protein distance in a space that combines multiple descriptors (*e.g.*, charge and mass).
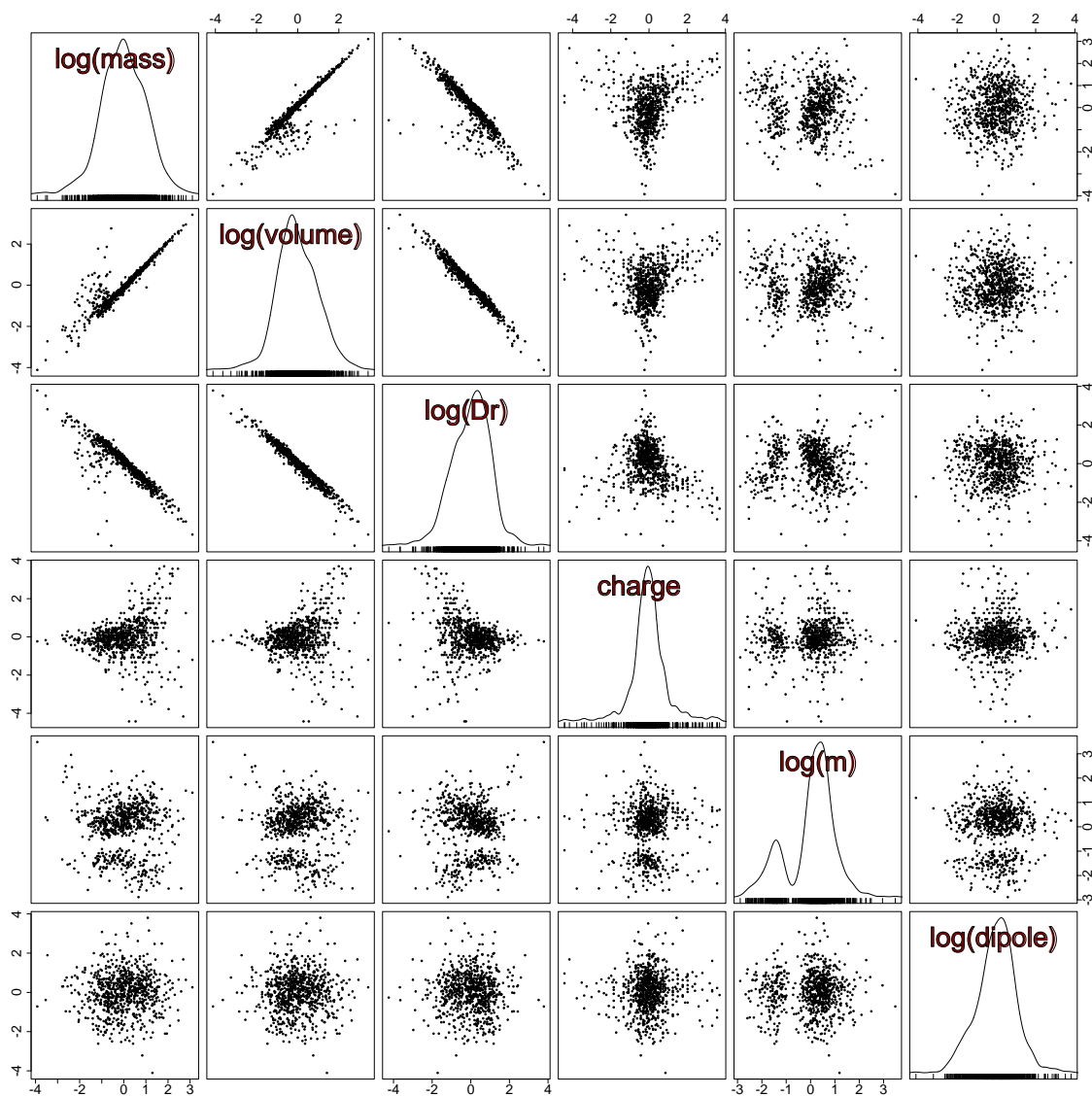
Figure 3-App.24 | A scatter plot matrix showing the relationships between the log-normalized quantities in Fig. 3-App.23. Mass, volume, and rotational diffusion constant show a high-degree of correlation; however charge, the length-to-diameter ratio (*m*), and the dipole moment show little correlation with any other descriptor.
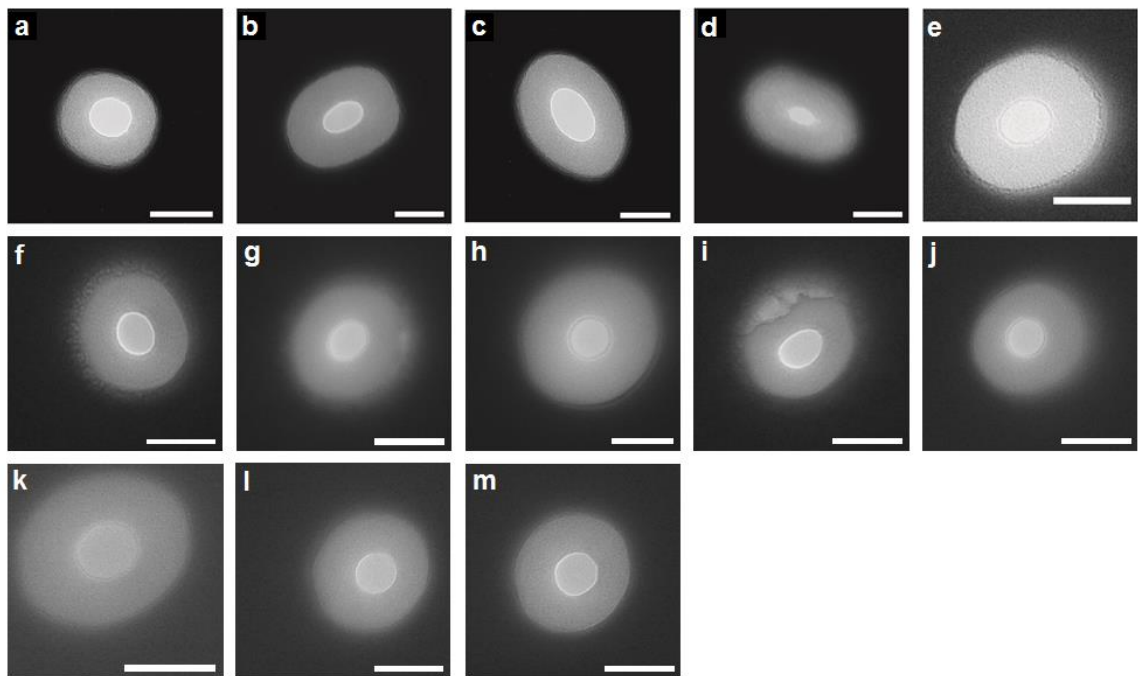
Figure 3-App.25 | Transmission electron micrographs of the nanopores used in this work. The brightest part in the center of each image depicts the shape and size of the nanopore and the surrounding circle with reduced brightness reflects the channel leading to the nanopore[20,31]. All scale bars are 50 nm. Nanopores shown are pore 1(a), pore 2 (b), pore 3 (c), pore 4 (d), pore 5 (e), pore 6 (f), pore 7 (g), pore 8 (h), pore 9 (i), pore 10 (j), pore 11 (k), pore 12 (l), and pore 13 (m). Using Image J, we measured the area of the nanopore (bright spot in the center) to determine the corresponding radius of a perfect circle with identical area, r_P (nm), and we determined the length, l_P (nm), of the nanopore from measurements of the electrical resistance of the nanopore[20]. The dimensions of the nanopores (in units of nm) without the lipid bilayer coating were: for pore 1 $r_P = 16.1$ and $l_P = 21.3$; for pore 2 $r_P = 16.4$ and $l_P = 17.3$; for pore 3 $r_P = 22.7$ and $l_P = 16.2$; for pore 4 $r_P = 9.6$ and $l_P = 18.0$; for pore 5 $r_P = 16.0$ and $l_P = 15.0$; for pore 6 $r_P = 14.2$ and $l_P = 10.0$; for pore 7 $r_P = 14.0$ and $l_P = 15.4$; for pore 8 $r_P = 17.8$ and $l_P = 15.5$; for pore 9 $r_P = 14.7$ and $l_P = 18.0$; for pore 10 $r_P = 13.6$ and $l_P = 14.0$; for pore 11 $r_P = 16.0$ and $l_P = 12.0$; for pore 12 $r_P = 14.5$ and $l_P = 10.0$; for pore 13 $r_P = 15.7$ and $l_P = 12.0$; and for pore 14 (not depicted) $r_P = 21.3$ and $l_P = 19.7$.

Appendix Tables

Table 3-App.1 | Values of fitting parameters determined from fitting the convolution model to the empirical distributions of ΔI values (Fig. 3.3 and 3.4 in the main text) as well as the resulting calculations of protein volume, Λ, and shape parameter, m.

Experiment	$E^{\dagger}$ (MV m^{-1})	ΔI_{min} (pA)	ΔI_{max} (pA)	σ (pA)	μ (D)	R^2	Λ^* (nm^3)	m^*
IgG$_1$, Pore 1	-1.5	329	678	58	596	0.998	292	0.37
IgG$_1$, Pore 2	-1.6	258	1,320	65	1,911	1.000	223	0.13
Intra-event (Fig. 4)	-1.6	281	938	48	302	0.998	232	0.21
IgG$_1$, Pore 3	-0.6	164	483	21	2,020	0.997	319	0.24
IgG$_1$, Pore 8	-1.4	266	1132	64	1,493	0.999	217	0.16
GPI-AChE, Pore 3	-1.0	280	375	14	3,530	0.999	278 *or* **306**	0.64 *or* **1.8**
GPI-AChE, Pore 5	-1.3	279	451	40	1,712	0.999	222 *or* **259**	0.50 *or* **3.1**
Fab, Pore 6	-2.1	178	231	11	972	1.000	71 *or* **77**	0.67 *or* **1.6**
β-PE, Pore 6	-0.8	181	302	31	2,125	0.999	**192** *or* 227	**0.48** *or* 3.5
G6PDH, Pore 7	-1.0	178	264	12	3,590	0.999	193 *or* 220	0.56 *or* **2.3**
G6PDH, Pore 14	-1.1	169	254	58	2,822	0.997	181 *or* 207	0.55 *or* **2.4**
L-LDH, Pore 8	-0.8	195	296	16	2,802	0.999	**267** *or* 307	**0.54** *or* 2.5
BSA, Pore 7	-1.9	165	258	17	1,263	0.998	**91** *or* 105	**0.52** *or* 2.7
BSA, Pore 9	-1.7	165	276	13	2,925	0.998	**110** *or* 130	**0.48** *or* 3.5
α-Amylase, Pore 10	-1.6	157	196	5	1,243	1.000	92 *or* **99**	0.71 *or* **1.5**
BChE, Pore 11	-1.7	150	364	18	1,007	1.000	82	0.30

† The electric field intensity was calculated according to the following equation: $E = V_A * R_p / (R_{total} * l_p)$, where R_p is the resistance of the pore, R_{total} is the total resistance of the circuit, and l_p is the length of the pore. * Values of Λ and m shown in bold are those corresponding to the correct shape (i.e. the shape that matches the crystal structure).

Table 3-App.2 | Estimated hydration shell thickness of proteins detected in this work. Based on the difference between the volume that we measured and the volume determined from crystal structures, we estimated the thickness of the hydration shell and the average number of water molecules required in this ordered water layer. The average hydration shell thickness is 0.34 ± 0.14 nm, which closely matches reported values that range from 0.3 to 0.5 nm[119-122].

Protein	Volume, Λ (nm³)		Hydration Shell Thickness	
	Measured	Crystal Structure	(nm)	N_{H2O}[†]
IgG$_1$	278	174	0.37	1.3
GPI-Acetylcholinesterase	283	145	0.53	1.9
Fab Fragment	77	56	0.21	0.8
β-Phycoerythrin	192	139	0.29	1.0
Glucose-6-Phosphate Dehydrogenase	220	135	0.40	1.4
L-Lactate Dehydrogenase	267	160	0.46	1.6
Bovine Serum Albumin	101	78	0.19	0.7
α-Amylase	99	65	0.29	1.0
Butyrylcholinesterase	82	69	0.12	0.4
Streptavidin	110	61	0.53	1.9

[†] The number of water molecules was calculated by dividing the thickness by the diameter of a water molecule (~0.28 nm)[123].

Table 3-App.3 | Estimated volume of proteins detected in this work from dynamic light scattering (DLS) measurements. The volume of each protein was estimated from the hydrodynamic radius that was obtained *via* DLS; the estimate of volume required assuming that the protein was a perfect sphere (*i.e.*, spherical) or was spheroidal (*i.e.*, ellipsoidal).

Protein	Hydrodynamic Radius from DLS, r_H (nm)	Spherical Volume from DLS (nm^3)	Spheroidal Volume from DLS (nm^3)[a]		Measured Volume from Nanopore Experiments (nm^3)
			Measured m	Reference m	
IgG$_1$	5.29	620	391	339–548	278
GPI-AChE	4.59	405	330	300	283
Fab	3.29	149	141	136–138	77
β-PE	3.83	235	205	179	192
G6PDH	3.95	257	214	206	220
L-LDH	4.07	282	256	261	267
BSA	3.38	162	143	150	101
α-Amylase	2.90	102	97	93	99
Streptavidin	2.82	94	N/A	94	110

[a] To calculate the volume of a spheroid particle that would return the hydrodynamic radius measured in DLS experiments, we set the value of *m* to those determined in nanopore experiments (measured *m*) or to those determined from crystal structures of the proteins (reference *m*).

Table 3-App.4 | Average volumes, length-to-diameter ratios, $m = A/B$, most probable dipole moments, rotational diffusion coefficients, and charges of proteins determined by analysis of resistive pulses and other methods.

Protein	Volume, Λ (nm³)		Length-to-Diameter Ratio, m		Rotational Diffusion Coef., $D_R^{[a]}$ (rad² s⁻¹)		Dipole Moment, $\mu^{[a]}$ (D)		Charge, z	
	Meas.	Ref.	Meas.	Ref.	Meas.	Ref.	Meas.	Ref.	Meas.	Ref.
IgG$_1$[b]	278	DLS[c]: 391 Theor.[d]: 266 Lit.: 347 ± 15[124]	0.25	0.2–0.5[125,126]	6,634	4,500[e]	855 @ pH 7.4	840[f,g]	-4.2 @ pH 7.4[h]	-4.6[h,i]
GPI-AChE[b]	283	DLS: 330 Theory: 195	2.4	2.9[j]	11,732	10,000[88]	731 @ pH 6.5	730[k]	-4.5 @ pH 6.5	--
Fab	77	DLS: 141 Theor.: 97 Lit.: 140[127] 170 ± 31[20]	1.6	1.7[j] 1.8[128]	22,505	27,000[e]	570[b] @ pH 7.4	630[f] 550[k]	-4.3 @ pH 7.4[h]	-3.9[h,i]
β-PE	192	DLS: 205 Theor.: 194	0.48	0.35[j]	8,595	8,400[e]	390 @ pH 7.4	395[l]	6.8 @ pH 5.1	10.5[i,m,n]
G6PDH	220	DLS: 214 Theor.: 222	2.3	2.5[j]	--	--	188 @ pH 6.1	203[l]	9.6 @ pH 6.1	15.0[i,m,n]
L-LDH	267	DLS: 256 Theory: 220	0.54	0.58[j]	--	--	267 @ pH 7.4	206[l]	-5.5 @ pH 6.1	-11.7[i,m,n]
BSA[b]	101	DLS: 143 Theor.: 111 Lit: 109[129] 123[130]	0.50	0.57[j]	--	--	522 @ pH 5.2	410[f]	-6.4 @ pH 5.2	-3.4[113,i,n]
α-Amylase	99	DLS: 97 Theor.: 89	1.5	1.8[j]	32,643	27,300[e]	375 @ pH 7.4	484[l]	-5.3 @ pH 7.4	-10.6[i,m,n]
BChE	82	Theor.: 103	0.30	0.47[j]	20,653	26,700[e]	992 @ pH 7.4	1,420[k]	-3.5 @ pH 7.4	--
Streptavidin	110 ± 25[o]	DLS: 94 Theor.: 88 Lit.: 94 ± 18[20] 105 ± 3[131]	1[k]	1.1[j]	--	--	--	--	-0.8 @ pH 7.4	-2.8[i,m,n]

[a] Most probable values determined from intra-event fitting; see Section 3-App.S6 for details. [b] Values were calculated from two or more experiments (Table 3-App.1). [c] Calculated from the hydrodynamic radius measured *via* DLS; see Section 3-App.2 for details. [d] An estimate of the volume of the hydrated protein

determined from the crystal structure of the protein using the software HydroPRO. [e] An estimate of the rotational diffusion coefficient determined from the crystal structure of the protein using the software HydroPRO; we accounted for the reduction in D_R due to the lipid anchor by multiplying by the known value of GPI-AChE and dividing by its theoretical estimate (a factor of 199). [f] Measured *via* dielectric impedance spectroscopy. [g] This value should be used as a loose approximation due to the low signal-to-noise ratio of the measurement. [h] Results from Yusko *et al.*[20]. [i] Values were reduced by 1 to account for the charge of the lipid anchor. [j] Estimated from the crystal structure of the protein. [k] Calculated from the crystal structure of the protein using the software HydroPRO. [l] Calculated from the crystal structure of the protein using the Weizmann server (http://bioinfo.weizmann.ac.il/dipol/). [m] Estimated using the PROPKA web interface (http://propka.ki.ku.dk/)[114-117]. [n] Values were reduced by 0.93 to account for the reaction of a primary amine on the protein surface with an NHS ester on the crosslinker molecule to form an amide bond[118]. All estimates were done in the absence of ligands except for G6PDH. [o] Since the distribution of ΔI values due to streptavidin translocations was unimodal and Normal, we assumed that streptavidin had a spherical shape, and therefore $m = 1$; to calculate the excluded volume of streptavidin, we solved equation (3-App.1) with γ set to a value of 1.5.

References

1. Picotti, P. & Aebersold, R. Selected reaction monitoring-based proteomics: Workflows, potential, pitfalls and future directions. *Nat. Methods.* **9**, 555-566 (2012).
2. Herr, A. E. Disruptive by design: A perspective on engineering in analytical chemistry. *Anal. Chem.* **85**, 7622-7628 (2013).
3. Rusk, N. Disruptive nanopores. *Nat. Methods.* **10**, 35-35 (2013).
4. Sali, A., Glaeser, R., Earnest, T. & Baumeister, W. From words to literature in structural proteomics. *Nature* **422**, 216-225 (2003).
5. Debye, P. *Polar molecules*. 1 edn, (Dover Publications Inc., 1929).
6. Oncley, J. L. The investigation of proteins by dielectric measurements. *Chem. Rev.* **30**, 433-450 (1942).
7. Felder, C. E., Prilusky, J., Silman, I. & Sussman, J. L. A server and database for dipole moments of proteins. *Nucleic Acids Research* **35**, W512-W521 (2007).
8. Chari, R., Jerath, K., Badkar, A. V. & Kalonia, D. S. Long- and short-range electrostatic interactions affect the rheology of highly concentrated antibody solutions. *Pharm. Res.* **26**, 2607-2618 (2009).
9. Hughes, J. P., Rees, S., Kalindjian, S. B. & Philpott, K. L. Principles of early drug discovery. *Br. J. Pharmacol.* **162**, 1239-1249 (2011).
10. Antosiewicz, J., Wlodek, S. T. & McCammon, J. A. Acetylcholinesterase: Role of the enzyme's charge distribution in steering charged ligands toward the active site. *Biopolymers* **39**, 85-94 (1996).
11. De Pascalis, A. R. *et al.* Binding of ferredoxin to ferredoxin: Nadp+ oxidoreductase: The role of carboxyl groups, electrostatic surface potential, and molecular dipole moment. *Protein Sci.* **2**, 1126-1135 (1993).
12. Movileanu, L., Howorka, S., Braha, O. & Bayley, H. Detecting protein analytes that modulate transmembrane movement of a polymer chain within a single protein pore. *Nat. Biotechnol.* **18**, 1091-1095 (2000).
13. Siwy, Z. *et al.* Protein biosensors based on biofunctionalized conical gold nanotubes. *J. Am. Chem. Soc.* **127**, 5000-5001 (2005).
14. Han, A. *et al.* Sensing protein molecules using nanofabricated pores. *Appl. Phys. Lett.* **88**, 093901 (2006).
15. Dekker, C. Solid-state nanopores. *Nat. Nanotechnol.* **2**, 209-215 (2007).
16. Wei, R., Gatterdam, V., Wieneke, R., Tampe, R. & Rant, U. Stochastic sensing of proteins with receptor-modified solid-state nanopores. *Nat. Nanotechnol.* **7**, 257-263 (2012).
17. Qin, Z. P., Zhe, J. A. & Wang, G. X. Effects of particle's off-axis position, shape, orientation and entry position on resistance changes of micro Coulter counting devices. *Meas. Sci. Technol.* **22** (2011).
18. Fologea, D., Ledden, B., David, S. M. & Li, J. Electrical characterization of protein molecules by a solid-state nanopore. *Appl. Phys. Lett.* **91**, 053901 (2007).
19. Sexton, L. T. *et al.* Resistive-pulse studies of proteins and protein/antibody complexes using a conical nanotube sensor. *J. Am. Chem. Soc.* **129**, 13144-13152 (2007).

20. Yusko, E. C. *et al.* Controlling protein translocation through nanopores with bio-inspired fluid walls. *Nat. Nanotechnol.* **6**, 253-260 (2011).
21. Raillon, C. *et al.* Nanopore detection of single molecule RNAP-DNA transcription complex. *Nano. Lett.* **12**, 1157-1164 (2012).
22. Soni, G. V. & Dekker, C. Detection of nucleosomal substructures using solid-state nanopores. *Nano. Lett.* (2012).
23. Di Fiori, N. *et al.* Optoelectronic control of surface charge and translocation dynamics in solid-state nanopores. *Nat. Nanotechnol.* **8**, 946-951 (2013).
24. Yusko, E. C. *et al.* Single-particle characterization of Ab oligomers in solution. *ACS Nano* **6**, 5909-5919 (2012).
25. Stefureac, R. I., Kachayev, A. & Lee, J. S. Modulation of the translocation of peptides through nanopores by the application of an ac electric field. *Chemical Communications* **48**, 1928-1930 (2012).
26. Bonincontro, A. & Risuleo, G. Dielectric spectroscopy as a probe for the investigation of conformational properties of proteins. *Spectrochim. Acta. A.* **59**, 2677-2684 (2003).
27. Mehl, J. W., Oncley, J. L. & Simha, R. Viscosity and the shape of protein molecules. *Science* **92**, 132-133 (1940).
28. McLoughlin, S. Y., Kastantin, M., Schwartz, D. K. & Kaar, J. L. Single-molecule resolution of protein structure and interfacial dynamics on biomaterial surfaces. *Proc. Natl. Acad. Sci.* (2013).
29. Weiss, S. Fluorescence spectroscopy of single biomolecules. *Science* **12**, 1676-1683 (1999).
30. 30 Berge, L. I., Feder, J. & Jossang, T. in *Particle size analysis* (eds N.G. Stanley-Wood & R. W. Lines) 374-383 (The Royal Society of Chemistry, 1992).
31. Li, J. *et al.* Ion-beam sculpting at nanometre length scales. *Nature* **412**, 166-169 (2001).
32. Uram, J. D., Ke, K. & Mayer, M. Noise and bandwidth of current recordings from submicrometer pores and nanopores. *ACS Nano* **2**, 857-872 (2008).
33. Rosenstein, J. K., Wanunu, M., Merchant, C. A., Drndic, M. & Shepard, K. L. Integrated nanopore sensing platform with sub-microsecond temporal resolution. *Nat. Methods.* **9**, 487-492 (2012).
34. Plesa, C. *et al.* Fast translocation of proteins through solid state nanopores. *Nano. Lett.* **13**, 658-663 (2013).
35. Golibersuch, D. C. Observation of aspherical particle rotation in Poiseuille flow via the resistance pulse technique. Part 1. Application to human erythrocytes. *Biophys. J.* **13**, 265-280 (1973).
36. Golibersuch, D. C. Observation of aspherical particle rotation in Poiseuille flow via the resistance pulse technique. Part 2. Application to fused sphere dumbbells. *J. Appl. Phys.* **44**, 2580-2584 (1973).
37. Deblois, R. W. & Wesley, R. K. A. Viral sizes, concentrations, and electrophoretic mobilities by nanopar analyzer. *Biophys. J.* **16**, A178-A178 (1976).
38. Smythe, W. R. Flow around a spheroid in a circular tube. *Phys. Fluids* **7**, 633-638 (1964).
39. Woodside, M. T. *et al.* Direct measurement of the full, sequence-dependent folding landscape of a nucleic acid. *Science* **314**, 1001-1004 (2006).

40. Velick, S. & Gorin, M. The electrical conductance of suspensions of ellipsoids and its relation to the study of avian erythrocytes. *Journal of General Physiology* **23**, 753-771 (1940).
41. Fricke, H. The electric permittivity of a dilute suspension of membrane-covered ellipsoids. *J. Appl. Phys.* **24**, 644-646 (1953).
42. Fricke, H. A mathematical treatment of the electrical conductivity of colloids and cell suspensions. *Journal of General Physiology* **6**, 375-384 (1924).
43. Hernandez-Ainsa, S. *et al.* Lipid-coated nanocapillaries for DNA sensing. *Analyst*, 104-106 (2013).
44. Rousseeuw, P. J. & Kaufman, L. *Finding groups in data: An introduction to cluster analysis*. (John Wiley & Sons, Inc., 1990).
45. Venkatesan, B. M. & Bashir, R. Nanopore sensors for nucleic acid analysis. *Nat. Nanotechnol.* **6**, 615-624 (2011).
46. Li, W. *et al.* Single protein molecule detection by glass nanopores. *ACS Nano* **7**, 4129-4134 (2013).
47. Kasianowicz, J. J., Brandin, E., Branton, D. & Deamer, D. W. Characterization of individual polynucleotide molecules using a membrane channel. *Proc. Natl. Acad. Sci. U. S. A.* **93**, 13770-13773 (1996).
48. Manrao, E. A. *et al.* Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat. Biotechnol.* **30**, 349-353 (2012).
49. Branton, D. *et al.* The potential and challenges of nanopore sequencing. *Nat. Biotechnol.* **26**, 1146-1153 (2008).
50. Nivala, J., Marks, D. B. & Akeson, M. Unfoldase-mediated protein translocation through an [alpha]-hemolysin nanopore. *Nat. Biotechnol.* **31**, 247-250 (2013).
51. Jijin, Y. *et al.* Rapid and precise scanning helium ion microscope milling of solid-state nanopores for biomolecule detection. *Nanotechnology* **22**, 285310 (2011).
52. Issaq, H. J. & Veenstra, T. D. Two-dimensional polyacrylamide gel electrophoresis (2d-page): Advances and perspectives. *Biotechniques* **44**, 697 - 700 (2008).
53. Dietz, H., Douglas, S. M. & Shih, W. M. Folding DNA into twisted and curved nanoscale shapes. *Science* **325**, 725-730 (2009).
54. Jin, R. *et al.* Photoinduced conversion of silver nanospheres to nanoprisms. *Science* **294**, 1901-1903 (2001).
55. Cecchini, M. P. *et al.* Rapid ultrasensitive single particle surface-enhanced raman spectroscopy using metallic nanopores. *Nano. Lett.* **13**, 4602-4609 (2013).
56. Auyeung, E. *et al.* Synthetically programmable nanoparticle superlattices using a hollow three-dimensional spacer approach. *Nat Nano* **7**, 24-28 (2012).
57. Alivisatos, A. P. Semiconductor clusters, nanocrystals, and quantum dots. *Science* **271**, 933-937 (1996).
58. Kalsin, A. M. *et al.* Electrostatic self-assembly of binary nanoparticle crystals with a diamond-like lattice. *Science* **312**, 420-424 (2006).
59. Yoo, J. & Aksimentiev, A. In situ structure and dynamics of DNA origami determined through molecular dynamics simulations. *Proc. Natl. Acad. Sci. U. S. A.* **110**, 20099-20104 (2013).
60. Pedone, D., Firnkes, M. & Rant, U. Data analysis of translocation events in nanopore experiments. *Anal. Chem.* **81**, 9689-9694 (2009).

61. Talaga, D. S. & Li, J. L. Single-molecule protein unfolding in solid state nanopores. *J. Am. Chem. Soc.* **131**, 9287-9297 (2009).
62. Bermudez, O. & Forciniti, D. Aggregation and denaturation of antibodies: A capillary electrophoresis, dynamic light scattering, and aqueous two-phase partitioning study. *J. Chromatogr. B* **807**, 17-24 (2004).
63. Jossang, T., Feder, J. & Rosenqvist, E. Photon-correlation spectroscopy of human-IgG. *J. Protein Chem.* **7**, 165-171 (1988).
64. Goodsell, D. *Acetylcholinesterase. June 2004 molecule of the month.*, <http://www.rcsb.org/pdb/101/motm.do?momID=54> (2004).
65. Sigma-Aldrich. *Product information: Acetylcholinesterase from human erythrocytes (c0663)*, <http://www.sigmaaldrich.com/catalog/product/sigma/c0663> (2012).
66. Tan, R. C., Truong, T. N., McCammon, J. A. & Sussman, J. L. Acetylcholinesterase: Electrostatic steering increases the rate of ligand binding. *Biochemistry* **32**, 401-403 (1993).
67. Luk, W. K. W., Chen, V. P., Choi, R. C. Y. & Tsim, K. W. K. N-linked glycosylation of dimeric acetylcholinesterase in erythrocytes is essential for enzyme maturation and membrane targeting. *Febs J.* **279**, 3229-3239 (2012).
68. Porschke, D. *et al.* Electrooptical measurements demonstrate a large permanent dipole moment associated with acetylcholinesterase. *Biophys. J.* **70**, 1603-1608 (1996).
69. Parihar, M. S. & Hemnani, T. Alzheimer's disease pathogenesis and therapeutic interventions. *J. Clin. Neurosci.* **11**, 456-467 (2004).
70. Rath, A. & Deber, C. M. Correction factors for membrane protein molecular weight readouts on sodium dodecyl sulfate-polyacrylamide gel electrophoresis. *Anal. Biochem.* **434**, 67-72 (2013).
71. Skinner, G. M., van den Hout, M., Broekmans, O., Dekker, C. & Dekker, N. H. Distinguishing single- and double-stranded nucleic acid molecules using solid-state nanopores. *Nano. Lett.* **9**, 2953-2960 (2009).
72. Maxwell, J. C. *A treatise on electricity and magnetism*. 3rd edn, 435-441 (Clarendon Press, 1904).
73. Deblois, R. W. & Bean, C. P. Counting and sizing of submicron particles by resistive pulse technique. *Rev. Sci. Instrum.* **41**, 909-915 (1970).
74. Grover, N. B., Naaman, J., Ben-sasson, S. & Doljansk, F. Electrical sizing of particles in suspensions. I.Theory. *Biophys. J.* **9**, 1398-1414 (1969).
75. Han, A. P. *et al.* Label-free detection of single protein molecules and protein-protein interactions using synthetic nanopores. *Anal. Chem.* **80**, 4651-4658 (2008).
76. Ito, T., Sun, L. & Crooks, R. M. Simultaneous determination of the size and surface charge of individual nanoparticles using a carbon nanotube-based Coulter counter. *Anal. Chem.* **75**, 2399-2406 (2003).
77. Hurley, J. Sizing particles with a Coulter counter. *Biophys. J.* **10**, 74-79 (1970).
78. DeBlois, R. W., Uzgiris, E. E., Cluxton, D. H. & Mazzone, H. M. Comparative measurements of size and polydispersity of several insect viruses. *Anal. Biochem.* **90**, 273-288 (1978).
79. Soni, G. V. & Dekker, C. Detection of nucleosomal substructures using solid-state nanopores. *Nano. Lett.* **12**, 3180-3186 (2012).

80. Lan, W.-J., Kubeil, C., Xiong, J.-W., Bund, A. & White, H. S. Effect of surface charge on the resistive pulse waveshape during particle translocation through glass nanopores. *J. Phys. Chem. C.* **118**, 2726-2734 (2014).
81. Fan, R. *et al.* DNA translocation in inorganic nanotubes. *Nano. Lett.* **5**, 1633-1637 (2005).
82. Chang, H. *et al.* DNA-mediated fluctuations in ionic current through silicon oxide nanopore channels. *Nano. Lett.* **4**, 1551-1556 (2004).
83. Das, S., Dubsky, P., van den Berg, A. & Eijkel, J. Concentration polarization in translocation of DNA through nanopores and nanochannels. *Physical Review Letters* **108**, 138101 (2012).
84. Japrung, D. *et al.* Single-molecule studies of intrinsically disordered proteins using solid-state nanopores. *Anal. Chem.* **85**, 2449-2456 (2013).
85. Osborn, J. A. Demagnetizing factors of the general ellipsoid. *Physical Review* **67**, 351-357 (1945).
86. Vogel, S. *Life in moving fluids: The physical biology of flow*. (Princeton University Press, 1994).
87. Axelrod, D., Koppel, D. E., Schlessinger, J., Elson, E. & Webb, W. W. Mobility measurement by analysis of fluorescence photobleaching recovery kinetics. *Biophys. J.* **16**, 1055-1069 (1976).
88. Yuan, Y. & Axelrod, D. Subnanosecond polarized fluorescence photobleaching - rotational diffusion of acetylcholine-receptors on developing muscle-cells. *Biophys. J.* **69**, 690-700 (1995).
89. Timbs, M. M. & Thompson, N. L. Slow rotational mobilities of antibodies and lipids associated with substrate-supported phospholipid monolayers as measured by polarized fluorescence photobleaching recovery. *Biophys. J.* **58**, 413-428 (1990).
90. Liu, S., Yuzvinsky, T. D. & Schmidt, H. Effect of fabrication-dependent shape and composition of solid-state nanopores on single nanoparticle detection. *ACS Nano* **7**, 5621-5627 (2013).
91. Jr., I. T., Sauer, K., Wang, J. C. & Puglisi, J. D. *Physical chemistry: Principles and applications in biological sciences.*, (Pearson Education, 2002).
92. Freedman, K. J., Haq, S. R., Edel, J. B., Jemth, P. & Kim, M. J. Single molecule unfolding and stretching of protein domains inside a solid-state nanopore by electric field. *Sci. Rep.* **3** (2013).
93. Oukhaled, A. *et al.* Dynamics of completely unfolded and native proteins through solid-state nanopores as a function of electric driving force. *ACS Nano* **5**, 3628-3638 (2011).
94. Howard, J. *Mechanics of motor proteins and the cytoskeleton*. (Sinauer Associates Inc. , 2001).
95. Ishijima, A. *et al.* Multiple- and single-molecule analysis of the actomyosin motor by nanometer piconewton manipulation with a microneedle: Unitary steps and forces. *Biophys. J.* **70**, 383-400 (1996).
96. Lewalle, A., Steffen, W., Stevenson, O., Ouyang, Z. & Sleep, J. Single-molecule measurement of the stiffness of the rigor myosin head. *Biophys. J.* **94**, 2160-2169 (2008).

97. Mathé, J., Aksimentiev, A., Nelson, D. R., Schulten, K. & Meller, A. Orientation discrimination of single-stranded DNA inside the alpha-hemolysin membrane channel. *Proc. Natl. Acad. Sci. U. S. A.* **102**, 12377-12382 (2005).

98. Tao, T. Time-dependent fluorescence depolarization and brownian rotational diffusion coefficients of macromolecules. *Biopolymers* **8**, 609-632 (1969).

99. García de la Torre, J., Huertas, M. L. & Carrasco, B. Hydronmr: Prediction of NMR relaxation of globular proteins from atomic-level structures and hydrodynamic calculations. *Journal of Magnetic Resonance* **147**, 138-146 (2000).

100. Gauthier, M. G. & Slater, G. W. Exactly solvable ogston model of gel electrophoresis. Ix. Generalizing the lattice model to treat high field intensities. *The Journal of Chemical Physics* **117**, 6745-6756 (2002).

101. Davenport, M. *et al.* The role of pore geometry in single nanoparticle detection. *ACS Nano* (2012).

102. Ai, Y. & Qian, S. Direct numerical simulation of electrokinetic translocation of a cylindrical particle through a nanopore using a poisson-boltzmann approach. *Electrophoresis* **32**, 996-1005 (2011).

103. Lines, R. W. in *Particle size analysis* (eds N. G. Stanley-Wood & R. W. Lines) 352 (The Royal Society of Chemistry, 1992).

104. Atkins, P. & Paula, J. d. *Elements of physical chemistry*. 357 (Oxford University Press, 2012).

105. Chari, R., Singh, S. N., Yadav, S., Brems, D. N. & Kalonia, D. S. Determination of the dipole moments of rnase sa wild type and a basic mutant. *Proteins: Structure, Function, and Bioinformatics* **80**, 1041-1052 (2012).

106. Cheng, Z., Chaikin, P. M. & Mason, T. G. Light streak tracking of optically trapped thin microdisks. *Physical Review Letters* **89** (2002).

107. Sakmann, B. & Neher, E. (Springer, 2009).

108. van Lengerich, B., Rawle, R. J. & Boxer, S. G. Covalent attachment of lipid vesicles to a fluid-supported bilayer allows observation of DNA-mediated vesicle interactions. *Langmuir* **26**, 8666-8672 (2010).

109. Solomentsev, Y. & Anderson, J. L. Electrophoresis of slender particles. *J. Fluid Mech.* **279**, 197-215 (1994).

110. Kerr, M. K. & Churchill, G. A. Bootstrapping cluster analysis: Assessing the reliability of conclusions from microarray experiments. *Proc. Natl. Acad. Sci.* **98**, 8961-8965 (2001).

111. Lawrence, A.-M., Besir, H. & seyin. Staining of proteins in gels with coomassie g-250 without organic solvent and acetic acid. *J Vis Exp*, e1350 (2009).

112. Daniel, Y. L. & Xinsheng Sean, L. On the distribution of DNA translocation times in solid-state nanopores: An analysis using schrödinger's first-passage-time theory. *Journal of Physics: Condensed Matter* **25**, 375102 (2013).

113. Yadav, S., Shire, S. & Kalonia, D. Viscosity analysis of high concentration bovine serum albumin aqueous solutions. *Pharm. Res.* **28**, 1973-1983 (2011).

114. Li, H., Robertson, A. D. & Jensen, J. H. Very fast empirical prediction and rationalization of protein pka values. *Proteins: Structure, Function, and Bioinformatics* **61**, 704-721 (2005).

115. Bas, D. C., Rogers, D. M. & Jensen, J. H. Very fast prediction and rationalization of pka values for protein–ligand complexes. *Proteins: Structure, Function, and Bioinformatics* **73**, 765-783 (2008).

116. Olsson, M. H. M., Søndergaard, C. R., Rostkowski, M. & Jensen, J. H. Propka3: Consistent treatment of internal and surface residues in empirical pka predictions. *Journal of Chemical Theory and Computation* **7**, 525-537 (2011).

117. Søndergaard, C. R., Olsson, M. H. M., Rostkowski, M. & Jensen, J. H. Improved treatment of ligands and coupling effects in empirical calculation and rationalization of pka values. *Journal of Chemical Theory and Computation* **7**, 2284-2295 (2011).

118. Gitlin, I., Mayer, M. & Whitesides, G. M. Significance of charge regulation in the analysis of protein charge ladders. *J. Phys. Chem. B.* **107**, 1466-1472 (2003).

119. Fogarty, A. C., Duboue-Dijon, E., Sterpone, F., Hynes, J. T. & Laage, D. Biomolecular hydration dynamics: A jump model perspective. *Chem. Soc. Rev.* **42**, 5672-5683 (2013).

120. Merzel, F. & Smith, J. C. Is the first hydration shell of lysozyme of higher density than bulk water? *Proc. Natl. Acad. Sci.* **99**, 5378-5383 (2002).

121. Austin, R. H., Beeson, K. W., Eisenstein, L., Frauenfelder, H. & Gunsalus, I. C. Dynamics of ligand binding to myoglobin. *Biochemistry* **14**, 5355-5373 (1975).

122. Steinhoff, H. J., Kramm, B., Hess, G., Owerdieck, C. & Redhardt, A. Rotational and translational water diffusion in the hemoglobin hydration shell: Dielectric and proton nuclear relaxation measurements. *Biophys. J.* **65**, 1486-1495 (1993).

123. Arrigo, J. S. Screening of membrane surface charges by divalent cations: An atomic representation. *American Journal of Physiology - Cell Physiology* **235**, C109-C117 (1978).

124. Schneider, S. W., Larmer, J., Henderson, R. M. & Oberleithner, H. Molecular weights of individual proteins correlate with molecular volumes measured by atomic force microscopy. *Pflugers Arch.* **435**, 362-367 (1998).

125. Ozinskas, A. J. in *Topics in fluorescence spectroscopy* Vol. 4 (ed Joseph R. Lakowicz) 487 (Kluwer Academic Publishers, 1994).

126. Carrasco, B. *et al.* Crystallohydrodynamics for solving the hydration problem for multi-domain proteins: Open physiological conformations for human IgG. *Biophysical Chemistry* **93**, 181-196 (2001).

127. Janeway, C. A. *Immunobiology: The immune system in health and disease*. 5th edn, (Garland Publishing, 2001).

128. de la Torre, J. G. & Carrasco, B. Hydrodynamic properties of rigid macromolecules composed of ellipsoidal and cylindrical subunits. *Biopolymers* **63**, 163-167 (2002).

129. Singh, M., Chand, H. & Gupta, K. C. The studies of density, apparent molar volume, and viscosity of bovine serum albumin, egg albumin, and lysozyme in aqueous and rbi, csi, and dtab aqueous solutions at 303.15 k. *Chemistry & Biodiversity* **2**, 809-824 (2005).

130. El Kadi, N. *et al.* Unfolding and refolding of bovine serum albumin at acid ph: Ultrasound and structural studies. *Biophys. J.* **91**, 3397-3404 (2006).

131. Neish, C. S., Martin, I. L., Henderson, R. M. & Edwardson, J. M. Direct visualization of ligand-protein interactions using atomic force microscopy. *Br. J. Pharmacol.* **135**, 1943-1950 (2002).

132. Menon, M. K. & Zydney, A. L. Effect of ion binding on protein transport through ultrafiltration membranes. *Biotechnol. Bioeng.* **63**, 298-307 (1999).

CHAPTER 4

Conclusions and Potential Avenues for Future Research

This dissertation introduced two distinct nanopore-based methods to characterize and identify single proteins. Chapter 2 described the development and use of dual-pore glass chips for recording single-ion-channel activity in cells, a major stride toward the development of an automated planar patch-clamp platform capable of performing single-channel recordings in high throughput. As single-channel recordings offer a wealth of information that is unattainable *via* whole-cell recordings, the development of such a platform should help to identify ion channels that serve as biomarkers and drug targets more rapidly and accurately. Chapter 3 presented theory for determining the shape, volume, charge, rotational diffusion coefficient, and dipole moment of single non-spherical proteins in solution based on resistive-pulses obtained using lipid-coated nanopores. By expanding the number of parameters that one can measure beyond merely volume and charge, this work greatly enhanced the ability of resistive-pulse sensing to characterize, distinguish, and identify native proteins; however, further advancements are needed in order to analyze complex mixtures (*e.g.*, human plasma) with reasonable accuracy. Nonetheless, the research presented here lays the groundwork for achieving this goal. In the following sections, I describe the shortcomings of the work summarized above and provide suggestions for future research.

4.1 Improving and extending the capabilities of the dual-pore platform

4.1.1 Enhancing the rate of gigaseal formation

Although the dual-pore glass chips introduced in Chapter 2 yielded some of the highest seal resistances of any planar patch-clamp platform to date, the rate of gigaseal formation was only fair (*e.g.*, the highest rate reported hitherto is 92 percent[1]; here, we obtained a rate of 61 percent) and therefore has room for improvement. A high rate of gigaseal formation is important for maximizing throughput and minimizing cost. As discussed in Chapter 2, a gigaseal may fail to form if the positioning pore does not situate a cell in suitable contact with the recording pore; in support of this hypothesis, the initial increase in resistance upon positioning a cell was generally lower when a gigaseal did not form. One potential solution to this problem is to reposition a poorly situated cell by applying alternating pulses of positive and negative pressure to the positioning pore (while the recording pore maintains an outward flow to keep its surface clean). Incorporation of optical feedback may help for ascertaining whether a cell is improperly situated if monitoring the change in resistance is insufficient. Another approach for improving the seal magnitude and rate of gigaseal formation is to optimize further the geometry of the dual-pore chips as well as the protocols for positioning a cell and forming a seal. Fig 2.4 shows that certain chips outperformed others, indicating that the geometry was not optimal in all cases. Once the ideal geometry is known, one could either avoid using chips with suboptimal geometries (based on either resistance measurements or SEM images) or improve the reproducibility of the machining process (*e.g.*, by using a more stable laser) to ensure all chips conform to the same standards following fabrication.

4.1.2 Reducing noise

Even though the dual-pore platform yielded the lowest noise ever reported for a planar platform in the cell-attached configuration, the noise varied considerably between experiments (*e.g.*, the RMS current ranged from 0.46 to 1.29 pA) and was higher than that of a conventional patch-clamp setup in approximately 50 percent of cases. The noise present in a single-channel recording dictates the maximum practical bandwidth for analysis in addition to the minimum detectable event magnitude at a given bandwidth; thus, the quality of data obtained by the dual-pore platform strongly depends on its noise performance. As mentioned in Chapter 2, we expect the total noise of the dual-pore platform to be dominated by the distributed *RC*-noise of the recording pore in the presence of a gigaseal. Therefore, reducing either the access resistance or capacitance of the recording pore should decrease the total noise most effectively. One approach for reducing the access resistance is to increase the dimensions of either the recording pore (Fig. 2.2b) or the L-shaped channel leading to the pore (Fig. 2.2c); however, we previously found that this approach tends to make the chips so fragile that they break during fabrication or handling. A more promising alternative is to use standard microfabrication techniques for producing a thin layer of Ag/AgCl that leads to the backside of the recording pore and serves as the command electrode, thereby circumventing the resistance contribution of the L-shaped channel. To reduce capacitance, on the other hand, one strategy is to increase the thickness of the substrate either by coating the chip with PDMS (excluding the pore openings) or by using thicker glass, which will increase the duration and difficulty of the fabrication process. Using quartz as a substrate instead of borosilicate glass will also reduce the capacitance in

addition to the dielectric noise; unfortunately, quartz is relatively expensive and the machining parameters would likely need modification.

4.1.3 Addressing the serial nature of fabrication

One major limitation of the work described in Chapter 2 is that the process for fabricating dual-pore glass chips is serial in nature (*i.e.*, we fabricated the chips one-by-one) and is therefore time-consuming and costly. While one could use a beam splitter to fabricate multiple chips in parallel using the same femtosecond laser, the number of times the beam can be divided is limited by the power output of the laser source. Furthermore, this approach requires expensive equipment (*e.g.*, optical microscope, nanostage, etc.) for each chip machined in parallel. Fig. 4.1 presents an alternate chip design that should be far easier to fabricate in large quantities while preserving the most important functional features of the dual-pore scheme, the single-shot aperture and inclusion of a second channel for positioning a cell. In this design, only the single-shot aperture must be fabricated using femtosecond laser ablation. It is possible to fabricate the remainder of the pore in parallel using photolithography and chemical etching since only one pore is present in the glass substrate. The design employs a PDMS scaffold that contains a lateral microchannel for positioning a cell *via* suction, similar to the microfluidic platform developed by Lau *et al.*[2] The main challenge associated with this design is placing the PDMS scaffold in the appropriate location on the surface of the substrate, which will require micrometer precision. If this challenge is overcome, however, this design would have a far greater potential for commercialization than the existing dual-pore platform.
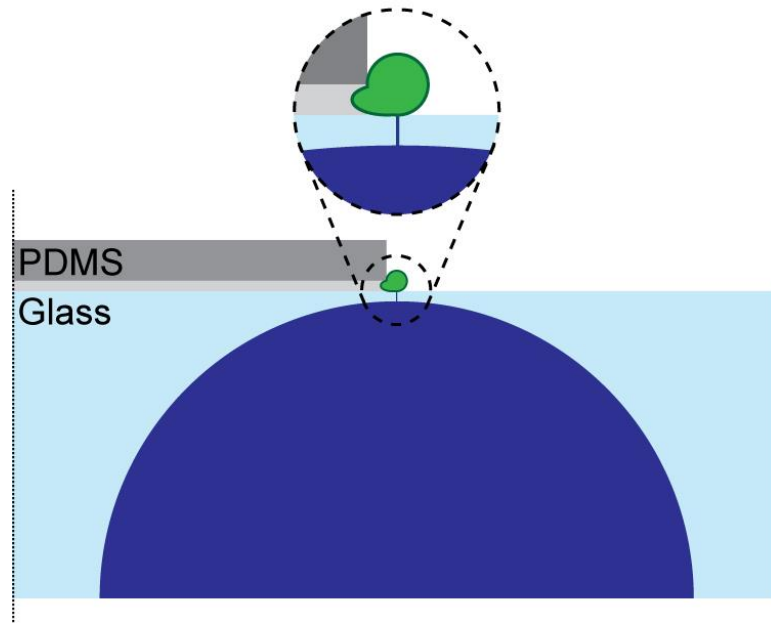
Figure 4.1 | Alternate chip design for high-throughput screening of single ion channels. This design involves fabricating a single-shot aperture on the surface of a glass substrate (light blue) and etching the backside of the substrate with hydrofluoric acid to make a through channel (the analog of the recording channel in the dual-pore design). A PDMS scaffold (grey) generated *via* soft lithography serves to position a cell (green) in close contact with the single-shot aperture by means of a lateral microchannel (the analog of the positioning channel in the dual-pore design).

4.1.4 Other applications

Thus far, this dissertation has only discussed using the dual-pore platform for performing single-channel recordings in cells. Nevertheless, with certain extensions, the dual-pore design may be amenable to several additional applications. Using the existing dual-pore chips, we found it nearly impossible to transition to the whole-cell configuration by applying pulses of suction or voltage, likely due to the small size of the recording pore. Consequently, simply etching a chip in hydrofluoric acid will likely make it suitable for performing whole-cell recordings; however, enlarging the recording pore may also increase noise and hinder seal formation. In addition to performing whole-cell recordings, it may also be possible to achieve the first on-chip excised patch recordings with the dual-pore platform. For instance, one could apply pulses of suction or voltage

with the positioning pore to rupture the non-patched membrane while a gigaseal is maintained with the recording pore, thereby gaining access to the intracellular side of the patch. Since the dual-pore platform enables quick exchange of both the bath and pore solutions, this on-chip excised patch would provide the functionality of both the inside-out and outside-out configurations used in conventional patch-clamp. The ability to establish all four of the primary patch-clamp configurations on-chip would greatly increase the utility of automated platforms for assessing ion channel activity.

Fig. 4.2 presents two additional extensions to the dual-pore platform that we briefly introduced in Chapter 2. The first extension involves fabricating two sets of dual-pores (*i.e.*, four total pores) in the same chip for performing the first on-chip gap junction recordings (Fig 4.2a). A gap junction is a specialized cell-to-cell interface that facilitates intercellular communication by permitting the passive diffusion of small molecules (*e.g.*, inorganic ions, second messengers, oligonucleotides, short linear peptides)[3,4]. Currently, dual whole-cell patch-clamp is the only method available for directly recording the electrical activity of gap junctions, which involves patching two adjacent cells each with a micropipette[5]. Not surprisingly, this technique is experimentally challenging; thus, a significant need exists to develop an automated platform for recording gap junction currents in a high-throughput manner, such as the design proposed here (Fig 4.2a). The second extension to the dual-pore platform involves fabricating an array of single-shot apertures for concentrating organelles (or small cells) in the vicinity of the recording pore (Fig 4.2b). Organellar membranes, like cell membranes, contain ion channels that play a role in a variety of cellular processes (*e.g.*, apoptosis, regulation of intracellular calcium, volume regulation, etc.)[6]; however, certain organelles (*e.g.*, non-enlarged lysosomes) are
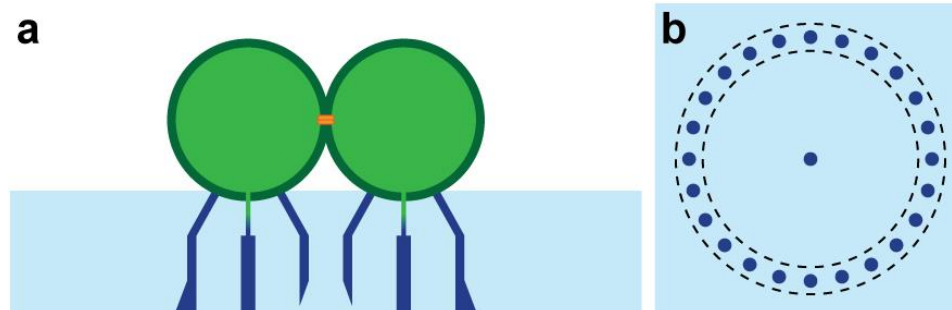
Figure 4.2 | Two proposed extensions to the dual-pore platform. a) Illustration depicting a cross-section of a chip with two sets of dual-pores (*i.e.*, four total pores) for performing gap junction recordings. This concept involves positioning and patching two cells in direct contact. In order to record gap junction activity, the whole-cell configuration must be established with each cell; hence, it may be necessary to etch the chip with hydrofluoric acid to increase the dimensions of the recording pores. b) Top view of a chip with an array of single-shot apertures for concentrating organelles in close proximity to the recording pore. After positioning a number of organelles, suction can be applied to the recording pore in order to aspirate one of these organelles for patching. For ease of fabrication, the array of single-shot apertures could connect to a common channel below the surface of the glass.

too small to be patched by standard patch pipettes and existing planar platforms[7]. On the other hand, the single-shot apertures used in this work are small enough to accommodate the size of most organelles. Therefore, the design shown in Fig. 4.2b may enable single-channel recordings in native organellar membranes that are not possible using current techniques.

4.2 Enhancing the capabilities of resistive-pulse sensing for characterizing single proteins

4.2.1 Increasing translocation times

In order to identify a protein in a complex mixture based on a single translocation event, the accuracy of the methods presented in Chapter 3 must be improved. One approach for improving accuracy is to increase translocation times such that each event provides more data reflecting additional rotation of the protein in the nanopore. Increasing the length of the pore will easily accomplish this goal; however, the magnitude

of the resistive-pulse will also decrease, reducing the signal-to-noise ratio (SNR). On the other hand, increasing the viscosity of the lipid coating, which dictates the diffusion coefficient of a lipid-anchored protein[8], will result in longer translocation times without reducing the SNR. Yusko *et al.* have previously shown that incorporating cholesterol in the coating increases its viscosity[8]. Furthermore, one could increase the viscosity by incorporating either lipids with long acyl chains or bipolar Archaeal lipids (see Fig. 4.3), which form monolayers that are relatively viscous compared to a conventional bilayer[9]. Linking proteins to multiple lipid anchors would have a similar effect to increasing the viscosity of the coating, although this approach may prohibit protein rotation in the
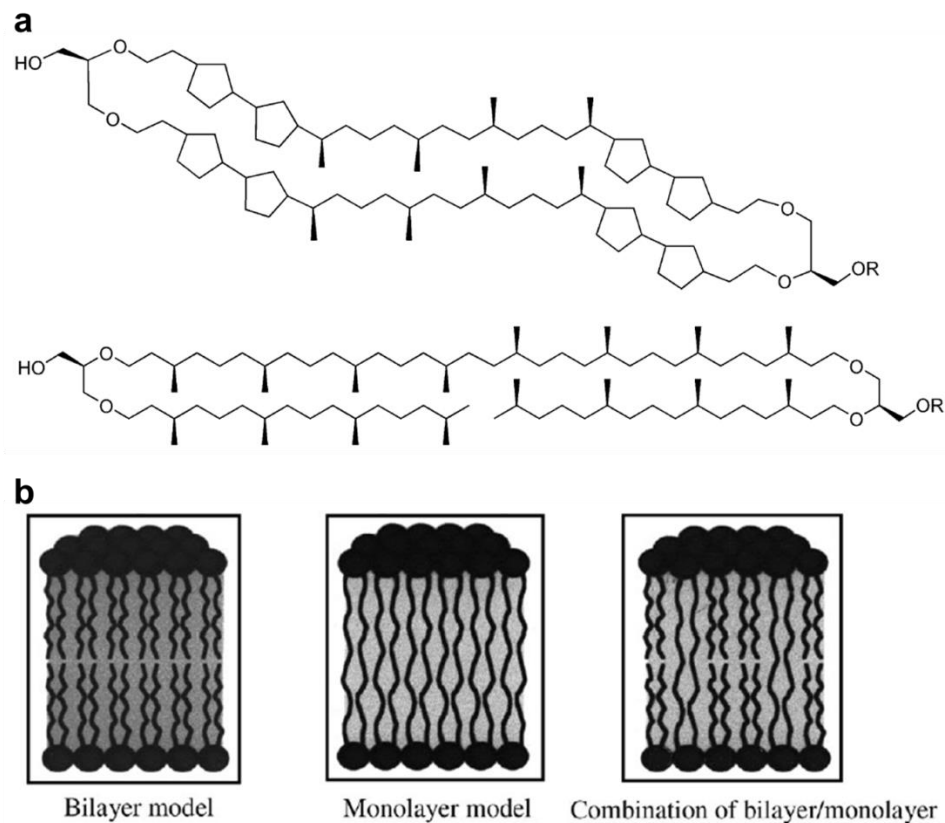
Figure 4.3 | Bipolar Archaeal lipids. a) Two examples of Archaeal lipids. b) Illustration depicting how standard (*i.e.*, monopolar) lipids form a bilayer (left) whereas bipolar Archaeal lipids form a more viscous monolayer (center). A mixture of monopolar and bipolar lipids forms a hybrid structure (right). Adapted from (10).

nanopore (as shown in Chapter 3). Nevertheless, both of these strategies introduce a major limitation as they will decrease the delivery rate of proteins to the nanopore and therefore reduce throughput for a given protein concentration. One solution to this dilemma is to generate a temperature gradient across the nanopore such that the *cis* side (*i.e.*, the side where protein is added) is at a high temperature in comparison to the *trans* side. In this scenario, the viscosity of the bilayer would increase from the *cis* to the *trans* side of the nanopore, slowing protein translocation while minimally affecting delivery rates to the pore. Alternatively, tethering the bottom leaflet of the bilayer coating to the surface of the pore walls could serve to increase viscosity locally in the pore.

While the above strategies serve to increase translocation times to some degree, the ideal approach should allow each translocating protein to be held in the pore for any given duration (*e.g.*, long enough to obtain a prescribed level of confidence in a measured parameter). Fig. 4.4 presents a strategy for achieving this lofty goal. While this approach is straightforward, it presents a number of technical challenges. First, an event must be detected with minimal delay (*i.e.*, on the order of tens of μs) in order to initiate the feedback control scheme illustrated in Fig. 4.4c. Next, the feedback control scheme must be able to respond even more rapidly (*i.e.*, on the order of a few μs) to keep a translocating particle from exiting the pore. Since this scheme relies on monitoring ΔI to determine when the protein is approaching the pore exit, one must be careful to avoid false positives or negatives resulting from electrical noise. Additionally, changing the potential during an event will elicit a capacitive transient, which must be cancelled appropriately (this is standard practice in patch-clamp experiments). Finally, this approach may promote clogging if a protein resides in the pore for too long. Nonetheless,
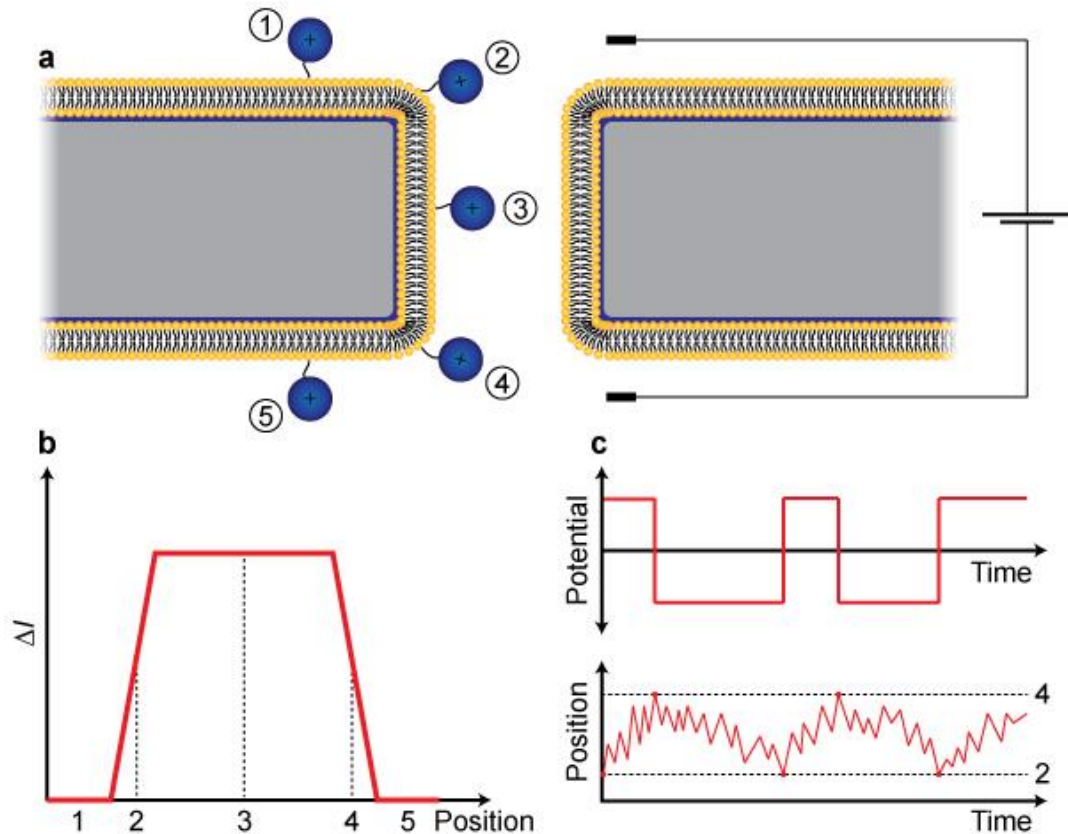
Figure 4.4 | Proposed strategy to keep a translocating protein in the nanopore for an indefinite duration. a) Illustration depicting various locations where a lipid-anchored protein may reside relative to the nanopore. b) A hypothetical plot depicting the magnitude of the resistive-pulse, ΔI, versus the position of the protein. The x-labels correspond to the numbers shown in pane (a). c) Protocol for preventing a translocating protein from exiting the nanopore. As the protein approaches the pore exit, which results in a reduction in ΔI (positions 2 and 4 in pane (b)), the patch-clamp amplifier switches the polarity of the applied potential (top) in order to bias the motion of the protein in the opposite direction (bottom). In addition to yielding more ΔI values, this protocol also provides multiple translocation times per event since the translocating protein passes from one end of the pore to the other with each change in the polarity of the applied potential.

this approach has the potential to greatly enhance the accuracy of the methods presented

in Chapter 3.

4.2.2 Reducing noise

In addition to increasing translocation times, reducing noise will also improve the

accuracy of the methods presented in Chapter 3. As predicted by theory, the dominant

source of noise in a resistive-pulse sensing experiment can be either thermal noise generated by the pore, dielectric noise generated by the substrate, or headstage and amplifier noise[11]. Reducing thermal noise requires increasing the resistance of the pore by either decreasing its diameter or increasing its length; however, the maximum detectable particle size is constrained by the pore diameter and increasing the pore length reduces the signal magnitude. Both dielectric noise and headstage and amplifier noise scale with the capacitance of the substrate. Thus, the most effective strategy for reducing noise may be to reduce capacitance by increasing the thickness of the chip (*e.g.*, *via* a PDMS coating) or by employing a substrate material with a lower dielectric constant such as quartz. Moreover, using quartz instead of silicon would reduce dielectric noise due to its relatively low dielectric loss. Researchers have also observed significant $1/f$ noise that is not predicted by theory and is believed to result from incomplete wetting of the nanopore or absorption of particles to the pore walls[11–15]. Piranha cleaning (as employed in this work) and plasma oxidation can be used to eliminate $1/f$ noise, although Beamish *et al*. have shown that high electric fields can reduce this type of noise when other methods are insufficient[14]. Finally, Yusko *et al*. have shown that lipid coatings can introduce additional noise at frequencies below 2 kHz[8]. Therefore, it may be useful to investigate different lipid coatings in order to find one that minimizes noise.

In 2012, Rosenstein *et al*. developed a CMOS-integrated nanopore platform that yielded the lowest noise and highest bandwidth of any platform to date (*i.e.*, SNR > 5 for DNA at 1 MHz)[15]. Using this platform in conjunction with the methods developed in Chapter 3 would not only improve accuracy due to the reduction in noise but would likely eliminate the need to tether proteins of interest to lipids in the nanopore coating

due to the increased bandwidth. While it remains to be seen whether the pores in this platform can be coated with a lipid bilayer, such a combination holds much promise.

4.2.3 Increasing the magnitude of a resistive-pulse

Increasing the magnitude of a resistive-pulse, ΔI, will also serve to improve the accuracy of the methods presented in this dissertation. Decreasing the diameter or length of the nanopore will accomplish this goal; however, reducing the pore length will also decrease translocation times. As indicated in the preceding section, the minimum usable pore diameter is dictated by the size of the particles to be detected. Researchers typically avoid using a pore diameter that is smaller than twice the maximum particle diameter since such a scenario requires the use of a correction factor[8]. Nonetheless, Qin *et al.* have shown that the correction factor developed by Smythe accurately describes simulated translocation events[16]. Therefore, this correction factor could likely be used to describe the magnitude of a resistive-pulse obtained using a pore diameter that is smaller than the conventional limit. On the other hand, such a small pore may introduce steric issues (*e.g.*, restrict certain protein orientations in the pore), increase the probability of clogging, and be difficult to coat with a lipid bilayer. In addition to reducing the pore dimensions, increasing the applied potential will also increase ΔI; however, this strategy will decrease translocation times and may increase noise[11] and promote protein unfolding. Finally, increasing the conductivity of the recording buffer will increase ΔI, as well, although this approach is limited by the solubility limit of the salt used. Moreover, increasing the ionic strength will likely reduce protein stability.

4.2.4 Possible extensions

In September 2014, Ivankin *et al*. introduced an optical method for detecting DNA using a solid-state nanopore and showed that the resulting optical signal closely resembled the electrical signal obtained from the same pore[17]. A few months later, Anderson *et al.* presented a refined version of this method capable of achieving sub-µs response times as predicted by theory[18]. While this optical method yields a lower signal-to-noise ratio than the conventional electrical method, it does not require each nanopore in a parallel array to be electrically isolated[17,18]. By performing optical and electrical detection in parallel, it may be possible to improve the accuracy of the methods presented in this dissertation.

In 2013, Hyun *et al*. developed an apparatus for threading a single DNA molecule through a solid-state nanopore using a tuning fork based force sensing probe tip[19]. This apparatus is capable of sub-nm spatial resolution unlike optical traps that tend to have resolutions of greater than 10 nm[19]. By attaching a protein of interest to the probe tip *via* a long flexible crosslinker, one could conceivably use this same apparatus to position the protein in an optimal location of the nanopore (*e.g.*, the narrowest constriction) for a prolonged duration to acquire a large volume of high quality data. This strategy would also enable the study of proteins that are too highly charged to be time resolved in a conventional resistive-pulse sensing experiment. Furthermore, this approach may present fewer technical challenges in comparison to the strategy depicted in Fig. 4.4.

One of the assumptions underlying the methods presented in Chapter 3 is that the dipole moment of a translocating protein lies parallel to a principal axis. Even though this

assumption resulted in low error (see Fig. 3.4e), future research should seek to take into account the direction of the dipole moment to improve accuracy. Furthermore, the direction of the dipole moment could serve as a new characteristic by which proteins are identified and distinguished from one another.

4.3 Concluding remarks

Proteins exhibit more diversity than any other class of biomolecules and serve innumerable physiological roles; thus, it comes as no surprise that proteins have tremendous potential as biomarkers and drug targets. This dissertation presented two different nanopore-based techniques aimed at characterizing and identifying single proteins in a high-throughput manner. In comparison to ensemble measurements that convolute the behavior or properties of many molecules, these single-molecule methods are better-suited to study heterogeneous populations, examine protein dynamics, and identify anomalies. With further refinements (such as those described above), the techniques presented in this work may ultimately expedite biomarker and drug discovery, enable the construction of personal proteomes, and improve our understanding of proteins and protein complexes in the context of health and disease.

References

1. Van Stiphout, P., Knott, T., Danker, T. & Stett, A. 3D Microfluidic Chip for Automated Patch-Clamping. *Mikrosystemtechnik Kongr.* 435–438 (2005).
2. Lau, A. Y., Hung, P. J., Wu, A. R. & Lee, L. P. Open-access microfluidic patch-clamp array with raised lateral cell trapping sites. *Lab. Chip* **6,** 1510–1515 (2006).
3. Abbaci, M., Neverauskas, V. & Skeberdis, V. A. Diversity and properties of connexin gap junction channels. *Med. Kaunas Lith.* **46,** 1–12 (2009).
4. Pfenniger, A., Wohlwend, A. & Kwak, B. R. Mutations in connexin genes and disease. *Eur. J. Clin. Invest.* **41,** 103–116 (2011).

5. Abbaci, M., Barberi-Heyob, M., Blondel, W., Guillemin, F. & Didelon, J. Advantages and limitations of commonly used methods to assay the molecular permeability of gap junctional intercellular communication. *BioTechniques* **45,** 33–52, 56–62 (2008).

6. Schieder, M., Rötzer, K., Brüggemann, A., Biel, M. & Wahl-Schott, C. Planar patch clamp approach to characterize ionic currents from intact lysosomes. *Sci. Signal.* **3,** pl3 (2010).

7. Schieder, M., Rötzer, K., Brüggemann, A., Biel, M. & Wahl-Schott, C. A. Characterization of Two-pore Channel 2 (TPCN2)-mediated Ca2+ Currents in Isolated Lysosomes. *J. Biol. Chem.* **285,** 21219–21222 (2010).

8. Yusko, E. C. *et al.* Controlling protein translocation through nanopores with bio-inspired fluid walls. *Nat. Nanotechnol.* **6,** 253–260 (2011).

9. Jarrell, H. C., Zukotynski, K. A. & Sprott, G. D. Lateral diffusion of the total polar lipids from Thermoplasma acidophilum in multilamellar liposomes. *Biochim. Biophys. Acta BBA - Biomembr.* **1369,** 259–266 (1998).

10. Benvegnu, T., Lemiègre, L. & Cammas-Marion, S. Archaeal Lipids: Innovative Materials for Biotechnological Applications. *Eur. J. Org. Chem.* **2008,** 4725–4744 (2008).

11. Uram, J. D., Ke, K. & Mayer, M. Noise and Bandwidth of Current Recordings from Submicrometer Pores and Nanopores. *ACS Nano* **2,** 857–872 (2008).

12. Tabard-Cossa, V., Trivedi, D., Wiggin, M., Jetha, N. N. & Marziali, A. Noise analysis and reduction in solid-state nanopores. *Nanotechnology* **18,** 305505 (2007).

13. Smeets, R. M. M., Keyser, U. F., Dekker, N. H. & Dekker, C. Noise in solid-state nanopores. *Proc. Natl. Acad. Sci.* **105,** 417–421 (2008).

14. Beamish, E., Kwok, H., Tabard-Cossa, V. & Godin, M. Precise control of the size and noise of solid-state nanopores using high electric fields. *Nanotechnology* **23,** 405301 (2012).

15. Rosenstein, J. K., Wanunu, M., Merchant, C. A., Drndic, M. & Shepard, K. L. Integrated nanopore sensing platform with sub-microsecond temporal resolution. *Nat. Methods* **9,** 487–492 (2012).

16. Qin, Z., Zhe, J. & Wang, G.-X. Effects of particle's off-axis position, shape, orientation and entry position on resistance changes of micro Coulter counting devices. *Meas. Sci. Technol.* **22,** 045804 (2011).

17. Ivankin, A. *et al.* Label-Free Optical Detection of Biomolecular Translocation through Nanopore Arrays. *ACS Nano* **8,** 10774–10781 (2014).

18. Anderson, B. N. *et al.* Probing Solid-State Nanopores with Light for the Detection of Unlabeled Analytes. *ACS Nano* **8,** 11836–11845 (2014).

19. Hyun, C., Kaur, H., Rollings, R., Xiao, M. & Li, J. Threading Immobilized DNA Molecules through a Solid-State Nanopore at >100 μs per Base Rate. *ACS Nano* **7,** 5892–5900 (2013).