TRANSDUCING SIGNALS AND PRE-CONCENTRATING MOLECULES FOR ENHANCED SOLID-STATE NANOPORE BIOSENSING

ZACHARY ROELEN

Thesis submitted to the University of Ottawa
in partial fulfillment of the requirements for the
Doctorate in Philosophy Physics

Department of Physics
Faculty of Science
University of Ottawa

Table of Contents

Abstract

Single-molecule biosensors offer distinct advantages over their ensemble-averaged counterparts by being able to extract information related to rare targets and specific molecular configurations within a sample. In particular, solid-state nanopores embody a promising single-molecule technique that is based on detecting target molecules by the amount of ionic current they block as they pass through a nanoscale aperture across a thin membrane. In this thesis, I present extensions of the basic nanopore system aimed at addressing some of its main limitations at present, namely: 1) the low rates at which nanopores capture molecules from a bulk volume, which restricts their ability to work with dilute ($\lesssim$ nM) samples, and 2) the difficulty in using nanopores to distinguish small or closely related molecules by their direct current blockage signatures alone.

I begin by describing the design and construction of a nanopore-based instrument that integrates an optical detection channel in parallel with ionic current sensing. A particular emphasis was placed on minimizing the electrical noise contributions of the added optical equipment on the original ionic current channel. Measuring the optical signals of translocating molecules together with their current blockages can improve the discrimination of two fluorescently labelled targets (or two configurations of a single target) that normally produce similar ionic current signatures.

I next investigate the combination of nanopore sensing with target pre-concentration, specifically, by embedding a nanopore membrane within a fluidic cell that features an insulator-based dielectrophoretic (iDEP) trap. Applying large ($\gtrsim$ 100 V) AC voltages across the iDEP channels of the cells resulted in the accumulation of polarizable targets (dsDNA, polystyrene beads) at the locations of the membranes, thus pointing toward a convenient method for the detection of ultra-dilute target samples in future nanopore devices.

Finally, I introduce improved protocols for the synthesis and nanopore signal analysis of dsDNA-based molecular carriers. In a molecular carrier scheme, in order to enhance the target specificity of the system, target molecules are not sensed directly by a nanopore but instead interact specifically with secondary molecules (“carriers”) to recognizably alter the carrier translocation signals. Here, I present proof-of-principle analyses of DNA carrier

experiments that highlight the multiplexing capabilities of our carrier design, which are based on separating targets by their interactions with carriers of different lengths.

Developments of the nanopore sensing platform such as those presented in this work, which leverage the intrinsic versatility of solid-state nanopores to be integrated within complex devices and to detect a wide range of target molecules, will play an important role in continuing to increase the precision of single-molecule measurements into the future and to expand their breadth of potential applications.

Résumé

Les biocapteurs mono-moléculaires offrent des avantages distincts face à leurs homologues à mesure d'ensemble, puisqu'ils offrent la possibilité d'extraire l'information contenue par des cibles rares et par les configurations moléculaires spécifiques présentes dans un échantillon. En particulier, les nanopores d'état solide constituent une technique mono-moléculaire prometteuse, qui détecte des molécules cibles à partir des chutes de courant ionique induites par les molécules traversant un pore situé dans une membrane mince. Dans cette thèse, je développe et étends le système de nanopore de base, visant à améliorer certaines de ses limites principales actuelles, dont: 1) la faible fréquence à laquelle les nanopores capturent les molécules à partir d'un volume tridimensionnel, limitant leur capacité à travailler avec des échantillons dilués, et 2) la difficulté d'utiliser des nanopores pour distinguer des molécules de petites tailles ou de dimensions similaires, soit seulement par leurs signatures de blocage de courant.

Je commence par décrire la conception et la construction d'un instrument qui, en parallèle avec la détection du courant ionique d'un nanopore, inclut un canal de détection optique. Une attention particulière a été mise sur la minimisation des contributions de l'équipement optique au bruit électrique du canal de courant ionique original. La mesure des signaux optiques dus à la translocation de molécules, en combinaison avec celle des blocages de courant, peut améliorer la discrimination de cibles marquées par fluorescence (ou de deux configurations d'une seule cible) qui produisent normalement des signatures de courant ionique similaires.

J'examine ensuite le couplage de la détection par nanopores avec la préconcentration des cibles, en intégrant une membrane détenant un nanopore dans une cellule fluidique qui contient un piège diélectrophorétique à base d'isolant (iDEP). L'application de tensions alternatives élevées ($\gtrsim$ 100 V) aux canaux iDEP favorise l'accumulation de cibles polarisables (ADN à double brin, billes de polystyrène) à proximité des membranes contenant un nanopore, soutenant ainsi l'utilisation de cette technique d'iDEP comme étant une méthode avantageuse pour la détection de cibles ultra-diluées avec des futurs appareils de nanopore.

Enfin, j’introduis des protocoles optimisés pour la synthèse et l’analyse de signaux induits par des porteurs moléculaires à base d’ADN. Dans un tel système de porteur moléculaire, les molécules cibles ne sont pas détectées directement par un nanopore, mais interagissent spécifiquement avec des molécules secondaires (« porteurs ») pour modifier de manière reconnaissable les signaux de translocation du porteur, afin d'améliorer la spécificité du système pour des cibles différentes. Je présente ici une preuve de concept qui met en évidence les capacités de multiplexage de notre conception de porteurs, basée sur la séparation des cibles en fonction de leurs interactions avec des porteurs de différentes longueurs.

Les développements de la plateforme de détection à base de nanopores tels que ceux présentés dans ce travail, qui profitent de la polyvalence fondamentale des nanopores d'état solide pour être intégrés dans des appareils complexes et pour détecter une large gamme de molécules cibles, joueront un rôle important dans la poursuite de l'augmentation de la précision des mesures mono-moléculaires et de l’élargissement de l'étendue de leurs applications potentielles dans le futur.

Statement of Originality

The material in this thesis represents research carried out over the course of my PhD degree, outside of *Chapter 1*, which reviews the prior work of others as part of an introduction to the background concepts involved.

Chapters 3, *4*, and *6* are original works documenting this research. I conceived of and performed all experiments (unless otherwise noted) and wrote the original chapter drafts. I was aided by my supervisor Dr. Vincent Tabard-Cossa in the initial conceptualization of these projects and, later, in edits to the written drafts that resulted in the final text versions presented here. In contrast, *Chapters 2* and *5* are published articles written with additional collaborators beyond my supervisor.

Chapter 2 (published as "Instrumentation for Low Noise Nanopore-Based Ionic Current Recording under Laser Illumination" in *Review of Scientific Instruments*) was written in collaboration with other students and post-doctoral fellows in the Tabard-Cossa Lab, namely, José Bustamante, Dr. Autumn Carlsen and Dr. Aidan Baker-Murray. These co-authors played roles in helping to: design the instrument, perform initial experiments, and draft the manuscript. For my part, I performed the final versions of the experiments (that produced the data appearing in the article), created figures for the article, wrote large sections of the final text, and led the manuscript submission and revision processes to get the article published.

Chapter 5 (published as "Analysis of Nanopore Data: Classification Strategies for an Unbiased Curation of Single-Molecule Events from DNA Nanostructures" in *ACS Sensors*), is similar to *Chapters 3, 4, 6* in terms of my participation (conceiving of and performing all experiments, writing the initial draft) but also featured contributions from Dr. Kyle Briggs, who had previously developed the Nanolyzer software that our lab members and others use to analyze nanopore data. For this work specifically, he helped devise strategies for analyzing the specific datasets involved and provided valuable edits to the article manuscript with his background knowledge of nanopore data analysis.

List of Contributions

Publications

1. **Roelen Z**, Tabard-Cossa V. Synthesis of Length-Tunable DNA Carriers for Nanopore Sensing. *PLOS One*. 2023;18: e0290559.
2. **Roelen Z**, Briggs K, Tabard-Cossa V. Analysis of Nanopore Data: Classification Strategies for an Unbiased Curation of Single-Molecule Events from DNA Nanostructures. *ACS Sensors*. 2023;8: 2809–2823.
3. **Roelen Z**, Bustamante JA, Carlsen A, Baker-Murray A, Tabard-Cossa V. Instrumentation for Low Noise Nanopore-Based Ionic Current Recording under Laser Illumination. *Review of Scientific Instruments*. 2018;89: 015007.
4. Tabard-Cossa V, Carlsen A, Charron M, Karau P, **Roelen Z**, Waugh M. Nanopores: Electronic Tools for Single-molecule Biophysics and Bio-nanotechnologies. *Physics in Canada*. 2017;73: 97–101.

Conference Presentations

5. **Roelen Z**, Bustamante JA, Carlsen A, Briggs K, Tabard-Cossa V. Poster: Integration of Optical Instrumentation and Microfluidics for CBD-Fabricated Nanopore Sensors. *Biophysical Society of Canada Annual Meeting*, Montreal, QC, 2017.
6. **Roelen Z**, Briggs K, Carlsen A, Tabard-Cossa V. Poster: Fabrication of ZMW Nanopores by Controlled Breakdown for Single-Molecule Optical Detection. *Biophysical Society Annual Meeting,* Los Angeles, CA, 2016.

List of Figures

List of Tables

List of Acronyms

2D:	two-dimensional
3D:	three-dimensional
AC:	alternating current
bp:	base pair
CBD:	controlled breakdown
CF:	concentration factor
DAQ:	data acquisition
DC:	direct current
DEP:	dielectrophoresis
DNA:	deoxyribonucleic acid
dNTP:	deoxynucleoside triphosphate
dsDNA:	double-stranded DNA
ECD:	equivalent charge deficit
eDEP:	microelectrode-based DEP
EDTA:	ethylenediaminetetraacetic acid
EOF:	electro-osmotic flow
EP:	electrophoresis
FEM:	finite element method
iDEP:	insulator-based DEP
IV:	current–voltage characteristic
LOD:	limit of detection
ML:	machine learning
nDEP:	negative DEP
nt:	nucleotide
OA:	oligonucleotide assembly
PCR:	polymerase chain reaction
pDEP:	positive DEP
PDMS:	polydimethylsiloxane
PEG:	polyethylene glycol
PSD:	power spectral density
PTFE:	polytetrafluoroethylene
RMS:	root mean square
SE-DNA:	sticky-ended DNA
SNR:	signal-to-noise ratio
ssDNA:	single-stranded DNA
TBE:	Tris-borate-EDTA
TD:	λ-terminase digest
TrECD:	truncated ECD
UV:	ultraviolet

CHAPTER 1 – INTRODUCTION

1.1 – Nanopore Sensors

Nanopores are powerful single-molecule biosensors, capable of the label-free detection and characterization of a broad range of target analytes (e.g. nucleic acids, proteins, nanoparticles) [1,2]. The principle of their typical operation is depicted in Figure 1.1a [3]. A nanoscale aperture through an otherwise electrically insulating membrane is placed between two electrolyte-filled reservoirs, connecting the two chambers in a narrow fluidic path. A pair of electrodes (one in each reservoir) is then used to apply a potential difference across the membrane, creating a background flow of electrolyte ions (K^+ and Cl^- in Fig. 1.1a) through the pore that is measured as ionic current by a high-bandwidth current amplifier. If a charged analyte (with critical dimensions on the scale of the pore dimeter) is present in solution on one side of the membrane, it can also be electrophoretically driven though the pore by the applied voltage and, in doing so, it temporarily blocks a measurable portion of this background ionic current. Figure 1.1a illustrates this process with a negatively charged DNA molecule; the inset on the right shows a representative current trace of it passing through the pore, blocking ionic current. Beyond just the overall depth of the current blockages (ΔI), nanopore signals can also be parametrized with additional metrics (e.g. duration [4], event substructure [5]) to further help characterize and identify the molecular species (by shape, charge, etc.) that generated them by translocating the membrane [1,2].

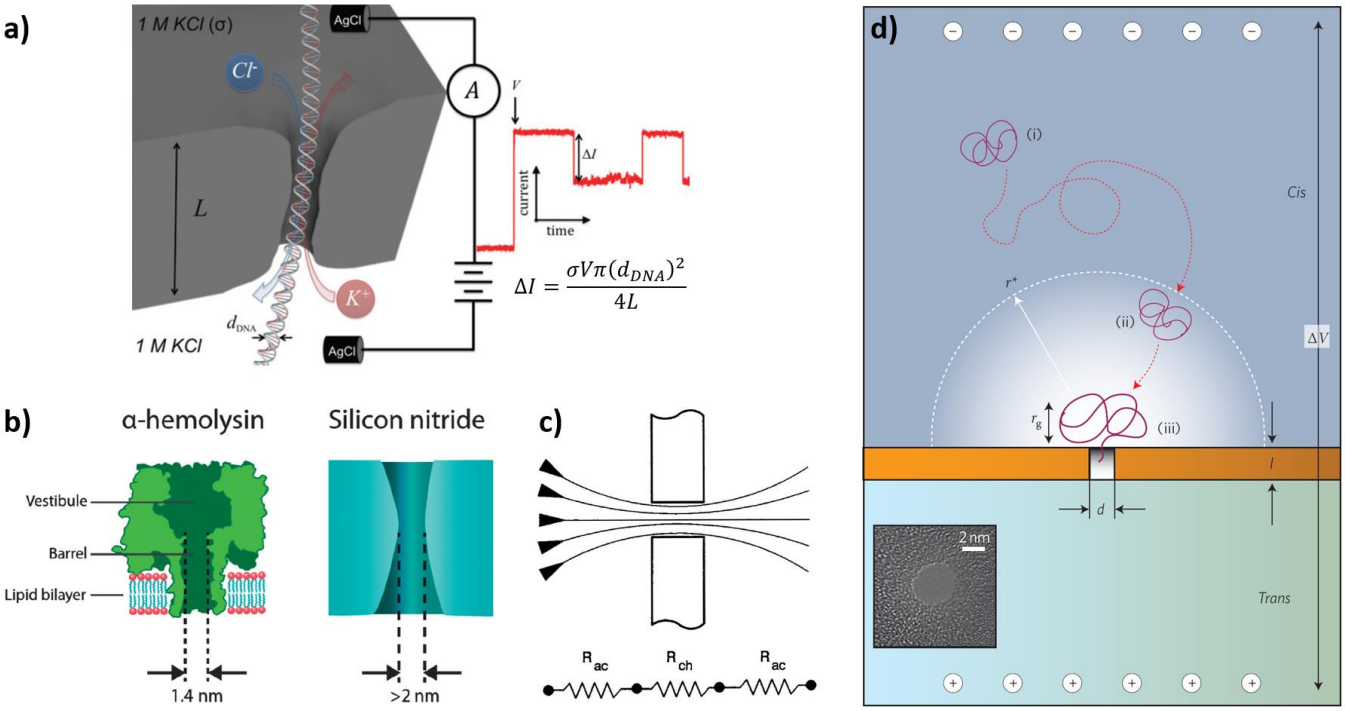


Figure 1.1: *Principles of nanopore sensing.* ***a)*** *Schematic of an experimental setup to capture and detect double-stranded DNA with a nanopore. The insets on the right show a sample current blockage signal, and an expression for the magnitude of the blockage depth as a function of the system variables. Image adapted from* [3]*, permission pending.* ***b)*** *Comparison of the internal structures of biological (left) and solid-state nanopores (right). Image adapted from* [6]*, with permission.* ***c)*** *Equivalent circuit model of the conductance through a nanopore. The total resistance of the system can be approximated as contributions in series from the bulk electrolyte near the pore entrance ("access resistance", R_{ac}) and from the pore volume itself ("channel resistance", R_{ch}). Image adapted from* [7]*, with permission.* ***d)*** *Illustration of the steps involved in target capture by a nanopore, involving: i) target diffusion to the "capture volume", ii) electrophoretic migration to the pore entrance, iii) target re-orientation to initiate translocation against an energy barrier. Image adapted from* [8]*, with permission.*

Nanopore sensors can be broadly divided into two categories: biological pores and solid-state pores (Figure 1.1b) [6]. Biological pores are based on transmembrane protein channels found in nature, where they serve various roles such as transporting important nutrients across bacterial membranes (e.g. MspA) and disrupting the membrane integrity of other cells as a toxin (e.g. α-hemolysin). Particular advantages to biological pores as sensors include their highly reproducible structures (set by their unique sequence of amino acids), their narrow ~1 nm constrictions (to help distinguish closely related analytes at this scale), and their ability to be engineered with additional functionalities such as by modifying their amino acid sequence or by acting as specific targets for other moieties. Biological pores have been successfully used in the past to sequence DNA [9], and to differentiate various unfolded peptide chains [2]. A major disadvantage to protein-based pores however results from the synthetic lipid bilayer that supports them (see Fig. 1.1b) – the eventual rupture of this delicate

substrate tends to limit the duration of pore sensing experiments [10] and these lipid bilayers can prove more difficult to integrate within complex microelectromechanical systems (MEMS). Solid-state nanopores, on the other hand, are conceptually simpler, typically consisting of a cylindrical or conical channel across a more robust dielectric membrane. While possessing a few downsides relative to their biological counterparts (e.g. difficulties with re-creating exact pore geometries, directly functionalizing their surfaces), solid-state nanopores can be fabricated in a large range of dimensions (length, diameter), and in a wide variety of materials (e.g. dielectrics: SiN_x [11], HfO_2 [12], monolayers: MoS_2 [13], graphene [14]) to best suit the needs of a particular target analyte. A particular low-cost and precise technique for fabricating solid-state nanopores (by "controlled dielectric breakdown") is discussed below in *Section 1.2*.

In order to better understand the ionic current signals from a nanopore sensor, it is useful to have a mathematical model of the pore's conductance states. A common model for the electrophoretic open-pore conductance through a cylindrical channel is represented by the following equation [15]:

$$G_{ch} = \frac{1}{R_{ch}} = \frac{\pi d^2}{4L}\left(\sigma + \frac{4q\mu}{d}\right) \tag{1.1}$$

where d and L are the pore diameter and length, respectively, σ is the electrolyte conductivity, q is the surface charge density inside the pore and μ is the electrophoretic mobility of the oppositely-charged counterions screening this surface charge. The first term represents the contribution to the current from the bulk electrolyte within the pore while the second term represents the contribution of the charged surface. In the typical, high-salt (≥ 1 M) nanopore sensing conditions used throughout this work, electrolyte conductivity is also high (specifically, $\sigma \gg q\mu/d$) and so the surface charge term can be essentially neglected. A further modification to the model involves including the effect of the decay in electric field strength just outside the pore entrance (see Figure 1.1c), and this is typically handled by adding "access resistance" terms in series with the main, cylindrical "channel resistance" [7]. A widely used mathematical form for the access resistance was derived by Hall [16],

using an analogy to the capacitance between a disk electrode (with the same diameter as the pore) and a hemispherical electrode at infinity, separated by an insulating material:

$$R_{ac} = \frac{1}{2\sigma d} \tag{1.2}$$

The final expression for the pore conductance in high salt conditions then becomes [17]:

$$\begin{aligned} G_{pore} &= (R_{ch} + 2R_{ac})^{-1} \\ &= \sigma\left(\frac{4L}{\pi d^2} + \frac{1}{d}\right)^{-1} \end{aligned} \tag{1.3}$$

As can be seen the above expression, including these access resistance terms becomes particularly important for nanopores that feature a large ratio of their diameter to their length (such that $d/L \gtrsim 4/\pi$), as would the case for wide pores (e.g. needed to accommodate large target analytes) or for pores in very thin materials (e.g. graphene).

Finally, for this section, we examine the process of target capture by the nanopore. As illustrated in Figure 1.1d [8], (and as approximated by the access resistance term discussed above), the electric field strength E generated by the potential difference between electrodes in each reservoir decays rapidly with distance away from the pore entrance. Since the velocity of particles under the electrophoretic mechanism used by the nanopore to capture them depends directly on this field strength ($v_{EP} = \mu_{EP}E$), the directed motion of charged analytes toward the pore becomes quickly washed out by random collisions with the bulk fluid (diffusive motion), even at moderate (~μm) distances away. This effect has been quantified with the concept of a hemispherical "capture volume" surrounding the pore, with a boundary that is defined by the condition that the magnitude of the electrophoretic velocity there be equal to a "diffusive velocity" ($v_D \sim D/r$) that represents the average speed with which a particle could migrate a distance r to that boundary from the pore by diffusion alone. By using an approximate, spherically symmetric expression for the distribution of the electric field outside the pore, this results in the following estimate for the radius r^* of this hemispherical capture volume [8,18]:

$$r^* = \left(\frac{d^2}{8L + 2\pi d}\right)\frac{\mu_{EP}\Delta V}{D} \tag{1.4}$$

where *d* and *L* are the diameter and length of the pore (as before), μ_{EP} and *D* are the electrophoretic mobility and diffusion constant of the target particle, and Δ*V* is the voltage across the membrane applied by the electrodes. The capture process of a charged target can then be broken down into the following steps [8,19,20], as depicted in Figure 1.1d: i) the target randomly diffuses within the bulk fluid until it arrives at the edge of the capture volume, ii) the target deterministically migrates along the electric field lines from the edge of the capture volume to the pore entrance, iii) the target re-orients itself at the pore entrance to overcome a free energy barrier and begin translocation. Typically, for a given set of experimental conditions (e.g. target molecule, nanopore geometry, transmembrane potential) the rate-limiting step to target capture and thus nanopore signal generation can either be waiting for the target to arrive to the capture volume ("diffusion limited regime") or waiting for the target to cross the energy barrier ("barrier limited regime"); transitioning between the two regimes is also possible, through modification of these experimental conditions (e.g. increasing the voltage) [8,21].

1.2 – Nanopore Fabrication by Controlled Breakdown

In this work, we used solid-state nanopores that were fabricated in-house just prior to running an experiment, through a technique called "controlled breakdown" (CBD) [22]. With this technique, thin (~nm thick) membranes are immersed in an electrolyte solution though which distant electrodes apply relatively large transmembrane potentials; the magnitude of this potential difference is set such that the electric field within the membrane is held near the dielectric strength of its insulating material. In this way, the applied voltage can quickly be removed moments (~ms) after the initial electrical breakdown of the material, resulting in a single nanoscale (as small as ~2 nm) perforation through the membrane. Before CBD was developed, solid-state nanopores were produced almost exclusively through sputtering of the membranes by focused beams of electrons or ions [23]. Benefits to CBD over this prior approach include considerably lowering the cost and

complexity of the fabrication apparatus (e.g. sidestepping the need for a multi-million dollar electron microscope) and forming the pores *in situ* under the same conditions that will be used during a sensing experiment (in contact with solution within a fluidic cell) – transitioning from nanopore fabrication to sensing is then as simple as flushing the fluidic reservoirs with a new electrolyte solution containing the target sample.

Figure 1.2a illustrates a basic set-up used to fabricate nanopores by CBD [22]. A solid-state chip (such as a TEM window) containing a thin dielectric membrane is first enclosed in a fluidic cell (typically made of an inert material such as PTFE or PEEK) whose reservoirs are filled with a conductive, pH-regulated buffer. Following this, a pair of non-polarizable electrodes (such as Ag/AgCl electrodes for use with chloride-based electrolytes), are inserted on each side of the membrane and used to apply breakdown voltages to the system, as supplied by a computer-controlled current amplifier. This current amplifier also simultaneously measures the leakage currents through the chip that are generated by the applied voltage – the magnitude of these currents suddenly rising past a particular threshold value is interpreted as a complete fluidic path (nanopore) being formed across the membrane and triggers a voltage shut-off. An overview of the mechanistic details proposed for the various stages of the CBD process is provided in Figure 1.2b-e [22].

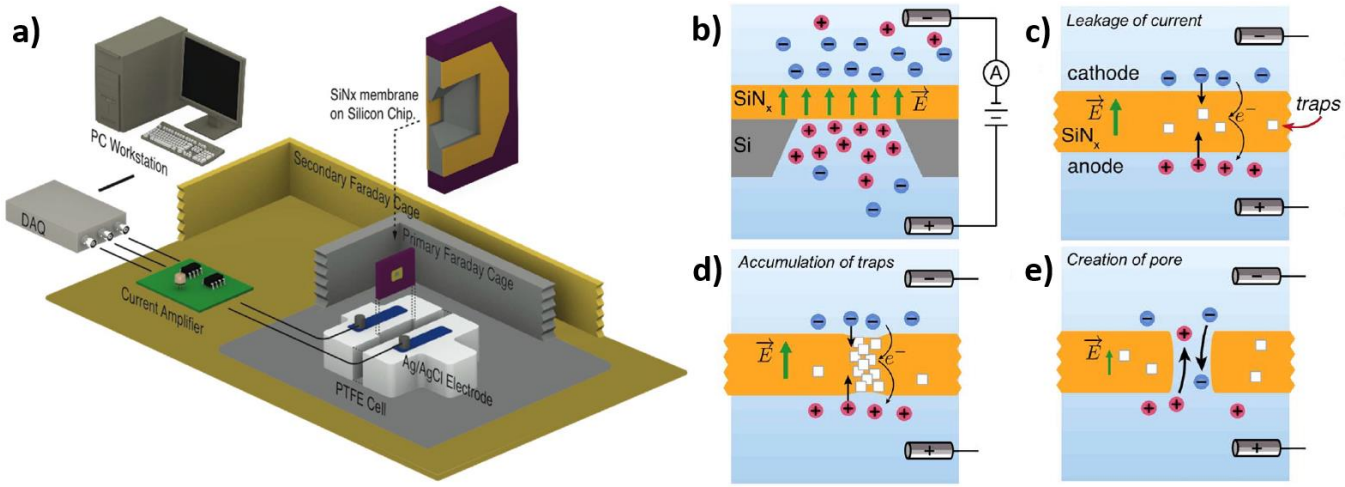


Figure 1.2: *Overview of the controlled dielectric breakdown method (CBD) of fabricating nanopores.* ***a)*** *Schematic of the basic CBD system, showing a SiN_x membrane (e.g. ~10-nm thick) embedded in a fluidic cell. A computer-controlled current amplifier applies elevated voltages (~10 V) across the membrane via Ag/AgCl electrodes in the cell, measures the resulting leakage current, and switches off the voltage when a large increase in current is detected (signaling the creation of the pore). The mechanism of voltage-assisted breakdown can be attributed to the following processes:* ***b)*** *generating an electric field inside the membrane and accumulating ions at the membrane-electrolyte interface,* ***c)*** *producing a leakage current though the membrane by charges tunneling across membrane defects ("charge traps"),* ***d)*** *increasing the density of traps along a conductive path via electric field-induced damage to the membrane,* ***e)*** *removal of membrane material at the current hotspot to create a single pore. Images adapted from* [22]*, with permission.*

Figure 1.3 presents measured outputs from each stage of a typical nanopore experiment, starting with the formation of the pore by CBD in Figure 1.3a. In this example, a ~12 nm thick SiN_x membrane was immersed in 1 M KCl buffer (pH 8) and subjected to a ramped voltage [24], increasing from 0 V to ~12 V over the course of ~7 minutes. As explained above, a threshold was set (grey dashed line in Fig. 1.3a, ~40 nA) such that when the leakage current across the membrane exceeded this value (in magnitude), the voltage was shut off to prevent further damage to the membrane beyond the desired formation of a single nanoscale hole. After pore formation, the solution in the reservoirs of the fluidic cell was exchanged for the buffer in which the eventual sensing experiment will take place, in this case 3.6 M LiCl, pH 8. Lithium chloride is known to slow down the passage of negatively-charged targets through a nanopore (relative to KCl) [25] through stronger binding of the smaller Li^+ counterions (further reducing the target's effective charge in solution), and thus can increase the time resolution of an experiment. Figure 1.3b presents the current response of the nanopore in this new buffer to a sweep in transmembrane voltage (I-V curve). The slope of the I-V curve corresponds to the conductance of the pore; through use of Equation 1.3, this conductance value can be translated into an estimate of the pore

diameter, by separately assuming the length of the pore *L* to be equal to the thickness of the membrane, and measuring the conductivity σ of the buffer. If the application for which the pore will be used requires a larger diameter than this conductance estimate indicates, the pore can be left in LiCl solution to naturally expand over the course of several days, or can also be subjected to moderate voltage pulses (roughly ~⅓ $\Delta V_{breakdown}$) to grow the pore on demand in a process referred to as "conditioning" [26].

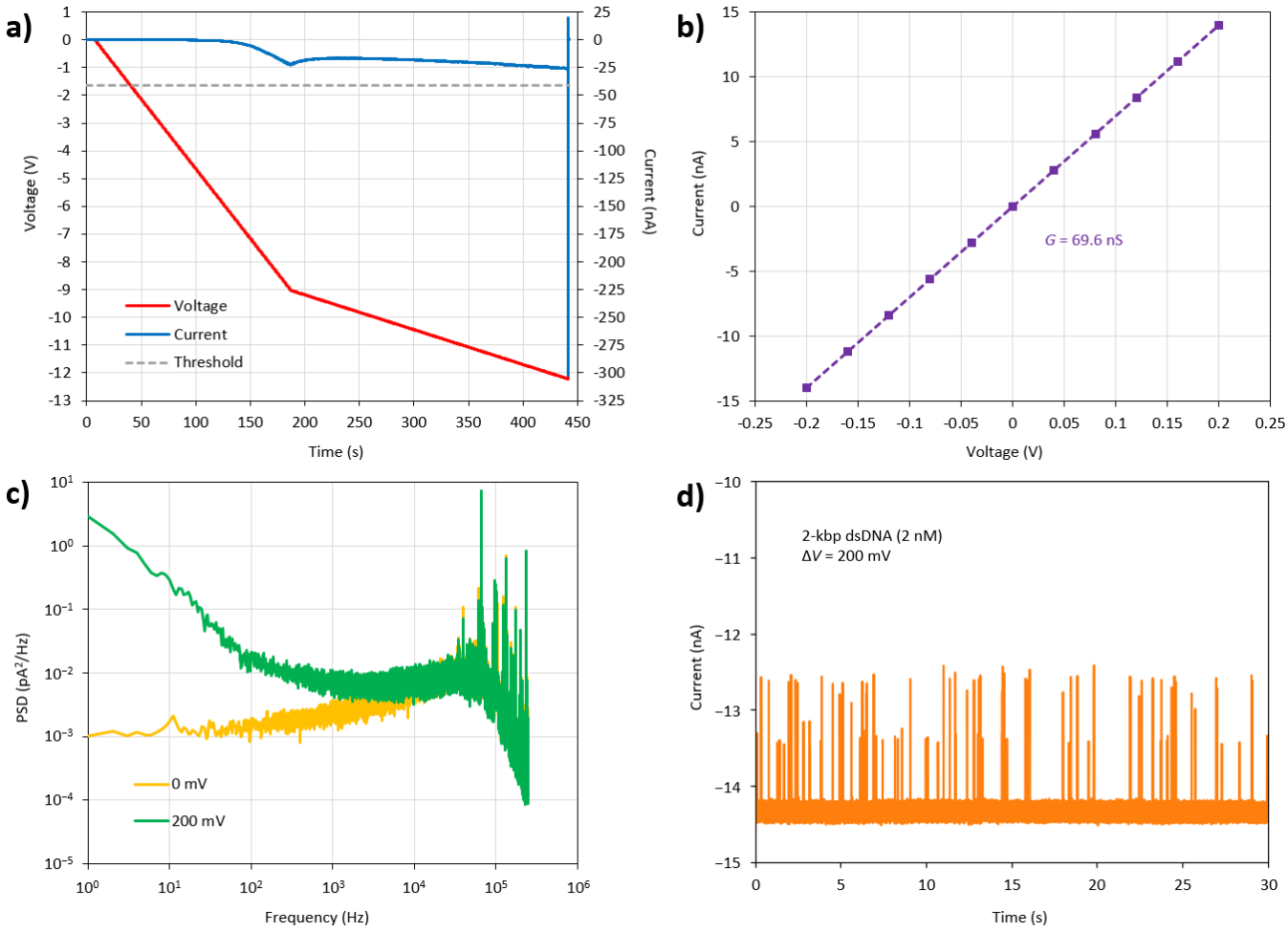


Figure 1.3: *Outline of a typical sequence used to fabricate a nanopore by CBD and use it in a biosensing application (in this case, with double-stranded DNA).* ***a)*** *Leakage current through a SiN_x membrane in response to an applied ramp in the transmembrane potential during CBD (in 1 M KCl, pH 8). The large spike in current (at ~440 s) signals the formation of a nanopore, at which point the voltage is switched off.* ***b)*** *Current-voltage characteristic of the pore fabricated in (a), showing a symmetric, linear response (in 3.6 M LiCl, pH 8). The conductance, G, though the pore can be used to provide an estimate of its diameter (in this case d ≈ 10 nm).* ***c)*** *Quantifying the signal noise in the combined nanpore-current amplifier system (Axopatch 200B) through power spectral densities (PSDs) of the measured current. Applying a voltage (200 mV) across the pore results in the appearance of an additional low-frequency contribution to the total noise ("flicker" or "1/f" noise).* ***d)*** *Recorded translocation signals of double stranded DNA (2 kbp), visible as transient reductions (in magnitude) of a background ionic current through the pore under an applied transmembrane potential (200 mV in 3.6 M LiCl). Panels (a)-(d) represent measurements from the same ~10-nm pore.*

Once a desired pore diameter is reached, the electrical noise present in recordings of ionic current through the pore is then typically quantified by examining its power spectral density (PSD), in order to ensure that overall noise levels are below the expected signal heights of translocating molecules ("signal-to-noise ratio", SNR > 1). *Chapter 2* presents a more detailed discussion of the origins of various noise contributions within the system, and at which parts of the frequency spectrum each contribution dominates. For now, we use the example of Figure 1.3c to note that an additional form of low-frequency noise ("flicker" or "1/*f*" noise) appears when a standard (~100 mV) target-capture voltage is applied across the pore [27]. Integrating the PSD produced under these representative conditions over the bandwidth of frequencies that will be used during the sensing experiment can then provide an estimate of the expected noise levels, usually expressed as a "root-mean-square" (RMS) value. If the noise through a pore is unacceptably high at this stage, additional "conditioning" pulses may be applied to try and stabilize the baseline current [26], or fabrication of a new nanopore on a different membrane can be attempted (facilitated by the ease and speed of the CBD technique).

Since Figure 1.3 has been tracking a nanopore that passed all these steps, the system is now ready for target analyte sensing. In this example, a sample of double-stranded DNA (2 kbp) suspended in the sensing buffer (2 nM DNA in 3.6 M LiCl) is added to one reservoir of the fluidic cell, and a voltage with the correct polarity is applied by a high-bandwidth current amplifier (e.g. Axopatch 200B) to translocate molecules through the pore and record their current signals. For negatively charged DNA, the reservoir in which the sample was placed ("*cis*") must be biased lower than the reservoir on the other side of the membrane ("*trans*") in order to draw molecules to the pore. Figure 1.3d shows a sample (30 s) current trace of the translocation signals produced under these conditions. The subsequent process of locating and analyzing such translocation signals within the full dataset of recorded ionic current, so as to extract information about the molecules that produced them, is explored in detail in *Chapter 5*.

1.3 – Increasing Nanopore Capture Rate/Lowering Limit of Detection

As outlined in *Section 1.1*, the process by which target molecules are captured by a nanopore sensor is associated with a finite duration (see Figure 1.1d), which usually sets the limiting rate at which sample-specific data can be acquired by the system. In particular, under typical sensing conditions (e.g. target sample, pore dimensions, applied voltages) nanopores often operate in the "diffusion-limited regime" referenced above – that is, the majority of time during an experiment is spent waiting for target molecules to randomly diffuse from a bulk reservoir volume (~µL–mL) into to a much smaller capture volume (~fL–pL) immediately outside the pore [28,29]. An estimate for this diffusion-limited rate of data acquisition can be made by picturing the boundary of the modelled hemispherical capture region (surface area $2\pi[r^*]^2$) to be perfectly absorbing, such that all molecules that arrive there by diffusion are eventually sensed by the pore [8,19]:

$$\begin{aligned} \frac{dn}{dt} &= 2\pi D r^* c_0 \\ &= \left(\frac{\pi d^2}{4L + \pi d}\right) \mu_{EP} \Delta V c_0 \end{aligned} \tag{1.5}$$

where dn/dt is the number of targets arriving per unit time, r^* is the radius of the capture volume (inserted from Equation 1.4), d and L are the diameter and length of the pore, ΔV is the voltage between the electrodes, and μ_{EP}, D, and c_0 are the electrophoretic mobility, diffusion constant, and bulk concentration of the target molecule, respectively.

If the event rate in a nanopore experiment is found to be too low (i.e. requiring impractically long experiment times to generate a statistically significant number of events), Equation 1.5 points to several ways through which this rate could be increased (dn/dt ↑), such as by using larger pores (d ↑), applying higher voltages (ΔV ↑), and raising the sample concentration (c_0 ↑). Not all of these solutions are always practical, however, or come without significant downsides to the overall experiment. For instance, larger pores also feature larger baseline currents, which can drown out the signals of translocating molecules (generated by blocking a small portion of this current), leading to low signal-to-noise ratios. Likewise, increasing the voltage across the membrane expands

the capture volume, but leads to faster electrophoretic velocities within the pore (shorter dwell times) [30], lowering the time resolution of the events. Raising the target concentration within a sample is of course the most straightforward approach to having more molecules arrive at the capture volume, but this cannot always be set as an independent variable within an experiment. In fact, a frequently envisioned application for nanopore sensors is to determine an initially unknown target concentration in a sample of interest, for instance using the biomarker count of a patient-collected specimen as part of a medical diagnosis [31]. Having a minimum concentration limit (~0.1 nM, as set by diffusion) to a practical nanopore experiment therefore restricts the breadth of biologically relevant samples that are accessible to the system.

Figure 1.4 highlights several examples of some more refined approaches to increasing the capture rate of solid-state nanopores (beyond just adjusting the variables of Eq. 1.5 from the start of an experiment) that have been implemented in the past. Many of these incorporate the idea of "active transport", that is, using a secondary mechanism to directly migrate targets into the capture volume of the pore, rather than passively waiting for diffusion to take place. The experimental setup and main results of each example are briefly outlined below.

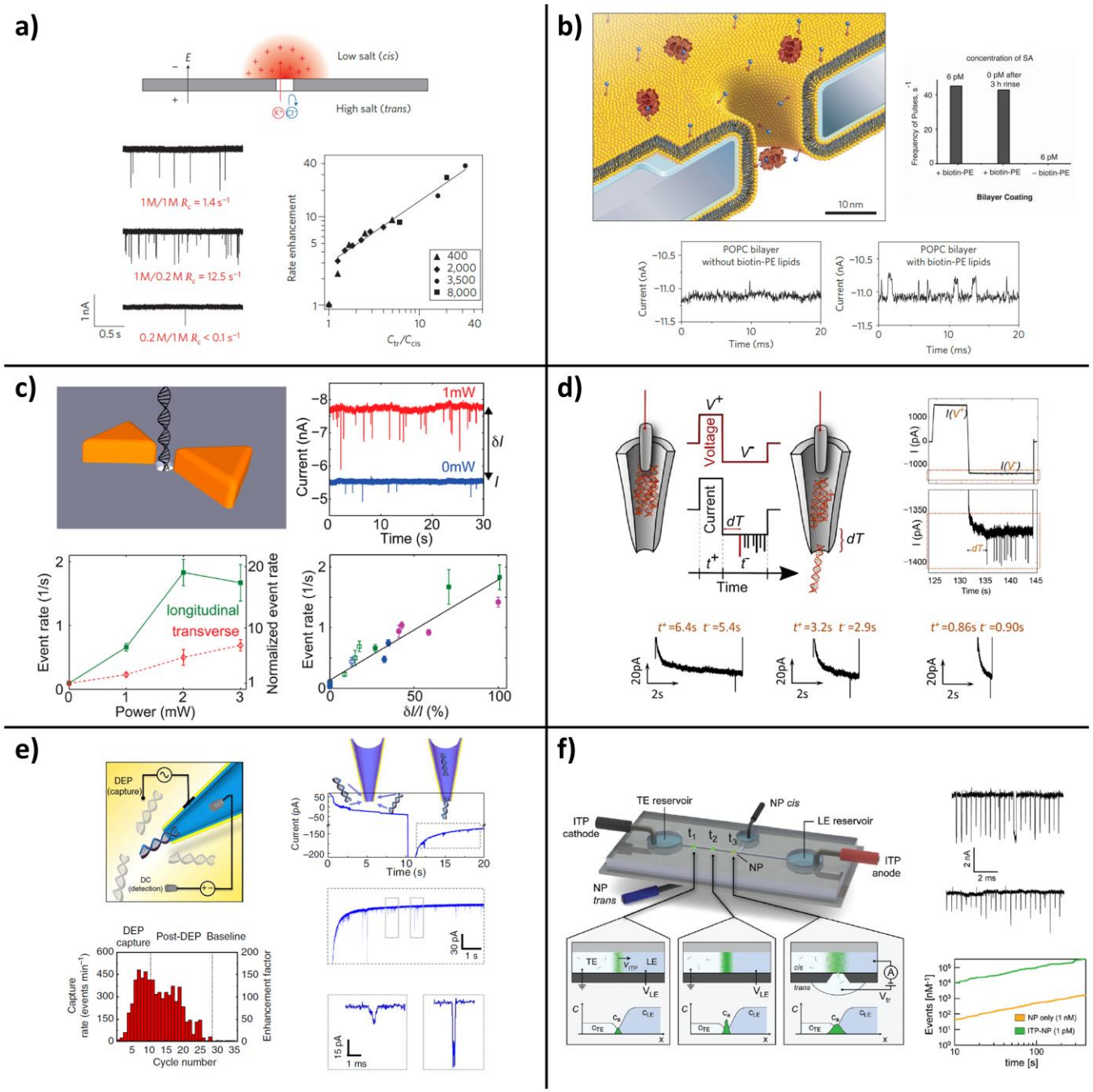


Figure 1.4: *Survey of existing experimental approaches to increase the capture rate of solid-state nanopores. See main text for a detailed description of each approach.* ***a)*** *Use of a trans-nanopore salt gradient to enlarge the extent of the electric field. Images adapted from* [8]*, with permission.* ***b)*** *Target capture by mobile ligands embedded in the lipid bilayer coating of a solid-state nanopore. Images adapted from* [32]*, with permission.* ***c)*** *Enhanced target delivery by thermophoresis within a nanopore/plasmonic nanoantenna system. Images adapted from* [33]*, with permission.* ***d)*** *Controlled delivery of targets inside a nanopipette using pulsed voltages. Images adapted from* [34]*, with permission.* ***e)*** *Dielectrophoretic target trapping with a metalized nanopipette. Images adapted from* [28]*, with permission.* ***f)*** *Microfluidic system combining isotachophoretic sample focusing and nanopore sensing. Images adapted from* [29]*, with permission.*

Transmembrane Salt Concentration Gradients (Fig. 1.4a)

In this technique, the two fluid reservoirs on either side of the membrane are held at different salt concentrations from each other. This has the effect of extending the electric field away from the pore within the low-salt reservoir and therefore expanding the capture volume there (relative to symmetric salt conditions, for a given value of ΔV) [20]. Wanunu et al. used this concept with small (< 5 nm) SiN_x pores in KCl buffer to successfully characterize dsDNA samples down to bulk concentrations of 3.8 pM [8]. For example, using a 20× salt gradient (0.2 M/4 M KCl), capture rates of 8 kbp DNA were observed to increase by ~30× when compared to symmetric salt conditions (1 M).

Membrane-Embedded Ligands (Fig. 1.4b)

In the work of Yusko et al. [32], SiN_x nanopores were coated with lipid bilayers containing a fraction of lipids that were functionalized to act as receptors for particular targets in solution (e.g. biotinylated lipids targeting free streptavidin molecules). Because these lipid-anchored receptors are free to move within the overall fluid bilayer, once targets from the bulk solution concentrate onto the membrane surface through the specific binding interactions, the target-receptor complexes are able to migrate to the pore by a 2D diffusion process (where they then translocate under an applied voltage). Using this technique, the authors detected small proteins (streptavidin) as well-resolved translocation signals at concentrations down to 6 pM. Capture rates of the 6 pM sample were enhanced by a factor of ~500× when compared to a control with a membrane-coated nanopore that lacked the mobile ligands (~20-nm pore in 2 M KCl).

Plasmonic Nanoantennas (Fig. 1.4c)

The system of Nicoli et al. consists of a gold bowtie antenna (triangular side lengths ~70 nm) centred on either side of a ~10-nm SiN_x nanopore (on the membrane surface) [33]. When the nanopore-bowtie is illuminated with polarized laser light (λ = 785 nm), plasmonic excitations in the nanoantenna lead to localized heating at its focus, resulting in higher baseline currents through the pore and higher blockage depths of translocating molecules. Relevant to the current discussion, plasmonic excitation also leads to increased nanopore capture rates by

delivering target molecules to the pore along the created temperature gradients, in a mechanism called thermophoresis. For instance, applying 2 mW of longitudinally polarized laser light resulted in a >10× rate enhancement of dsDNA translocations (48.5 kbp) in 2 M LiCl.

Controlled Delivery within Nanopipettes (Fig. 1.4d)

Glass nanopipettes represent a common alternative to membrane-based solid-state nanopores [35]. In the work of Ivanov et al. [34], rather than the typical configuration of capturing molecules from a bulk solution outside of the nanopipette, DNA molecules were first loaded on the interior side of the conical pipette and then delivered to the terminal nanopore with controlled voltage pulses (comprising a positive pulse drawing the negatively-charged DNA away from the tip followed by a longer negative pulse drawing it toward the tip). The confined geometry inside the nanopipette extends the electric field farther away from the pore than in a cylindrical, membrane-based nanopore of similar diameter, and dsDNA (5–10 kbp) could be reliably delivered and detected from pre-loaded concentrations as low as 3 pM (~25-nm pores in 0.1 M KCl).

Pre-concentration by Dielectrophoresis (Fig. 1.4e)

Freedman et al. coated a glass nanopipette with a thin (5 nm) gold layer and positioned it (submerged in 1 mM KCl) 10s of μm away from a second, planar electrode consisting of a similarly gold-coated glass slide [28]. When a large AC voltage (10–20 V_{pp}, 1 MHz) is applied between the two gold surfaces, dsDNA in solution is attracted to the sharp tip of the nanopipette though the mechanism of dielectrophoresis (DEP), which acts on the induced polarization of the target under an external electric field. By cycling between 10-s intervals of high-*V* DEP concentration and low-*V* nanopore sensing (500 mV applied by Ag/AgCl electrodes), the event rate of 10 kbp dsDNA translocations in subsequent sensing phases increased as more and more DNA was accumulated at the tip of the pore during the concentration phase. Capture rate enhancements of ~80× were observed (compared to without DEP), and this allowed for the efficient detection of DNA down to concentrations of 5 fM.

Pre-concentration by Isotachophoresis (Fig. 1.4f)

Isotachophoresis (ITP) is a technique used to concentrate samples at the interface between two electrolyte solutions of different electrophoretic mobilities as they migrate down a linear microfluidic channel. Spitzberg et al. integrated such a channel above one side of a SiN_x nanopore (acting as the *cis* chamber) so that when the focused sample region passed above the membrane, the large transverse ITP voltage (100 V) driving motion down the channel was switched off and a smaller transmembrane potential was applied (500 mV) to capture molecules with the pore from this zone of enhanced concentration [29]. Using their system to focus a 1 pM sample of 5 kbp dsDNA, a >300× increase in event rate was observed following a ~3 min ITP focusing time (6-nm pore in ~300 mM KCl).

Additional Strategies for Lowering the Limit of Detection (Fig. 1.5)

As an alternative to the above examples where nanopores are used in their standard mode of operation (directly translocating molecules of interest across a pore and recognizing the distinct current blockage signatures they generate), some authors have altered this basic scheme to achieve gains in the limit of detection (LOD) for particular targets. Figure 1.5 illustrates two such approaches based on first binding the targets with larger nanoparticles and then capturing the target-nanoparticle complexes with nanopores that taper to diameters that are too small to allow the particles to fully translocate. The end state of a successful target detection is then a partially clogged nanopore that passes less ionic current than in its open state – importantly, the experiments are designed such that permanent clogs only occur when a target-specific interaction with the nanoparticle has taken place.

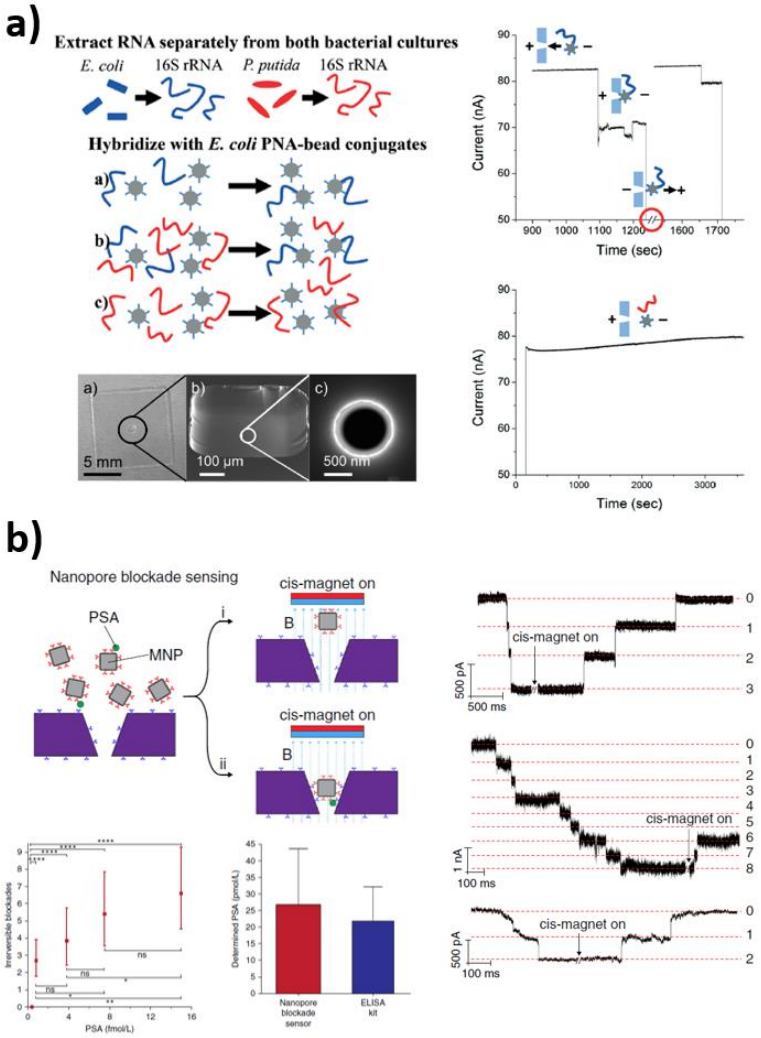


Figure 1.5: *Alternative, amplification-free approaches to lowering the limit of target detection of solid-state nanopores based on permanently blocking the baseline current with nanoparticles that are larger than the pore diameter.* ***a)*** *Probing charged nucleic acids with neutral, PNA-functionalized capture beads. Only when sequence-specific binding occurs between the target RNA and a PNA probe can the bead-RNA complex with net charge migrate to the pore by electrophoresis. Images adapted from* [36]*, with permission.* ***b)*** *Sandwich assay with antibody-functionalized magnetic nanoparticles and nanopore arrays. Proteins are captured by the magnetic nanoparticles and drawn to a nanopore-containing membrane with an external electromagnet. There, antibodies within the pore bind a separate region of the protein so that reversing the magnetic field does not clear the pore of the protein-nanoparticle complex. Images adapted from* [37]*, with permission.*

In the first example, by Koo et al. (Figure 1.5a) [36], polystyrene beads (820 nm diameter) were functionalized with peptide nucleic acid (PNA) probes for specific bacterial RNA sequences, in a way that left the final bead-PNA conjugates electrically neutral. Only when the bead conjugates have bound their negatively charged RNA targets (with as little as a single RNA molecule per bead) can the complex as a whole be electrophoretically captured by the electric field outside of a glass nanopore (1.5 V transmembrane potential, in 10 mM KCl). Using this set-up, *E. coli* ribosomal RNA was reliably detected down to 1 aM of sample concentration, even in a

background of 1 pM non-specific RNA (ribosomal RNA from a different bacterial species). The key to achieving such low LODs with this system seems to stem from using large (~750 nm) nanopores and high voltages, which enlarges the capture volume (see Eq. 1.5), as well as from needing only a single (permanent) blockage event per experiment to confirm the "yes/no" presence of the target analyte. Note, however, that this comes at the cost of a potentially reduced set of information when compared to a conventional translocation-based experiment, for example, regarding the exact target concentration (more concentrated samples result in higher translocation rates, see Eq. 1.5) and in the distinct current blockage signals generated by different translocating molecules – target specificity here has instead been shifted to the design of the PNA probes.

The second example in Figure 1.5, by Chuah et al. (Figure 1.5b) [37], is conceptually similar, this time involving magnetic nanoparticles (MNPs, 50 nm diameter) functionalized with protein capture antibodies. In this setup, an external magnetic field is used to actively transport the particles down to the surface of a nanopore-containing SiN_x membrane, where they can more easily be captured by the electric field though the pore, applied in tandem (100 mV in 100 mM KCl). Critically, the SiN_x surface is also functionalized with antibodies for the target protein (recognizing a different epitope than the first set on the MNPs) so that when a nanoparticle bound to a target protein arrives to a pore, it can form a "sandwich" complex with the membrane though the antibody-protein-antibody interaction (see Figure 1.5b). Reversing the direction of the magnetic field dislodges all of the "temporary" pore clogging events caused by free MNPs – permanent clogs are attributed to the specific formation of the sandwich complex and thus the presence of the target protein. Under this scheme, concentrations of prostate-specific antigen (PSA) down to 0.8 fM were successfully detected. While the pore diameters (130 nm tapering to 30 nm) and applied voltages used in this case are smaller than in the previous example (Figure 1.5a) the overall capture rate is increased by both the active transport of nanoparticles by the magnetic field and by using an array of nine nanopores (3 × 3) on each membrane – the fraction of permanently clogged pores within the array at the end of an experiment can also be used (somewhat) as a measure of target protein concentration (see bottom left panel of Figure 1.5b).

1.4 – Properties of DNA

Deoxyribonucleic acid (DNA) is a ubiquitous molecule in biosensing applications, due to its roles as both the genetic foundation of living things (encoding the amino acid sequence of proteins), and as a model linear polymer whose properties have been extensively studied [38,39] (and that can therefore be used to help characterize the sensing system itself). The basic chemical structure of DNA is illustrated in Figure 1.6a [40], and is built as a linear chain from individual building blocks called nucleotides. Each nucleotide consists of a pentose sugar attached to both a nitrogenous base (of one of four varieties: A, C, G, T) and a phosphate group. Within a single DNA strand (ssDNA), adjacent nucleotides are linked by covalent bonds between the 3′ hydroxyl group of one sugar molecule and the 5′ phosphate group of the next, forming the "backbone" of the strand. The nitrogenous bases of the strand, meanwhile, are free to form stable hydrogen bonds with other bases, either from the same strand or from a second strand, and do so in a way that is specific to their base variety ("base pairing"): A bonds with T, C bonds with G. Two DNA strands are said to be "complementary" if their sequences of bases are exactly mirrored by this base pairing rule (A ↔ T, C ↔ G) and if this sequence runs in the opposite direction, denoted by the positions of the phosphate and hydroxyl groups referenced earlier (5′→3′ ↔ 3′→5′).

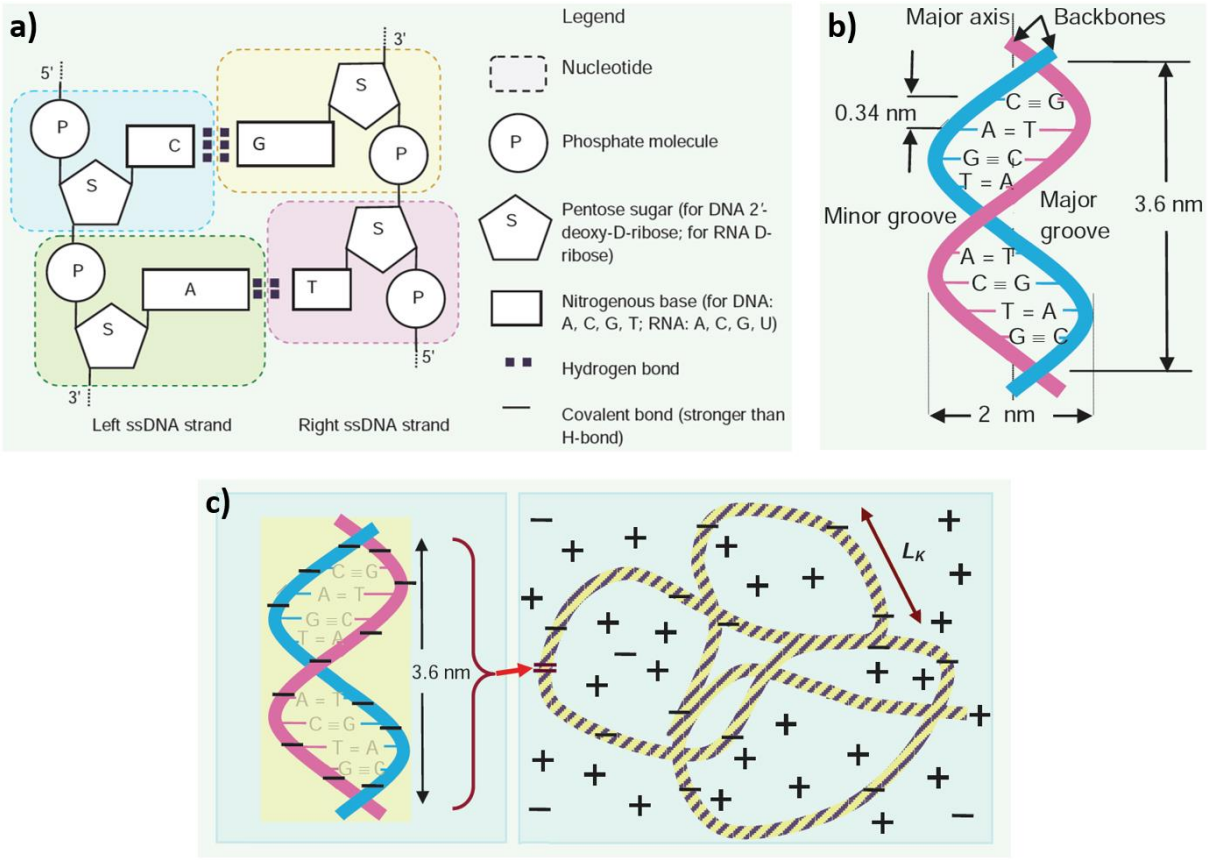


Figure 1.6: *Structure and conformation in solution of double-stranded DNA.* ***a)*** *Schematic of the chemical composition and base-pairing mechanism of dsDNA.* ***b)*** *Dimensions of the predominant double-helical form of dsDNA (B-DNA).* ***c)*** *Illustration of the conformation of dsDNA in ionic solution. Screening of the negatively charged sugar-phosphate backbone by positive counterions allows DNA to be arranged as a relatively compact random coil (worm-like chain). The step size of this random walk can be characterized by the Kuhn length (L_K), which is related to the polymer stiffness under these conditions. Images adapted from* [40]*, with permission.*

When two complementary strands encounter each other in solution, under appropriate ("renaturing" or "annealing") conditions of temperature and solvent composition, hydrogen bonding between their bases results in the spontaneous formation of a double-helical structure referred to as double-stranded DNA (dsDNA). Within the double helix, the negatively charged (hydrophilic) sugar-phosphate backbones are located on the outside of the molecule, in contact with the aqueous solvent, and the hydrophobic bases are located in the interior. Figure 1.6b [40] lays out the main dimensions of the predominant form of dsDNA found in nature ("B-DNA"), entailing a 0.34-nm spacing between base pairs, a 2-nm helix diameter, and a 3.6-nm length along the axis for the helix to twist back to its starting orientation. If the DNA-solvent system is shifted away from renaturing conditions, for instance by raising the temperature, the hydrogen bonds holding the double helix together can

be reversibly disrupted to expose individual bases over sections of the DNA or even to completely separate the two ssDNA strands from each other (DNA "denaturation" or "melting").

Figure 1.6c illustrates the conformation of a "long" (as defined below) dsDNA molecule in an aqueous electrolyte solution [40]. Over "short" distances, the stacking of the nitrogenous bases within the double helix give dsDNA a structural rigidity that tends to keep it aligned as a relatively inflexible rod (as in Figure 1.6b). However, over much larger distance scales, thermal fluctuations from the solvent tend to randomize the orientation of the molecule (aided by electrolyte counterions screening the backbone charge, suppressing repulsive Coulomb self-interactions), which minimizes the free energy of the system. This balance between the energetic cost to bending and the thermal energy of the system is characterized within the "wormlike chain" model of polymer conformation by a length scale called the persistence length L_P [38]:

$$L_P = \frac{YI}{k_B T} \tag{1.6}$$

where YI is the "flexural rigidity" of dsDNA (composed of its Young's modulus Y multiplied by its second moment of cross-sectional area I), k_B is the Boltzmann constant and T is the absolute temperature. A second model of polymer conformation ("freely-jointed chain") instead characterizes the molecule's path through space as a random walk of step size L_K called the Kuhn length [38]. A direct connection between the persistence length and the Kuhn length can be made by comparing predictions of the two models, resulting in the simple relation $L_K = 2\ L_P$. A long dsDNA molecule in solution (length $\gg L_P \approx 50$ nm $\rightarrow$ 150 bp) can thus be pictured as a series of connected, relatively straight segments (length L_K, see Figure 1.6c) that are oriented randomly from each other.

The picture of DNA as a random walk also leads to an estimate of the overall size of this random coil in solution, often characterized by the radius of gyration r_G [38]:

$$\sqrt{\langle {r_G}^2 \rangle} = \sqrt{\frac{L_C L_K}{6}} \tag{1.7}$$

where L_C is the contour length, that is, the extended end-to-end length of the molecule (equal to the number of base pairs N_{bp} multiplied by 0.34 nm/base pair for DNA). The radius of gyration is simply the root-mean-square distance of each monomer in the coil to its centre of mass. Notably, Equation 1.7 predicts that the size of a DNA molecule in solution scales with the square root of its linear length ($r \propto L_C^{1/2}$). This has important consequences for the diffusion of DNA, since this motion has been well-described by a "non-draining polymer" model (governed by "Zimm dynamics") in which the molecule and its surrounding solvent diffuse together as a solid object, mediated by the viscous coupling between the two components [38]. The Stokes-Einstein relation predicts that such a solid object (approximated as a sphere) migrates with diffusion constant D [38]:

$$D = \frac{k_B T}{6\pi\eta r} \tag{1.8}$$

where η is the dynamic viscosity of the solvent and r is the radius of the object. Longer DNA molecules are therefore predicted to have smaller diffusion constants, scaling as $D \propto r^{-1} \propto L_C^{-1/2}$. Simple modifications to the random walk model of polymer conformation that incorporate "self-avoidance", that is prevent two monomers of the molecule from occupying the same volume in space (prohibited by electrostatic repulsion), also modify the predicted scaling of polymer size vs. contour length to be $r \propto L_C^{3/5}$ [39] and thus predict $D \propto L_C^{-3/5}$. Experiments measuring the diffusion of fluorescently labelled dsDNA in solution have typically observed a scaling value that falls somewhere between the two above estimates, for instance $D \propto L_C^{-0.57}$ [39].

This dependence of dsDNA diffusion on fragment length greatly contrasts with its behaviour during electrophoresis, since the latter effect is better described by modelling DNA as a "free-draining polymer" (governed by "Rouse dynamics") [38]. In this model, the solvent molecules surrounding the charged polymer are unaffected by its electrophoretic motion and so do not mediate coupling between distant segments of the coil through viscous interactions (these are screened by oppositely charged counterions). In this way, each charged segment of the polymer electrophoreses essentially independently from each other, resulting in the prediction that the electrophoretic mobility ($\mu_{EP} = v_{EP}/E$) of sufficiently long dsDNA ($L_C \gg L_P$) is independent of its contour

length [38]. The length-independence of μ_{EP} has important implications for nanopore sensing as, referring back to Equation 1.5, we see that it should translate into a similar length-independence for the diffusion-limited capture rate. The dependence on dsDNA length of both free-solution electrophoresis and nanopore capture have been tested experimentally, and a length-independent regime ($N_{bp} \gtrsim 1000$) has been verified in each case [8,39].

We end this section by noting that the precise physicochemical properties of DNA outlined above have enabled the development of several powerful molecular biology techniques that helped revolutionize progress within the fields of medicine and biology. In particular, here we highlight the "polymerase chain reaction" (PCR) technique, which we make extensive use of in many of the thesis chapters that follow (see *Chapters 3, 4, 5*). The basic goal of PCR is to amplify (create many copies of) a specific sequence within a target DNA molecule in solution (called the "template"). This could be motivated by, for instance, producing a large amount of genetic material with a particular sequence for further study, or to detect a rare nucleic acid species (e.g. DNA or reverse-transcribed RNA) within a sample. The overall process is outlined in Figure 1.7 and it relies directly on both the precise base-pairing rules that govern hydrogen bonding between DNA strands as well as the ability to reversibly disrupt these bonds with small changes to the reaction conditions, in this case raising and lowering of the temperature.

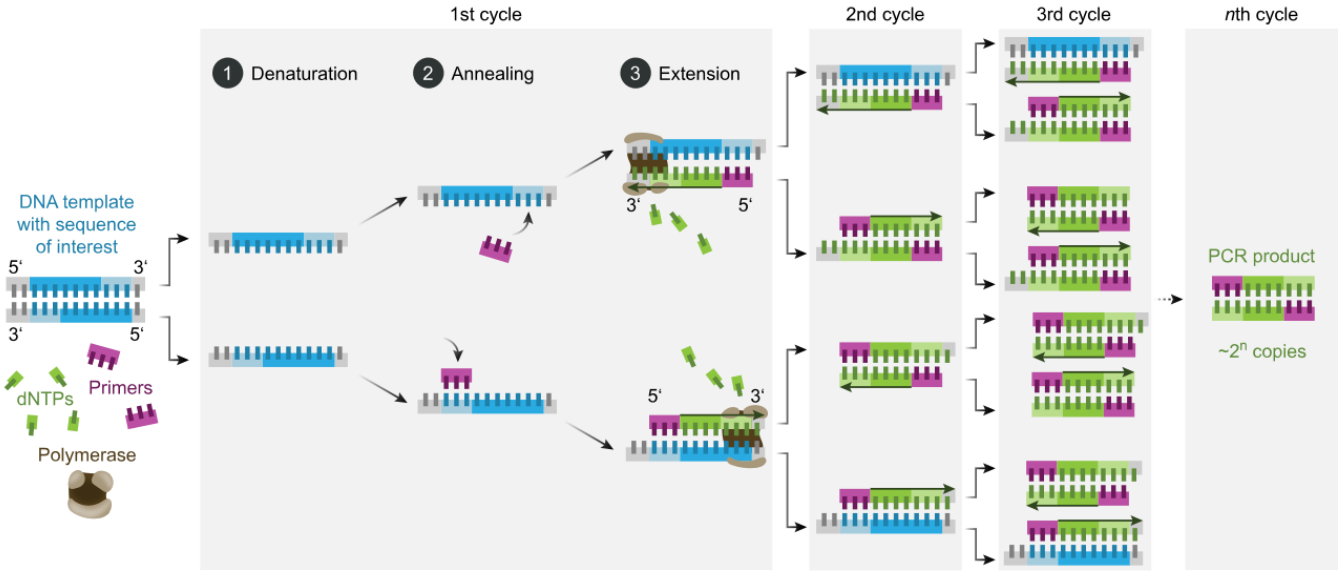


Figure 1.7: *Overview of the polymerase chain reaction (PCR). In each cycle, a dsDNA template is heat denatured to allow a short complementary oligonucleotide ("primer") to anneal to a specific single-stranded sequence in the template. A thermostable DNA polymerase then extends the growing primer strand by incorporating free dNTPs in solution base-by-base to the 3' end. The result is a doubling of dsDNA products (with sequences bounded by the two primers) after each cycle. Image by Enzoklop, "Schematic mechanism of PCR", 2020, Wikimedia Commons.*

In addition to the DNA template, the main reactants involved in PCR are: a pair of short ssDNA molecules ("primers", e.g. ~20-40 bases) that are complementary to regions just upstream (in the 5' direction) of the target sequence on both template strands, a heat-stable DNA polymerase enzyme (e.g. *Taq* polymerase, isolated from the thermophilic bacteria *Thermus aquaticus*), and a supply of triply phosphorylated nucleotides ("dNTPs" of each variety, A, C, G, T) that the polymerase can use as raw material to produce copies of the template sequence. Once all the components are combined in a reaction volume (~50 µL), the first step of the process consists of raising the temperature of the solution above the melting point of DNA, so as to expose the free bases of the template sequence (Step 1: "Denaturation" in Figure 1.7). Next, the temperature is lowered to be just beneath the melting point of the primers, which are then able to hybridize with their newly exposed complements on the template strands (Step 2: "Annealing"). Finally, the temperature is shifted once more, this time to a value that optimizes the performance of the polymerase, usually in between the temperatures of the first two steps; the enzyme acts by extending the 3' ends of the template-annealed primers (through covalent addition of free dNTPs) to create new strands that are complementary to the template (Step 3: "Extension"). This process of denaturation → annealing → extension (collectively referred to as a "cycle") is then repeated,

each time resulting in the number of DNA copies present in the reaction volume being roughly doubled; specifically, it is the portion of the template sequence bookended by the primers that is amplified (see Figure 1.7). The temperature cycling that powers PCR is typically automated in a machine specialized for this purpose ("thermocycler"), and the number of cycles performed is chosen such that a desired copy number for the target sequence is achieved.

1.5 – Structured Targets for Nanopore Sensing

The properties of dsDNA make it, in many ways, an ideal target for nanopore sensing. This includes the high levels of electric charge on its backbone that allow it to be readily captured by an electric field, its relatively rigid structure (e.g. compared to ssDNA [41]) that leads to prolonged translocation times as it sequentially threads through the pore, and its length-independent capture rate under certain conditions (discussed above) that enables the efficient detection of multiple DNA species in a single experiment. On the other hand, a notable downside with using dsDNA is that the nanopore signal it generates in its native state is also largely independent of the exact sequence of bases it contains, and instead mainly just reflects the current blocked by its uniform (~2 nm) diameter [11]. This has important consequences for sensing applications as it is often precisely this sequence information that distinguishes the significance of detecting particular DNA species in an experiment (e.g. as specific biomarkers for disease diagnosis) [42]. The situation is similar for the other main biomolecule type, proteins, which, in their fully folded states, often assume a compact, globular shape that results in relatively featureless and poorly time-resolved current blockage signals on a nanopore [43]. In particular, such signals tend to, again, be independent of the exact (amino acid) sequence or enzymatic function of the proteins that produced them.

There is therefore value to adding greater target "specificity" into the nanopore system, such that target molecules reliably produce signals containing information that is specific to their molecular identity, even if surrounded by closely related molecules. This would allow individual translocation signals to be easily associated to a target species (e.g. with a particular diagnostic value) and thus report on, for instance, the concentration of

that target within a mixture, or on interactions between the target and other components of the mixture. We return to this important topic in more detail in *Chapters 4* and *5*. In the meantime, the remainder of the current section highlights a particular approach to increasing the specificity of nanopore detection that has been implemented in the past, namely by engineering target molecules to contain particular physical or chemical substructures in their designs. The idea is that the *physical* substructures of the molecules are translated into identifiable *signal* substructures through the modulations in the current blockages they produce while translocating the pore [1]. Representative examples of such structured nanopore targets that are built from dsDNA are illustrated in Figure 1.8 and a brief description of each example follows below.

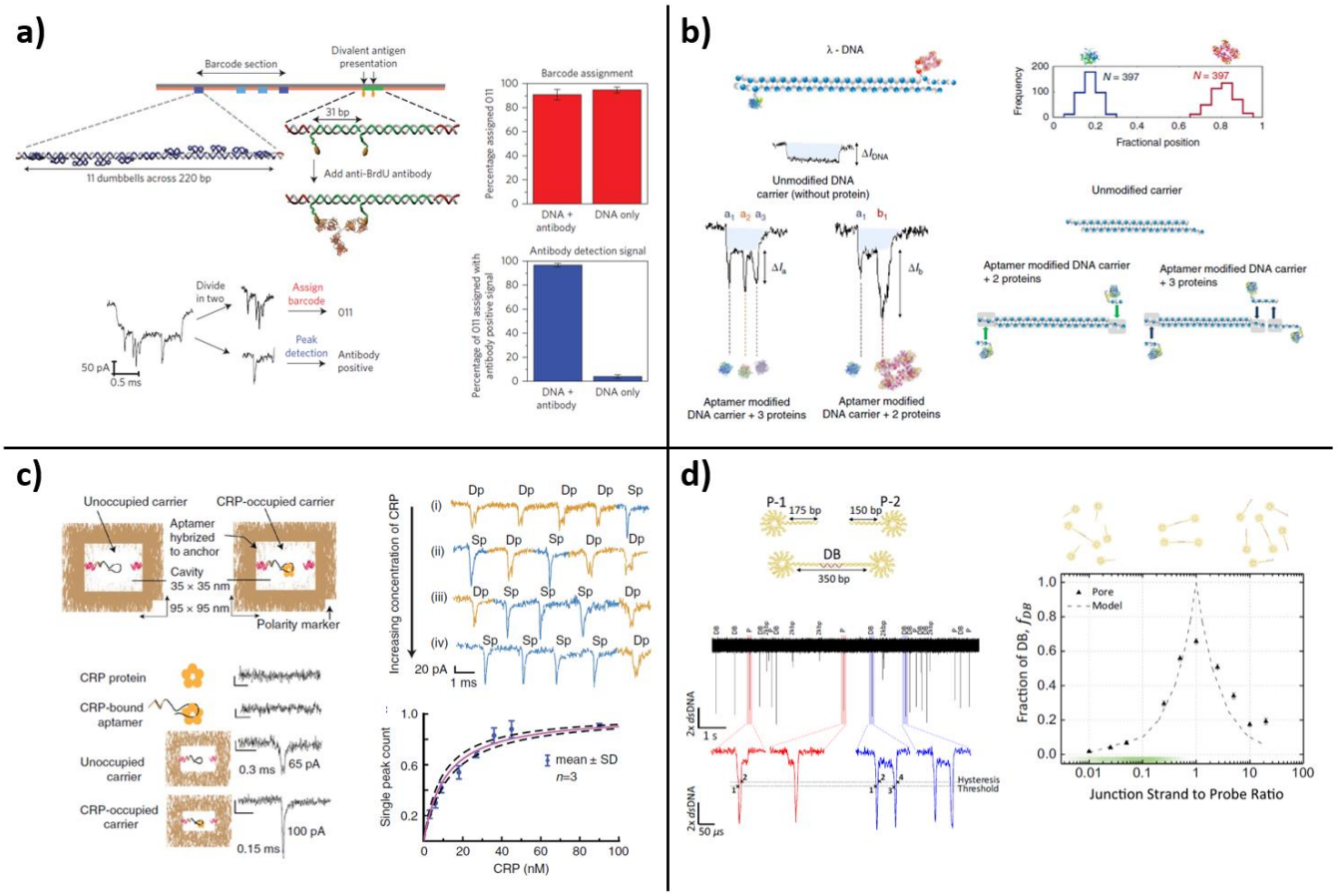


Figure 1.8: *Examples of DNA-based structured nanopore targets.* ***a)*** *"Barcoded" linear DNA carrier built from the single-stranded M13 bacteriophage genome and hundreds of complementary oligonucleotides. The barcode region contains protrusions ("dumbbells") that identify which protein is targeted at the separate binding region. Images adapted from* [44]*, with permission.* ***b)*** *Using the double-stranded λ bacteriophage genome, which features 12-nt ssDNA extensions on its ends, as a linear DNA carrier. The single-stranded overhangs are used to attach DNA aptamers for target proteins. Images adapted from* [45]*, with permission.* ***c)*** *Rectangular DNA-origami carrier constructed from an M13 scaffold, which features a DNA aptamer for a target protein at the centre of an open cavity. The binding of this protein changes the number of peaks in its nanopore translocation signal. Images adapted from* [46]*, with permission.* ***d)*** *DNA "shooting star" probes assembled from individual oligonucleotides into a 12-arm dsDNA multiway junction, where one of the arms is extended and features a terminal ssDNA overhang. A target oligonucleotide ("junction strand") binds to the overhangs of two separate probes and combines them into a single, nanopore-distinguishable construct. Images adapted from* [31]*, with permission.*

Bell et al., 2016 (Fig. 1.8a)

In this work [44], the authors produced a linear, 7-kbp dsDNA construct (referred to as a "carrier") by hybridizing hundreds of short (38 bp) complementary oligonucleotides to an ssDNA bacteriophage genome (M13), similar to the methods of "DNA origami" [47]. Importantly, individual oligos in the complement could also be replaced with modified variants to add certain functionalities at particular positions along the ssDNA scaffold. In this way, the dsDNA construct was essentially divided into two regions (Figure 1.8a). In the "binding site" region, two

adjacent oligos were modified to feature (poly-T) overhangs that terminated in antigens for a specific IgG antibody. When this antibody is present in solution, it can bind to the two antigens and thus be transported along with the carrier through a nanopore, producing an additional current blockage spike. Meanwhile, in the "barcode" region, oligos were replaced with a series of "dumbbell" hairpins (see upper left panel of Figure 1.8a) that could also be read as deeper blockages when passing through a pore. Specifically, the absence or presence of each of three dumbbell sections was used to encode a digital ("0" or "1"), 3-bit substructure in the translocation signals (the "barcode"). Linear dsDNA carriers of this design were used to demonstrate multiplexed antibody detection with a nanopore. To do so, several carrier versions were produced with antigens targeting different antibodies at their binding site – in each case, binding of the antibody produced an extra peak in the carrier signal, however, these peaks were not distinguishable among the different antibodies. Instead, the carrier for a particular antibody was also designed with a unique barcode at its second region. Detecting a specific antibody in their system is then possible by capturing an antibody-bound carrier with a pore (signalled through the presence of an extra peak from the occupied binding site) and decoding the signal of its barcode region to assign the antibody identity.

Sze et al., 2017 (Fig. 1.8b)

This system also makes use of linear carriers, this time using the dsDNA genome of bacteriophage λ, which features 12-nt ssDNA extensions at each of its ends. Here, the authors hybridized DNA aptamers that bind specific proteins to these ends by extending the apatmer sequences to include complements to the λ overhangs [45]. Similar to Bell et al. (Figure 1.8a), translocation of a protein-bound carrier through a nanopore resulted in the appearance of additional spikes in its current blockage signal when compared to that of just the free carrier (λ-DNA + aptamers). Carriers were constructed to bind as many as three proteins (thrombin) at a time by further extending the aptamer sequence at one end to include a region for an additional aptamer to hybridize to (separated by 96 bases). Despite the close spacing of these aptamers, the binding of two thrombin molecules within this region was reproducibly signalled by the presence of two distinguishable peaks at one end of the carrier blockages; the increase in resolution necessary for this was attributed to a slowdown in the translocation

velocity when bulky proteins occupy the confined space of the nanopore. Finally, the authors also demonstrated the simultaneous detection of two different proteins (thrombin and AChE) with a single carrier by functionalizing each carrier end with a different aptamer. The identities of the bound proteins, and thus the orientations of the carriers during translocation ("head-to-tail" or "tail-to-head"), were determined by the larger protein (AChE) producing current blockage sub-peaks with larger amplitudes and longer durations (see bottom left panel of Figure 1.8b).

Raveendran et al., 2020 (Fig. 1.8c)

Here, DNA-based carriers are again involved in transporting proteins to/across a nanopore but instead of linear carriers like the previous examples, DNA origami is used to produce a hollow, rectangular carrier [46]. In particular, the carrier designs were built from an M13 ssDNA scaffold and folded with staple strand oligos into a $\sim 95 \times 95$ nm^2 square with a central $\sim 35 \times 35$ nm^2 central cavity (see Figure 1.8c). Protruding into this cavity was an ssDNA "anchor stub" to which a DNA aptamer to bind human C-reactive protein (CRP) could be hybridized. According to the authors, the advantages of a symmetric, rectangular carrier system are that they avoid the issues of folding (having DNA hairpins in the pore act as false positives of carrier occupancy) and ambiguity in the direction of translocation (passing head first or tail first) that affect linear carriers. When the central aptamer binds CRP, the translocation signals of the carriers noticeably shift from containing two shallower sub-peaks ("unoccupied carrier") to containing a single deeper peak ("CRP-occupied carrier"), likely from the bound protein filling part of the cavity volume. Tracking the proportion of single-peaked events over a range of CRP concentrations resulted in a monotonic trend transitioning from mostly double-peaked (unoccupied) events at low CRP concentrations to a saturated response of mostly single-peaked (occupied) events at high concentrations (see bottom right panel of Figure 1.8c). By fitting a curve to this trend, the authors were able to extract an estimate for the dissociation constant of the CRP aptamer; this could also presumably be used as a calibration curve for an unknown CRP sample by measuring the fraction of single-peaked carrier events it produces and interpolating the estimated concentration value that this corresponds to.

He et al., 2021 (Fig. 1.8d)

Finally, we examine a slightly different system where, rather than sensing a target protein directly with a nanopore (either on its own or attached to a carrier), a proxy molecule is generated to signal the presence of the target. In this example, the structured reporter molecules are a pair of 12-arm DNA stars ("probes") made up of twelve dsDNA fragments – eleven short (25 bp) and one long (150/175 bp) – connected at a central junction, and assembled from a set of ~50-nt oligonucleotides [31]. The long arm of each probe also features an additional 25-nt ssDNA overhang (either in the 5′ or 3′ direction) so that the two star types can be linked together by hybridization with a complementary 50-nt oligo ("junction strand") that spans the two overhangs (see Figure 1.8d). Crucially, this junction strand is released as the output of a separate sandwich immunoassay step (similar to that depicted earlier in Figure 1.5b) for the target protein (TSH): TSH is bound by both a magnetic bead-linked antibody and junction strand-linked antibody, the complexes are collected by an external magnetic field, washed of background molecules, and finally the junction strands are released by UV photocleavage and collected. The result is (ideally) one released junction strand copy for every bound protein; these proxy molecules then translate into nanopore-readable signals by linking two star probes together. As in the previous example (Figure 1.8c), positive and negative probe signals are distinguished by the number of sub-peaks they contain, corresponding to the number of bulky star moieties present in the translocated analyte: two for junction-linked probes, one for free probes (see bottom left panel of Figure 1.8d). Also as before, a calibration curve relating the fraction of double-peak events to protein sample concentration can be constructed, enabling the concentration of an unknown TSH sample to be estimated by interpolation along this curve. Separating the full protein assay into two steps (sandwich assay and nanopore sensing) is useful because it allows for the conditions (e.g. salt concentration) of each step to optimized separately and for the dynamic range of the assay to be increased by inserting an amplification step (one protein ⇒ many junction strands) in between.

1.6 – Overview of Thesis

In the thesis chapters that follow, we present work undertaken to address some key issues with the basic solid-state nanopore system that were highlighted in the above sections. Specifically, this work represents efforts that fall under the major themes of: A) increasing the target specificity of the system (to extract additional information from specific analytes of interest), and B) improving its event detection rate (to enable the efficient analysis of lower-concentration samples).

First, in *Chapter 2*, we present an integrated instrument for collecting synchronous electrical and optical translocation signals though a nanopore. By adding a second detection channel in parallel, analytes that generate similar ionic current blockages but that differ in their optical signals can be better distinguished [48], increasing the overall target specificity of the sensor. The main focus of the chapter is to minimize the electrical noise contributions of the introduced optical equipment (e.g. microscope, illumination source) on the original ionic current channel – keeping signal-to-noise ratios high over a variety of experimental conditions (salt concentration, pore diameter, etc.) expands the range of potential single-molecule targets that can be analyzed with the system.

Next, *Chapter 3* describes attempts to combine solid-state nanopores with an active target delivery mechanism (in this case insulator-based dielectrophoresis or "iDEP"), in order to increase the capture rates of the pores beyond their diffusion-limited values. Here, SiN_x membranes are embedded within microfluidic cells featuring constricted channel geometries – these constrictions serve to focus the electric field lines through the system at the location of the nanopore and thus generate field gradients to accumulate particles by dielectrophoresis. We separately evaluate the ability of our system to both concentrate and sense double-stranded DNA, and discuss challenges that remain for pre-concentrating analytes within the device prior to single-molecule nanopore sensing.

In *Chapter 4*, we develop simplified protocols to synthesize "carrier" molecules of the sort described in Figure 1.8 for specific analyte detection on a nanopore. Our carriers are based on linear, double-stranded DNA

with single-stranded extensions on each end ("sticky ends") through which target-specific functionalities can be integrated. In particular, the synthesis methods presented incorporate the ability to vary the overall length of the carriers on demand – this has important implications for the multiplexing potential of the system since molecule length represents another variable by which carriers mapping to different targets can be distinguished on a pore [49].

Chapter 5 builds on *Chapter 4* by using molecules produced by those methods as a model system to describe the signal analysis of structured nanopore analytes. Since the target-specific information of a structured molecule is contained in its precise sequence of *signal* substructures, more refined analysis approaches are needed to efficiently sort translocation events by the molecular species that produced them, beyond the typical technique of grouping signals on whole-event parameters (e.g. max blockage, dwell time). Our analysis approach places particular emphasis on reducing both the number of translocation events rejected as false negatives and the amount of selection bias present in the dataset from doing so. The issue of event rejection also impacts our major theme of increasing nanopore detection rates, since this is set not just by the capture rate of the pore but also by what fraction of captured molecules produce signals that survive event filtration during analysis.

Finally, *Chapter 6* provides a brief summary of the entire thesis, pointing to future directions in which to expand the current work, and presents preliminary results toward achieving a few of these goals.

1.7 – References

1. Xue L, Yamazaki H, Ren R, Wanunu M, Ivanov AP, Edel JB. Solid-state nanopore sensors. Nat Rev Mater. 2020;5: 931–951. doi:10.1038/s41578-020-0229-6
2. Ying YL, Hu ZL, Zhang S, Qing Y, Fragasso A, Maglia G, et al. Nanopore-based technologies beyond DNA sequencing. Nat Nanotechnol. 2022;17: 1136–1146. doi:10.1038/s41565-022-01193-2
3. Tabard-Cossa V, Carlsen A, Charron M, Karau P, Roelen Z, Waugh M. Nanopores: Electronic Tools for Single-molecule Biophysics and Bio-nanotechnologies. Physics in Canada. 2017;73: 97–101.
4. Storm AJ, Storm C, Chen J, Zandbergen H, Joanny J-F, Dekker C. Fast DNA Translocation through a Solid-State Nanopore. Nano Lett. 2005;5: 1193–1197. doi:10.1021/nl048030d

5. Roelen Z, Briggs K, Tabard-Cossa V. Analysis of Nanopore Data: Classification Strategies for an Unbiased Curation of Single-Molecule Events from DNA Nanostructures. ACS Sens. 2023;8: 2809–2823. doi:10.1021/acssensors.3c00751

6. Miles BN, Ivanov AP, Wilson KA, Doğan F, Japrung D, Edel JB. Single molecule sensing with solid-state nanopores: novel materials, methods, and applications. Chem Soc Rev. 2013;42: 15–28. doi:10.1039/C2CS35286A

7. Vodyanoy I, Bezrukov SM. Sizing of an ion pore by access resistance measurements. Biophys J. 1992;62: 10–11. doi:10.1016/S0006-3495(92)81762-9

8. Wanunu M, Morrison W, Rabin Y, Grosberg AY, Meller A. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. Nat Nanotechnol. 2010;5: 160–165. doi:10.1038/nnano.2009.379

9. Bayley H. Nanopore Sequencing: From Imagination to Reality. Clin Chem. 2015;61: 25–31. doi:10.1373/clinchem.2014.223016

10. Zaitseva E, Obergrussberger A, Weichbrodt C, Boukhet M, Bernhard F, Hein C, et al. Electrophysiology on Channel-Forming Proteins in Artificial Lipid Bilayers: Next-Generation Instrumentation for Multiple Recordings in Parallel. In: Dallas M, Bell D, editors. Patch Clamp Electrophysiology. Springer; 2021. pp. 67–92. doi:10.1007/978-1-0716-0818-0_4

11. Li J, Gershow M, Stein D, Brandin E, Golovchenko JA. DNA molecules and configurations in a solid-state nanopore microscope. Nat Mater. 2003;2: 611–615. doi:10.1038/nmat965

12. Larkin J, Henley R, Bell DC, Cohen-Karni T, Rosenstein JK, Wanunu M. Slow DNA Transport through Nanopores in Hafnium Oxide Membranes. ACS Nano. 2013;7: 10121–10128. doi:10.1021/nn404326f

13. Liu K, Feng J, Kis A, Radenovic A. Atomically Thin Molybdenum Disulfide Nanopores with High Sensitivity for DNA Translocation. ACS Nano. 2014;8: 2504–2511. doi:10.1021/nn406102h

14. Schneider GF, Kowalczyk SW, Calado VE, Pandraud G, Zandbergen HW, Vandersypen LMK, et al. DNA Translocation through Graphene Nanopores. Nano Lett. 2010;10: 3163–3167. doi:10.1021/nl102069z

15. Smeets RMM, Keyser UF, Krapf D, Wu M-Y, Dekker NH, Dekker C. Salt Dependence of Ion Transport and DNA Translocation through Solid-State Nanopores. Nano Lett. 2006;6: 89–95. doi:10.1021/nl052107w

16. Hall JE. Access resistance of a small circular pore. J Gen Physiol. 1975;66: 531–532. doi:10.1085/jgp.66.4.531

17. Kowalczyk SW, Grosberg AY, Rabin Y, Dekker C. Modeling the conductance and DNA blockade of solid-state nanopores. Nanotechnology. 2011;22: 315101. doi:10.1088/0957-4484/22/31/315101

18. Davenport M, Healy K, Pevarnik M, Teslich N, Cabrini S, Morrison AP, et al. The Role of Pore Geometry in Single Nanoparticle Detection. ACS Nano. 2012;6: 8366–8380. doi:10.1021/nn303126n

19. Grosberg AY, Rabin Y. DNA capture into a nanopore: Interplay of diffusion and electrohydrodynamics. J Chem Phys. 2010;133: 165102. doi:10.1063/1.3495481

20. Charron M, Philipp L, He L, Tabard-Cossa V. Elucidating the dynamics of polymer transport through nanopores using asymmetric salt concentrations. Nano Res. 2022;15: 9943–9953. doi:10.1007/s12274-022-4886-3

21. Charron M, Briggs K, King S, Waugh M, Tabard-Cossa V. Precise DNA Concentration Measurements with Nanopores by Controlled Counting. Anal Chem. 2019;91: 12228–12237. doi:10.1021/acs.analchem.9b01900

22. Kwok H, Briggs K, Tabard-Cossa V. Nanopore Fabrication by Controlled Dielectric Breakdown. Hall A, editor. PLoS One. 2014;9: e92880. doi:10.1371/journal.pone.0092880

23. Briggs K, Waugh M, Tabard-Cossa V. Fabricating Solid-State Nanopores for Single-Molecule Sensing. In: Leburton J-P, editor. Solid State Nanopores. Cham: Springer International Publishing; 2023. pp. 1–33. doi:10.1007/978-3-031-42336-9

24. Waugh M, Briggs K, Gunn D, Gibeault M, King S, Ingram Q, et al. Solid-state nanopore fabrication by automated controlled breakdown. Nat Protoc. 2020;15: 122–143. doi:10.1038/s41596-019-0255-2

25. Kowalczyk SW, Wells DB, Aksimentiev A, Dekker C. Slowing down DNA Translocation through a Nanopore in Lithium Chloride. Nano Lett. 2012;12: 1038–1044. doi:10.1021/nl204273h

26. 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. 2012;23: 405301. doi:10.1088/0957-4484/23/40/405301

27. Tabard-Cossa V, Trivedi D, Wiggin M, Jetha NN, Marziali A. Noise analysis and reduction in solid-state nanopores. Nanotechnology. 2007;18: 305505. doi:10.1088/0957-4484/18/30/305505

28. Freedman KJ, Otto LM, Ivanov AP, Barik A, Oh S-H, Edel JB. Nanopore sensing at ultra-low concentrations using single-molecule dielectrophoretic trapping. Nat Commun. 2016;7: 10217. doi:10.1038/ncomms10217

29. Spitzberg JD, van Kooten XF, Bercovici M, Meller A. Microfluidic device for coupling isotachophoretic sample focusing with nanopore single-molecule sensing. Nanoscale. 2020;12: 17805–17811. doi:10.1039/D0NR05000H

30. Carson S, Wilson J, Aksimentiev A, Wanunu M. Smooth DNA Transport through a Narrowed Pore Geometry. Biophys J. 2014;107: 2381–2393. doi:10.1016/j.bpj.2014.10.017

31. He L, Tessier DR, Briggs K, Tsangaris M, Charron M, McConnell EM, et al. Digital immunoassay for biomarker concentration quantification using solid-state nanopores. Nat Commun. 2021;12: 5348. doi:10.1038/s41467-021-25566-8

32. Yusko EC, Johnson JM, Majd S, Prangkio P, Rollings RC, Li J, et al. Controlling protein translocation through nanopores with bio-inspired fluid walls. Nat Nanotechnol. 2011;6: 253–260. doi:10.1038/nnano.2011.12

33. Nicoli F, Verschueren D, Klein M, Dekker C, Jonsson MP. DNA Translocations through Solid-State Plasmonic Nanopores. Nano Lett. 2014;14: 6917–6925. doi:10.1021/nl503034j

34. Ivanov AP, Actis P, Jönsson P, Klenerman D, Korchev Y, Edel JB. On-Demand Delivery of Single DNA Molecules Using Nanopipets. ACS Nano. 2015;9: 3587–3595. doi:10.1021/acsnano.5b00911

35. Steinbock LJ, Otto O, Chimerel C, Gornall J, Keyser UF. Detecting DNA Folding with Nanocapillaries. Nano Lett. 2010;10: 2493–2497. doi:10.1021/nl100997s

36. Koo B, Yorita AM, Schmidt JJ, Monbouquette HG. Amplification-free, sequence-specific 16S rRNA detection at 1 aM. Lab Chip. 2018;18: 2291–2299. doi:10.1039/C8LC00452H

37. Chuah K, Wu Y, Vivekchand SRC, Gaus K, Reece PJ, Micolich AP, et al. Nanopore blockade sensors for ultrasensitive detection of proteins in complex biological samples. Nat Commun. 2019;10: 2109. doi:10.1038/s41467-019-10147-7

38. Kirby BJ. Micro- and Nanoscale Fluid Mechanics: Transport in Microfluidic Devices. Cambridge University Press; 2010.

39. Nkodo AE, Garnier JM, Tinland B, Ren H, Desruisseaux C, McCormick LC, et al. Diffusion coefficient of DNA molecules during free solution electrophoresis. Electrophoresis. 2001;22: 2424–2432. doi:10.1002/1522-2683(200107)22:12<2424::AID-ELPS2424>3.0.CO;2-1

40. Bakewell D, Vergara-Irigaray N, Holmes D. Dielectrophoresis of Biomolecules. JSM Nanotechnology & Nanomedicine. 2013;1: 1–14.

41. Kowalczyk SW, Tuijtel MW, Donkers SP, Dekker C. Unraveling Single-Stranded DNA in a Solid-State Nanopore. Nano Lett. 2010;10: 1414–1420. doi:10.1021/nl100271c

42. Kong J, Zhu J, Keyser UF. Single molecule based SNP detection using designed DNA carriers and solid-state nanopores. Chemical Communications. 2017;53: 436–439. doi:10.1039/C6CC08621G

43. Plesa C, Kowalczyk SW, Zinsmeester R, Grosberg AY, Rabin Y, Dekker C. Fast Translocation of Proteins through Solid State Nanopores. Nano Lett. 2013;13: 658–663. doi:10.1021/nl3042678

44. Bell NAW, Keyser UF. Digitally encoded DNA nanostructures for multiplexed, single-molecule protein sensing with nanopores. Nat Nanotechnol. 2016;11: 645–651. doi:10.1038/nnano.2016.50

45. Sze JYY, Ivanov AP, Cass AEG, Edel JB. Single molecule multiplexed nanopore protein screening in human serum using aptamer modified DNA carriers. Nat Commun. 2017;8: 1552. doi:10.1038/s41467-017-01584-3

46. Raveendran M, Lee AJ, Sharma R, Wälti C, Actis P. Rational design of DNA nanostructures for single molecule biosensing. Nat Commun. 2020;11: 4384. doi:10.1038/s41467-020-18132-1

47. Rothemund PWK. Folding DNA to create nanoscale shapes and patterns. Nature. 2006;440: 297–302. doi:10.1038/nature04586

48. Cai S, Sze JYY, Ivanov AP, Edel JB. Small molecule electro-optical binding assay using nanopores. Nat Commun. 2019;10: 1797. doi:10.1038/s41467-019-09476-4

49. Bell NAW, Muthukumar M, Keyser UF. Translocation frequency of double-stranded DNA through a solid-state nanopore. Phys Rev E. 2016;93: 022401. doi:10.1103/PhysRevE.93.022401

CHAPTER 2 – INSTRUMENTATION FOR LOW NOISE IONIC CURRENT RECORDING UNDER LASER ILLUMINATION

This chapter is published in *Review of Scientific Instruments* as:

"Instrumentation for Low Noise Nanopore-Based Ionic Current Recording under Laser Illumination" (2018)

by Zachary Roelen, José A. Bustamante, Autumn Carlsen, Aidan Baker-Murray, and Vincent Tabard-Cossa

2.1 – Abstract

We describe a nanopore-based optofluidic instrument capable of performing low-noise ionic current recordings of individual biomolecules under laser illumination. In such systems, simultaneous optical measurements generally introduce significant parasitic noise in the electrical signal, which can severely reduce the instrument sensitivity, critically hindering the monitoring of single-molecule events in the ionic current traces. Here, we present design rules and describe simple adjustments to the experimental setup to mitigate the different noise sources encountered when integrating optical components to an electrical nanopore system. In particular, we address the contributions to the electrical noise spectra from illuminating the nanopore during ionic current recording and mitigate those effects through control of the illumination source and the use of a PDMS layer on the SiN_x membrane. We demonstrate the effectiveness of our noise minimization strategies by showing the detection of DNA translocation events during membrane illumination with a signal-to-noise ratio (SNR) of ~10 at 10 kHz bandwidth. The instrumental guidelines for noise minimization that we report are applicable to a wide range of nanopore-based optofluidic systems and offer the possibility of enhancing the quality of synchronous optical and electrical signals obtained during single-molecule nanopore-based analysis.

2.2 – Introduction

Nanopore sensors have emerged as a promising technology for single-molecule analysis. Traditionally, electrical measurements are taken by monitoring pA to nA current levels as ions in a voltage-biased electrolyte solution pass through a nanopore, defined as a nanometer-sized aperture in a thin, insulating membrane. When a molecule of interest enters the confined geometry of the nanopore, information about molecular size, shape, and charge may be obtained.[1–3] The potential for nanopore sensors in sequencing applications has been widely recognized.[1,4,5] However, a number of fundamental features of the nanopore translocation process cannot be assessed from traditional ionic current measurements alone. These include: i) the trajectory, conformation, and behavior of the DNA strand as it enters the capture radius of the pore and attempts translocation[6], ii) the mechanism behind nanopore clogging events[7], and iii) the true source of anomalous events, in which current blockages cannot be attributed to simple DNA translocation based on available conductance models.[8,9] Since these three factors influence event rate and duration,[10,11] device throughput and longevity, and ease of data analysis, respectively, enhancing our understanding is essential to the successful implementation of many applications and more generally the viability of nanopore systems as a disruptive omic technology.

The addition of an optical channel to the conventional nanopore setup offers a means to visually assess DNA molecules in the vicinity of the nanopore, complementing the electrical (ionic-current) measuring capabilities of the nanopore itself. Fluorescence measurements of individual DNA molecules could provide some of the information required to address the phenomena described above, namely the behavior of DNA molecules during capture, clogging, and anomalous events. Other potential benefits of the optical system include the implementation of photon blockage schemes for parallel measurements through an array of nanopores,[12,13] the facilitation of plasmonic effects for improved local fluorescence[14,15], local heating[14,16–18] and trapping,[19] or the use of laser illumination to slow down translocation kinetics.[20]

Synchronous optical and electrical detection of biomolecules traversing through a solid-state nanopore was originally achieved by employing Total Internal Reflection Fluorescence to reduce optical background and obtain

sufficient optical signal-to-noise ratio (SNR).[21] However, the use of the fluorophore Alexa 647 to covalently label DNA[21,22], which is compatible with the levels of electrolyte typically employed in nanopore experiments (1 M KCl), did not provide sufficient spatial resolution to enable observation of the conformation of individual DNA molecules. Subsequent studies[7,23] employed DNA-intercalating dyes, such as YOYO-1, which reveal DNA conformation but require lower salt conditions (0.1 M) to remain bound to the DNA. Such low salt concentrations reduce the electrical signal-to-noise ratio, meaning less sensitive detection of DNA by the nanopore. In addition, several groups have employed zero-mode waveguides (ZMWs) – subwavelength apertures in thin metal films – to confine the effective laser illumination volume to within the close vicinity of a nanopore, thus greatly suppressing background fluorescence and cross-talk between molecules[24,25]. In this way, multiple fluorescently-labeled DNA molecules have been simultaneously detected by optical means through arrays of ZMW-nanopores, driven either by pressure gradients[24] or electrophoresis[25].

A recently introduced alternative strategy uses Ca^{2+} ions on one side of the pore and Ca^{2+}-activated fluorophores on the other to detect the passage of unlabeled DNA both optically and electrically.[12,13] Since higher salt concentrations (0.4 to 1 M KCl) are permitted, the electrical signal-to-noise ratio is greatly improved over the low-salt measurements of intercalating dyes. However, it is not possible to observe the conformation of the label-free DNA used in this approach, as it is the distribution of Ca^{2+} ions near the pore and not the DNA itself that is being fluorescently probed.

As a first step to carrying out similar measurements to those described above, either performing synchronous optical and electrical detection, or electrical recording under illumination, researchers must design an instrument that maintains a low enough signal-to-noise ratio in the electrical channel (described in detail below) to resolve events while the optical channel is active. To assist researchers in that task, we have identified instrumental noise sources related to the integration of a standard nanopore setup with a wide-field fluorescence inverted microscope equipped with different types of illumination. Here, we report valuable

guidelines for minimizing these noise sources in the ionic current recordings with the objective of promoting advancement in the field of synchronous nanopore-based optical and electrical measurement.

2.3 – Description of the Optofluidic Nanopore-Based Instrument

This section describes an optofluidic nanopore-based instrument (Figure 2.1) following optimization for low-noise ionic current recording under illumination. Essentially, light is directed through an inverted microscope to a nanopore in a silicon chip encased in a PDMS flow cell filled with electrolyte solution. A voltage bias applied to the electrolyte solution filling the flow cell pulls negatively charged DNA toward the pore. Upon entering the illuminated area surrounding the pore, DNA can produce an optical signal due to fluorescence that is recorded by a camera. In parallel, passage of the DNA through the nanopore and its access regions causes a temporary drop in the ionic current, as measured by a current amplifier. We describe each element of the instrument in greater detail below.

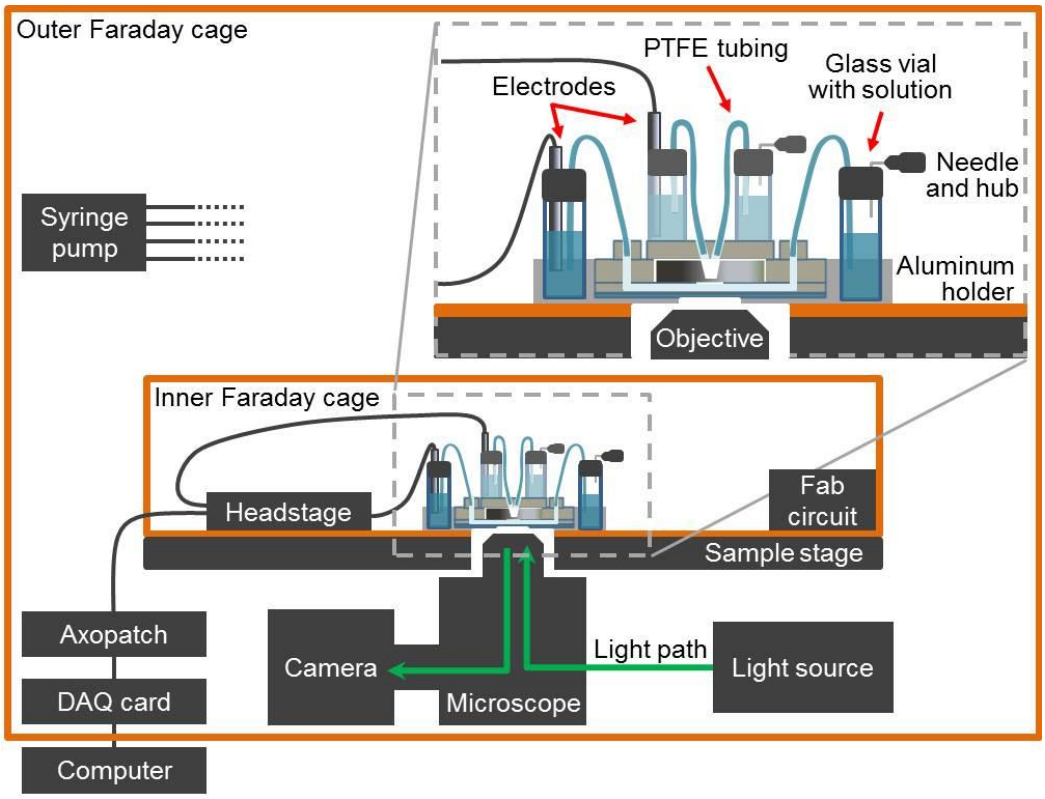


Figure 2.1: *Schematic diagram of optofluidic nanopore-based instrument. Light is directed from the light source through the microscope to the nanopore, where it illuminates fluorophores. The emission is filtered by a specific filter cube inside the microscope and collected by a camera. Using a computer-controlled DAQ card, an electric potential difference is applied to electrodes submerged in electrolyte solution inside 1.5-mL glass vials. The electrodes are electrically linked to the nanopore through the electrolyte solution flowing from the glass vial through PTFE tubing to the microfluidic device. The flow of electrolyte solution from the glass vial through the microfluidic device can be controlled by connecting a syringe actuated by a syringe pump to the hub of the needle protruding from the vial. Ionic current is detected using an Axopatch 200B connected to the computer-controlled DAQ card. Two Faraday cages are used to shield the nanopore from electromagnetic interferences. Inset shows detail of microfluidic device with electrodes, PTFE tubing and 1.5-mL glass vials filled with electrolyte solution, all embedded in an aluminum holder secured on the sample stage over the microscope objective.*

2.3.1 – Fluidic Cell and Circuitry

At the heart of the optofluidic nanopore-based system is a 200 µm thick, 3 mm wide Si chip with a 50×50 μm^2 free-standing SiN_x membrane, 10 or 30 nm thick in its center. A schematic of the fluidic device used to mount the chip in liquid is shown in Figure 2.2. The chip (Norcada, Inc. product # NT005X or NT005Z) is encapsulated in a custom-made fluidic device composed of three PDMS layers (Sylgard 184, Dow Corning) and a glass cover slip (22-mm, #1). The 4-part fluidic device (Figure 2.2) uses PDMS channels, either defined by soft

lithography or hand punched, to provide independent fluidic access to both sides of the SiN_x membrane and to allow positioning of the nanopore within the working distance of the microscope objective.

The bottom PDMS layer, which is bonded to a 22×22 mm^2 glass coverslip, is formed by spin coating (600 rpm for 90 s with a ramp speed of 133 rpm/s)[26] a thin, 100 µm layer of PDMS (10:1 base to curing agent) onto a Si wafer patterned with 50 µm high, 400 µm wide SU-8 features defined by UV photolithography to produce fluidic access channels to the bottom of the Si chip (see Figure 2.2a). Up to 7 devices can fit onto a single 100-mm wafer. After soft-baking the coated wafer at 80°C for two hours then treating it with a 60-W air plasma for 60 seconds, a similarly plasma-treated, 1 mm thick PDMS layer with three pillars (each punched with a 0.75 mm hole) and a 3-mm circular cut-out (to accommodate the Si support chip) is lowered manually onto the thin wafer-bound PDMS layer using tweezers, with careful alignment of the pillars and channels (Figure 2.2a). To prepare the thick layer, PDMS (mixed at 5:1 base to curing agent) is poured into aluminum moulds (containing three 2 mm deep, 4 mm diameter wells to form the pillars) which are then also cured at 80°C for two hours. The punched pillars will allow fluidic access to the channels in the thin PDMS layer, and the 3-mm cut-out for the Si support chip lies at the intersection of these channels. After the binding of the thick PDMS layer (with the pillars) to the thin wafer-bound PDMS layer (with the channels), a scalpel is used to separate the two from the wafer. Through the 3-mm cut-out, a 1-mm hole is punched at the intersection of the channels in the thin PDMS layer to later provide fluidic access to the nanopore chip. Reusable stainless steel biopsy punches of the appropriate diameter (i.e. 0.75 mm, 1 mm, 3 mm) are used to cut all holes. A plasma-treated glass coverslip is then bonded to the flat underside of the patterned PDMS layer to provide stability for further handling and imaging. Finally, the nanopore chip is plasma-bonded into the plasma-treated, 3-mm cut-out then sealed into place with a thick, plasma-treated PDMS cap bearing two closely spaced 0.75-mm channels punched in its center. A cross-sectional view (Figure 2.2b) shows independent fluidic access to the etch pit side of the membrane through the two channels punched in the cap on top of the device, while fluidic access to the flat side of the membrane chip is obtained through the three pillars (only two are visible on the schematic). When the microfluidic device is fully assembled, the SiN_x membrane containing the nanopore is then positioned 100 µm above the top glass coverslip

interface (i.e. the thickness of the thin PDMS layer), which is well within the working distance of the 60× water immersion or 100× oil immersion microscope objectives used for single-molecule measurements, of ca. 170 µm and 130 µm, respectively.

As seen in Figure 2.1, one end of a short section of thin-wall PTFE tubing (#30, Cole-Parmer) is inserted into each of the three pillars of the middle PDMS layer and the two channels of the top PDMS cap while the opposite end of the tubing is inserted through a septum into the liquid-filled portion of 1.5-mL glass vials containing ~1 mL of electrolyte solution. Ag/AgCl electrodes, consisting of a ~5 cm length of 1-mm dia. Ag wire (Alfa Aesar, 99.999%), treated with bleach using protocols described previously[27], are also inserted through the septa and lowered into contact with the electrolyte solution in one vial connected to the top side and in another connected to the bottom side of the membrane. Note that the redundant vials, which do not contain electrodes during current recordings, offer a means to confirm proper fluidic and electrical connection throughout the device, enable improved mixing, and offer access to an alternative channel in the thin, bottom PDMS layer if one of the others is blocked. Syringe needles are then inserted through the septa into the air-filled portion of each vial, with the needle hub remaining available for attachment to a manual syringe or a computer-controlled syringe pump located next to the microscope frame. Additional air pressure applied via the syringe needle to the air-filled portion of the vial will force liquid into the submerged PTFE tubing leading to the fluidic device. In our instrument, a manual 1-mL syringe is typically used for initial device wetting. During experiments, a syringe pump (Harvard Apparatus 70-3009) connected to the needle hubs by air-filled PTFE tubing can be employed to precisely control the flow (1 to 5 µL/s) of solution through the fluidic device.

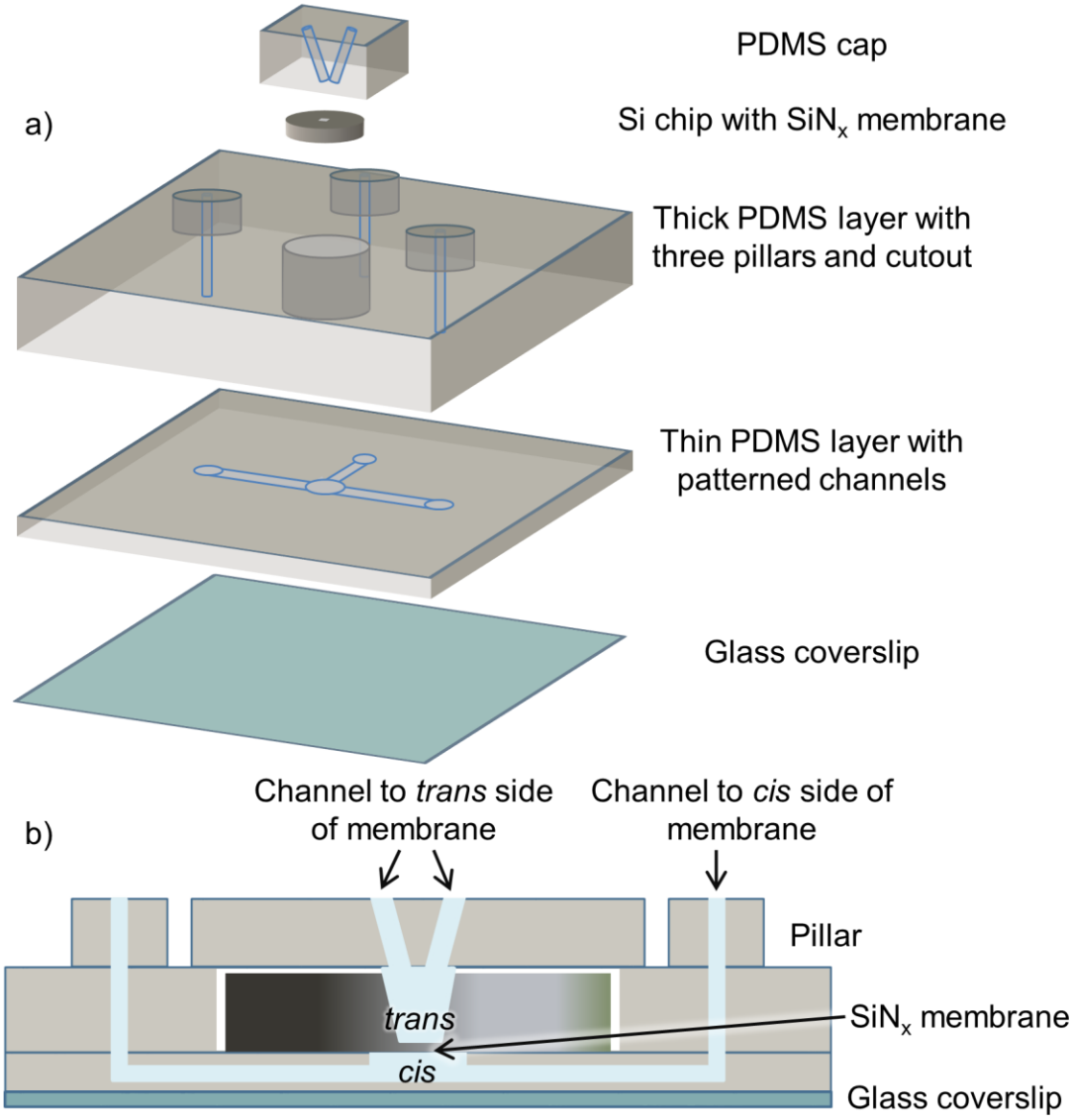


Figure 2.2: *Schematic diagrams of fluidic device.* ***a)*** *Overview of the assembly process. A glass coverslip offers stability to a thin PDMS layer with patterned channels which are aligned with the three pillars in a thick PDMS layer. A Si chip with freestanding SiN_x membrane is lowered into a cutout punched in the thick PDMS layer then sealed in place with a PDMS cap bearing two channels.* ***b)*** *Cross-sectional view of fluidic device. Channels punched in the pillars offer fluidic access to the bottom side of the SiN_x membrane, while the two channels in the PDMS cap offer fluidic access to the top side of the membrane.*

As detailed below, an additional ~10 µm thin layer of PDMS was applied to the surface of some of the Si chips before placement into the fluidic cells, in an effort to reduce electrical noise during ionic current recording[28]. This process begins by mixing PDMS at a ratio of 10:1 and then placing the mixture under vacuum to

remove trapped air bubbles. Under a stereomicroscope, the SiN_x membrane sides of the chips are then hand painted with PDMS in a two-stage process using a custom single-bristle brush, made by fixing with epoxy a bristle in a pipette tip[29]. First, two opposing sections of the chip that run just adjacent to (or slightly on top of) the free-standing membrane are painted, as illustrated in Figure 2.3a. The chips are then transferred to a glass slide on top of a hot plate set to 80 °C for 5 min, in order to cure these painted sections. The process of painting followed by curing is then repeated on the other two sides, resulting in a chip coated by PDMS except for a $\sim$100×100 μm^2 central region surrounding the free-standing SiN_x membrane (Figure 2.3b). By painting the chips in two stages, it reduces the risk of a continuous PDMS ring surrounding the central unpainted region from closing in (due to surface tension effects of the viscous liquid), covering the membrane before it is completely set and rendering the device unusable. An additional step of partially curing the PDMS drop used for painting with a few minutes on the hot plate can also be employed to increase the paint's viscosity and thus lower the distance it is able to spread before setting.

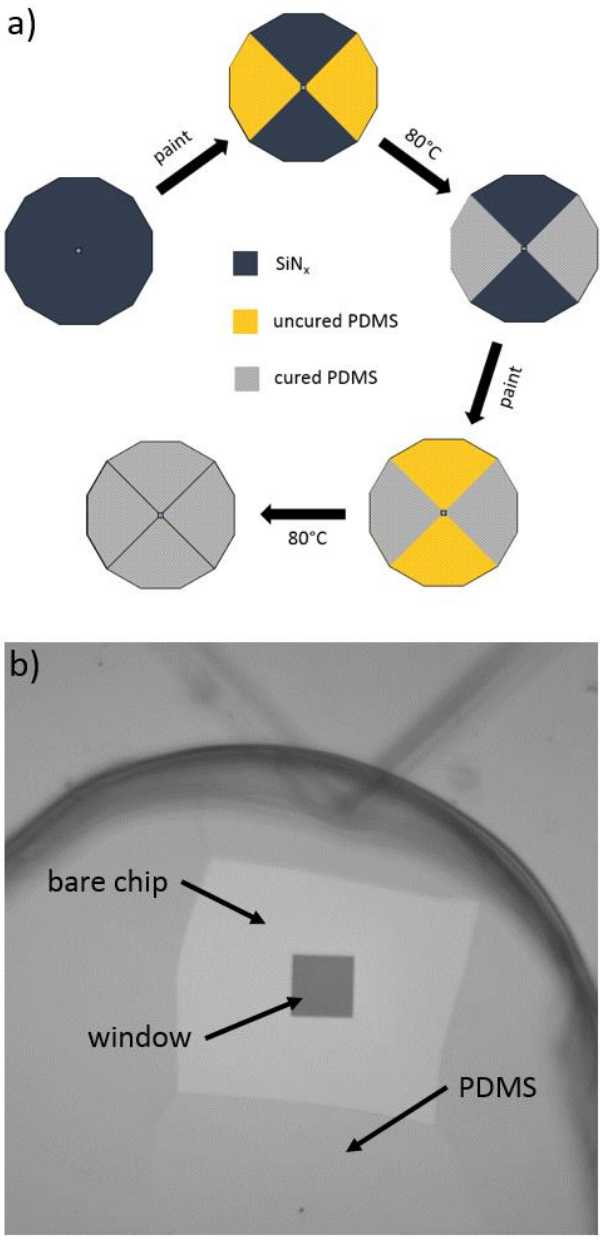


Figure 2.3: *PDMS painting.* ***a)*** *Schematic of the steps employed in painting the top surface of a Si chip with PDMS to reduce electrical noise. Only two of the four sides surrounding the central SiN_x membrane are painted at a time to reduce the risk of the uncured PDMS running over the membrane. A 5 min curing step using a hotplate set to 80°C follows the each of two painting stages.* ***b)*** *Optical image (10× objective) of a painted Si chip (50×50 μm^2, 30 nm thick SiN_x membrane) mounted into a fluidic cell, displaying the extent of PDMS coverage. Painting close to the membrane and thus isolating as much of the top surface from the surrounding electrolyte as possible is preferred. The patterned channels of the fluidic cell are visible in the background.*

2.3.2 – Electrical Assembly

To minimize external electrostatic and electromagnetic noise, a nanopore system is typically enclosed in a grounded metal box or Faraday cage. In our setup, a complete Faraday cage made of metal mesh encloses all components except the computer (Figure 2.1). Within this outer cage is a smaller, copper Faraday cage, which is complete except for a 13-mm circular opening in the bottom for optical access by the microscope objective. Although we have not observed this opening to contribute significant noise in the ionic current RMS below 10 kHz (in comparison with a typical nanopore setup comprising a pair of complete, nested Faraday cages[30,31]), it should be possible in principle to complete the electrical shield by using an ITO-coated glass cover slide, to eliminate any line-frequency pick-up noise (hum) and its harmonics if necessary for higher bandwidth measurements. The PDMS microfluidic device is secured into an aluminum holder with a matching circular opening and slots to hold electrolyte-filled vials in close proximity. When the aluminum holder is fixed into place in the inner Faraday cage, the fluidic device is aligned with the openings, and the membrane can easily be centered directly over the microscope objective. The inner Faraday cage also houses the low-noise current amplifier headstage (Molecular Devices, Axopatch 200B) and a custom-made circuit to perform nanopore fabrication by controlled breakdown (CBD). With this circuit, following successful wetting of a SiN_x membrane within the PDMS fluidic device, a nanopore is fabricated *in situ* using the previously described CBD approach.[30,32,31,33] Briefly, CBD consists of applying a strong electric field (> 0.5 V/nm) across the membrane and monitoring the resulting leakage current. A significant current spike signals the onset of ionic current and thus the formation of a nanopore. The Ag/AgCl electrodes used during CBD fabrication are then disconnected from the custom circuit and connected to the Axopatch 200B headstage for higher bandwidth ionic current measurements. The signal from the Axopatch 200B is then digitized and recorded by a computer controlled data acquisition card (National Instruments, USB-6351 DAQ) with the help of a custom LabView program.

2.3.3 – Optical Assembly

The optical detection system is composed of an inverted microscope (Nikon Ti-E) equipped with a 60× objective (Nikon 60× WI 1.27NA), a motorized XY Stage with 0.01 µm step size (Prior H117P2NN), and an emCCD camera (Andor iXon Ultra 897). Two different light sources were employed to investigate the noise on the ionic current signal under different illumination conditions: a mercury vapor arc lamp (X-Cite 120 PC, Excelitas) and a 488-nm laser diode (50 mW, Coherent OBIS). The former is a broad-spectrum lamp designed to fit into one of the microscope ports. For the latter, a custom optical system was designed to integrate the laser with the commercial microscope setup. To illuminate the full field of view with an optimally collimated beam, the laser beam is expanded from a diameter of 0.7 mm to 12 mm using two lenses: a biconcave lens (focal distance of 6 mm, N-SF11, Thorlabs) to make a divergent beam, and a biconvex lens (focal distance of 100 mm, N-BK7, Thorlabs) to render the rays parallel again. The two lenses are mounted in an optical cage, separated by a distance of 94 mm.

Illumination intensity can be adjusted with an attenuator as in Di Fiori et al.[20] Briefly, a Glan-Thompson polarizer (extinction ratio 1:10000, GTH10M, Thorlabs) is used to produce a highly linearly polarized beam. The polarization of the light is then rotated with a half-wave plate (WPH10M-488, Thorlabs) and passed through a second polarizing beam splitter cube (CM05-PBS201, Thorlabs). The resulting attenuation is described by Malus' Law, $I = I_0 \cos^2\theta$, where I is the final intensity, I_0 is the initial intensity and θ is the angle between the polarization axis of the light exiting the half-wave plate and that of the second polarizer.

The 488-nm laser diode illumination source can be used to excite YOYO-1 intercalating dyes, which then emit at 509 nm when complexed with DNA, or Ca^{2+} sensitive fluorescent dyes (such as Fluo-4) as they bind to calcium ions driven through a nanopore. In the latter case, the passage of a biomolecule through the pore blocks a corresponding volume of calcium ions and produces a characteristic drop in optical intensity.[12,13] In either case, the emission is filtered by a FITC filter cube inside the microscope and collected by the camera. Subsequent analysis of the images is performed in software (Nikon, NIS-Elements).

2.4 – Noise Analysis and Minimization

In order to identify potential noise sources, we analyze the electrical characteristics of the ionic current signal under different experimental configurations and illumination conditions. Based on our findings, we describe useful methods to minimize the superfluous electrical noise typically occurring in optofluidic nanopore-based instruments with the objective of advancing the development of similar instruments designed to combine high-resolution fluorescence microscopy and low-noise, high-bandwidth electrical detection through nanopores and nanochannels.

To evaluate the impact of electrical noise on nanopore-based analysis, consider the ionic conductance signal from a cylindrical nanopore and the blockage produced by the passage of a single biomolecule. The conductance of a nanopore can be described by the equation[8]

$$\mathrm{G_{pore}} = \frac{I_{open}}{V_{applied}} = \sigma\left(\frac{4L_{mem}}{\pi d_{pore}^2} + \frac{1}{d_{pore}}\right)^{-1} \tag{2.1}$$

where I_{open} is the open nanopore ionic current, $V_{applied}$ is applied voltage across the membrane, L_{mem} is the thickness of the membrane, d_{pore} is the diameter of the pore, and σ is the conductivity of the bulk electrolyte solution. This simplified equation does not take into account the effects of surface charge, which can be neglected in high salt conditions (concentrations > 1 M). Based on extensive analysis of experimental data, we assume the pore length for CBD-fabricated nanopores to be roughly equal to the nominal membrane thickness when using thin membranes (≤ 10 nm).[31] CBD-fabricated pores in thicker membranes will deviate from a pure cylindrical geometry and have effective pore length below the nominal membrane thickness.[34]

The conductance blockage caused by passage of a DNA molecule through the pore can be estimated by considering the bulk conductance blockage caused by the area exclusion of the DNA, and is to first order given by:[8]

$$\Delta\mathrm{G} = \frac{\Delta I_{block}}{V_{applied}} = \sigma\frac{\pi d_{DNA}^2}{4\,L_{mem}} \tag{2.2}$$

where ΔI_{block} is the ionic current blockage amplitude, d_{DNA} is the diameter of the DNA. Using a DNA diameter estimate of 2.2 nm and the appropriate conductivity and membrane thickness, we can calculate the expected change in conductance during DNA translocation (ΔG*)* under different conditions. These DNA-induced current blockages must be greater than the peak-to-peak noise (I_{pp}) to be distinguishable from background noise (SNR = $\Delta I_{block} / I_{pp} > 1$). Based on Gaussian statistics, I_{pp} is roughly 6 times the root mean square current (I_{rms})[35]. Values of the maximum acceptable RMS noise in the ionic current signal for various experimental conditions are given in Table 2.1. From these values, we see that the ability to work in lower salt conditions or with thicker membranes is highly dependent on the noise performance of the nanopore device.

Table 2.1: *Maximum acceptable RMS noise current calculated at $V_{applied}$ = 200 mV, assuming a low 1/f noise nanopore device and translocation of DNA with d_{DNA} = 2.2 nm. Acceptable noise levels scale linearly in voltage, solution conductivity and inverse linearly in membrane thickness as per equation 2.2.*

Solution	**σ** **(nS/nm)**	**ΔI_{block} (nA)**		**Maximum I_{rms} (pA)**	
		L_{mem} = 10 nm	**L_{mem} = 30 nm**	**L_{mem} = 10 nm**	**L_{mem} = 30 nm**
400 mM KCl	4.6	0.35	0.12	~55	~15
1 M KCl	10	0.76	0.25	~125	~40
3.6 M LiCl	16	1.22	0.41	~200	~65

Before describing the noise sources associated with the addition of optofluidic components, we consider the mechanisms of noise production in the ion current signal of the standard nanopore system itself. The power spectral density (PSD) of electrical current noise, *S*(*f*), of a nanopore system provides information about the relative contributions of various noise sources, and can be expressed by the following equation:

$$S(f) = \frac{a_0}{f} + a_1 + a_2 f + a_3 f^2 \ \ [A^2/Hz] \tag{2.3}$$

where f is the frequency in hertz, and the coefficients a_0, a_1, a_2, and a_3 represent the respective weight of the contribution of the following types of noise: Flicker (1/f) and Thermal (typically dominating in the low-frequency regime, < 1 kHz); Dielectric and Capacitive in the high-frequency regime. Each of these noise types is produced by a different mechanism and dominates in a different range of the frequency spectrum as discussed in more detail elsewhere[27].

In addition to these inherent noise types, superfluous electrical noise can be picked up by the system and incorporated into the signal, thereby increasing I_{rms} and reducing SNR. As described in the following sections, care must be taken in the instrument design to minimize pickup, acoustic and illumination noise when a fluorescence microscope is coupled to a nanopore-based electrical detection system.

2.4.1 – Minimizing pickup and acoustic noise

Pickup noise consists of all external electrical noise sources that contaminate the electrical response of the nanopore system. The contamination occurs through two principal mechanisms: external electrical signals infiltrate the system through its power cords (i.e. noise coupled through the common ground of the power line), or through radiative electromagnetic interferences captured by induction. The first type of contamination is reduced by removing all possible noise sources from the electrical network in proximity of the system, i.e., powering down all adjacent unnecessary equipment or moving AC-DC converters of other instruments as far away as possible. For sensitive recordings, this may include laboratory light bulbs, microscope halogen lamps, the syringe pump, and even the microscope body, in order to power down the motors for the objective, filter cubes and the IR focusing laser employed in auto focus systems. The current amplifier internal ground should also be connected to the inner Faraday cage for shielding so as to avoid ground loops in the power lines.[27]

To address the second type of contamination – radiative pickup noise – two complete, nested Faraday cages are traditionally used.[27] In our setup, the inner copper Faraday cage is complete except for a small opening in the bottom to allow optical access to the nanopore as described above. To further isolate the system, the inner Faraday cage and all other instrumentation (except for the computer) are situated inside a second, outer

Faraday cage composed of wire mesh. Since the microscope is inside the outer, complete Faraday cage, its electronic features are shut down during data collection, in particular to shut off the IR laser of the auto focus system, which depending on the microscope type may be continuously emitting light even when it is not used to track the sample height. However, the Axopatch 200B, DAQ card, camera, and light source also produce electromagnetic radiation and must remain within the outer Faraday cage. Moreover, as described earlier, thin-walled PTFE tubing connects glass vials containing electrolyte solution to the fluidic device. When these tubes are filled with conducting solution, they behave as antennas, picking up electromagnetic interferences. Since controlling flow directly with a fluid-filled syringe would necessitate longer lengths of conductive tubing and thus be more vulnerable to this radiative noise pickup, we control the liquid flow using air-filled tubing to connect the glass vials inside the primary Faraday cage to the syringes located outside the cage. Further, the photosensitive AgCl electrodes are not placed directly into the fluidic device, as this could generate electrical noise and current offsets, even under scattered illumination conditions. By connecting the fluidic device to the electrodes via electrolyte-filled tubing, we can remove the electrodes from the immediate vicinity of the pore and prevent light exposure. Keeping tube length minimal is advised to limit both the antenna effect and the value of the fluidic tubing access resistance to a few tens of kΩ (or below a few % of the pore resistance).

Another noise source is related to the mechanical motion of the Ag/AgCl electrodes; acoustical vibrations of the electrodes in the solution-filled glass vials can appear in the electrical current signal in the low-frequency (< 100 Hz) regime of the PSD. To minimize this noise type, we mount the nanopore instrument on top of a vibration isolation table, following protocols from previously published work.[27] Clamping the electrodes firmly to the lids of the glass vials further reduces the potential for vibrations.

2.4.2 – Minimizing noise produced by the illumination source

A. Illumination source stability

Illuminating a standard nanopore chip during ionic current recording can significantly increase the electrical noise level. This is attributed to surface charge[20,36] and photoconductive effects in the Si support structure and

SiN_x membrane.[37–39] As a noticeable example, when laser light first strikes the SiN_x membrane, either by powering on the laser or aligning the optical path, a transient, decaying spike is observed in the ionic current trace, followed by a current spike in the opposite direction when illumination is removed (see Figure 2.5c). Together, these current spikes offer clear indication of coupling between the incident light and the ionic current.

Upon illuminating the chip with a mercury arc lamp, we observed unexpectedly high, persistent noise at multiple frequencies in the power spectrum of the ionic current signal. We attribute this to the operating principle behind the lamp, specifically its pulse width modulation (50% duty cycle at 150 Hz),[40] which implies that the bulb illuminates intermittently rather than continuously. In effect, the lamp is rapidly cycling on and off. Since fluctuations in the incident light intensity will cause corresponding fluctuations in the ionic current signal, an optofluiduic nanopore system optimized for low noise electrical recording demands a constant intensity light source.

We thus selected a continuous wave solid-state laser, a 488 nm diode laser with a maximum operating power of 50 mW (OBIS from Coherent). In many applications, it is important to control the intensity of the laser beam illuminating the sample, for instance to avoid rapid photobleaching of fluorophores or possible damage to CCD camera components. This can be achieved either through optical attenuation or by reducing the electrical power supplied to the laser. We have observed that the means by which the intensity is attenuated influences the noise on the electrical signal.

The effects of reducing the laser intensity by varying the selected operating power electrically in software (with the half-wave plate set to full passage) on the nanopore ionic current signal can be seen in the left column of Figure 2.4. Power spectrum densities of the ionic current signal were recorded while illuminating at 0 – 50 mW a Si chip supporting a 10-nm thick, 100×100 μm^2 SiN_x membrane immersed in 1M KCl, pH 10 mounted inside the fluidic cell described above (Figure 2.4a). Overall, relative to the 0 mW case (laser system turned on, but not lasing), laser-illuminated PSDs are characterized by a broad increase in noise at low frequencies ($\sim 10^1$-10^3 Hz) as seen in Figure 2.4b. Superimposed over this broad noise, substantial spikes can also be observed, indicating sources of noise of specific frequencies (ca. 30Hz, 60Hz, 75Hz, 100Hz, etc.). Interestingly,

both of these noise contributions increase significantly as laser power is electrically reduced from 50 mW (full operating power) to 5 mW. In addition, another broad peak centered around ~3 kHz becomes more pronounced at lower operating powers. Here, the noise from 0, 5, and 50 mW laser illumination is displayed – intermediate power values, (e.g at 40, 20, 10 mW) resulted in steadily increasing noise, with their spectra bounded by those of 50 and 5 mW (data not shown). Inspection of the RMS current values as a function of bandwidth (Figure 2.4c) reveals the relative contributions from these different noise sources. In particular, we see that various low frequency components (up to ~200 Hz) contribute substantially to the overall noise increase and that noise performance is best under full power illumination (~35 pA I_{rms} for 50 mW vs. ~88 pA I_{rms} for 5 mW at 10 kHz bandwidth). Significant photo-induced electrical noise is thus present in our system. In particular, the laser diode performance is dependent on the selected illumination power, with optimum stability and minimal noise increase achieved at full operating power (50mW). For ionic current measurements then, the laser should be operated at full power and optically attenuated down to a desired experiment-specific value, using the system of polarizers and filters described earlier. This results in noise comparable to or lower than that of the 50 mW (red) traces in Figure 2.4 over a range of effective illumination intensities.

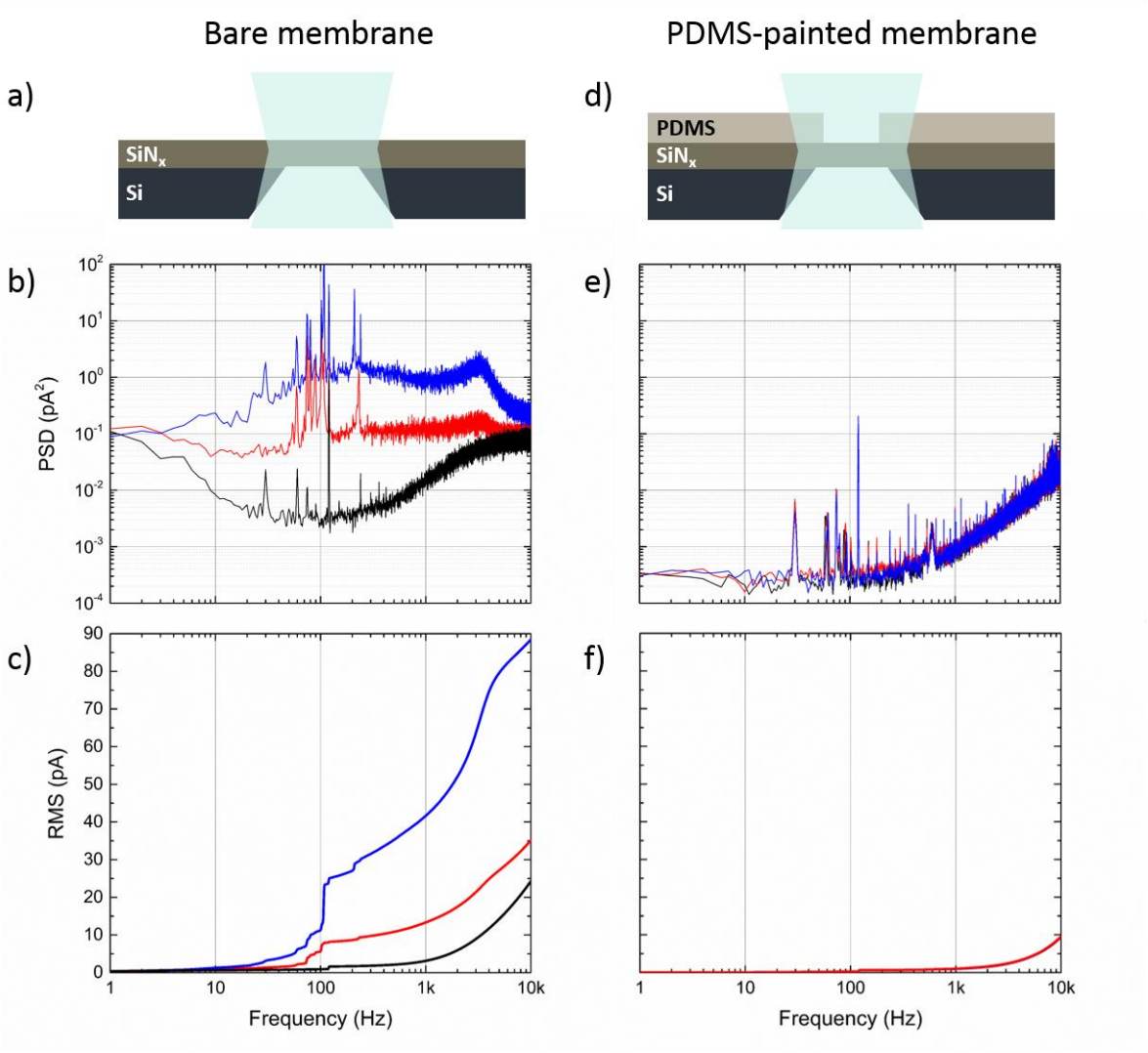


Figure 2.4: *Comparison of noise produced by various laser power settings in bare membranes (left column) versus PDMS-painted membranes (right column).* ***a)*** *Schematic diagram of silicon support chip with bare SiN_x membrane. The laser is illuminating the membrane and part of the supporting Si structure.* ***b)*** *PSD of the system with different power settings: 0 mW (black), 5 mW (blue), and 50 mW (red).* ***c)*** *I_{rms} calculated as a function of frequency for various laser powers (same color code) for the same device.* ***d)*** *Schematic diagram of silicon support chip with SiN_x membrane closely painted with PDMS. The laser is illuminating the PDMS layer, in addition to the SiN_x membrane and part of the supporting Si structure.* ***e)*** *PSD of noise spectra produced by PDMS-painted chip under 50-mW laser illumination (red curve) compared with the PSD for the same system with 5-mW (blue), and 0-mW (black) laser illumination.* ***f)*** *Integrated I_{rms} values of the PSDs in (e) with bandwidth. The similarity of the traces demonstrates the successful decoupling of illumination effects from device noise. In all cases, data collected using 10-nm thick, 100×100 μm^2 membranes in 1M KCl pH 10 and low-pass filtered at 100kHz with an Axopatch 200B.*

B. PDMS-painting of chip surface

Despite an overall reduction in noise with the laser running at maximum operating power, there is still additional noise in the ionic current signal when comparing to a standard nanopore setup. We attribute this illumination-related noise to two principal mechanisms. First, illumination of SiN_x (and the underlying Si support) with visible light can produce electron-hole pairs[37–39], which increase the conductivity of the substrate and undermine its dielectric properties. Such an effect would be expected to result in increased dielectric noise under illumination, which is proportional to both the dissipation factor D, and capacitance C_{chip} of the nanopore chip[27]. The visibly larger slope in the PSDs of illuminated devices above ~1 kHz (up to the 3-kHz peak) may in fact be a manifestation of this increased dielectric noise (Figure 2.4b); compare also the changes in the RMS curves, Figure 2.4c, over this range ($\Delta I_{RMS}|_{1\ kHz}^{3\ kHz} = 6$ pA, 8 pA, 22 pA for 0 mW, 50 mW and 5 mW laser powers, respectively). Secondly, these photoinduced electron-hole pairs are also thought to promote electrochemical reactions at the chip-electrolyte interface[36,41]. Through such reactions or otherwise, we propose that even small fluctuations in the illumination intensity modulate the charge density of the SiN_x surface exposed to liquid electrolyte[20,36], which affects the distribution of ions in the electrostatic double layer, much like a capacitor charging/discharging. Indeed, when the laser is switched on or off, a feature very similar to a voltage-induced capacitive spike is seen in the current signal, indicative of this effect, albeit more pronounced[41,42] (Figure 2.5c).

It has previously been shown[28] that coating the chip in PDMS isolates the vast majority of the surface, reducing chip capacitance and corresponding high-frequency noise. Here, as seen in the right column of Figure 2.4, painting PDMS on the SiN_x membrane side of the chip (Figure 2.4d, described in detail above) similarly improves our noise performance, reducing the coupling between the noise and the illumination intensity, and bringing the fully-illuminated noise spectrum on par with the illumination-free PSD (Figure 2.4e). Note that although a few adverse spikes remain in the noise spectra at low frequencies, none integrate to contribute substantially (< 1 pA) to the overall RMS noise (Figure 2.4f). Illuminating a PDMS-painted membrane (10-nm, 100x100 μm^2 SiN_x in 1M KCl, pH 10) at 50 mW and at a current-reduced 5 mW, we observe no degradation of noise performance (I_{rms} ~9.5 pA at 10 kHz), demonstrating that the coupling between the laser fluctuations and

the ionic current can be reduced by simultaneously minimizing device capacitance and limiting the surface area (of the substrate exposed to electrolyte) available for interfacial electrochemical reactions/surface charge modulation. In this way, the PDMS coating would influence both potential noise mechanisms (dielectric and surface charge fluctuation) described earlier. Reduction in photo-induced electrical noise has previously been achieved by replacing the Si substrate with high-resistivity borosilicate glass[36]; our results represent a similar decoupling of noise from laser power while maintaining the same underlying chip design. However, the magnitude of the chip capacitance will also have a significant effect as the bandwidth of the current recording increases[27].

2.5 – DNA Translocation under Membrane Illumination

The primary objective of the instrument improvements described herein is to reduce background noise (I_{rms}) to a level that permits electrical detection of DNA passage through a nanopore under laser illumination. The implementation of these instrument improvements has reduced the I_{rms} to ~10 pA at 10 kHz, making possible the detection of dsDNA under different conditions as described in Table 2.1. In the following section, we confirm the success of our noise minimization strategies with successful electrical detection of 10-kbp dsDNA through a laser-illuminated nanopore (see Figure 2.5a).

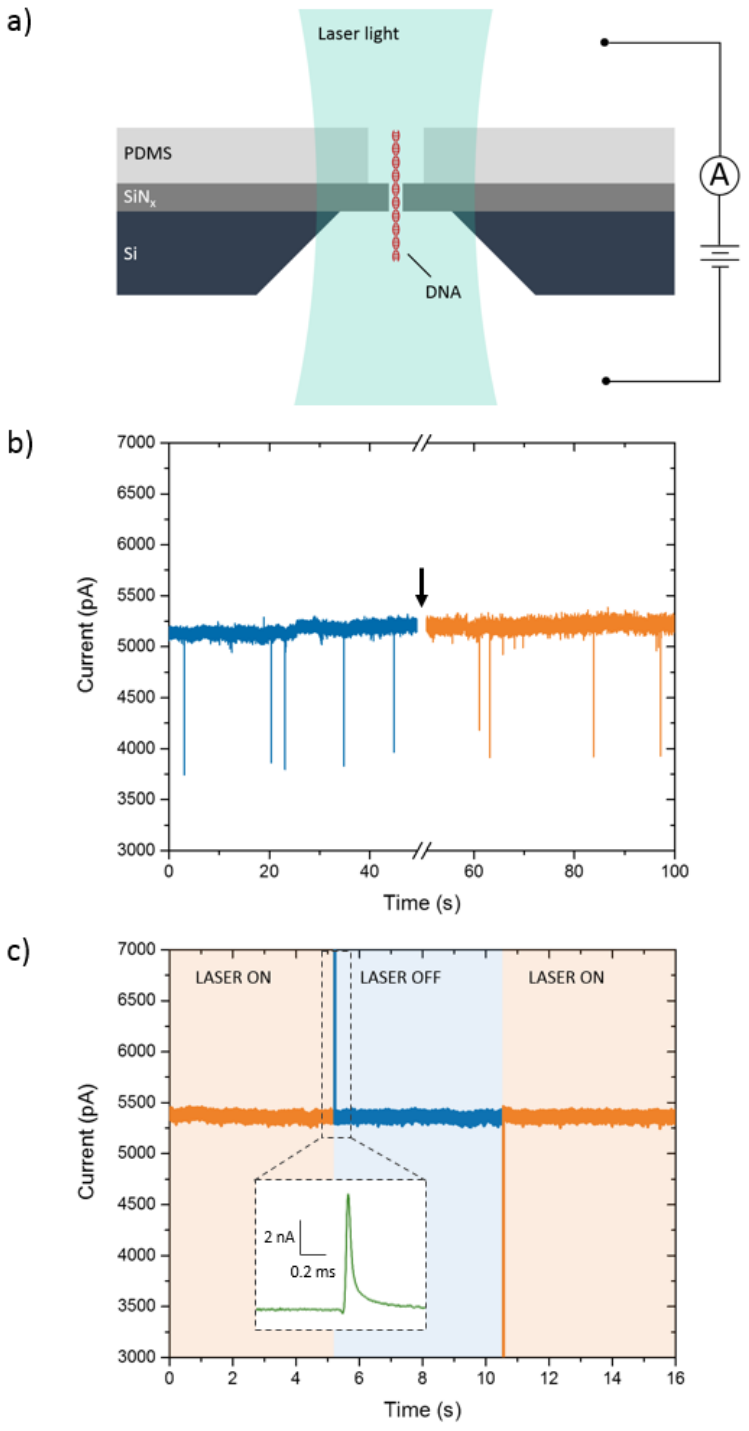


Figure 2.5: *Electrical detection of unlabelled 10-kbp DNA with and without laser illumination.* ***a)*** *Schematic of fluidic device under illumination, depicting layers of low-noise chip design and DNA passage through the nanopore.* ***b)*** *Typical trace of ionic current data collected from a nanopore under illumination from a 50-mW laser (orange) and with the laser switched off (blue); the position of the arrow indicates a transition between illumination conditions. Data collected from a 5.4 nm pore in a 10 nm membrane in 3.6 M LiCl at 200 mV.* ***c)*** *Response of ionic current to the addition or removal of incident laser light, attributed to membrane charging effects. Data taken with 200 mV transmembrane bias, and 50 mW laser power (under "laser on" condition).*

In this experiment, unlabelled 10-kbp dsDNA (1 nM) was introduced to the electrolyte solution (3.6M LiCl, pH 8) in a glass vial connected by PTFE tubing to the *membrane* side of a PDMS-painted membrane containing a CBD-fabricated pore. Using an Axopatch 200B amplifier, a potential difference of 200 mV was then applied across the Ag/AgCl electrodes to electrophoretically drive the negatively charged DNA through the nanopore while the ionic current was measured. Typical traces featuring DNA translocations with and without 50 mW laser illumination are displayed in Figure 2.5b. Noticeably, both the overall noise and the DNA blockage depths remain consistent between illumination conditions. Transient spikes in the ionic current that are observed when the laser is switched on or off (Figure 2.5c) are attributed to a capacitive response as described above, and confirm the illumination of the membrane during the experiment. By employing DNA as a molecular sized ruler, and entering the average translocation characteristics into equations 2.1 and 2.2, the nanopore dimensions can be extracted. These are displayed in Table 2.2. The experimental blockage amplitude of 1.27 nA compares favourably with that expected for dsDNA (Table 2.1) after factoring in the slightly larger conductivity of the electrolyte used here (16.4 S/m). Meanwhile an effective pore length of 9.9 nm is also within expectations[31], given possible pore deviations from perfectly cylindrical geometry and a vendor-reported nominal membrane thickness of 10 nm. Employing the noise reduction strategies described above our system is shown to electrically detect DNA passage through a nanopore under membrane illumination, with an SNR ~10 at a 10 kHz bandwidth.

Table 2.2: *Characterization of DNA translocations for the experiment pictured in Figure 2.5.*

σ (nS/nm)	ΔI_{block} (nA)	d_{pore} (nm)	L_{eff} (nm)	I_{rms} (pA)	SNR
16.4	1.27	5.4	9.9	21.0	10.0

2.6 – Conclusions

In summary, we have presented the design and instrumentation of an integrated optofluidic solid-state nanopore sensor capable of low-noise electrical detection of single biomolecules under laser illumination. We identified major sources of electrical noise in nanopore systems and described design adjustments made to minimize them, with a focus on the noise resulting from chip illumination. In particular, we find that PDMS-painting of the SiN_x membrane brings the noise performance of the illuminated device on par with its non-illuminated performance, even eliminating the otherwise adverse effects of operating the laser below its nominal power value. The success of our efforts is demonstrated by results showing detection of DNA translocation events during membrane illumination with a signal-to-noise ratio (SNR) of $\sim$10.

As discussed in *Section 2.4*, reducing electrical noise plays an important role in expanding the range of experimental conditions (e.g. salt concentration, analyte diameter, membrane thickness) available to single-molecule nanopore sensors. Many of the noise reduction strategies implemented here are widely applicable and may help enable future nanopore-based instruments to achieve high quality synchronous optoelectronic sensing.

Acknowledgments

The authors would like to thank M. Godin for generously sharing his microfabrication equipment, R. Bhardwaj for his valuable assistance in calibrating the illumination sources, and L. Andrzejewski for valuable technical support. We thank H. Kwok and K. Briggs for many valuable discussions. We thank N. Yelle and M. Windeler for helping in various aspects of the construction of the instrument.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada (NSERC), and by the National Secretary for Higher Education and Science of Ecuador SENESCYT. ZR acknowledges the financial support provided by the OGS program for postgraduate fellowships.

2.7 – References

[1] M. Wanunu, Phys. Life Rev. **9**, 125 (2012).

[2] C. Dekker, Nat. Nanotechnol. **2**, 209 (2007).

[3] E.C. Yusko, B.R. Bruhn, O.M. Eggenberger, J. Houghtaling, R.C. Rollings, N.C. Walsh, S. Nandivada, M. Pindrus, A.R. Hall, D. Sept, J. Li, D.S. Kalonia, and M. Mayer, Nat. Nanotechnol. **12**, 360 (2016).

[4] D. Branton, D.W. Deamer, A. Marziali, H. Bayley, S. a Benner, T. Butler, M. Di Ventra, S. Garaj, A. Hibbs, X. Huang, S.B. Jovanovich, P.S. Krstic, S. Lindsay, X.S. Ling, C.H. Mastrangelo, A. Meller, J.S. Oliver, Y. V Pershin, J.M. Ramsey, R. Riehn, G. V Soni, V. Tabard-Cossa, M. Wanunu, M. Wiggin, J. a Schloss, and others, Nat. Biotechnol. **26**, 1146 (2008).

[5] H. Bayley, Clin. Chem. **61**, 25 (2015).

[6] S.C. Vollmer and H.W. de Haan, J. Chem. Phys. **145**, 154902 (2016).

[7] G. Ando, C. Hyun, J. Li, and T. Mitsui, ACS Nano **6**, 10090 (2012).

[8] R.M.M. Smeets, U.F. Keyser, D. Krapf, M.-Y. Wu, N.H. Dekker, and C. Dekker, Nano Lett. **6**, 89 (2006).

[9] S.W. Kowalczyk, A.Y. Grosberg, Y. Rabin, and C. Dekker, Nanotechnology **22**, 315101 (2011).

[10] L. Bo, F. Albertorio, D.P. Hoogerheide, J.A. Golovchenko, and B. Lu, Biophys. J. **101**, 70 (2011).

[11] C. Plesa, L. Cornelissen, M.W. Tuijtel, and C. Dekker, Nanotechnology **24**, 475101 (2013).

[12] A. Ivankin, R.Y. Henley, J. Larkin, S. Carson, M.L. Toscano, and M. Wanunu, ACS Nano **8**, 10774 (2014).

[13] B.N. Anderson, O.N. Assad, T. Gilboa, A.H. Squires, D. Bar, and A. Meller, ACS Nano **8**, 11836 (2014).

[14] F. Nicoli, D. Verschueren, M. Klein, C. Dekker, and M.P. Jonsson, Nano Lett. **14**, 6917 (2014).

[15] S. Nam, I. Choi, C. Fu, K. Kim, S. Hong, Y. Choi, A. Zettl, and L.P. Lee, Nano Lett. **14**, 5584 (2014).

[16] S. Pud, D.V. Verschueren, N. Vukovic, C. Plesa, M.P. Jonsson, and C. Dekker, Nano Lett. 150903072257007 (2015).

[17] C.R. Crick, P. Albella, B. Ng, A.P. Ivanov, T. Roschuk, M.P. Cecchini, F. Bresme, S.A. Maier, and J.B. Edel, Nano Lett. (2014).

[18] J.E. Reiner, J.W.F. Robertson, D.L. Burden, L.K. Burden, A. Balijepalli, and J.J. Kasianowicz, J. Am. Chem. Soc. **135**, 3087 (2013).

[19] M. Belkin, S.-H. Chao, M.P. Jonsson, C. Dekker, and A. Aksimentiev, ACS Nano (2015).

[20] N. Di Fiori, A. Squires, D. Bar, T. Gilboa, T.D. Moustakas, and A. Meller, Nat. Nanotechnol. **8**, 946 (2013).

[21] G. V Soni, A. Singer, Z. Yu, Y. Sun, B. McNally, and A. Meller, Rev. Sci. Instrum. **81**, 14301 (2010).

[22] B. McNally, A. Singer, Z. Yu, Y. Sun, Z. Weng, and A. Meller, Nano Lett. **10**, 2237 (2010).

[23] V. Kurz, E.M. Nelson, J. Shim, and G. Timp, ACS Nano **7**, 4057 (2013).

[24] T. Auger, J. Mathé, V. Viasnoff, G. Charron, J.M. Di Meglio, L. Auvray, and F. Montel, Phys. Rev. Lett. **113**, 1 (2014).

[25] J. Larkin, M. Foquet, S.W. Turner, J. Korlach, and M. Wanunu, Nano Lett. **14**, 6023 (2014).

[26] W.Y. Zhang, G.S. Ferguson, and S. Tatic-Lucic, in *17th IEEE Int. Conf. Micro Electro Mech. Syst. Maastricht*

MEMS 2004 Tech. Dig. (IEEE, 2004), pp. 741–744.

[27] V. Tabard-Cossa, in *Eng. Nanopores Bioanal. Appl.*, edited by J.B. Edel and T. Albrecht, First Edit (Norwich, N.Y. : William Andrew ; Oxford : Elsevier Science distributor, 2013), pp. 59–93.

[28] V. Tabard-Cossa, D. Trivedi, M. Wiggin, N.N. Jetha, and A. Marziali, Nanotechnology **18**, 305505 (2007).

[29] N.N. Jetha, M. Wiggin, and A. Marziali, in *Micro Nano Technol. Bioanal.*, edited by J.W. Lee and R.S. Foote (Totowa, N.J. : Humana Press ; New York: Springer Science distributor, 2009), pp. 113–127.

[30] K. Briggs, H. Kwok, and V. Tabard-Cossa, Small **10**, 2077 (2014).

[31] H. Kwok, K. Briggs, and V. Tabard-Cossa, PLoS One **9**, e92880 (2014).

[32] E. Beamish, H. Kwok, V. Tabard-Cossa, and M. Godin, Nanotechnology **23**, 405301 (2012).

[33] K. Briggs, M. Charron, H. Kwok, T. Le, S. Chahal, J. Bustamante, M. Waugh, and V. Tabard-Cossa, Nanotechnology **26**, 84004 (2015).

[34] R. Tahvildari, E. Beamish, K. Briggs, S. Chagnon-Lessard, A.N. Sohi, S. Han, B. Watts, V. Tabard-Cossa, and M. Godin, Small **13**, 1 (2017).

[35] R. Sherman-Gold, editor , *The Axon Guide - Electrophysiology and Biophysics Laboratory Techniques*, 3rd ed. (Molecular Devices, 2012).

[36] W.H. Pitchford, H.-J. Kim, A.P. Ivanov, H.-M. Kim, J.-S. Yu, R.J. Leatherbarrow, T. Albrecht, K.-B. Kim, and J.B. Edel, ACS Nano **9**, 1740 (2015).

[37] J. Robertson and M.J. Powell, Appl. Phys. Lett. **44**, 415 (1984).

[38] J. Robertson, W.L. Warren, and J. Kanicki, J. Non. Cryst. Solids **187**, 297 (1995).

[39] S.M. Sze and K.K. NG, *Physics of Semiconductor Devices* (Willey-Interscience, 2007).

[40]*X-Cite 120 PC User's Guide* (EXFO Photonic Solutions Inc., 2004).

[41] Y. Li, C. Chen, S. Kerman, P. Neutens, L. Lagae, G. Groeseneken, T. Stakenborg, and P. Van Dorpe, Nano Lett. **13**, 130306120838001 (2013).

[42] G. Danda, P. Masih Das, Y.-C. Chou, J.T. Mlack, W.M. Parkin, C.H. Naylor, K. Fujisawa, T. Zhang, L.B. Fulton, M. Terrones, A.T.C. Johnson, and M. Drndić, ACS Nano **11**, 1937 (2017).

CHAPTER 3 – MICROFLUIDIC DEVICES FOR IDEP PRE-CONCENTRATION AND SOLID-STATE NANOPORE SENSING

3.1 – Introduction

Solid state nanopores are capable of performing highly sensitive, single-molecule measurements on biological analytes, but this high sensitivity generally comes paired with relatively low throughputs and detection efficiencies [1]. This is because the rate of target passage through a nanopore is typically limited by the rate at which molecules are able to randomly diffuse to be within a short distance from its opening, where capture forces from a localized electric field dominate molecular motion [2] (e.g. a ~μm capture radius ⇒ ~fL–pL capture volume is typical in experiments involving double-stranded DNA, compared to ~μL–mL loaded sample volumes). Operating within this "diffusion-limited" regime can therefore lead to prolonged experiment times while waiting for a sufficient number of detection events (from which to draw statistically valid conclusions about a sample) to be generated. One way of increasing the absolute rate of detection is to use samples at elevated (bulk) target concentrations, but this is not always feasible for particular diagnostically relevant targets (e.g. at < nM), especially if direct detection within a minimally processed sample (e.g. human serum) is desired. For instance, using nanopore capture rates on the order of 1 Hz/nM (as observed with ~10 nm pores capturing dsDNA at ~100 mV [3]), the time to acquire a typical dataset (hundreds of translocation events in size) is predicted to rise from a few minutes at sample concentrations in the nM range to > 1 day as the concentration is lowered into the fM range and below.

Instead of relying on passive diffusion to deliver targets, in this work we investigate an alternative approach of raising the detection rate by actively transporting molecules to the capture volume of the nanopore. Specifically, we use dielectrophoresis (DEP), an electrokinetic technique that operates on a molecule's induced polarization in solution under an external electric field. Advantages of the DEP method to our particular application include its universality (all molecules, even uncharged, have some intrinsic degree of polarizability), its selectivity (the DEP force / frequency response experienced by a molecule depends strongly on its size and dielectric properties), and its ease of integration with existing nanopore infrastructure (e.g. fluidic cells that already

incorporate microchannels, electrodes, etc. to enable the electrophoretic translocation of molecules though the pore). Here, we first discuss some basic design principles involved in combining insulator-based DEP (iDEP) pre-concentration with nanopore sensing in a single device. We then present preliminary results of (separately) concentrating and sensing double-stranded DNA in our realized iDEP-nanopore system. Finally, we identify and discuss some key remaining challenges to iDEP integration; surmounting these challenges will enable improved nanopore capture rates in similar devices moving forward.

3.2 – Pre-Concentration with Dielectrophoresis

3.2.1 – Target Polarization

Dielectrophoresis (DEP) is the motion of an electrically polarizable object within a non-uniform electric field [4]. The left panels of Figure 3.1a illustrate the effect for the case of a neutral, homogeneous sphere. In a uniform electric field, separated charges (positive and negative) in the polarized object feel equal and opposite electric forces ($F_{EP} = qE$), resulting in no net motion of the particle. In a non-uniform field, however, separated charges in the region of higher E-field feel a stronger force than their oppositely charged counterparts, and so the particle as a whole is drawn along the gradient in field strength toward the highest-$|E|$ regions (for "positive DEP" specifically, see below). Note that this net force is independent of the *direction* of the electric field; switching the voltage polarity in Figure 3.1a reverses the orientation of the external field lines and the charge distribution within the polarized sphere, but separated charges near the top of the sphere (now positive) still feel a greater force (from the higher $|E|$ values), again moving the particle toward the top electrode (now negative). One important impact of this property is that DEP motion can be actuated by alternating (AC) electric fields, which allows for its decoupling from other electrokinetic effects whose forces are linear with E (electrophoresis, electroosmosis [4]).

The expected dielectrophoretic force on such a uniformly polarizable, uncharged sphere has a relatively simple expression [5]:

$$\vec{F}_{DEP} = 2\pi\varepsilon_M r^3 \mathrm{Re}\{f_{CM}\}\vec{\nabla}\left|\vec{E}\right|^2 \tag{3.1}$$

where r is the particle radius, ε_M is the permittivity of the surrounding medium, and $\vec{E}$ is the external electric field. The parameter f_{CM} is called the "Clausius-Mossotti factor" and depends on the permittivities (ε) and conductivities (σ) of both the particle (P) and the medium (M) through the "complex permittivity", ε^*, of each:

$$f_{CM} = \frac{\varepsilon_P^* - \varepsilon_M^*}{\varepsilon_P^* + 2\varepsilon_M^*} \tag{3.2}$$

$$\varepsilon^* = \varepsilon - i\frac{\sigma}{\omega} \tag{3.3}$$

where ω is the (angular) frequency of the external AC field and $i = \sqrt{-1}$.

Two characteristic phenomena involved in DEP manipulation can be brought to light by examining the Clausius-Mossotti factor in detail. The first is that the overall sign of F_{DEP} is determined by the relative polarizabilities of the particle and the medium (through their difference in ε^*). A particle that is more polarizable than its surrounding medium has a positive value for $\mathrm{Re}\{f_{CM}\}$ and thus moves along the gradient in E^2 toward regions of highest field strength ("positive DEP", pDEP). Conversely, a particle that is less polarizable than its surrounding medium has negative $\mathrm{Re}\{f_{CM}\}$ and moves opposite $\vec{\nabla}E^2$ toward regions of lowest field strength ("negative DEP", nDEP). The right panels of Figure 3.1a summarize the two situations, which are differentiated by which side of the particle-medium interface contains the greater density of net charge induced by the external field.

The second phenomenon encapsulated by the Clausius-Mossotti factor is that, even if the electrical properties (ε, σ) of particle and medium are approximated to be independent of the driving frequency of the electric field, the direct appearance of ω in the expression of complex permittivity (Equation 3.3) makes DEP itself intrinsically frequency-dependent. This can affect the overall strength of F_{DEP}, and even whether positive or negative DEP is observed. A typical response (e.g. for a polystyrene bead in a low-σ medium) is for DEP to be positive at low frequencies and transition to being negative at high frequencies, passing through a critical crossover frequency where $\mathrm{Re}\{f_{CM}\} = 0$ in the process [6]. Selecting an electric field frequency that is appropriate to the desired

outcome of a given experiment (with its particular DEP targets and medium) therefore becomes an important consideration.

In the context of integrating DEP pre-concentration with nanopore sensors, the theory of target polarization is considerably more complex for the prototypical nanopore analyte of double-stranded DNA (dsDNA) than for the case of a simple uncharged sphere (as summarized by Eq. 3.1). DNA is a polyelectrolyte, with regularly-spaced negative charge located along its sugar-phosphate backbone [6]. When placed in an ionic solution (such as those utilized in both DEP manipulation and nanopore sensing) the free ions in solution tend to screen this net negative charge on the DNA by accumulating a cloud of excess positive charge (counter-ions) at its surface [7]; this charge excess decays with distance away from the DNA surface over a scale called the "Debye length", λ_D (see Figure 3.1b). Higher salt concentrations in solution lead to this screening effect taking place over shorter distances, with the Debye length predicted to scale with the inverse square root of the bulk ion concentration ($\lambda_D \propto c_\infty^{-1/2}$) [5]. For instance, the Debye length in a monovalent salt solution is predicted to shrink from ~10 nm to ~0.3 nm as the salt concentration is raised from 1 mM to 1 M.

In the most commonly invoked model of DNA polarization, counterions are pictured to move freely along a characteristic subunit length of the DNA's longitudinal axis in response to the parallel component of an external electric field, thus creating an induced dipole moment in this direction ("counterion fluctuation model" [6]). A particular phase of "condensed" counterions that are strongly, but non-locally, bound to the backbone charge of DNA (see Figure 3.1b) are thought to be especially important to the process. When estimates of the polarization subunit length are extracted from experiments (on the dielectric properties of DNA), they roughly scale with the Kuhn length (see *Chapter 1*) of DNA in those conditions, indicating that the "breaks" in the free movement of counterions may arise from the worm-like chain conformation of DNA in solution. While the agreement between theory and experiment for DNA polarization is often less than quantitative [8], the counterion fluctuation model generally predicts larger DNA polarizabilities for longer molecules (more subunits to contribute to a net dipole) and this trend has been verified experimentally by several authors [9,10].

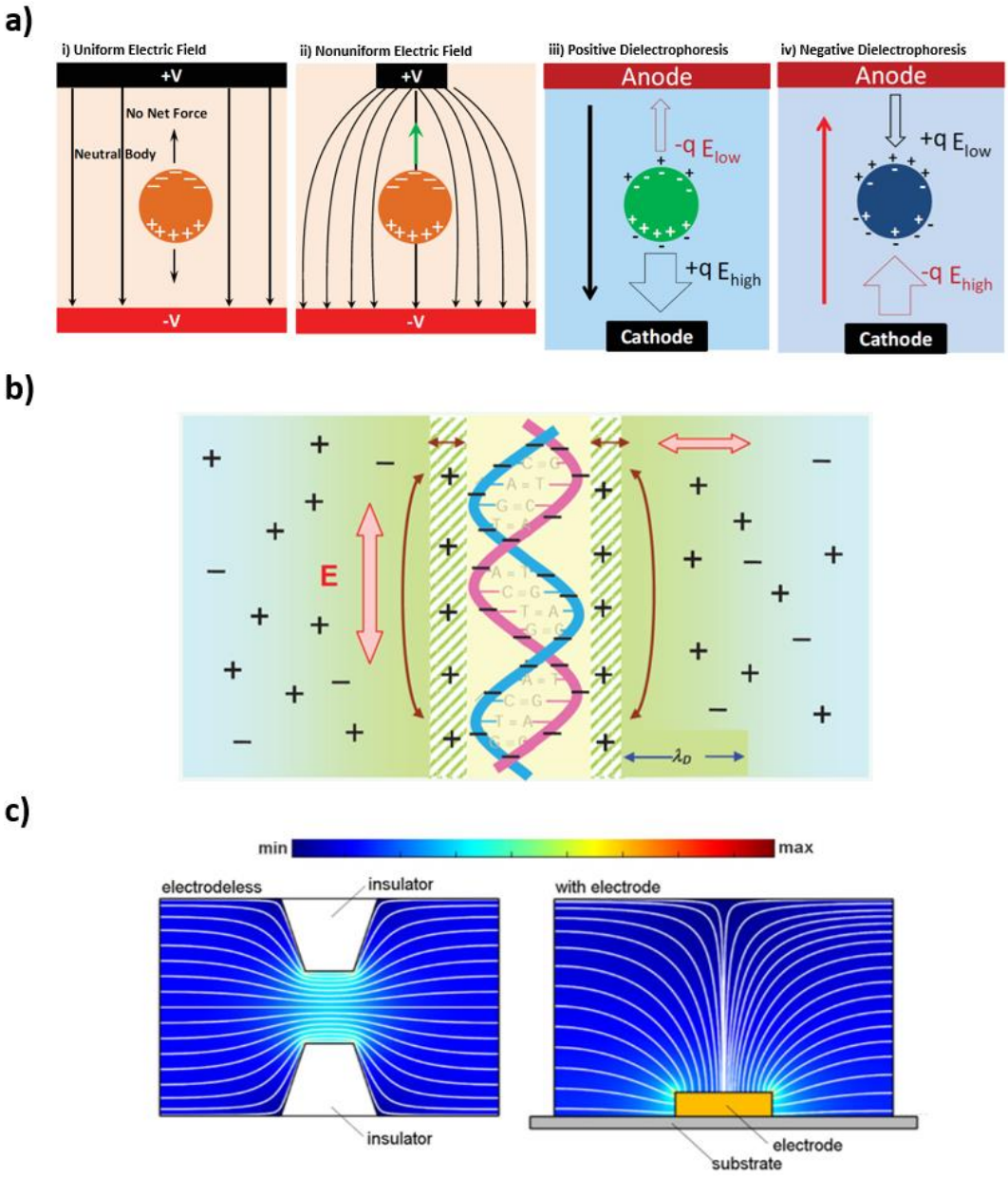


Figure 3.1: *Principles of dielectrophoresis.* ***a)*** *Left: net electric forces (green arrows) on a homogeneous, uncharged sphere placed in uniform (i) and non-uniform external electric fields (ii). Right: comparison of positive (iii) and negative (iv) dielectrophoresis conditions, based on the relative polarizabilities of a target sphere and its surrounding medium. The net motion of the particle in each case is indicated by the solid arrow, either toward the cathode (in black) or the anode (in red). Figure adapted from* [11]*, with permission.* ***b)*** *Illustration of the distribution of counterions around a negatively-charged dsDNA fragment. Two phases of positive counterions are shown: a tightly-bound ("condensed") layer at the DNA surface and a lower-density ("diffuse") layer outside of this that decays in excess charge with characteristic length λ_D away from the DNA. Longitudinal motion of the condensed counterions in response to an external electric filed are thought to play a large role in the polarization of DNA. Figure adapted from* [6]*, with permission.* ***c)*** *Comparison of insulator-based ("iDEP", left) and microelectrode-based ("eDEP", right) DEP schemes. Each panel shows the distribution of electric field lines (in white) and field gradient strengths (∇E^2, colourbar range overtop) around a representative channel geometry. Figure adapted from* [12]*, with permission.*

3.2.2 – Insulator-Based Dielectrophoresis

Two main approaches are commonly used to generate the non-uniform electric fields needed to power DEP within a microfluidic device (Figure 3.1c) [4]. The initial and standard method is to embed micro-patterned electrodes within the channels of the device, at the position where the intended target manipulation is to take place ("microelectrode-based DEP", "eDEP"). The alternative method is to place macroscale electrodes far from each other (at the ends of the channel), and instead create inhomogeneities in the field by varying the cross-sectional area along the channel between them, e.g. by incorporating posts, constrictions, etc. made out of the same electrically insulating material of the channel walls. This second approach is called "insulator-based DEP" or "iDEP". Advantages to iDEP over eDEP include: a less complex fabrication procedure (not involving metal deposition), no electrode surface fouling or electrochemical side effects at the location of the manipulated targets (electrodes far away), and electric field inhomogeneities that span the entire height of the channel and thus affect targets throughout (in contrast to eDEP where the fields rapidly decay away from the electrode surface, see Figure 3.1c). Another benefit to iDEP is that targets can be readily delivered to their main manipulation zone by electrokinetic effects (electrophoresis, electroosmosis) simply by adding a DC component to the potential difference between the electrodes; eDEP devices usually require the use of additional pressure regulators to actuate particle migration by pressure-driven flow, adding complexity to the experimental set-up [4]. On the other hand, a major downside to iDEP is that, because of the large electrode separation, it typically requires higher voltages relative to eDEP in order to generate DEP forces of similar strength. Higher voltages can also reduce the range of potential *E*-field frequencies that can be applied to an iDEP system due to the limited slew rates of the high-V power supplies used to create them [4].

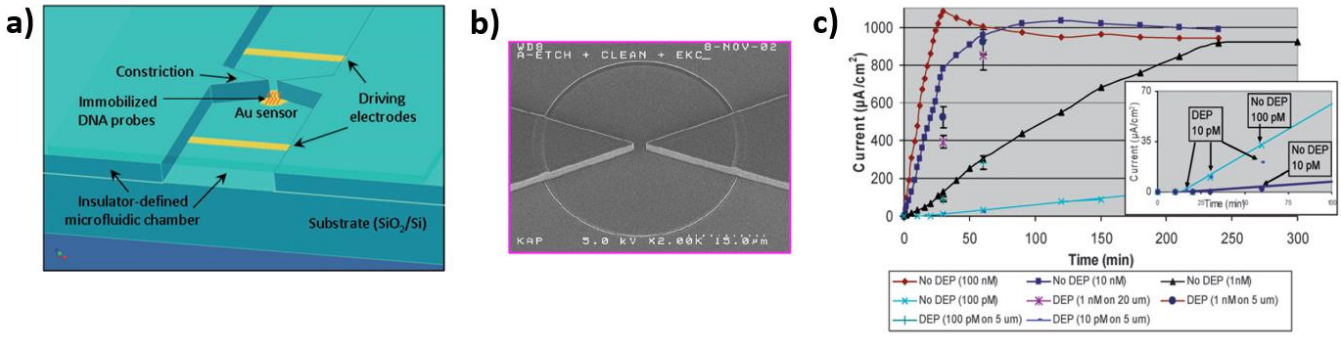
For our application of combining DEP pre-concentration with a nanopore, an added challenge exists of integrating and carefully positioning a secondary (nanopore) sensor within the DEP microfluidic device. Two examples of realized hybrid DEP-sensor systems are highlighted in Figure 3.2, both featuring iDEP. In the first (by Swami et al. [13]), a DNA hybridization sensor was centred underneath a 1 μm wide, triangular constriction in PDMS (narrowing from 500 μm) that was used to draw single-stranded DNA (ssDNA) to the sensor at elevated

concentrations by DEP trapping (Figure 3.2a-c). The sensor itself consisted of a circular gold electrode functionalized with complementary ssDNA capture probes (Figure 3.2a,b); hybridization of target DNA to the immobilized capture probes was signalled by increased current levels as measured in a "two-potential electrochemical detection platform" [13]. Over the range of target ssDNA tested (10–1000 pM), when hybridization detection was preceded by a DEP concentrating step (150 V/cm, 500 Hz AC potential for 30–60 min in 50 mM NaCl) it resulted in signal levels that were comparable to trials with only passive target diffusion (no DEP) at ~10× the concentration, see Figure 3.2c.

Similarly, in the second example (by Puttaswamy et al. [14]), a sensor consisting of a poly-silicon nanowire field effect transistor (Poly-Si NWFET) was placed underneath the centre of an insulating constriction, this time made of polyethylene glycol diacrylate (PEG-DA) (Figure 3.2d-e). The critical dimensions for the construction in this case were 70 µm wide at the channel inlets/outlets, narrowing to 4 µm wide in the sensing zone. When short dsDNA (1 nM 20 bp, in 10 mM sodium phosphate buffer) was introduced into the channel, a noticeable effect of DEP concentrating voltages (30 V_{pp}, 500 Hz for 5 min) was observed in the shift of the FET characteristics (drain current vs. gate voltage) and this was attributed to the accumulation of DNA on the nanowires; removing the DEP voltage resulted in the return of I_D-V_G back to baseline (Figure 3.2f).

This most direct DEP pre-concentration design illustrated by the above examples, namely of placing a high-sensitivity sensor directly at the focus of the electric field lines of a constriction-based iDEP channel (so as to transport low-concentration targets into its sensing region) was pursued in our own work with nanopore sensors that follows. In our case, it is a thin (~10-nm) silicon nitride membrane that will eventually contain the pore that must be carefully aligned within the system. Once an iDEP-SiN_x membrane device is constructed, pore fabrication followed by immersing the channels in high salt buffer (1 M KCl) and applying relatively large transmembrane potentials (~10 V) to create a single nano-scale hole by "controlled dielectric breakdown" (CBD) [15].

Swami et al, *Lab on a Chip*, 2009



Puttaswamy et al, *Sensors and Actuators B*, 2013

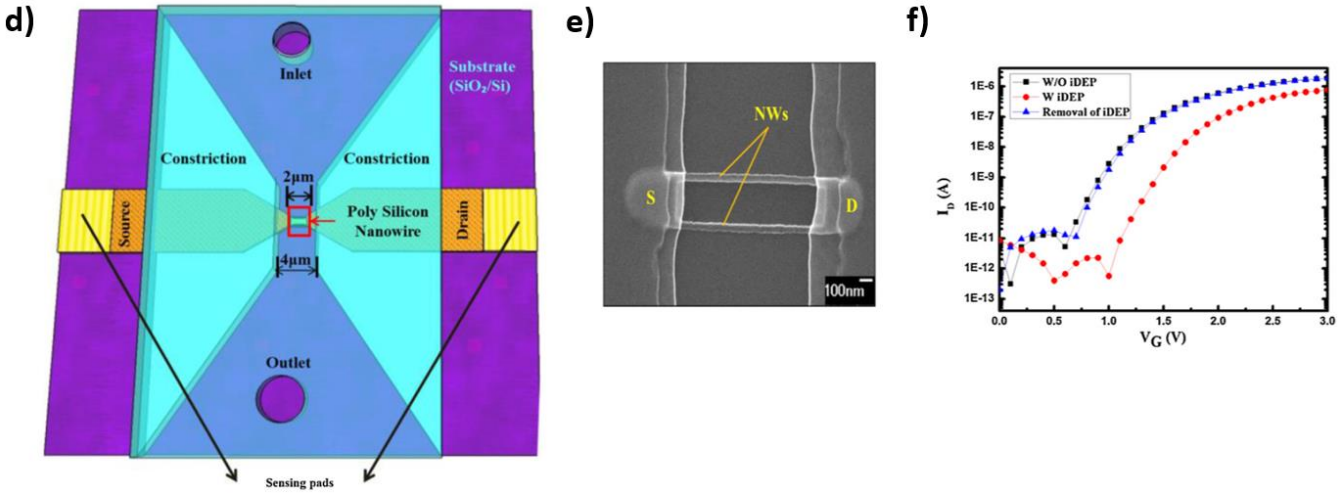


Figure 3.2: *Examples of combined iDEP-sensor microfluidic systems.* ***a)*** *Schematic of a constriction-based PDMS iDEP channel that incorporates a DNA hybridization sensor at its centre. The sensor consists of a circular gold electrode functionalized with ssDNA capture probes.* ***b)*** *Electron microscope image showing a detailed view of the constriction/gold sensor geometry of (a).* ***c)*** *Plot of target-probe DNA hybridization signals ("two-potential electrochemical sensing") using the system in (a) at various target concentrations (10 pM – 100 nM), with and without a DEP concentrating step (150 V/cm, 500 Hz V_{AC} for 30–60 min in 50 mM NaCl).* ***d)*** *Schematic of a constriction-based PEG-DA iDEP channel that incorporates a Poly-Si NWFET sensor at its centre.* ***e)*** *SEM image of Poly-Si NWFET showing its basic structure of two Si nanowires connected to source and drain electrodes* ***f)*** *FET characteristics of Poly-Si NWFET sensor embedded in a dsDNA-filled iDEP channel (1 nM 20 bp, in 10 mM sodium phosphate buffer), with and without an applied DEP voltage (30 V_{pp}, 500 Hz for 5 min). Panels (a-c) adapted from* [13]*, panels (d-f) adapted from* [14]*, with permission.*

3.3 – Considerations for iDEP Cell Design

In this work, we tested a bowtie/triangular design for our iDEP channels, similar to examples from the previous section [13,14]. A schematic of the basic system is shown in Figure 3.3a, where a rectangular channel (starting at ~1 mm wide) narrows down linearly to a central opening (~10 μm wide) over a relatively short distance (~100 μm) along the longitudinal channel axis. As explained above, when an electric potential is applied across

such a channel, its geometry has the effect of focusing the resultant electric field lines though this opening, creating a gradient of electric field strength near the constriction that can drive dielectrophoretic motion. A sensing element, in this case a nanopore-containing membrane, is then envisaged to be placed somewhere in the vicinity of the gradient-generating structure so that the DEP forces can deliver target molecules to within the capture radius of the pore at enhanced concentration.

While the overall layout of our iDEP channel is as depicted in Figure 3.3a, questions remain around the exact dimensions to choose for its key features (so as to optimize its target-concentrating power), as well as how best to assess the relative performance of alternate sets of dimensions. From Equation 3.1, above, the force on a polarizable molecule/particle is predicted to scale with the gradient of the electric field strength squared ($\nabla|E|^2$) and so, generally speaking, higher field gradients should lead to increased trapping capacities. Attempts have been made in the past by others to quantitatively translate the distribution of field gradients generated by particular layouts into estimates of the spatial extents of their trapping regions, usually by comparing the magnitudes of the resultant DEP forces to those of competing effects present in the system (e.g. from diffusion, other electrokinetic motion, etc.) [1,9,16]. For instance, the Edel group, working with gold-coated nanopipettes to concentrate 10-kbp dsDNA in 1 mM KCl, estimated a critical field gradient strength of $\nabla|E|^2 \approx 10^{16}\ V^2/m^3$ to overcome thermal motion under their conditions [1].

Figure 3.3b presents the result of a finite element simulation (COMSOL Multiphysics 5.1) of the $\nabla|E|^2$ distribution near the central constriction of our channel design for a particular set of dimensions (e.g. 1-mm main channel width, 30-µm constriction, 141 V applied across the electrodes), where the maximum of the colourbar range has been set to match the $10^{16}\ V^2/m^3$ limit of Edel & co. If the region, in red, where $\nabla|E|^2$ is above this threshold is taken to map directly to the trapping volume of the device, it predicts the capture of targets out to ~10 µm from the tips of the aperture. Figure 3.3c-d, on the other hand, show a set of fluorescence images of a device constructed in PDMS with a similar channel geometry (further detailed in *Section 3.4 – Fluorescence Measurements*) and where a surrounding distribution of dsDNA (SYBR Green-stained λ-DNA in 0.5× TBE) is shown before (Figure 3.3c) and after (Figure 3.3d) the application of a 400 V_{pp}, 1 kHz sine wave potential

(141 V_{RMS}) across the constriction. Notably, the regions that increase in fluorescence intensity after the voltage is applied (corresponding to elevated DNA concentrations) are observed to: 1) be localized to the constriction tips, and 2) extend into the channel by distances on the order of ~10 µm, matching the general predictions of the FEM modelling. Although the conditions (e.g. DNA length, buffer conductivity) of the experiment in Figure 3.3c-d are not strictly identical to those in the derivation of the $\nabla|E|^2 \approx 10^{16}$ V^2/m^3 threshold, the agreement seen here of at least the broad distribution patterns found in simulations of the electric field and in experiments involving DNA concentration supports the use of $\nabla|E|^2$ as a predictor of DEP trapping.

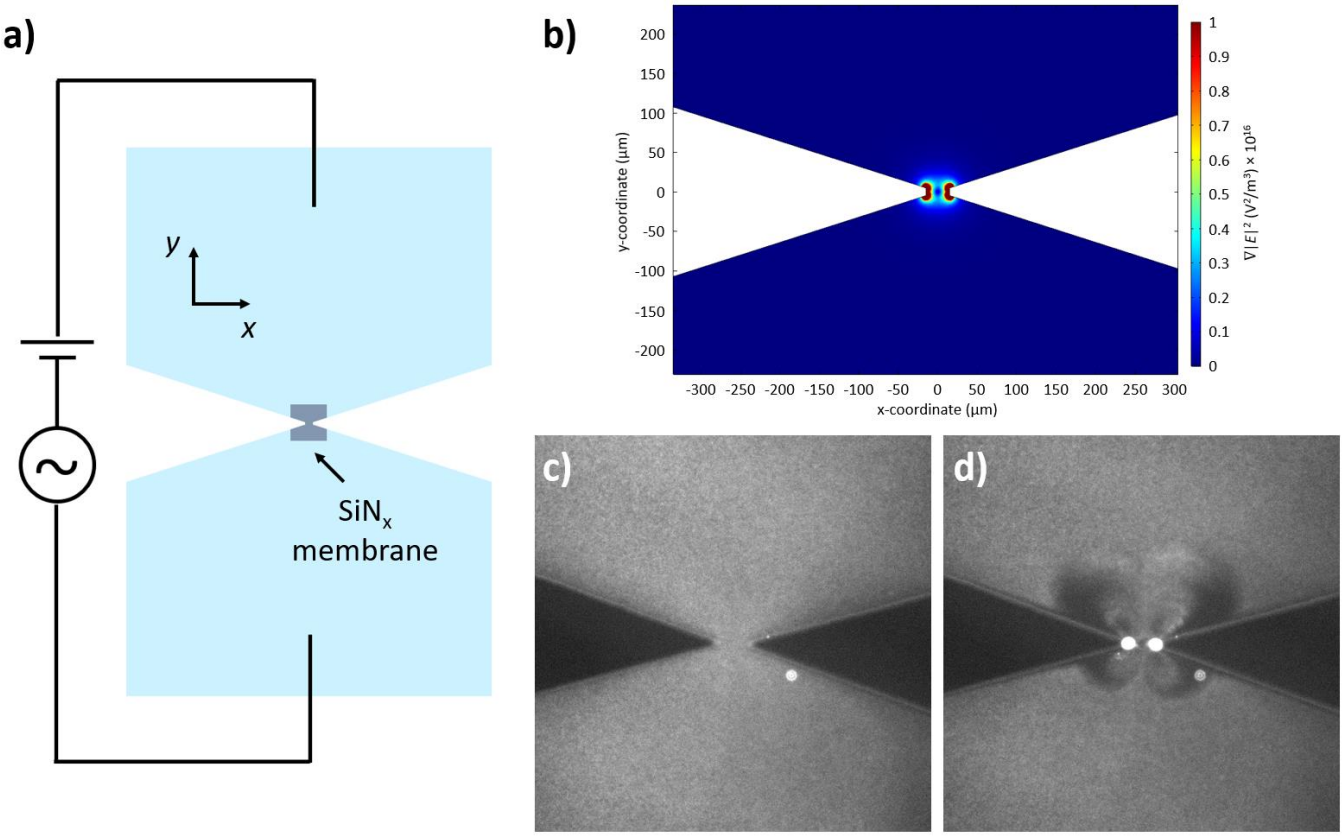


Figure 3.3: *Overview of the iDEP channel system.* ***a)*** *Schematic of an iDEP-nanopore concentrator device. The gap between the insulating walls of a microfluidic channel is narrowed over a relatively short (µm-scale) distance, focusing the electric field lines generated by an applied voltage across this constriction. The resulting electric field gradients draw polarizable targets to regions of highest field strength (dielectrophoresis), where a nanopore-containing membrane is placed to sample from the elevated target concentration.* ***b)*** *COMSOL simulation of the electric field gradient ($\nabla|E|^2$) near an insulating constriction (30 µm wide) after applying a potential difference (141 V) across distant electrodes situated up- and down-stream (1-mm main channel width).* ***c)*** *Image of the distribution of fluorescently labelled DNA (SYBR Green I stain, λ-DNA in 0.5× TBE) around a PDMS constriction matching the geometry of the simulation in (b), prior to the application of a voltage across the constriction.* ***d)*** *Image of the same device as in (c) after applying a 400 Vpp, 1 kHz AC voltage for ~1 min. The concentration of DNA is locally enhanced at the tips of the constriction, following the general pattern of the $\nabla|E|^2$ distribution in (b).*

Figure 3.4 illustrates the impact of varying the channel dimensions on the magnitude and distribution of $\nabla|E|^2$ in our system. The basic parameters that set the channel geometry here are outlined in Figure 3.4a and consist of: constriction length (ℓ), constriction width (w), constriction taper angle (ϕ), and channel width (w_{ch}). For each of these parameters, the electric field distribution around the constriction was simulated for a range of values (with the other parameters held fixed at intermediate values) and the magnitude of $\nabla|E|^2$ along a cut-line through the center of the channel ($x = 0$) was plotted, as in Figure 3.4b-e. From the plots, general trends emerge of increased peak $\nabla|E|^2$ magnitudes with: shorter and narrower constrictions, sharper taper angles, and wider main channel widths. Many of these effects can be understood from the perspective of the focusing of the electric field lines, e.g. by transitioning across greater extremes of field line density (increasing the ratio w_{ch} / w) or by shortening the length scale over which this transition occurs (reducing ϕ). Notably, not all parameters are observed to impact $\nabla|E|^2$ equally; doubling the channel width from 0.5 to 1 mm, for instance, can be seen to only increase the peak gradient magnitude by a factor of ~1.5, while, conversely, shrinking the constriction width in half from 4 to 2 μm grows $\nabla|E|^2$ by ~6×.

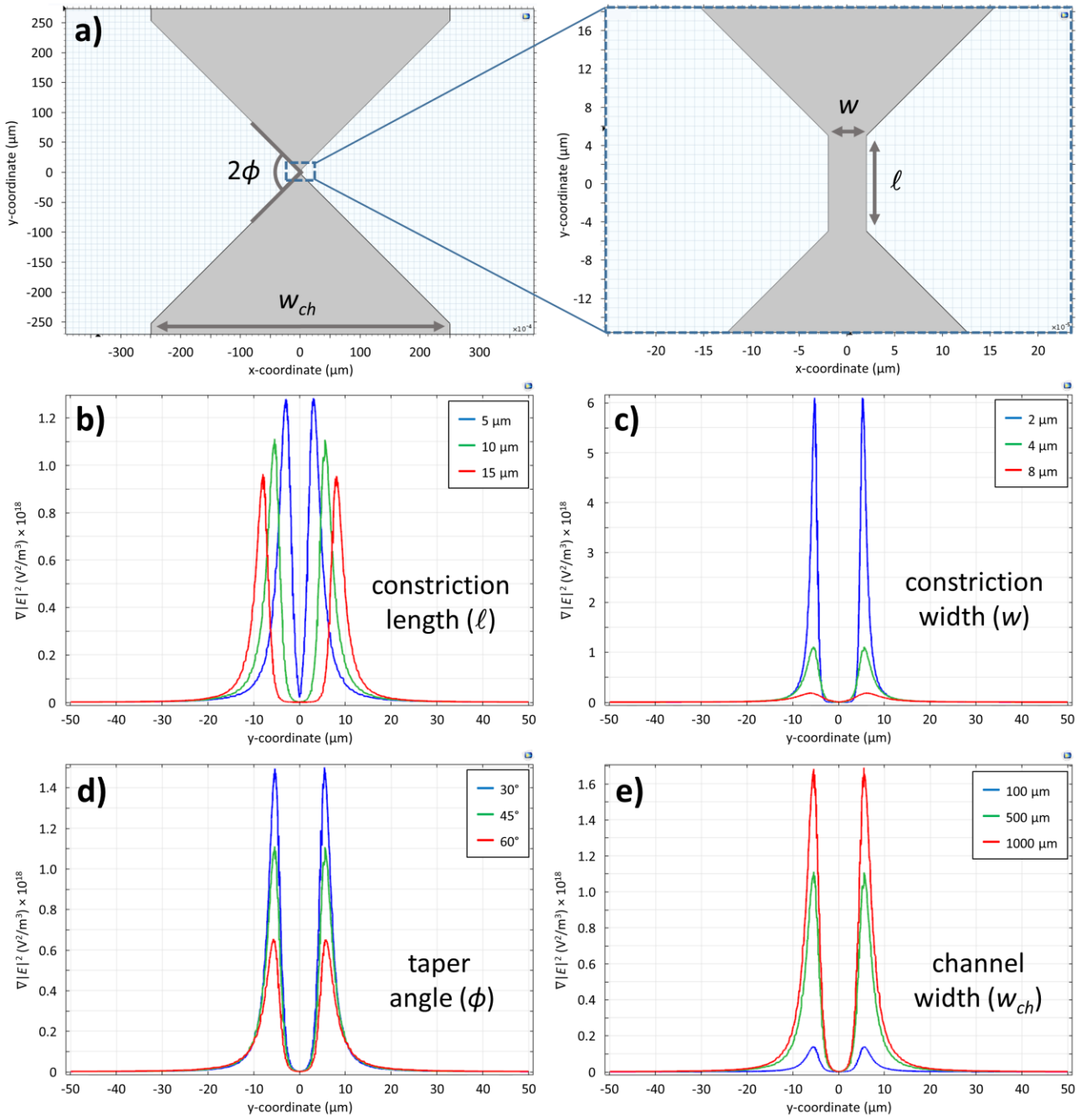


Figure 3.4: *COMSOL simulations of $\nabla|E|^2$ with changing channel geometry.* ***a)*** *Overview of geometric parameters that determine channel dimensions (ℓ – constriction length, w – constriction width, ϕ – constriction taper angle, w_{ch} – channel width). A potential difference of 200 V was applied across electrodes placed at the ends of the main channels.* ***b)*** *Magnitude of $\nabla|E|^2$ along cut-line (x = 0, y $\in$ [± 50 μm]) for multiple values of constriction length (w = 4 μm, ϕ = 45°, w_{ch} = 500 μm).*
c) *Magnitude of $\nabla|E|^2$ along cut-line for multiple values of constriction width (ℓ = 10 μm, ϕ = 45°, w_{ch} = 500 μm).*
d) *Magnitude of $\nabla|E|^2$ along cut-line for multiple values of constriction taper angle (w = 4 μm, ℓ = 10 μm, w_{ch} = 500 μm).*
e) *Magnitude of $\nabla|E|^2$ along cut-line for multiple values of channel width (w = 4 μm, ℓ = 10 μm, ϕ = 45°).*

From the results of Figure 3.4, selecting a channel design with optimized target-concentrating power involves shifting the values of each of the parameters that determine its geometry toward their extremes (e.g. smallest constrictions, narrowest taper angles, widest channels). In each case, however, there are practical considerations that arise during the assembly of physical devices that prevent following these guidelines to their limits. For instance, the photolithography methods employed in this work to create master moulds for the PDMS channels (outlined in *Section 3.4 – Fluorescence Measurements*) place constraints on the minimum feature sizes of these moulds; most importantly for iDEP channels, this limits how small the dimensions of the insulating constrictions (w, ℓ) can be made. Factors such as the resolution of the design on the photomask, the height of the gap between the mask and the photoresist during exposure, and, ultimately, the wavelength of the exposure light itself all combine to set the size of the smallest features reliably achieved by these methods [17]. Furthermore, each step toward moving to even smaller minimum feature sizes is often associated with rapidly increasing costs and complexities of the microfabrication techniques involved in doing so [18–21].

Similarly, there are practical upper limits to the width of the main channels feeding the constrictions (w_{ch}). On a basic level, making the channels too wide will lower the efficiency with which the entire channel volume can: 1) be wetted with solution initially (no air pockets) and 2) have its solution replaced with a one of a different composition as particular applications demand (no dead volumes). When combined with the flexible nature of the PDMS elastomer composing its walls, an excessively wide channel can also lead to channel collapse if a critical cross-sectional aspect ratio (of ~10:1, width × height) is exceeded. The height of the channel can of course be increased in concert with the width in order to maintain a fixed aspect ratio, but this can present downsides of its own (e.g. poor heat dissipation, a low fraction of the channel volume being sampled by the nanopore), as will be discussed in a later section (*Section 3.6 – Outstanding Challenges*). A satisfactory compromise implemented here has been to add interior posts within the channel design (e.g. as in the upper left corners of Figure 3.5a-b). This serves to support the "roof" of wide and shallow channels against collapse and can facilitate complete channel wetting, but at the cost of increased channel resistance (compared to the

same design without the posts) leading to slightly less of the voltage drop/field gradient powering dielectrophoresis being localized across the constriction.

Finally, Figure 3.5 examines some constraints placed on how thin the gradient-generating constrictions of our channels can be made, as viewed through the parametrization of taper angle (ϕ). One interfering effect is that photolithography tends to produce features in the developed photoresist that lack perfectly vertical sidewalls, e.g. with the "bases" of structures (attached to the substrate) being expanded relative to their "tops" (in contact with the photomask during exposure) [22]. For our purposes of fabricating moulds that are used to cast PDMS negatives, this would translate into channel cavities that widen from top to bottom. Figure 3.5a-b illustrates this effect with optical images of two iDEP channels of different taper angles (2ϕ = 30° and 2ϕ = 10°). In these images, the regions of darker contrast are sections of the "bottom" of the PDMS casts that are plasma-bonded to a glass slide, while the lighter (slightly-out of focus) outlines surrounding them correspond to the remainder of their PDMS structures that are found closer to the "roof" of the channel. From the observed distance of ~5 μm that the edges of each feature recede by from top to bottom, together with the ~50 μm height of the channels, we estimate a sidewall profile of ~84° (compared to a perfectly vertical 90°).

Pertinently to our discussion of taper angle, these sloped sidewalls can be seen to have a greater (negative) impact on the overall design fidelity of the 10° constriction than on the 30° constriction. This is best encapsulated by the noticeably wider opening at the bottom in the former case (compare Figure 3.5a and Figure 3.5b), despite both devices being designed with identical aperture widths (of 10 μm) on their photomasks. This phenomenon can easily be understood from simple geometric considerations as seen in Figure 3.5c, where uniformly shrinking two sets of triangles with different taper angles by a fixed distance s results in a larger gap Δx appearing between the narrower (smaller ϕ) triangles: $\Delta x = 2s \cdot \cos\phi \,/\, \tan\phi$. From the simulations of Figure 3.4, however, we note that increasing the overall $\nabla|E|^2$ strength (and thus DEP concentration power) of a channel design appears to depend much more strongly on decreasing the width of the aperture (in this case $w \approx \Delta x$) than on decreasing its taper angle ϕ. Therefore attempts to enhance the field gradients by thinning the constrictions are expected to quickly run into diminishing returns and may even work

against the intended effect (through the expansion of w) beyond certain intermediate values of ϕ. Making the constrictions too thin can also have a negative impact on the mechanical strength and stability of these flexible PDMS structures, as illustrated in Figure 3.5d which shows an image of a delicate 10° constriction that was damaged at some stage of device assembly.

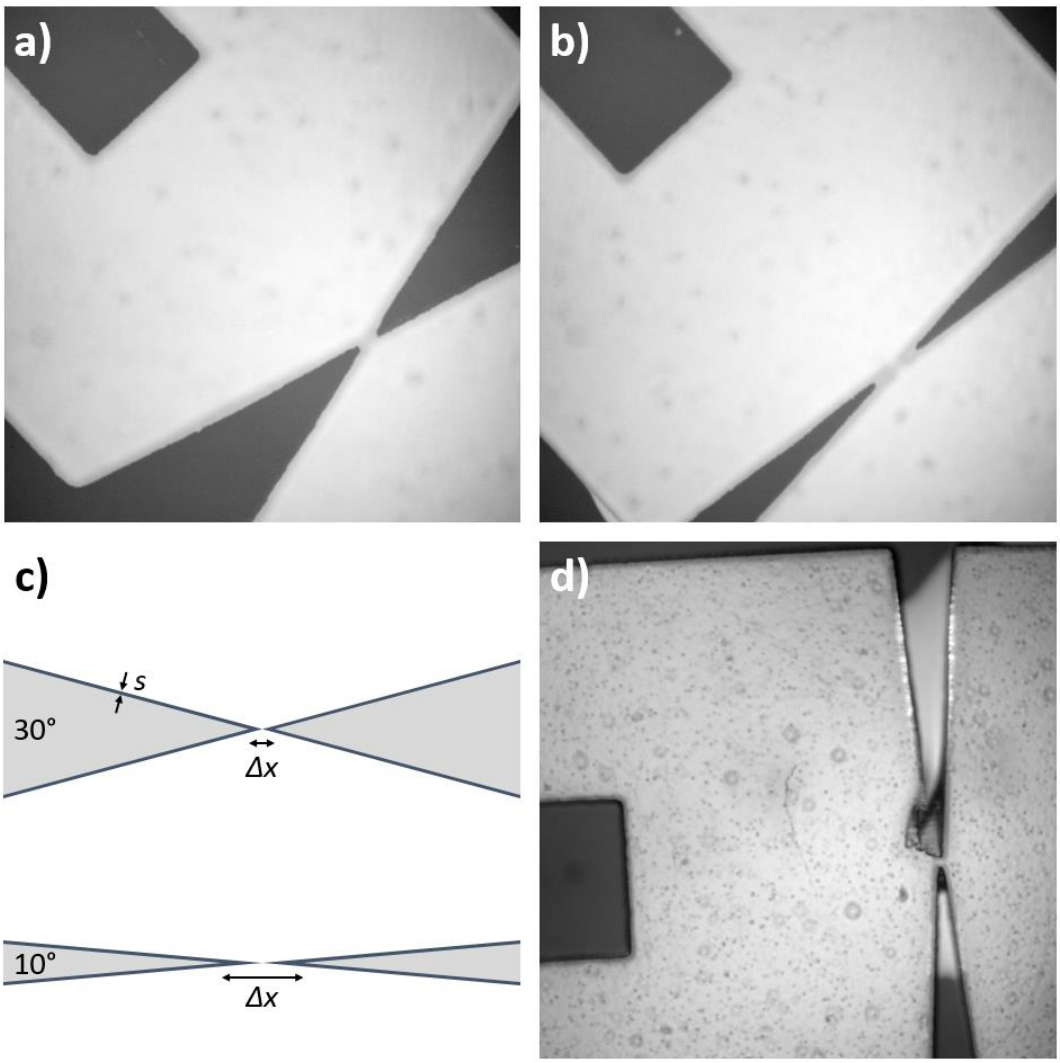


Figure 3.5: *Considerations for the tapering angle of PDMS constrictions during the design and assembly of iDEP devices.* ***a)*** *Image of a PDMS channel featuring a 30° constriction taper (ϕ = 15°, as parametrized in Figure 3.4) plasma-bonded to a glass slide.* ***b)*** *Image of a PDMS channel featuring a 10° constriction taper (ϕ = 5°) plasma-bonded to a glass slide.* ***c)*** *Illustration of the effect of a uniform shrink factor (fixed width s) on the overall dimensions of triangular constrictions. The gap size Δx between the two triangles comprising a constriction is predicted to increase more for those with narrower taper angles (Δx = 2s ·cosϕ / tanϕ) under such an effect.* ***d)*** *Image of a PDMS iDEP channel that suffered from mechanical deformation / damage of its thin insulating constriction (10° taper) during some stage of device assembly.*

In summary, the exact geometry of the insulating channels used in a DEP concentrator device plays an important role in determining its effectiveness, as quantified through the strength and distribution of the electric field gradients ($\nabla|E|^2$) it can generate. As we saw, however, tailoring the channel dimensions toward values that optimize $\nabla|E|^2$ must be balanced against practical considerations that arise during the design and fabrication of real devices. In subsequent sections, we make use of iDEP channels that were designed around these principles, and the specific dimensions selected for relevant devices will be detailed as they are presented throughout the chapter.

3.4 – Fluorescence Measurements

Initially, fluorescence microscopy was used to test the ability of the iDEP channels of the previous section to manipulate labelled targets. The basic protocol to build a device for this purpose starts with casting the channel designs into a slab of PDMS using positive moulds. The moulds are produced on 4-inch silicon wafers by UV photolithography with SU-8, a negative photoresist, using either emulsion or chrome photomasks containing the channel patterns. Next, inlets and outlets are created in the slabs by punching out cylindrical cores at the ends of the mould-defined channels; these inlets/outlets connect to fluid reservoirs made out of a second PDMS layer plasma bonded on top (see *Section 3.5*, later, for details). Finally, a glass slide is bonded on the bottom of the device to create the floor of the channel, and it is though this glass slide that the fluorescent targets are imaged using an inverted microscope (Nikon Ti-E).

To set up a fluorescence experiment, the channels of the assembled fluidic cell are first coated with PEG silane in an attempt to reduce any electroosmotic flows (EOF) that could compete with the desired DEP interactions on the targets (arising from counterions to the charged glass/PDMS surfaces moving under an applied electric field) [23]. For this, the channels are initially exposed to air plasma (e.g. 30s at 45 W) in order to generate Si-OH (silanol) groups at their surfaces; when a PEG silane solution (e.g. 3 mM SIL-PEG_8 in 10 mM HCl) is subsequently flushed in and incubated in the channels, the PEG silanes are able to covalently attach to the surfaces via reactions with these groups [24]. After PEG treatment, the channels are generously flushed with deionized

water and the fluorescent sample is then introduced, suspended in its imaging buffer, typically 0.5× Tris/Borate/EDTA (TBE). Gold electrodes are placed in reservoirs connecting to opposite sides of the constriction, allowing signals from a voltage amplifier (A400, FLC Electronics) to be transmitted into the system.

The initial targets tested in our system were fluorescent polystyrene beads (Dragon Green fluorophore, Bangs Laboratories), chosen due to their availability in multiple diameters (to investigate the effect of particle size on DEP force), and the bright, photobleaching-resistant signals of their internal dyes. One complication of working with polystyrene beads is that their specific mass ($\sim$1.05 g/cm^3) often causes them to sink to the bottom of channels filled with dilute aqueous solutions (density $\sim$0.997 g/cm^3), and so to bring the beads closer to neutral buoyancy, glycerol (density $\sim$1.26 g/cm^3) is added in a 1:4 ratio to our 0.5× TBE buffer.

Figure 3.6 illustrates flow patterns through the iDEP constrictions for multiple bead sizes and applied DC voltages. When large ($\gtrsim$ 100 V) voltages were applied to 2.1 μm beads, for instance, their distribution in the channel changed from being uniform (no voltage, Figure 3.6a) to being focused into a narrower stream by the fields near the constrictions (voltage applied, Figure 3.6b). This is referred to as "streaming DEP" and occurs when the DEP forces in the system are large enough to overcome particle diffusion but not large enough to overcome the electrokinetic forces driving the particles through the channel [25]. Larger 7.3 μm particles also experienced streaming DEP, this time even at moderate voltages ($\gtrsim$ 10 V, Figure 3.6c). However, in the case of the 7.3 μm beads, when the voltage was further increased to levels used to observe streaming with the 2.1 μm beads, a new pattern emerged of beads migrating up to the edges of the constriction (by EP) and becoming immobilized there (Figure 3.6d). When DEP forces are large enough to overcome both diffusion and electrophoresis in this way, it is referred to as "trapping DEP" [25]. Interestingly, individual 7.3 um beads that were trapped at the constriction tended to organize themselves into groups of long chains aligned with the electric field (Figure 3.6d). This phenomenon has been studied before (as "pearl chaining") and is believed to result from an attractive force between polarized particles due to the local perturbations in the electric field they create when surrounded by a medium of different polarizability [26].

The overall trends of Figure 3.6 serve to illustrate the strong dependence of F_{DEP} on particle size, as predicted by Equation 3.1. Despite the modest ~3.5× increase in diameter between the two bead sizes (2.1 μm and 7.3 μm) tested here, the factor of r^3 that appears in Eq. 3.1 predicts DEP forces that are ~44× times greater on the larger particles (with all other factors kept equal). This is supported by both the smaller voltages (and thus field gradients $\nabla|E|^2$) that were needed to generate a similar streaming DEP effect on the larger 7.3 μm beads, as well as the additional manipulation mode (trapping DEP) available to these beads within the given voltage range.

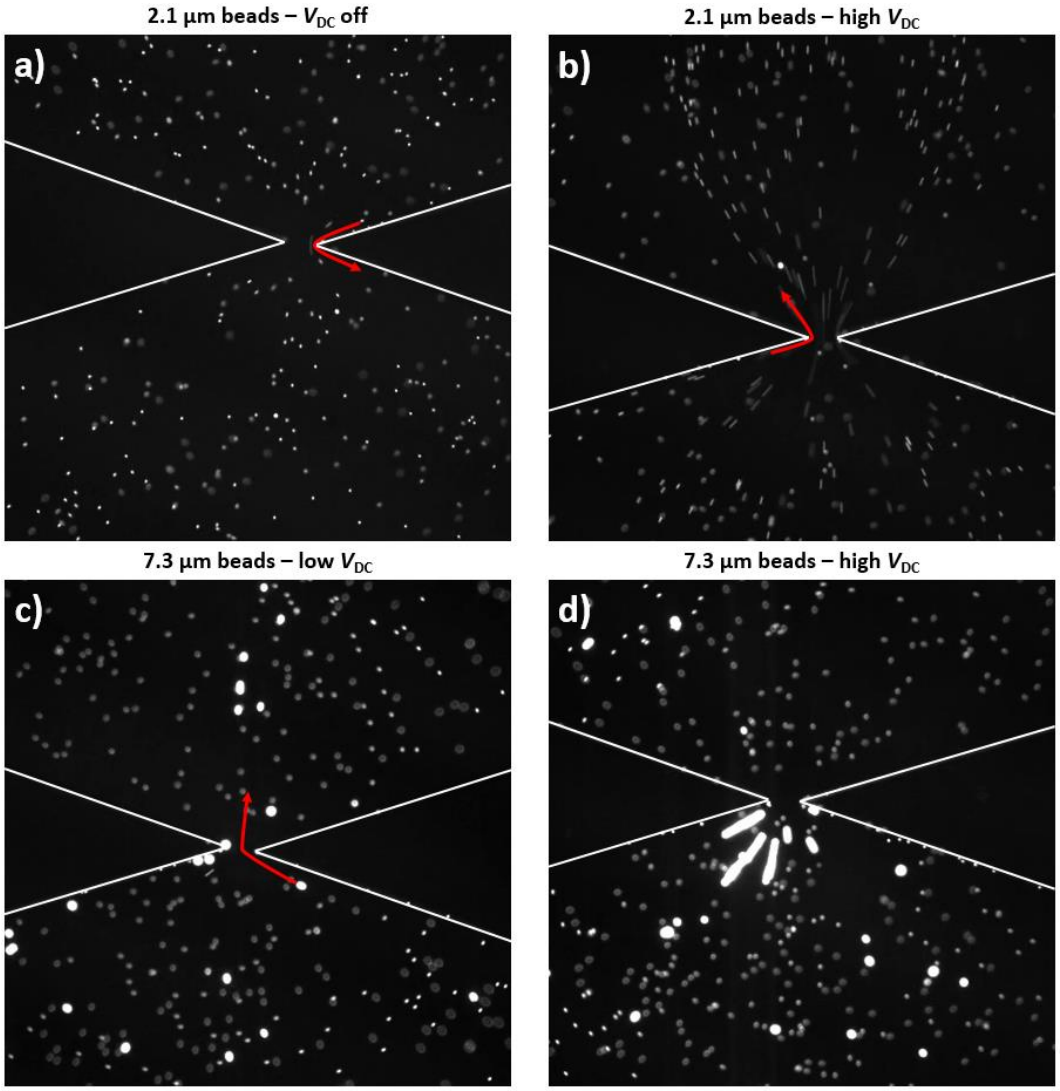


Figure 3.6: *Manipulating polystyrene beads with iDEP channels.* ***a)*** *Fluorescence image of 2.1 µm beads migrating through a PDMS constriction (35° angle, ~30 µm opening) in a 1:4 mixture of glycerol and 0.5× TBE buffer. The dark triangular constriction has been outlined in white. Here, no voltage is applied across the channel and beads move through the constriction carried by a weak background flow (likely arising from slight height imbalances in the fluid reservoirs). The red arrow traces the path of an individual particle along a streamline in the channel.* ***b)*** *Migration of 2.1 µm beads under a large applied potential (≳ 100 V) driving electrophoresis. As the beads pass through the constriction (a region of high DEP forces), their paths are focused into a narrower stream. The red arrow traces the path of a sample particle that starts near the outside edge of the channel (before the constriction) and is deflected toward the centre (after the constriction).* ***c)*** *Migration of 7.3 µm beads under a lower applied potential (≳ 10 V), showing a similar "streaming DEP" effect as in (b). (Sedimented 2.1 µm beads from a prior experiment are also visible out of focus on the channel floor).* ***d)*** *Migration of 7.3 µm beads under a large applied potential (≳ 100 V). Here the beads are trapped before the constriction, aligning themselves in groups of "pearl chains" oriented along the electric field lines.*

After the sensible results with the polystyrene beads, we next moved on to attempts to manipulate our presumptive nanopore sensing target, double-stranded DNA. For these experiments, DNA samples were labelled with SYBR Green I intercalating dye, loaded into the cell in 0.5× TBE buffer, and observed under various applied AC voltages. The typical effect is illustrated in Figure 3.7a (see also Figure 3.3d) where, upon application of the

voltage, DNA becomes trapped at elevated concentrations in the two regions of the constriction where the largest electric fields are generated (here using a 400 V_{pp}, 1 kHz sine wave with λ-DNA, 48.5 kbp). Shutting off the voltage with the system in this state results in the DNA being released from the trap, migrating off with weak background flows (e.g. from slight height imbalances in the reservoirs) as a diffusing cloud of fluorescent particles (Figure 3.7b). From comparing the fluorescence intensities of the concentrated regions with those of the channel bulk (both background-subtracted), peak concentration factors of ~200× are estimated (averaging ~60× within the entire ~30-um concentrated regions), and these are achieved ≲ 1 min from the onset of the voltage.

As with the polystyrene beads, however, a strong dependence of trapping efficiency on target size was observed. Figure 3.7c shows a typical fluorescence image of shorter dsDNA fragments (here 7 kbp) being concentrated by our system. In the case of the 7 kbp molecules, smaller peak concentration factors were recorded (of ~100× channel bulk) when compared to λ-DNA (48.5 kbp) and these required longer wait times (≲ 5 min) under the applied voltage to reach. As discussed in *Section 3.2*, longer DNA molecules are expected to be more polarizable (creating greater charge separation under an external electric field) and should thus experience stronger concentrating DEP forces in a given (non-uniform) spatial distribution of field strength. Additionally, we note that the frequency response of DNA molecules under an AC field is also expected to depend on DNA length [10,27]. This means that if trapping parameters like AC frequency are optimized for a particular target molecule, the system is unlikely to remain optimized when moving to a new target of markedly different size, shape, charge distribution, etc. (or when trapping mixtures of these molecules). This effect could also be used advantageously if, for instance, only one species within a mixture represented a desired nanopore target (e.g. a fully assembled DNA nanostructure mixed with shorter contaminants) by setting trapping conditions that overwhelmingly favour its selective capture over the others.

Finally, we highlight the importance of the step implemented here to suppress electroosmotic flows (EOF) within the channels using a surface coating. Up to this point, all of the fluorescence results presented have been in fluidic cells whose channels were treated with a PEG silane (6-9 PEG repeats, "SIL-PEG_8"), as described above.

PEG, or polyethylene glycol, is an extended hydrophilic polymer and is thought to work at quenching EOF in microfluidic systems by bridging across the channel counterions (Debye layer) and disrupting the viscous coupling between these mobile ions and the bulk solution [28]. Figure 3.7d shows the effect of a strong AC voltage (1 kHz, 400 V_{pp}) on the distribution of λ-DNA in an iDEP channel that had not been treated with PEG. In this case, no significant pre-concentration was observed at the tips of the constriction, although a similar streaming effect that of to the polystyrene beads is visible where DNA is deflected from near the walls by forces at the constriction's focus. When this same cell was subsequently treated with SIL-PEG_8, DNA concentration patterns (akin to those of Figure 3.3d and Figure 3.7a,c) were restored (not shown), presumably due to the suppression of some competing hydrodynamic effects in the system. Despite this main reversal, it is not claimed that the coating molecule or treatment protocol used here has been fully optimized. In fact, the presence of swirling vortices at the constriction during DNA concentration (see Figure 3.3d and Figure 3.7a) may point to residual hydrodynamic forces that remain significant [29,30]. In prior work by others, various alternative EOF-suppressing coatings have been tested in DEP channels including POP-6 [23,27], triblock-copolymers [24,31], and other PEG-silanes [24], and an improvement on our protocol likely exists among these options or elsewhere.

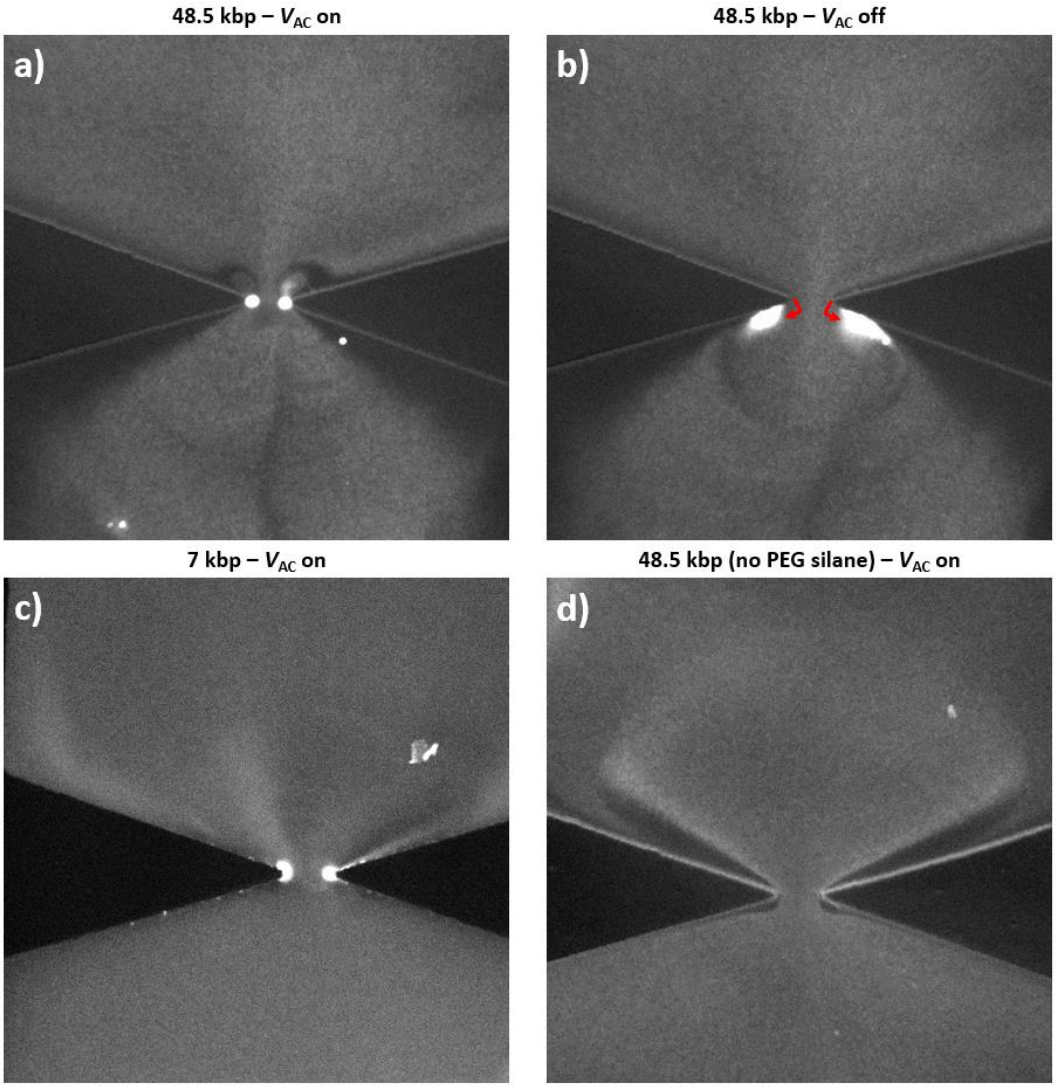


Figure 3.7: *Capturing double-stranded DNA with iDEP channels.* ***a)*** *Concentration of SYBR Green I-stained λ-DNA in 0.5× TBE at a PDMS constriction (35° angle, ~30 μm opening) under an applied trans-constriction AC voltage (400 Vpp, 1 kHz, applied for $\lesssim$1 min.).* ***b)*** *Distribution of λ-DNA moments after shutting off the voltage in (a). The concentrated DNA formations at the tips of the constriction are released downstream of a weak background flow (likely due to slight height imbalances in the fluid reservoirs) and begin to dissipate their augmented concentration by diffusion.* ***c)*** *Concentration of SYBR Green I-stained 7-kbp DNA (in 0.5× TBE) under an applied AC voltage.* ***d)*** *Distribution of SYBR Green I-stained λ-DNA in 0.5× TBE under an applied AC voltage (400 Vpp, 1 kHz) in a channel that had not been treated with PEG-silane. DNA did not significantly accumulate at the constriction tips as in (a) and (c) (see also Figure 3.3d), even after several minutes.*

In summary, using fluorescence microscopy, our iDEP channel designs were shown to be capable of manipulating target particles by applying elevated (± 200 V) voltages across reservoirs at the channel ends, far away from their main insulating features. We also saw evidence of the expected dependence of trapping efficiency on target size (polystyrene beads and dsDNA), which may have important implications for optimized

target selection in a combined iDEP-nanopore device (e.g., longer DNA molecules are more polarizable, but complicate nanopore sensing via additional molecular folding and pore clogging).

3.5 – Integrating Nanopore Membrane into Fluidic Device

Having validated the above iDEP channel designs in PDMS as being capable of concentrating targets (polystyrene beads, dsDNA) near their narrowest features, the next task was to incorporate such a channel as part of a nanopore-sensing device. Figure 3.8 presents an overview of the approach to this adopted here.

The overall microfluidic cell design is illustrated in Figure 3.8a and is based around five main channel layers. Two of the layers (L2 and L4) feature relatively short (~4 mm) channels that seal directly against either side of a 5×5 mm^2 SiN_x chip (L3), thereby providing direct fluidic connections to the membrane. One of these channels (L4) will be patterned with an iDEP geometry and so this layer must be carefully aligned with the underlying membrane during assembly (see also Figure 3.8c, detailed later) in order to have the eventual nanopore of the device sampling from channel regions that accumulate targets under DEP forces. Two additional layers of "access channels" (L1 and L5) then connect to these membrane-adjacent channels and serve to extend the fluidic access out to separately addressable fluid reservoirs. The lower panels of Figure 3.8a present the basic photomask designs used to develop the SU-8 moulds for each layer. The central (1 mm wide) transparent regions of the masks are responsible for creating the main channels, while the smaller surrounding elements are alignment markers that aid in correctly positioning neighbouring device layers on top of each other during assembly.

There are a few advantages to having the fluidic path between one side of the membrane and its reservoirs being split across two channel levels as in this design. For one, it allows the "inner" channel (L2/L4) to remain compact in its horizontal (*x*-*y*) dimensions, enabling the small, membrane-containing chip to serve directly as the channel "floor". An alternative approach, as implemented elsewhere [32,33], is to have an intermediate "microvia" layer underneath the channel that provides access to the membrane only through a single sub-mm hole. In addition to increasing the complexity of device assembly slightly (e.g. more stringent layer alignments,

difficulty in removing "films" over thin PDMS apertures), in our case a microvia layer also introduces a vertical gap (the height of the layer) between the membrane where the nanopore is found and the iDEP channel patterns that concentrate targets. These targets would then have to diffuse across this extra distance to reach the nanopore, negating some of the intended pre-concentration effect in the process.

A second advantage to a two-layer half-cell design is that it allows the height of each channel to be set independently. As will be discussed in *Section 3.6*, it is desirable for the channel with the iDEP constriction (L4, right above the membrane) to be relatively shallow (e.g. for improved heat dissipation, having a greater fraction of the channel volume sampled by the pore, etc). However, by designing the secondary access channels (L5) to be substantially taller than this, their electrical resistance is lowered relative to the constriction channel, thereby allowing for a greater proportion of the total voltage applied across the cell (at the inlets/outlets) to appear near the iDEP features. Figure 3.8d illustrates an example of this phenomenon with a simulation of the electric potential across the constriction for multiple channel designs. Each design differed in either the width of its constriction opening (w, see Figure 3.4) or in its number of channel layers. As expected, the designs with narrower constrictions (2 μm) have larger voltage transitions across this region relative to their counterparts with wider constrictions (10 μm); the slope of the voltage transition is of course the electric field strength along this direction ($E_y = -\partial V/\partial y$), which plays an important role in determining the DEP forces a target will experience when passing through the channel (see Eq. 3.1). Most notably, however, the designs that feature two channel layers (a 25-μm constriction channel and 100-μm access channels) are observed with substantially higher voltage drops near their constrictions than their one-layer counterparts of equal w (e.g. 62% vs. 34% of the total 1 kV applied, for w = 10 μm). This is due to the voltage divider effect from the lower resistance access channels, and will translate into higher trapping forces in a realized device.

Briefly, the device assembly process consists of the following steps. Liquid PDMS is first poured over the moulds to create negative casts of the SU-8 patterns, where the total mass/volume of PDMS added sets the final layer height (targeted here at ~1 mm). Each mould consists of nine replicates (3 × 3 grid) of a single SU-8 pattern (those at the bottom of Figure 3.8a) developed on top of a 4-inch silicon wafer. Once the PDMS is cured (e.g.

overnight on a 40°C hotplate), the nine replicates of each layer are cut out as a 20 × 20 mm^2 squares using a scalpel. In order to create the necessary connections between channels from separate layers (otherwise isolated from each other by PDMS gaps), a 1-mm biopsy punch is used at this stage to produce vertical holes at relevant positions in the PDMS squares (see Figure 3.8a for an overview of the affected layers). Note that punch holes in layers L2 and L4 must bridge an air gap created by the height of the SiN_x chip to connect to each other, and so PDMS spacers that roughly match the 200 µm chip thickness are added in between. These are made by cutting out small squares from a thin layer of PDMS spin-coated onto a blank silicon wafer.

The individual PDMS layers that will make up a completed device are now ready to be bonded together. To do so, each layer first has its surface activated by air plasma (e.g. 60 s at 45 W RF power) and then neighboring layers are brought into contact to form covalent (Si-O-Si) links between their surfaces ("plasma bonding"). For all bonding steps but one (L3-to-L4) the relatively large (~mm) dimensions of the channels/punch holes make it so that alignment can be performed relatively easily by hand under a stereomicroscope. For the L3-to-L4 step specifically, since this involves positioning a ~10 µm wide constriction opening above a 40 × 40 μm^2 membrane, the micropositioning stage and microscope of a mask aligner system are used for more accurate placement (Figure 3.8c shows a sample view of the microscope camera during an alignment). Finally, in addition to the five carefully aligned, mould-cast layers presented in Figure 3.8a, two others are bonded by hand to complete a single device – a glass slide underneath to act as a stable support, and a top reservoir layer above the inlets/outlets constructed from punching wide (4-mm) holes in a blank PDMS slab. An image of a fully assembled iDEP-nanopore cell is shown in Figure 3.8b.

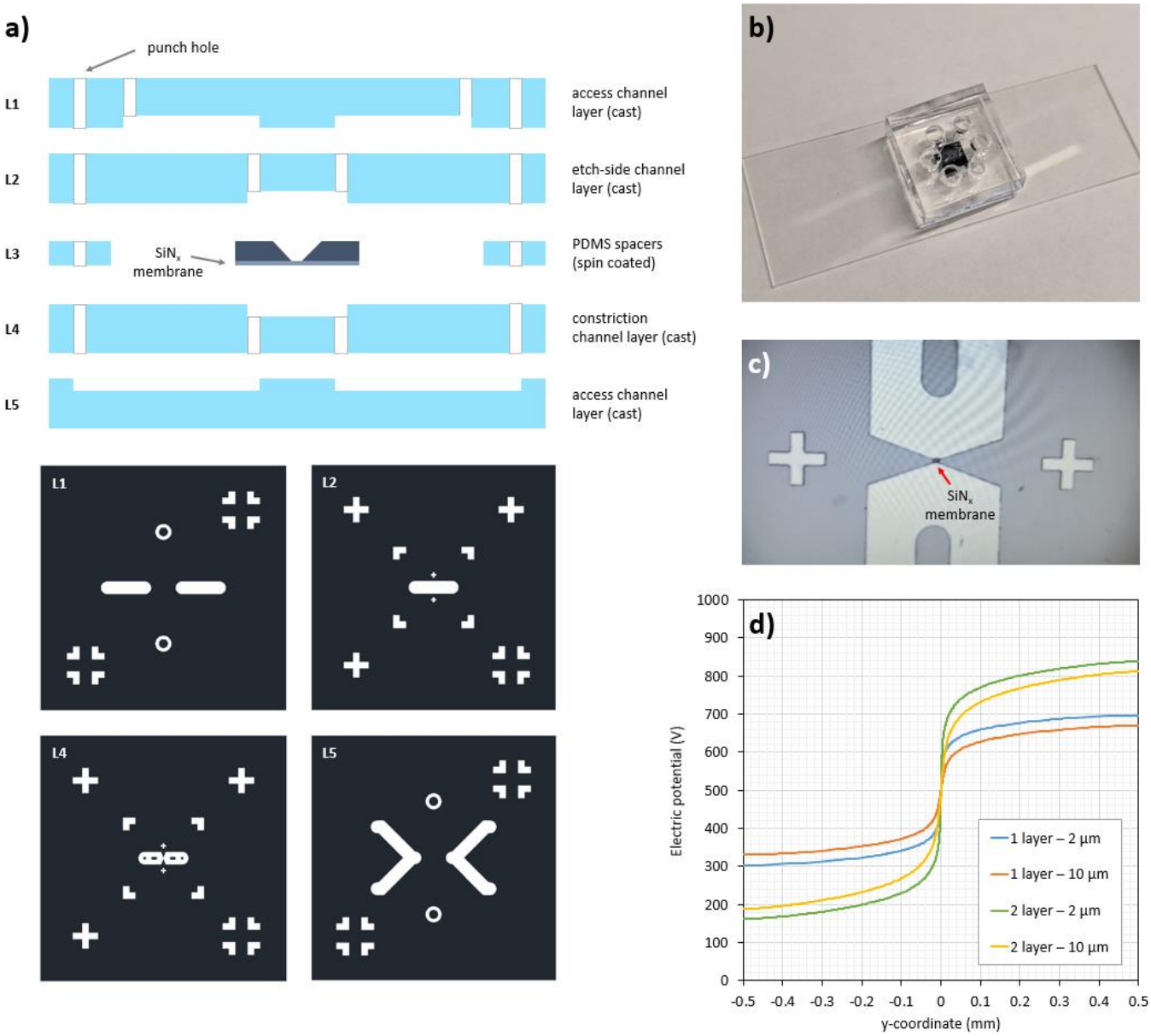


Figure 3.8: *Overview of the nanopore-iDEP PDMS fluidic cell design.* ***a)*** *Schematic diagram of the main layers comprising a cell. Each layer is cast in PDMS from its SU-8 mould and bonded to adjacent layer(s) by surface activation with air plasma. The bottom panels show the photomask designs used to develop the SU-8 moulds for each layer.* ***b)*** *Fully assembled cell showing the addition of a bonded reservoir layer overtop of the punched inlets/outlets, as well as a glass slide providing support as a base.* ***c)*** *Microscope image of a constriction channel layer (L4) bonded to a SiN_x chip (L3). The free-standing membrane of the chip is visible in the image as a darker square and has been aligned near the centre of the constriction opening.* ***d)*** *Simulation (COMSOL Multiphysics 5.1) of the electric potential across the insulating constriction (along the x=0 cut-line, see Figure 3.4) for multiple cell designs that differ in constriction width (2 or 10 µm) or in number of channel layers (1 or 2). The 2-layer design is as depicted in (a), with a 100 µm tall access channel (L5) leading to a 25 µm tall constriction channel (L4), while the 1-layer design reproduces this overall (x-y) fluidic path in a single layer (e.g. 100 µm tall). A 1 kV potential difference is applied across the inlets/outlets of the simulated cell.*

After assembly, a fluidic cell following the design of Figure 3.8 contains an intact SiN_x membrane embedded within layered PDMS channels. To now form a nanopore in this membrane, we use the controlled breakdown method (CBD), presented in detail in *Chapter 1*. In short, the channels are filled with electrolyte (e.g. 1 M KCl) and a potential difference (~10 V) is applied across the membrane (by biasing reservoirs connected to L1 and L5) until a large spike in current through the membrane is detected, signalling the formation of a nm-scale hole that connects both sides of the chip in a fluidic path.

Figure 3.9 highlights some example attempts at testing the suitability of nanopores fabricated by CBD in these PDMS cells for DNA sensing. After a pore is made, an estimate of its diameter can be extracted by measuring its conductance (*G*) in a solution of known conductivity and fitting this value to a mathematical model [34]. Figure 3.9a shows the current-voltage characteristics ("I-Vs") of two pores, fabricated in two slightly different cell designs: one matching the general layout of Figure 3.8a ("with constriction") and one where layers L4 and L5 have been replaced with duplicates of L2 and L1, respectively, so that both sides of the membrane face straight, featureless channels ("without constriction"). In either case, the I-Vs are quite linear, implying a symmetric flow of ions in both voltage polarities as well as pore shapes that are roughly cylindrical [35]. From the measured conductances (*G* = 13.8 nS and 16.2 nS, in 1.8 M LiCl) both pores also feature relatively similar estimated pore diameters of 4.6 nm and 5.0 nm for the cells with and without a constriction, respectively.

Once the nanopores had been fabricated and characterized electrically, a sample of purified dsDNA (15 kbp, NoLimits, ~1 nM) was added to the *cis* chambers (reservoirs connecting to L5) of both devices to confirm the presence of telltale DNA translocation signals upon application of a small transmembrane voltage (200 mV). Figure 3.9b presents a representative 30-s trace of the blockage current during such an experiment with each of the pores, and indeed, transient interruptions of the open-pore baseline current are visible as DNA molecules enter the pore and block a portion of the background flow of ions (Li^+ and Cl^- here). Also presented is a magnified view of a single translocation event through each of the pores (overlaid with fitted step-function sublevels in black) showing the "2-1" blockage sequence that is characteristic of folded DNA [36]. Finally, in Figure 3.9c, all of the events from the two experiments depicted in Figure 3.9b (~600 events each) are plotted

on axes of “maximum blockage” (the deepest fitted blockage level of each event) vs. “dwell time” (the total event duration). Similar distributions are observed in the scatterplot for both cell designs, with main populations that can be easily associated with unfolded ($\Delta I_{max} \approx 0.85$ nA) and folded ($\Delta I_{max} \approx 1.7$ nA) dsDNA translocations. The exact blockage depths do differ slightly between the two pores (most noticeably for folded events), which may be attributable to differences in their specific geometries, such as pore length [37].

As an interesting sidenote, we observed that SiN_x membranes mounted within the PDMS cell designs presented here produced pores that remained remarkably stable in size over time and over number of experimental trials. Figure 3.9d shows a plot of pore diameter over time for pores in five PDMS devices (featuring straight channels on either side of the membrane), where each point represents the result of an I-V curve, usually taken right before applying moderate voltages for extended periods (e.g. 30 min at 200 mV) as part of a DNA sensing experiment. Remarkably, over the span of weeks and with two different SiN_x chip designs (with and without a 60-nm oxide layer) growth remained low, with rates ~0.1 nm/day observed. This represents approximately an order of magnitude improvement when compared to controls with the same chips mounted in traditional, larger volume PEEK cells (~1 nm/day, data not shown). Here, we speculate that part of this improvement in stability comes from having channels that direct liquid flow tangentially over the membrane surface (Figure 3.8a) rather than pointed perpendicularly at it, as in some previous designs [38,39]. This may apply fewer stresses to the membrane in the form of dynamic pressure during solution exchanges. Probing this connection between channel design and pore stability further as part of a future investigation could reveal important considerations for the optimal design of cells produced by milling/3D printing and soft lithography alike.

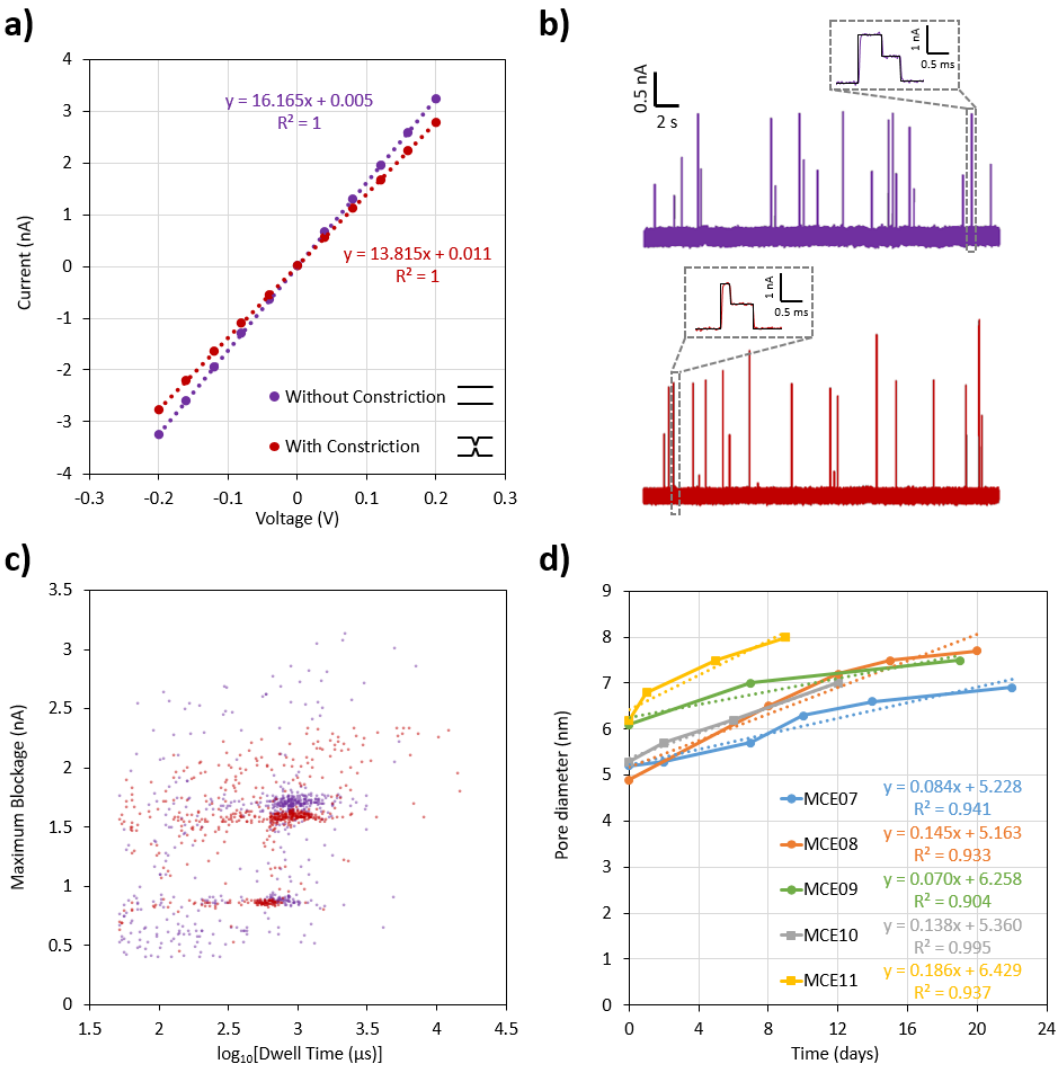


Figure 3.9: *Nanopore sensing in PDMS fluidic cells.* ***a)*** *Current-voltage characteristics of two CBD-fabricated pores (~5 nm in diameter) in 1.8 M LiCl. One of the pores (G = 16.2 nS ⇒ d = 5.0 nm) is mounted in a "symmetric" cell featuring a straight, 100 µm tall channel on both sides of the SiN_x chip, while the other (G = 13.8 nS ⇒ 4.6 nm) has one of these channels replaced with a constriction channel (50 µm tall) centred over the membrane.* ***b)*** *Sample current traces (30 s) of DNA translocations (15 kbp dsDNA, 200 mV, 1.8 M LiCl) through each of the pores in (a) – the colour code is carried over from that panel into (b) and (c). Insets show a magnified view of a representative (folded) event from each pore. Data collected with Axopatch 200B patch-clamp amplifer (hardware low-pass filtered at 10 kHz, digitally sampled at 500 kHz).* ***c)*** *Scatter plot of maximum blockage vs. log[dwell time] for all events from the translocation experiments in (b).* ***d)*** *Plot of pore growth over time for five different pores mounted in PDMS cells ("symmetric" straight channels). The devices plotted with square markers (MCE10-11) use slightly modified SiN_x chips featuring a 60-nm SiO_2 layer between the nitride film (outside of the membrane) and the silicon frame, in contrast with the other devices which lack this layer (MCE07-09, circle markers).*

In summary, microfluidic cells were designed and constructed out of PDMS that successfully integrated an iDEP concentration channel directly over a nanopore-compatible SiN_x membrane. Pores fabricated in these cells were found to behave similarly to those in more standard two-piece PEEK/Teflon designs (perhaps even with an increase in pore stability) and no appreciable downside to nanopore sensing was observed when placing a PDMS

constriction in close proximity to the membrane (compared to a straight, featureless channel placed farther away, see Figure 3.9). In the next section, we discuss some of the remaining obstacles to running complete experiments, involving sequential steps of DEP pre-concentration and nanopore sensing, using our design.

3.6 – Outstanding Challenges

3.6.1 – DNA Aggregation

Given the direct relationship between the DEP force experienced by a particle and the magnitude of the electric field gradient ($\nabla|E|^2$, see Equation 3.1) as well as the strong dependence observed of the size of the generated gradients on the width of the constriction openings in our system (w, see Figure 3.4c), being able to reduce the constriction widths should result in tangible benefits to particle concentration, either through increased target concentration factors (CFs) or lessening the voltage requirements for a given CF. Figure 3.10a shows a PDMS cast of an iDEP channel pattern from an SU-8 mould where attempts have been made to optimize for constriction width. Namely, the emulsion photomask used to transfer iDEP patterns onto the photoresist was replaced with a higher-resolution chrome mask (featuring w = 2 µm), and the gap between mask and photoresist during the UV exposure step was reduced, both by application of pressurized N_2 underneath the photoresist substrate ("hard contact") and by prior removal of the photoresist's "edge bead" after spin coating, in order to limit the expansion of these patterns during wafer development.

When these narrow constriction iDEP cells were subsequently tested with fluorescent DNA, superficially similar target concentration effects (Figure 3.10b-c) were observed at their high-field regions as compared to previous results in wider-w channels (see Figure 3.3d, Figure 3.7a). However, as estimated from fluorescence counts, noticeably higher CFs were achieved over similar < 1 min timescales in the present case (peak values $\gtrsim$ 1500×, average values $\gtrsim$ 400× inside the concentrated region). Additionally, rather than two separate regions of increased DNA concentration existing at each constriction tip like before, the closer tip proximity and overall larger CFs involved here resulted in a single overlapped region at the design's centre (Figure 3.10c). Having a single concentration region like this could be beneficial when integrating a nanopore sensor underneath the

trap, both by making the pore easier to predictably align within this region and by not lowering the efficiency of the device as a whole by having a secondary concentration region lying outside the pore's capture radius.

Perhaps the largest difference in performance between narrower and wider constrictions, however, was observed when the AC fields generating the DEP forces were removed. Instead of the concentrated DNA releasing as a freely-dispersing "cloud" as before (see Figure 3.7b), DNA exposed to the higher E-fields of the narrower constrictions drifted off as a single aggregated "ball" (Figure 3.10d). Such (semi-)permanent DNA aggregates are unlikely to be compatible with immediate nanopore sensing, as individual, molecule-by-molecule passage through the pore is required to read out the discrete information in their translocation signals, or to use the count of individual molecules per unit time as a proxy measure of target concentration. This is without mentioning the possibility of the pore also instantly and permanently clogging if subjected to a large DNA aggregate blob that rivals or exceeds its diameter in size. It should be noted that intercalating dyes (such as SYBR Green) are known to alter the physical properties of dsDNA relative to its native state [40], and thus could also be affecting the overall level of repulsion/attraction between molecules under these conditions.

Previous studies have similarly observed the formation of DNA aggregates in DEP concentration systems that expose molecules to high electric field strengths [1,41]. Notably, an article by Gallo-Villanueva *et al.* reported on the use of particular DNA "stabilizers" (1.5% Tween 20 and 1.5 M triethanolamine) to suppress permanent molecule aggregation at the high-field regions of an iDEP channel [41]. Validation of similar DNA stabilizer techniques in our own system could be a worthy avenue of future investigation, as this might enable us to retain the benefits of narrower iDEP constriction widths (larger E-field gradients) while suppressing one of the major downsides that come with this. We note however, the additional requirement that exists in our case that any solution additives incorporated into an experimental protocol be compatible with the later nanopore sensing stage. For instance, Tween 20 is known to have a "critical micelle concentration" above which additional molecules begin to organize themselves into spherical aggregates (micelles); these Tween micelles have been known to interfere with nanopore sensing and analysis through the additional translocation events they generate and through their destabilizing effects on the open-pore baseline [42]. A careful balance of

concentration considerations (DNA stabilization vs. micelle formation) would thus be necessary when attempting to introduce such a molecule into our system.

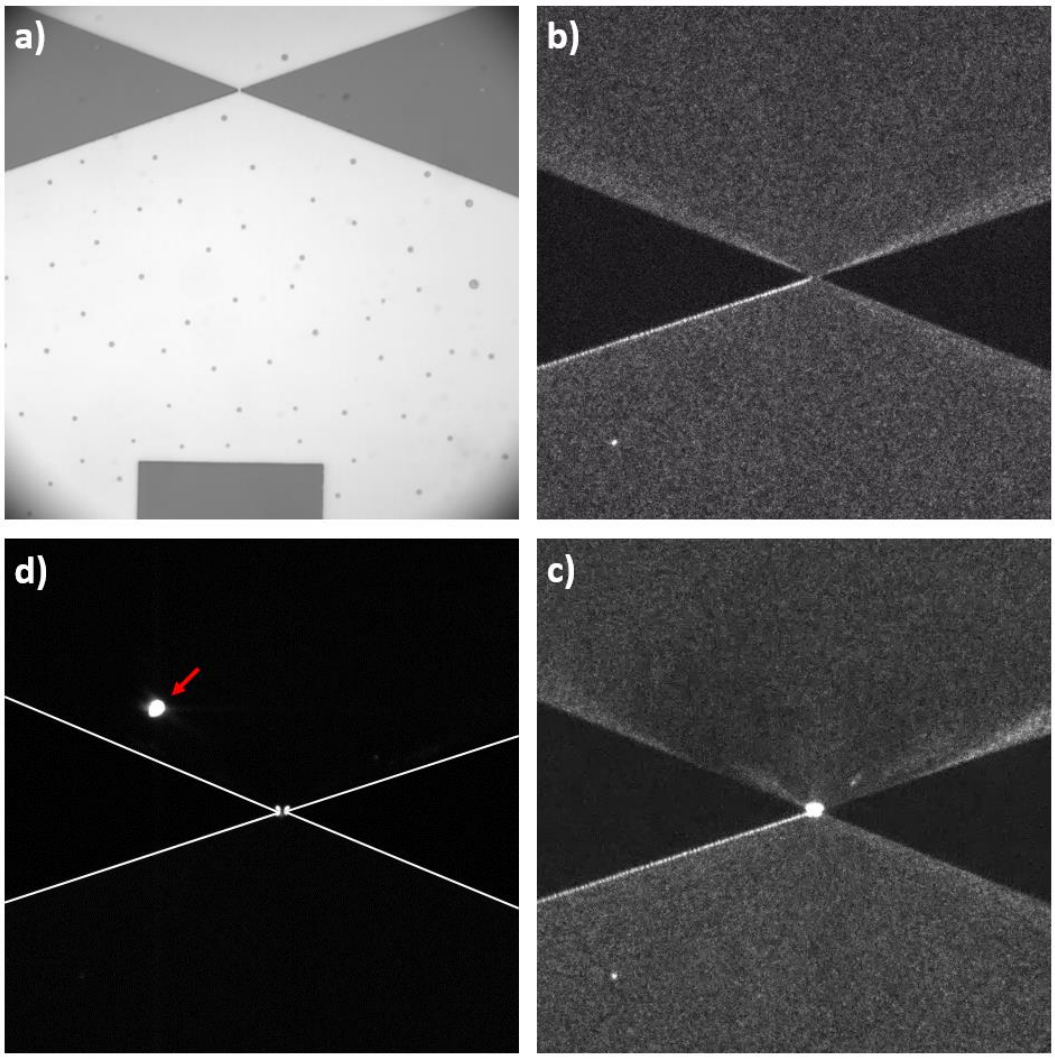


Figure 3.10: *DNA capture in iDEP cells with reduced constriction widths.* ***a)*** *Optical image of a PDMS iDEP pattern cast from a chrome mask-developed master mould (constriction width $\lesssim 10$ µm, taper angles $2\phi = 40°$). The darker regions of the channel (triangular constrictions, rectangular channel post) represent areas where the PDMS cast has been plasma-bonded to a glass coverslip.* ***b & c)*** *Fluorescence images of SYBR Green-stained λ-DNA in 0.5× TBE surrounding a narrow ($\lesssim 10$ µm) iDEP constriction before (b) and ~30 s into (c) the application of an AC potential (1 kHz, ~100 V_{pp}) across the constriction.* ***d)*** *Fluorescence image of a narrow iDEP constriction in between cycles of applied voltage. Newly-concentrated λ-DNA is visible at the tips of the constriction (triangular contours outlined in white) while DNA that had been accumulated in the previous cycle and subsequently released from the iDEP trap at voltage shutoff is visible in the image as an aggregated "ball" (at the position of the red arrow).*

3.6.2 – iDEP Concentration in High Salt

So far, our iDEP experiments have all made use of a standard, low conductivity electrophoresis buffer (0.5× TBE, $\sigma \sim 0.1$ mS/cm) as the medium surrounding our various polarizable molecules. As will be detailed in the next

subsection however ("*Nanopore Sensing in Low Salt*"), low-σ solutions tend to correspond to intrinsically poor conditions for nanopore sensing, during which signals are of course generated from blockages in the background flow of ions from these electrolyte solutions (typically, lower $\sigma \Rightarrow$ fewer ions). In a combined nanopore-iDEP system, there is thus a motivation to investigate how high the conductivity of the shared buffer can be raised while still maintaining acceptable DEP concentration functionality.

For its part, the type of positive DEP (pDEP) observed in the prior DNA concentration experiments (target migration toward areas of highest *E*-field) is predicted to have the opposite trend: lower medium conductivity leads to a stronger DEP effect. This can be easily seen in the form of the pre-factor in the expression for F_{DEP} on a spherical particle (Clausius-Mossotti factor, Equation 3.2). Here the magnitude of this factor scales with the difference in complex permittivities between the target particle and the medium ($\varepsilon^*_P - \varepsilon^*_M$), and since ε^* maps to σ in the low ω (DC) limit (Eq. 3.3), this means that lowering σ_M relative to σ_P in these conditions increases the strength of the DEP force such a particle feels.

Figure 3.11 shows a series of fluorescent DNA images that illustrates this effect. In each case, λ-DNA was suspended in a buffer of varying conductivity, injected into an iDEP cell, and its concentration patterns observed ~30 s into an AC voltage (1 kHz, 400 V_{pp}) being applied across the constriction. As expected, the strength of DEP interactions is greatest in the samples of lowest conductivity, as evidenced by both the larger spatial extents of the concentration regions and the magnitudes of the achieved concentration factors (e.g. compare Figure 3.11a and Figure 3.11c). Moreover, this trend is relatively indifferent to the exact ion species involved: the distribution in 0.1× TBE + 10 mM KCl (Figure 3.11d) has greater similarities with that of 2× TBE (Figure 3.11c) – both solutions $\sigma \sim 1$ mS/cm – than with that of the more similarly-composed, but lower conductivity 0.1× TBE sample (Figure 3.11a).

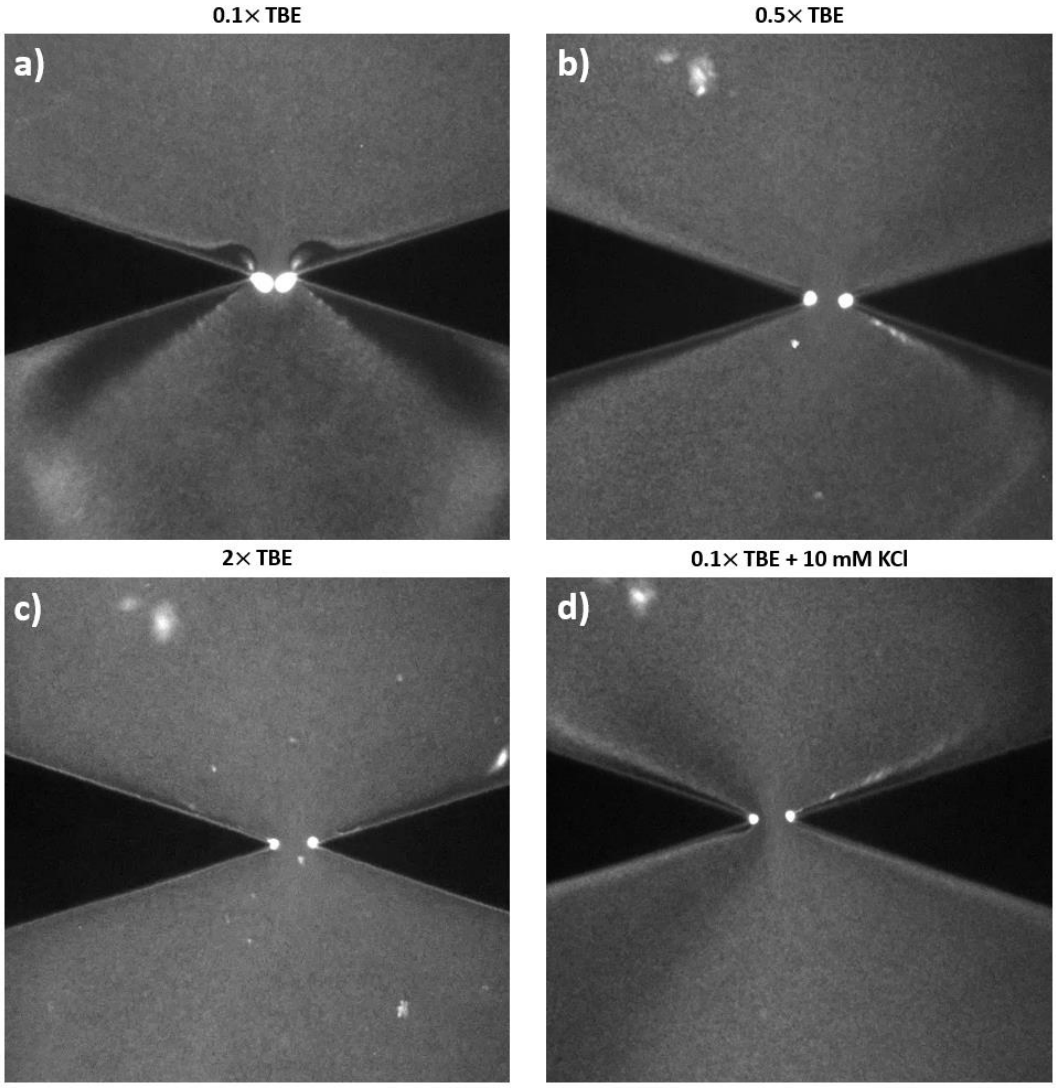


Figure 3.11: *iDEP DNA concentration with varying buffer conductivity.* ***a–d)*** *Distribution of SYBR Green-stained λ-DNA in various buffers around an iDEP constriction (~30 µm opening), approximately 30 s into an applied AC voltage (1 kHz, 400 V_{pp}). These buffers included: a) 0.1× Tris/Borate/EDTA (TBE), b) 0.5× TBE, c) 2× TBE, and d) 0.1× TBE + 10 mM KCl. A general trend emerged of reduced DNA trapping regions with increased buffer conductivity, regardless of the exact ion species (compare 0.1× TBE + 10 mM KCl, (d), with 0.1× TBE, (a), and 2× TBE, (c)). Additional hydrodynamic effects are also visibly more pronounced in the buffer with lowest conductivity (lowest surface charge screening) in the form of swirling vortices adjacent to the E-field hotspots at the constriction tips.*

While DEP concentration effects were observed to decrease over the range of buffer conductivities in Figure 3.11, conditions in typical nanopore experiments represent at least another order of magnitude increase in σ still (to ~1 – 10 mS/cm). Figure 3.12 examines the use of our iDEP constrictions at the lower end of these nanopore-compatible conditions, starting with Figure 3.12a-b, which shows the effect of applying moderate AC voltages (~100 V_{pp}, 1 kHz) to covalently-dyed (Alexa Fluor 488) 10-kbp DNA in 200 mM LiCl across a narrow ($\lesssim$ 10 µm) constriction. The covalent labelling, achieved by substituting modified dNTPs (aminoallyl-dUTP)

during a polymerase chain reaction that are able to subsequently bond with an amino-reactive dye (Alexa Fluor 488 TFP ester), was necessary due to the poor binding of intercalating dyes (such as SYBR Green) that occurs when their electrostatic interactions with DNA are effectively screened in high salt conditions. DNA was again concentrated at the highest *E*-field region of the iDEP channel (Figure 3.12b), similar to the pattern in lower salt (e.g. Figure 3.10c), but the effect is noticeably weaker here (peak CFs $\lesssim$ 30×, average CFs within hotspot $\lesssim$ 10×), as anticipated by a Claussius-Mossotti factor with a smaller conductivity gap between particle and medium.

Applying even larger voltages ($\gtrsim$ 200 V_{pp}) for extended times across an iDEP channel filled with high-σ solution invariably led to the formation of bubbles at its constriction tips (Figure 3.12c). This is perhaps the most striking symptom of the extensive Joule heating that is occurring in our system under these conditions, the generation term for which scales directly with medium conductivity ($dP/dV = \sigma_M E^2$) [29]. The electric field dependence of this term also explains why the bubbles are formed within the constriction, as this is where the highest *E*-fields (and thus current densities) are found, by design. Even below temperatures that reach the boiling point of the buffer however, the fact that the heat generated at a particular location in the channel depends on the strength of the electric field there generally results in a non-uniform temperature distribution throughout our system. This can then translate into additional hydrodynamic forces on our target ("electrothermal flow"), interfering with those of DEP, that arise from the temperature-dependent electrical properties (σ_M, ε_M) of the medium [12,29,30,43]. Beyond transient effects such as bubble formation and additional fluid flows that disperse when the electric field is released, permanent damage to the PDMS channel structures can also occur in high salt at the regions of highest temperature/field strength. Figure 3.12d shows an iDEP constriction that was subjected to elevated voltages ($\gtrsim$ 200 V_{pp}) for several minutes in 500 mM NaCl.

Being able to use our system with buffer conductivities that more closely match those of nanopore sensing might therefore be a matter of more efficiently dissipating the excess heat generated under these conditions. One way of doing so could be to shrink the channel dimensions so as to move the heat sinks provided by the channel walls closer to every region of the fluid bulk. This would have to be done with care, however, for reasons outlined in *Section 3.3 – Considerations for iDEP Cell Design*. For instance, the minimum constriction

width is limited by the method of channel fabrication (e.g. UV photolithography), while the width of the main channel feeding it needs to be much wider than this in order to generate sufficiently large field gradients. Channels made out of flexible PDMS are also at risk of collapsing if made too shallow, relative to their widths. Several previous studies have sidestepped some of these issues by having their iDEP channel designs etched directly in glass (though at significant increase in fabrication complexity/cost) [18,25,27,41]. This not only increases the channel rigidity, but, by replacing the three walls in PDMS with ones in glass (with its much higher thermal conductivity [44]), the heat dissipation capabilities of the system should be improved directly.

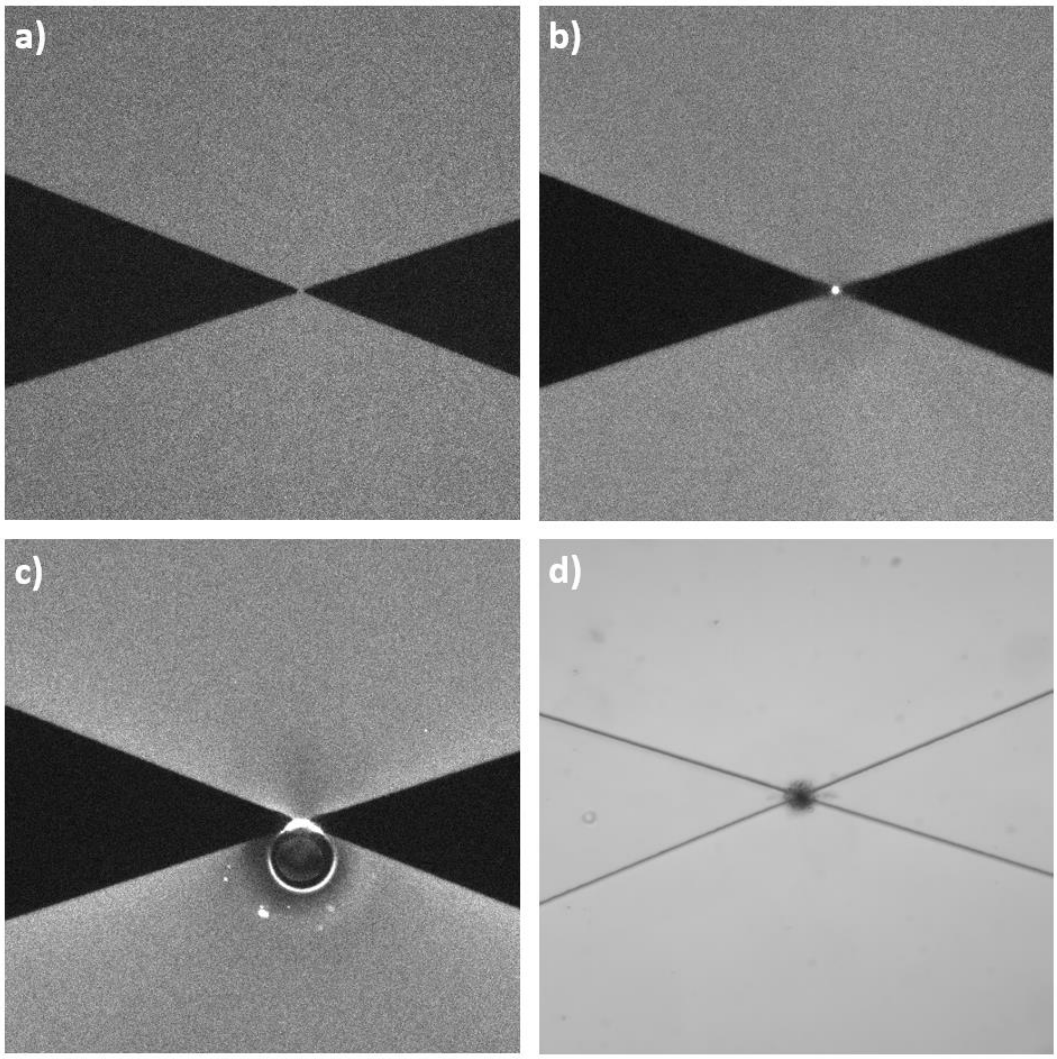


Figure 3.12: *iDEP DNA concentration in high salt buffers.* ***a & b)*** *Fluorescence images of covalently-labelled (Alexa Fluor 488) 10 kbp dsDNA distributions in 200 mM LiCl surrounding a narrow iDEP constriction (width ≲ 10 μm) before (a) and ~1 min into (b) the application of a trans-constriction AC potential (1 kHz, ~100 V_{pp}).* ***c)*** *Bubble generated from the sustained application of elevated (≳ 200 V_{pp}) AC voltages (same constriction/conditions as a,b).* ***d)*** *Optical image of a narrow iDEP constriction in the aftermath of an extended high-V (≳ 200 V_{pp}) AC potential in 500 mM NaCl. Noticeable damage to the underlying PDMS structures is visible at the region mapping to the highest field strengths under applied voltage.*

3.6.3 – Nanopore Sensing in Low Salt

As the counterpoint to raising the conductivity conditions of DEP experiments that are optimized at low-σ, we can also try lowering the conductivities in our nanopore experiments, typically run at high-σ, into a DEP-compatible range. The reasons for preferentially using nanopores in high salt are multifold. In addition to the increased signal strength (higher signal-to-noise ratio, SNR) mentioned earlier when more ions are present, most of the other main benefits are related to the additional screening of surface charge on both the nanopore membrane and polyelectrolyte target that occurs with increased concentrations of oppositely-charged

counterions in the system. For SiN_x pores and DNA (both negatively charged at e.g. pH 8), this leads to less electroosmotic flow [45] and electrostatic repulsion [46] opposing electrophoresis-driven capture and translocation, as well as less overlap in the Debye layers of pores and targets with closely matched sizes during transport [47], which simplifies analysis.

Figure 3.13 summarizes nanopore results in the lowest salt conditions that were successfully tested in the course of this work. The first panels (Figure 3.13a-b) deal with circumstances where each of the two fluid reservoirs on either side of the pore are filled with buffers of different conductivity ("salt asymmetry"). This allows for one side of the membrane to be held at low salt conditions (e.g. for DEP experiments) while maintaining higher ionic currents through the nanopore overall from the more conductive membrane-electrode path on the other side. Figure 3.13a shows a scatter plot of 20 kbp dsDNA translocations (average blockage vs. $\log_{10}$[dwell time]) through a single pore in three salt conditions. The DNA was introduced (on the *cis* side of the membrane) at consecutively lower concentrations of LiCl (720 mM, 360 mM, 180 mM), while the *trans* side was held fixed at 720 mM, creating salt gradients of 1:1, 2:1 and 4:1 from *trans* to *cis*. From the plot, we see that the overall blockage depths decreased with increasing degree of asymmetry, while the passage times through the pore remained more or less constant. A second benefit from running a nanopore under this polarity of salt asymmetry is that it leads to an increased target capture rate, due to the larger reach of the capture voltage away from the pore in the higher resistivity medium [48]. A plot of this effect is shown for our data in Figure 3.13b, where a capture rate increase of $\gtrsim 8\times$ was observed from the 1:1 to 4:1 conditions. While such results appear promising, it must be noted that a downside of salt asymmetries (and lower salt experiments in general) exists in the form of increased levels of pore clogging [49], with typical experiments generally lasting for shorter durations than their high symmetric salt counterparts.

Figure 3.13c-d contrasts these results with DNA translocations in *symmetric* low salt conditions. Figure 3.13c presents a 30 s current trace of 20 kbp events through two ~10 nm pores at 720 mM and 360 mM LiCl, respectively, and serves to illustrate the dramatic reduction in SNR that comes with halving the salt concentration in these circumstances. The events from each experiment are plotted in Figure 3.13d on

maximum blockage vs. $\log_{10}$[dwell time]. Here we see that in addition to having lower blockages (~100 pA at 360 mM vs. ~250 pA at 720 mM), the unfolded translocations in lower salt pass also through the pore with much shorter durations (~$10^{2.4}$ = 250 μs vs. ~$10^{2.7}$ = 500 μs). These two effects compound to further lower the time resolution of an experiment – the events intrinsically last shorter times and must also be more heavily low-pass filtered to have their signals (blockage depths) rise above the background noise levels in the current trace. Another interesting trend from Figure 3.13d is that the 360 mM events consist almost entirely of unfolded translocations while the 720 mM events also feature a healthy folded population (at ~500 pA). This may be the result of additional electroosmotic counter-flow stretching out the translocating molecules in lower salt [45], and/or may simply be related to the difficulty of fitting event substructure to the short folded subevents under the low time resolution described above.

As mentioned, the 360 mM LiCl experiment represented the lowest successfully tested symmetric salt condition. A common failure mode at salt concentrations below this was for the pore to instantly clog upon application of a transmembrane voltage, before any translocation events could be recorded. Other times, there would be no visible events in low salt (clean baseline), despite the same pore successfully passing the same DNA sample at higher salt concentrations. This may be related to the stronger electroosmotic/electrostatic effects opposing translocation in these conditions, or DNA successfully translocating but with a SNR value below 1, even after low-pass filtering. Previous papers have observed a critical salt concentration where DNA molecules passing through a pore transition from reducing the background ionic current to actually increasing the magnitude of the baseline through their co-transported counterions [50,51], and perhaps experiments without visible events represent conditions near the crossover point (ΔI = 0). Finally, some low salt experiments saw the appearance of current blockages only after DNA was added, but with signals that were chaotic and hard to interpret (in contrast to the recognizable “0-1-0” and “0-2-1-0” level patterns of unfolded and folded DNA typically observed, e.g. see Figure 3.9b), perhaps a sign of the increased importance of competing forces to electrophoresis in low salt, preventing smooth DNA transport.

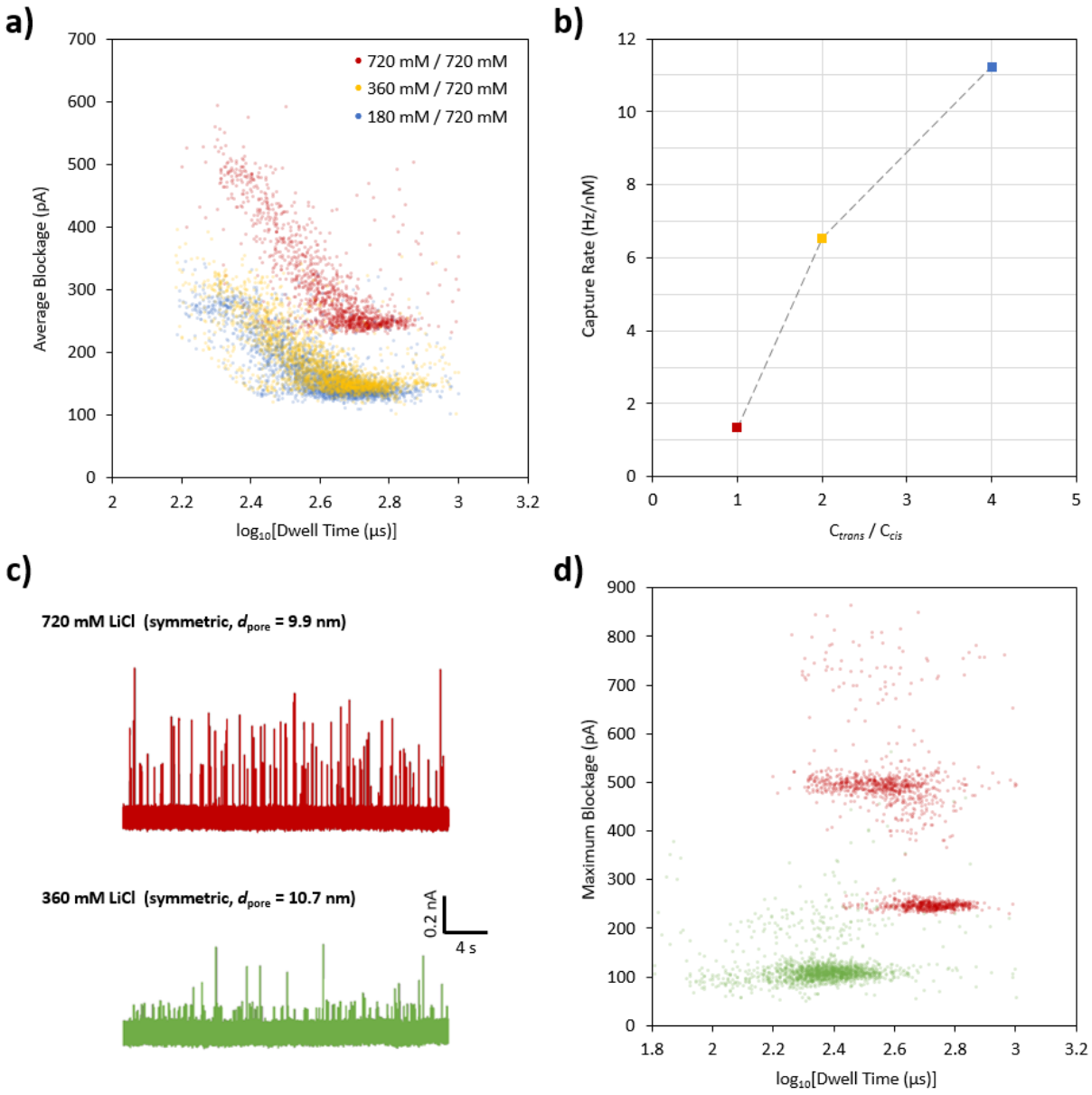


Figure 3.13: *Nanopore sensing in low salt buffers.* ***a)*** *Scatter plot of average blockage vs. log[dwell time] for 20 kbp dsDNA translocations through a single nanopore (~10 nm pore, 200 mV transmembrane potential) under symmetric and asymmetric salt conditions. In each case, the 'trans' reservoir of the fluidic cell contained 720 mM LiCl while the DNA sample was introduced into the 'cis' reservoir at three different LiCl concentrations.* ***b)*** *Normalized nanopore capture rate for the three salt conditions tested in (a) (the colour scheme is maintained from that panel), showing the increase in capture rate with degree of salt asymmetry (as parametrized by C_{trans}/C_{cis} where C_{trans} = 720 mM).* ***c)*** *Sample current trace (30 s) of 20 kbp dsDNA translocations through two ~10 nm nanopores (200 mV) at different symmetric LiCl concentrations. The 720 mM trace (in red) is from the same dataset depicted in (a) and (b), while the 360 mM trace (in green) is from a different, similarly sized pore.* ***d)*** *Scatter plot of max blockage vs. log[dwell time] for all translocations from the symmetric salt experiments depicted in (c) (the colour scheme is maintained from that panel).*

3.6.4 – Membrane Integrity under Transverse E-Fields

Figure 3.14 outlines a series of tests designed to probe the stability of nanopores fabricated within in our iDEP-NP cells under the application of AC voltages similar to those used to concentrate DNA in previous sections.

Having a pore remain stable over each envisioned stage (iDEP pre-concentration, nanopore sensing) of a full experiment involving these cells would be a necessary condition to its overall success. For these preliminary tests, the fluidic channels were assembled in one of two configurations: either with the iDEP constriction placed directly over the SiN_x membrane ("aligned", Figure 3.14a) or with the constriction and membrane slightly displaced from each other ("offset", Figure 3.14b).

After fabricating a nanopore (as described above, see *Section 3.5 – Integrating Nanopore Membrane into Fluidic Device*), the first step was to perform a DNA translocation experiment to ensure that the measured conductance through the membrane was the result of a single nm-scale hole. Figure 3.14c shows a sample current trace from such an experiment, where 15-kbp DNA was passed through a ~8-nm pore in 1.8 M LiCl ("aligned" cell). The general presence of the blockage events as well as their recognizable distributions along extracted statistics (e.g. the hyperbolic relationship between average blockage and dwell time, see inset) confirm the intended geometry of the nanopore.

Before applying transverse AC voltages across a pore-containing membrane, the salt concentration of its fluid reservoirs was first lowered down to values at the limits of prior successful nanopore/DEP experiments (~200 mM, see preceding sections). Figure 3.14d shows I-Vs from an iDEP cell-embedded nanopore at two different salt concentrations, 1.8 M and 0.225 M LiCl. When transitioning from high- to low-salt conditions, the current though the pore remained linear with voltage, but with a lower extracted conductance (slope). This new conductance value was, however, consistent with the pore maintaining its diameter from the high-salt measurement (~6 nm in both cases), after accounting for the reduced conductivity of the low salt buffer.

Once submersed in the correct low-σ buffer, a series of AC voltage pulses of increasing magnitude were applied overtop of the pore-containing membranes (across the iDEP constrictions), with I-V measurements taken before and after each voltage treatment in order to assess its impact on the pore through changes in its conductance. An overview of this process for a single pore is presented in the plot of Figure 3.14e. Each point represents the conductance value of a single I-V measurement, and the position of the arrows indicate in between which I-Vs a

transverse AC voltage was applied (1 kHz frequency, 30 s duration, amplitudes as labelled in the figure). At applied voltage amplitudes of 20 V_{pp} (between I-Vs #2 and #3 in Figure 3.14e), no major change in pore conductance was observed. As the amplitude was increased to 40 V_{pp}, however, discrete jumps in conductance appeared (e.g. between I-Vs #3 and #4). These jumps could not be attributed to transient effects (such as increased solution temperature) as subsequent I-Vs, taken several minutes later, displayed the same elevated conductance values, or higher. A similar trend was observed over multiple cells tested, both with membranes and constrictions "aligned" and "offset", where a critical voltage amplitude was reached (at 40–60 V_{pp}) that resulted in rapidly growing membrane conductance. Figure 3.14f shows an optical image of an iDEP-nanopore system that displayed a large ~μm-sized opening in its membrane after application of an elevated transverse voltage (20 s at 100 V_{pp}), lending further support to the idea that these increased conductance values map to physical paths through the membrane.

A previous article by the Meller group has also reported on the electrokinetic pre-concentration of DNA in a PDMS channel above a SiN_x nanopore, this time using a technique called "isotachophoresis" [52]. In that work, they applied large voltages (100 V_{DC}) across a straight, salt-filled (600 mM KCl) channel and noted that preventing deformation of their comparable (10–15 nm thick) SiN_x membranes and expansion of their nanopores under these conditions (from electromechanical stress) required the use of a "shunt" connection between the *cis* and *trans* sides of the membrane while the transverse voltage was applied. This was achieved by integrating a short fluidic side-channel at the position of the membrane (*cis* side) that led to an Ag/AgCl electrode connected, in turn, to a similar electrode on the etch-pit (*trans*) side via a solid-state relay, holding the two sides of the membrane at equal potentials. The relay is then opened during nanopore sensing so that a (small) transmembrane potential can be applied to translocate DNA. Implementing a similar shunt connection in our iDEP-NP system may alleviate some of the electromechanical stress present there and perhaps help suppress the runaway conductance growth, although this would need to be designed with care since much of our voltage drop occurs in the near vicinity of the constriction/membrane, by design. A successful implementation therefore likely requires some degree of membrane "offset" (Figure 3.14b) and a shunt electrode placed in close proximity.

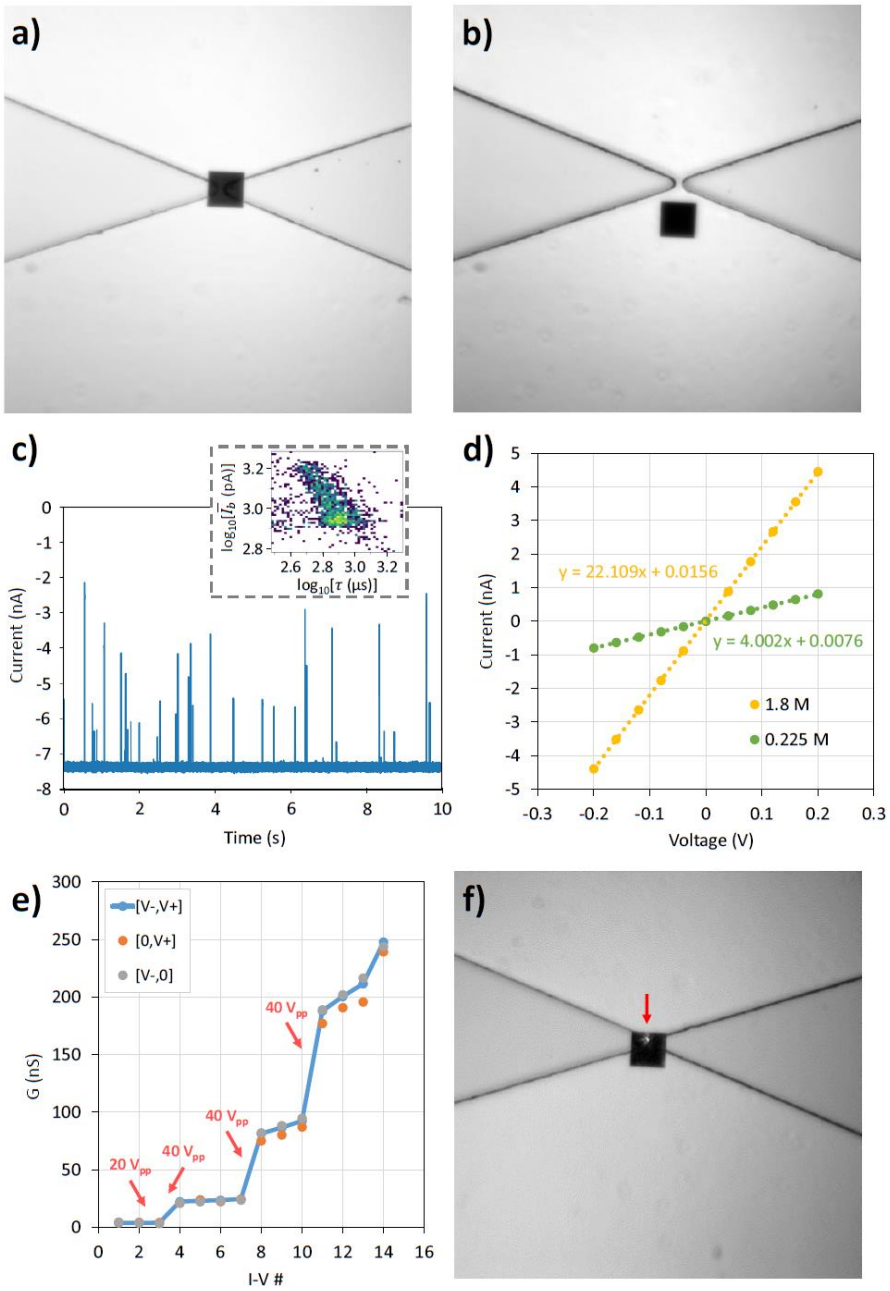


Figure 3.14: *Applying transverse AC voltages near pre-fabricated nanopores.* ***a)*** *&* ***b)*** *Optical images of PDMS constrictions (~10-μm openings, taper angles 2ϕ = 40°) bonded directly on top of ("aligned", a) and slightly offset to ("offset", b) a 40 × 40 μm² SiN_x membrane.* ***c)*** *Sample current trace of 15 kbp dsDNA translocation events through a nanopore embedded in a PDMS iDEP cell ("aligned" constriction, ~8-nm pore, 1.8 M LiCl, 200 mV). Inset: 2D histogram of average blockage vs. dwell time (log-log scale) for the full dataset, showing an inverse relationship (slope ~-1) between the two quantities (characteristic of DNA translocations with conserved ECD).* ***d)*** *Current-voltage (I-V) curves through an iDEP cell-embedded nanopore ("offset" constriction) at two LiCl concentrations. When accounting for the different conductivities of the two solutions, a consistent pore diameter of ~6-nm was estimated from the slope (conductance) of each curve.* ***e)*** *Calculated conductance through an iDEP-cell embedded SiN_x membrane as determined by sequential I-Vs (in 225 mM LiCl). In between particular I-V measurements (indicated by pink arrows) a 1-kHz voltage was applied for 30 s across the constriction at the labelled voltage amplitudes. Prior to these measurements, the membrane contained a single ~6-nm pore, as depicted in (d).* ***f)*** *Optical image of an iDEP cell-embedded membrane after extended application of a 100 V_{pp}, 1 kHz AC voltage across the constriction (in 225 mM LiCl). A bright spot of transmitted light (from an ambient source behind the cell) is visible within the extents of the otherwise dark membrane (see red arrow).*

3.7 – Conclusions and Outlook

In summary, this chapter presented a design for a microfluidic device incorporating both insulator-based DEP and solid-state nanopore functionalities, with the goal of widening the applicability of nanopore sensors to low-concentration samples by actively transporting targets in these samples to the pore so as to increase their rates of capture. Overall principles to the design and operation of the devices were discussed, and, in the course of this, necessary compromises were identified between idealized DEP geometries and practical fabrication concerns (e.g. in specific values of the channel dimensions) as well as between optimal conditions for DEP concentration and for nanopore sensing (e.g. in the conductivity of the buffer filling the channels). Preliminary experimental results with the devices were also shown, validating their ability to both effectively concentrate particles at the constriction tips (polystyrene beads, dsDNA) and detect molecules translocating through the nanopore, though these capabilities were demonstrated separately, not in a single, continuous experiment.

Despite these promising results, multiple hurdles to sequentially concentrating and then sensing targets in our system remain, as documented in *Section 3.6 – Outstanding Challenges*. By implementing some of the minor adjustments proposed in that section, however, such as adding a "shunt" connection to keep both sides of the SiN_x membrane at equipotential when DEP voltages are applied (to maintain the integrity of the membrane) and introducing DNA "stabilizers" (e.g. Tween 20) into our samples (to suppress DNA aggregation under high field strengths), a lightly modified version of the basic system can be envisioned that should still bring tangible benefits to overall target detection rate. For instance, if membrane deformation and DNA aggregation can be avoided, the results presented above suggest that moderate AC voltages ($\sim$100 V_{pp}) applied to DNA in a nanopore-compatible $\sim$200 mM LiCl salt concentration would result in reasonable concentration enhancements (CFs < 100×) at the site of the constriction (Figure 3.12b); if, simultaneously, the *trans* reservoir contained a higher concentration of LiCl (e.g. $\sim$800 mM), then asymmetric salt conditions could be expected to contribute a further $\sim$10× increase to the capture rate (Figure 3.13b), while also leading to improvements in the nanopore signal-to-noise ratio and translocation times (compare Figure 3.13a,d).

Going forward, there is also the opportunity to incorporate broader changes to the basic iDEP-nanopore device design presented here. For instance, rather than having a single constriction located directly in the sensing region of the system, some researchers have placed a full two-dimensional array of posts/constrictions upstream of this region [23,41]. Device operation in this scheme generally consists of the following sequence: 1) applying a relatively small DC voltage ($\gtrsim$ 10 V) to transport targets into the constriction array from the channel inlet by electrophoresis/electroosmosis, 2) applying a larger voltage ($\gtrsim$ 100 V, AC or DC) to trap/concentrate targets within the array, and 3) returning back to the low DC voltage to release the targets from the DEP traps and migrate them as a mass of enriched concentration toward the downstream sensing region. This arrangement could benefit our system by physically separating the delicate SiN_x membrane from the regions of highest *E*-field (at the constrictions), though we also note that some degree of dispersion of the concentrated target mass would be expected to take place as it travels downstream to the sensor, potentially resulting in lower capture rates than in the case where the pore is placed directly at the DEP hotspot. Pressure regulators could also be integrated into such a system as an alternative to EP/EOF for driving flows within the channels of the device (e.g. if linear electrokinectic effects were interfering with DEP trapping). In particular, one could imagine having a pair of fluid reservoirs at different ionic strengths and a series of pressure-actuated valves that controlled which reservoir solution was flowing into the main channel at a given point in the experiment. DEP targets could then be surrounded by (optimally) low-σ buffer during the DEP concentration stage and, subsequently, this could be replaced by high- σ buffer as the AC voltage is switched off and targets are transported downstream to the nanopore, thus alleviating some of the intrinsic conflict between ideal salt conditions of DEP concentration and nanopore sensing.

Finally, we note that the geometries of our iDEP devices have remained somewhat unoptimized due to the simple fabrication techniques adopted at this early prototyping stage (2D channels cast in PDMS from SU-8 moulds). In addition to incorporating some of the more complex fabrication techniques referred to earlier [18,19,21], a more conceptual approach to increasing the DEP forces in our system, for instance, could be to locally reduce the height of our channels, in concert with to their width, at site of the constriction. This would

represent a second, independent dimension over which to focus the electric field lines in the channel and thus generate larger field gradients (see Eq. 3.1). Constructing fully-3D constrictions in this way could be also achieved without additional advanced cleanroom equipment, such as by aligning multiple 2D layers on top of each other, by "reflowing" the photoresist of narrow structures to thin their heights [33], by directly fabricating channels with maskless photolithography systems [14], or by using 3D-printed channel moulds [53]. Regardless of the exact geometries ultimately chosen, we see both nanopore sensors and electrokinetic concentrators as important components of future integrated "lab-on-a-chip" (LOC) or "micro total analysis systems" (μTAS) to rapidly detect and analyze targets of biological importance from low-volume samples.

3.8 – References

1. Freedman KJ, Otto LM, Ivanov AP, Barik A, Oh S-H, Edel JB. Nanopore sensing at ultra-low concentrations using single-molecule dielectrophoretic trapping. Nat Commun. 2016;7: 10217. doi:10.1038/ncomms10217
2. Grosberg AY, Rabin Y. DNA capture into a nanopore: Interplay of diffusion and electrohydrodynamics. J Chem Phys. 2010;133: 165102. doi:10.1063/1.3495481
3. Charron M, Briggs K, King S, Waugh M, Tabard-Cossa V. Precise DNA Concentration Measurements with Nanopores by Controlled Counting. Anal Chem. 2019;91: 12228–12237. doi:10.1021/acs.analchem.9b01900
4. Regtmeier J, Eichhorn R, Viefhues M, Bogunovic L, Anselmetti D. Electrodeless dielectrophoresis for bioanalysis: Theory, devices and applications. Electrophoresis. 2011;32: 2253–2273. doi:10.1002/elps.201100055
5. Kirby BJ. Micro- and Nanoscale Fluid Mechanics: Transport in Microfluidic Devices. Cambridge University Press; 2010.
6. Bakewell D, Vergara-Irigaray N, Holmes D. Dielectrophoresis of Biomolecules. JSM Nanotechnology & Nanomedicine. 2013;1: 1–14.
7. Manning GS. Limiting Laws and Counterion Condensation in Polyelectrolyte Solutions I. Colligative Properties. J Chem Phys. 1969;51: 924–933. doi:10.1063/1.1672157
8. Hölzel R. Dielectric and dielectrophoretic properties of DNA. IET Nanobiotechnol. 2009;3: 28. doi:10.1049/iet-nbt.2008.0014
9. Tuukkanen S, Kuzyk A, Toppari JJ, H?kkinen H, Hyt?nen VP, Niskanen E, et al. Trapping of 27 bp?8 kbp DNA and immobilization of thiol-modified DNA using dielectrophoresis. Nanotechnology. 2007;18: 295204. doi:10.1088/0957-4484/18/29/295204

10. Regtmeier J, Eichhorn R, Bogunovic L, Ros A, Anselmetti D. Dielectrophoretic trapping and polarizability of DNA: The role of spatial conformation. Anal Chem. 2010;82: 7141–7149. doi:10.1021/ac1005475

11. Jubery TZ, Srivastava SK, Dutta P. Dielectrophoretic separation of bioparticles in microdevices: A review. Electrophoresis. 2014;35: 691–713. doi:10.1002/elps.201300424

12. Viefhues M, Eichhorn R. DNA dielectrophoresis: Theory and applications a review. Electrophoresis. 2017; 1483–1506. doi:10.1002/elps.201600482

13. Swami N, Chou C-F, Ramamurthy V, Chaurey V. Enhancing DNA hybridization kinetics through constriction-based dielectrophoresis. Lab Chip. 2009;9: 3212–20. doi:10.1039/b910598k

14. Puttaswamy SV, Lin C-H, Sivashankar S, Yang Y-S, Liu C-H. Electrodeless dielectrophoretic concentrator for analyte pre-concentration on poly-silicon nanowire field effect transistor. Sens Actuators B Chem. 2013;178: 547–554. doi:10.1016/j.snb.2013.01.016

15. Kwok H, Briggs K, Tabard-Cossa V. Nanopore Fabrication by Controlled Dielectric Breakdown. Hall A, editor. PLoS One. 2014;9: e92880. doi:10.1371/journal.pone.0092880

16. Yeo W-H, Kopacz AM, Kim J-H, Chen X, Wu J, Gao D, et al. Dielectrophoretic concentration of low-abundance nanoparticles using a nanostructured tip. Nanotechnology. 2012;23: 485707. doi:10.1088/0957-4484/23/48/485707

17. Kumar V, Pallapa M, Rezai P, Selvaganapathy PR. Polymers. Reference Module in Materials Science and Materials Engineering. Elsevier; 2016. pp. 1–62. doi:10.1016/B978-0-12-803581-8.00522-1

18. Liao K-T, Chou C-F. Nanoscale Molecular Traps and Dams for Ultrafast Protein Enrichment in High-Conductivity Buffers. J Am Chem Soc. 2012;134: 8742–8745. doi:10.1021/ja3016523

19. Li S, Ye Z, Hui YS, Gao Y, Jiang Y, Wen W. On-chip DNA preconcentration in different media conductivities by electrodeless dielectrophoresis. Biomicrofluidics. 2015;9. doi:10.1063/1.4932177

20. Chiou C-H, Chien L-J, Kuo J-N. Nanoconstriction-based electrodeless dielectrophoresis chip for nanoparticle and protein preconcentration. Applied Physics Express. 2015;8: 085201. doi:10.7567/APEX.8.085201

21. Zhang P, Liu Y. DC biased low-frequency insulating constriction dielectrophoresis for protein biomolecules concentration. Biofabrication. 2017;9: 045003. doi:10.1088/1758-5090/aa82d6

22. Zaouk R, Park BY, Madou MJ. Introduction to Microfabrication Techniques. Microfluidic Techniques. New Jersey: Humana Press; 2010. pp. 3–16. doi:10.1385/1-59259-997-4:3

23. Regtmeier J, Thanh TD, Eichhorn R, Anselmetti D, Ros A. Dielectrophoretic manipulation of DNA: Separation and polarizability. Anal Chem. 2007;79: 3925–3932. doi:10.1021/ac062431r

24. Hellmich W, Regtmeier J, Duong TT, Ros R, Anselmetti D, Ros A. Poly(oxyethylene) Based Surface Coatings for Poly(dimethylsiloxane) Microchannels. Langmuir. 2005;21: 7551–7557. doi:10.1021/la0510432

25. Cummings EB, Singh AK. Dielectrophoresis in microchips containing arrays of insulating posts: Theoretical and experimental results. Anal Chem. 2003;75: 4724–4731. doi:10.1021/ac0340612

26. Pohl HA, Pollock K, Rivera H. The electrofusion of cells. Int J Quantum Chem. 1984;26: 327–345. doi:10.1002/qua.560260734

27. Chou CF, Tegenfeldt JO, Bakajin O, Chan SS, Cox EC, Darnton N, et al. Electrodeless dielectrophoresis of single-and double-stranded DNA. Biophys J. 2002;83: 2170–2179. Available: http://linkinghub.elsevier.com/retrieve/pii/S0006349502739775

28. Hickey OA, Harden JL, Slater GW. Molecular Dynamics Simulations of Optimal Dynamic Uncharged Polymer Coatings for Quenching Electro-osmotic Flow. Phys Rev Lett. 2009;102: 108304. doi:10.1103/PhysRevLett.102.108304

29. Hawkins BG, Kirby BJ. Electrothermal flow effects in insulating (electrodeless) dielectrophoresis systems. Electrophoresis. 2010;31: 3622–3633. doi:10.1002/elps.201000429

30. Sridharan S, Zhu J, Hu G, Xuan X. Joule heating effects on electroosmotic flow in insulator-based dielectrophoresis. Electrophoresis. 2011;32: n/a-n/a. doi:10.1002/elps.201100011

31. Davalos R V., McGraw GJ, Wallow TI, Morales AM, Krafcik KL, Fintschenko Y, et al. Performance impact of dynamic surface coatings on polymeric insulator-based dielectrophoretic particle separators. Anal Bioanal Chem. 2008;390: 847–855. doi:10.1007/s00216-007-1426-5

32. Tahvildari R, Beamish E, Tabard-Cossa V, Godin M. Integrating nanopore sensors within microfluidic channel arrays using controlled breakdown. Lab Chip. 2015;15: 1407–1411. doi:10.1039/C4LC01366B

33. Tahvildari R, Beamish E, Briggs K, Chagnon-Lessard S, Sohi AN, Han S, et al. Manipulating Electrical and Fluidic Access in Integrated Nanopore-Microfluidic Arrays Using Microvalves. Small. 2017;13: 1–7. doi:10.1002/smll.201602601

34. Albrecht T, Gibb T, Nuttall P. Ion Transport in Nanopores. First Edit. Engineered Nanopores for Bioanalytical Applications. First Edit. Elsevier; 2013. pp. 1–30. doi:10.1016/B978-1-4377-3473-7.00001-7

35. Siwy ZS. Ion-Current Rectification in Nanopores and Nanotubes with Broken Symmetry. Adv Funct Mater. 2006;16: 735–746. doi:10.1002/adfm.200500471

36. Kumar Sharma R, Agrawal I, Dai L, Doyle PS, Garaj S. Complex DNA knots detected with a nanopore sensor. Nat Commun. 2019;10: 4473. doi:10.1038/s41467-019-12358-4

37. Kowalczyk SW, Grosberg AY, Rabin Y, Dekker C. Modeling the conductance and DNA blockade of solid-state nanopores. Nanotechnology. 2011;22: 315101. doi:10.1088/0957-4484/22/31/315101

38. Tabard-Cossa V. Instrumentation for Low-Noise High-Bandwidth Nanopore Recording. First Edit. In: Edel JB, Albrecht T, editors. Engineered Nanopores for Bioanalytical Applications. First Edit. Norwich, N.Y. : William Andrew ; Oxford : Elsevier Science distributor; 2013. pp. 59–93. doi:10.1016/B978-1-4377-3473-7.00003-0

39. Waugh M, Briggs K, Gunn D, Gibeault M, King S, Ingram Q, et al. Solid-state nanopore fabrication by automated controlled breakdown. Nat Protoc. 2020;15: 122–143. doi:10.1038/s41596-019-0255-2

40. Lerman LS. Structural considerations in the interaction of DNA and acridines. J Mol Biol. 1961;3: 18-IN14. doi:10.1016/S0022-2836(61)80004-1

41. Gallo-Villanueva RC, Rodríguez-López CE, Díaz-de-la-Garza RI, Reyes-Betanzo C, Lapizco-Encinas BH. DNA manipulation by means of insulator-based dielectrophoresis employing direct current electric fields. Electrophoresis. 2009;30: 4195–4205. doi:10.1002/elps.200900355

42. Li X, Hu R, Li J, Tong X, Diao JJ, Yu D, et al. Non-sticky translocation of bio-molecules through Tween 20-coated solid-state nanopores in a wide pH range. Appl Phys Lett. 2016;109. doi:10.1063/1.4964117

43. Chaurey V, Rohani A, Su Y-H, Liao K-T, Chou C-F, Swami NS. Scaling down constriction-based (electrodeless) dielectrophoresis devices for trapping nanoscale bioparticles in physiological media of high-conductivity. Electrophoresis. 2013;34: 1097–1104. doi:10.1002/elps.201200456

44. Erickson D, Sinton D, Li D. Joule heating and heat transfer in poly(dimethylsiloxane) microfluidic systems. Lab Chip. 2003;3: 141. doi:10.1039/b306158b

45. Ermann N, Hanikel N, Wang V, Chen K, Weckman NE, Keyser UF. Promoting single-file DNA translocations through nanopores using electro-osmotic flow. J Chem Phys. 2018;149: 163311. doi:10.1063/1.5031010

46. Jeon BJ, Muthukumar M. Polymer capture by α-hemolysin pore upon salt concentration gradient. Journal of Chemical Physics. 2014;140. doi:10.1063/1.4855075

47. Zanjani MB, Engelke RE, Lukes JR, Meunier V, Drndić M. Up and down translocation events and electric double-layer formation inside solid-state nanopores. Phys Rev E. 2015;92: 022715. doi:10.1103/PhysRevE.92.022715

48. Wanunu M, Morrison W, Rabin Y, Grosberg AY, Meller A. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. Nat Nanotechnol. 2010;5: 160–165. doi:10.1038/nnano.2009.379

49. Charron M, Philipp L, He L, Tabard-Cossa V. Elucidating the dynamics of polymer transport through nanopores using asymmetric salt concentrations. Nano Res. 2022;15: 9943–9953. doi:10.1007/s12274-022-4886-3

50. Smeets RMM, Keyser UF, Krapf D, Wu M-Y, Dekker NH, Dekker C. Salt Dependence of Ion Transport and DNA Translocation through Solid-State Nanopores. Nano Lett. 2006;6: 89–95. doi:10.1021/nl052107w

51. Steinbock LJ, Lucas A, Otto O, Keyser UF. Voltage-driven transport of ions and DNA through nanocapillaries. Electrophoresis. 2012;33: 3480–3487. doi:10.1002/elps.201100663

52. Spitzberg JD, van Kooten XF, Bercovici M, Meller A. Microfluidic device for coupling isotachophoretic sample focusing with nanopore single-molecule sensing. Nanoscale. 2020;12: 17805–17811. doi:10.1039/D0NR05000H

53. Comina G, Suska A, Filippini D. PDMS lab-on-a-chip fabrication using 3D printed templates. Lab Chip. 2014;14: 424–430. doi:10.1039/C3LC50956G

CHAPTER 4 – PRODUCTION OF VARIABLE-LENGTH STICKY-ENDED DNA CARRIERS

Elements of this chapter are shared with an article published in *PLOS ONE*:
"Synthesis of Length-Tunable DNA Carriers for Nanopore Sensing" (2023)
by Zachary Roelen and Vincent Tabard-Cossa

4.1 – Introduction

4.1.1 – Molecular Carriers

Molecular carriers are a sub-category of an increasingly popular trend in nanopore sensing to use designed, structured molecules to characterize a sample of biological interest[1] or to probe the fundamentals of polymer physics[2]. Under the molecular carrier scheme, instead of sensing a target analyte directly with a nanopore, a proxy molecule ("carrier") that interacts with this target (or with intermediate reporters of the target) is sensed[3,4]. The key requirement of this approach is for the nanopore signals generated by the translocating carriers to be recognizably different depending on whether this interaction with the target has occurred or not (Figure 4.1a). For instance, a target could bind to a specific subregion of a carrier, resulting in this subregion blocking additional ionic current when passing through the pore (as compared to without the target attached)[5]. In this way, the presence of the target analyte can be observed indirectly through its effect on the carrier.

Using molecular carriers in place of direct sensing can provide a number of benefits, depending on the application. For one, it can increase the sensitivity of a nanopore experiment for a particular target, such as in cases where the target molecule on its own is not easily captured (e.g. weakly charged) or detected (e.g. small size, fast translocations) by the pore. It also offers a means of imparting target specificity to an experiment, without the need for complex functionalization of the nanopore itself. Many potential targets (e.g. globular proteins) can have very different functions (or diagnostic values in an assay) from each other, all while producing fairly similar (fast, structureless) signals when passing through a pore[6]. If, however, only one of the targets

interacts in a unique way with a molecular carrier designed for this purpose, then its specific presence in a mix of otherwise indistinguishable molecules can be established.

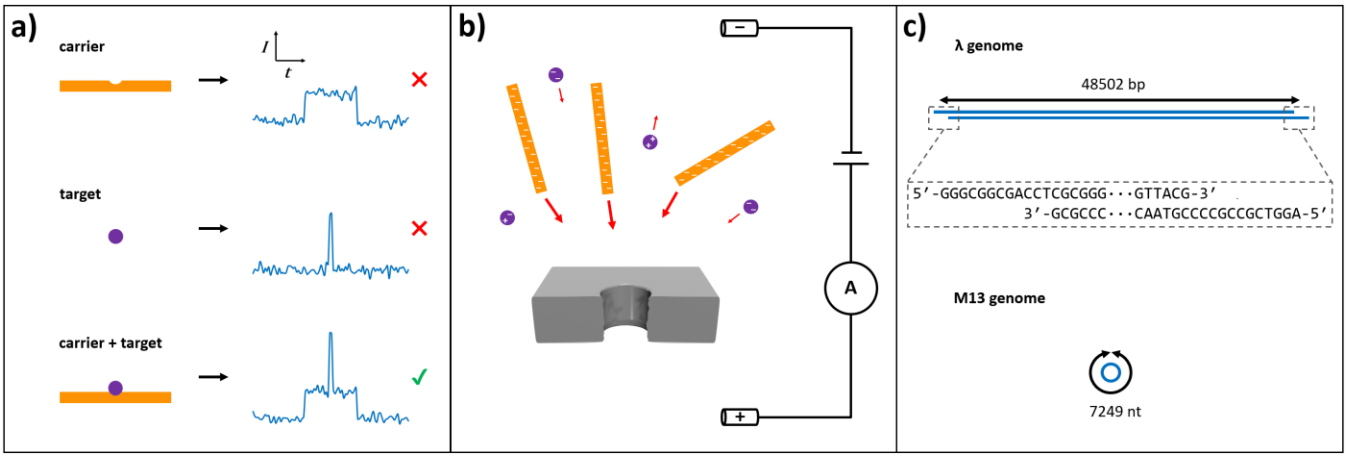


Figure 4.1: a) *Illustration of characteristic nanopore signals (right) generated by carrier and target molecules (left). Carriers are selected to produce extended, well-characterized signals, while smaller targets may suffer from fast translocations that are more difficult to locate and fit. Adding the two pieces together into a single hybrid molecule results in signals that combine the higher frequency and detectability of the carrier signals with the recognizable information content encoded by the attached target.* ***b)*** *Schematic of a nanopore experiment featuring carrier molecules (orange) and potential assay targets (purple) in solution. The carrier molecules are depicted here as being extended in size and highly negatively charged (as in the case of DNA) and thus experience strong electrophoretic forces under an applied transmembrane potential. Conversely, the target molecules are depicted as being more compact and coming with a variety of net charges (positive, negative, and neutral), e.g. as could result from a mix of different globular proteins. These properties (small size and variable charge) would make such targets harder to consistently capture and detect on their own with a nanopore.* ***c)*** *Structures of two viral genomes commonly used as the basis of DNA carriers (and DNA nanotechnology generally). Lambda DNA is a linear, double-stranded genome with 12-nt cohesive, single stranded ends, while M13 DNA is a circular, single-stranded genome. The relative contour lengths of the long (48.5 kbp) λ-DNA and the shorter (7.25 knt) M13-DNA are drawn to scale.*

Finally, molecular carriers can allow for nanopore sensors to probe a wide variety of target types without heavily modifying a base sensing protocol. Figure 4.1b illustrates an example of such a system, where a collection of target molecules exists with varying values (in magnitude and in sign) of net electric charge. This could again be representative of globular proteins, for example, whose surface charges are pH-dependent and determined by the sequence of amino acids that compose them[7]. However, if every target was combined with a molecular carrier (depicted in Fig. 4.1b as highly-negatively charged) whose individual contribution dominated the net electrophoretic mobility of the target-carrier pair, then similar experimental conditions such as transmembrane potential could be used to detect each species. Single molecular carriers (or a mix of carriers) can also be

engineered to have interaction probes for multiple targets, allowing a single nanopore experiment to report on each of these targets simultaneously ("multiplexing")[8,9].

4.1.2 – Double-Stranded DNA

A primary candidate for a carrier molecule is double-stranded DNA (dsDNA), which features high levels of regularly-spaced negative charge along its backbone and which has been extensively characterized inside the nanopore system[10]. In addition to being readily captured under an applied electric field, the rigid, double-helical structure of dsDNA results in it having a relatively open, extended conformation in solution[11], allowing for positional information (such as the attachment of a target molecule along its contour) to be read out in its translocation signal as it threads sequentially though the pore. To build a molecular carrier based on DNA, commercially-produced viral genomes are often used as the starting point. These have the advantages of being widely available and of possessing known sequences around which their target-selective functionalities can be designed. Here, we highlight two existing approaches to implementing DNA carriers based on viral genomes.

A first approach, following work by the Keyser group[12], starts with the circular, single-stranded genome of the M13 bacteriophage (7.25 knt in length, see Figure 4.1c), typically linearized first by enzymatic cleavage. Inspired by the staple strands of DNA origami[13], to this single-stranded genome are attached hundreds of shorter (~40 bp) complementary oligonucleotides in order to form a double-stranded construct. By modifying the sequence or chemical functionalization of particular oligos in the complement, the capacity to interact with specific targets can be integrated at precise locations along the contour of the completed carrier. Carriers built by this method have been used to detect and measure the concentration of multiple proteins[14], to distinguish two DNA strands differing by a single base[15], and to store digital data in the form of carrier structure[16–18]. An innovative way of signalling target specificity in this system is with the concept of "barcoding" where the particular target interaction being probed by a carrier is labelled by a nanopore-readable structure ("barcode") at a dedicated region in the sequence[9,18]. In this way, two targets that generate somewhat similar signals at the

separate, interaction regions of their respective carriers can be reliably distinguished by the different readouts of their barcodes, while maintaining the aforementioned benefits of the carrier system as a whole.

A second, possibly simpler DNA carrier approach, as developed by the groups of Edel and Ivanov[8], is based on λ DNA which is a linear, double stranded genome (48.5 kbp) that features short, cohesive 12-nucleotide single-stranded extensions at its ends (see Figure 4.1c). It is to these cohesive ends that functionalized oligos imparting carrier capabilities to the final constructs can be attached, in analogy with the methods of Keyser & co. Carriers based on λ DNA have been used successfully to probe fluorescent DNA oligos[19], biotin-streptavidin binding[19], and protein-aptamer interactions[8,19]. Although the location of carrier activity within the DNA substrate is more restricted in this system (when compared to approaches involving fully single-stranded genomes and their 100+ oligo complements) each of the two cohesive ends can be effectively extended with linker oligos as needed and the simultaneous integration of up to three homo- and heterogeneous protein-aptamer pairs on a single λ-carrier has been demonstrated[8].

While molecular carriers produced from viral genomes can be versatile, there are some downsides associated with such sources of DNA. For one, an extended genome such as λ (48.5 kbp) represents a long length for a nanopore analyte (typically $\lesssim$ 10 kbp[10,20,21]). The longer a DNA molecule is, the more prone it is to entering the nanopore in a folded conformation[22,23] which could the cause the local, intra-event signals generated by substructures in the molecule (e.g. as from attached targets) to be read outside of their expected order[17,24]. Greater proportions of events going unrecognized in this way can necessitate longer experiment times (to generate sufficient statistics from the non-rejected events) and may even lead to experimental bias if correlations exist between the folding frequency of a carrier and its state of target attachment. Longer molecules also tend to irreversibly stall more often during a translocation attempt ("pore clogging"), potentially ending a single-pore experiment early before all sample conditions can be tested.

Additional issues can also arise from sourcing purified versions of these viral templates commercially. While widely available from popular vendors in several molecular configurations and order sizes, these purified

genomes may become costly to purchase in the quantities needed to continuously run nanopore experiments, especially if losses are incurred in their subsequent processing into their final, modified forms. There is also variability that can exist (e.g. batch-to-batch, vendor-to-vendor) in the quality of a particular template sample, for instance in the fraction of normally circular genomes (e.g. M13) that can arrive randomly sheared (and therefore linearized) partway through their sequence[12]. This can translate into *experimental* variability if, after processing, the defective copies of the molecule present in the sample lead to altered nanopore signals.

4.1.3 – Chapter Overview

The overall goal of this work was therefore to develop protocols for the in-lab synthesis of DNA carriers that are well-suited for nanopore sensing. In particular, we focused on examples that followed the general structure of λ-DNA (see Fig. 4.1c), namely, linear, double-stranded molecules that terminate in single-stranded overhangs ("sticky ends") onto which targets may be attached by sequence-specific DNA hybridization. Two main objectives were prioritized when developing these protocols. The first was to have any methods used be achievable with standard, countertop molecular biology lab equipment (e.g. thermocycler, gel electrophoresis system, microcentrifuge). This lowers the barrier for other researchers to make use of and iterate on these methods, without the need to procure access to highly-specialized, institution-shared equipment (e.g. HPLC system, ultracentrifuge). Note, however, that restricting ourselves to these common tools still allows for plenty of flexibility in the design and functionality of the molecular outputs. For instance, the powerful polymerase chain reaction (PCR) technique combines the thermocyler with a handful of widely-available consumables (including a thermostable DNA polymerase) to repeatably amplify DNA fragments with specific targeted sequences and can readily incorporate functional groups or labels into these fragments by simple substitutions in the starting materials (e.g. using modified dNTPs or primers)[25,26].

The second priority when designing our synthesis protocols was to allow the overall length of the sticky-ended carriers produced through these methods to be highly adjustable. As discussed above, a DNA fragment that is overly long can be problematic for nanopore sensing through its increased tendencies to adopt complex,

difficult to analyze conformations and to clog pores. Conversely, there are also complications to sensing very short DNA fragments, arising from the finite bandwidth of the system used to measure the small nanopore currents (centred around a low-noise amplifier) – short fragments may translocate the membrane too quickly to be reliably detected or to have the details of their signals be well resolved by this system. Beyond the need for an intermediate fragment length between these two extremes however, there is additional utility in the context of molecular carriers to being able to finely adjust the product length within this range. This is because carrier length represents another variable by which carriers mapping to different assay targets could be separated by a nanopore, specifically through the dwell times or equivalent charge deficits of the translocation events they produce[20,27,28]. Moreover, many of the techniques for carrier identification highlighted above (such as barcoding or end-labelling) can be applied simultaneously with length selection on the same molecule, creating a multiplicative effect that supports high levels of multiplexing from a small number of variants in each mode.

In this chapter, we present two methods for synthesizing sticky-ended DNA molecules that adhere to the objectives outlined above and that were investigated in some detail. In each case, the basic protocol is first outlined and validated. Modifications to this basic protocol are then introduced that allow for the length of the product to be adjusted as desired. Following on this, the two methods are compared and key advantages/ disadvantages to each are considered.

4.2 – Results

4.2.1 – Method 1: Oligonucleotide Assembly

The first synthesis method is illustrated in the schematic of Figure 4.2a. It is based on a DNA origami-like approach of using a relatively long (~knt) piece of single-stranded DNA ("scaffold") combined with many short (10-100 nt) DNA oligonucleotides ("staples") that bind to the scaffold by sequence-specific DNA hybridization to form an extended double-stranded structure[12].

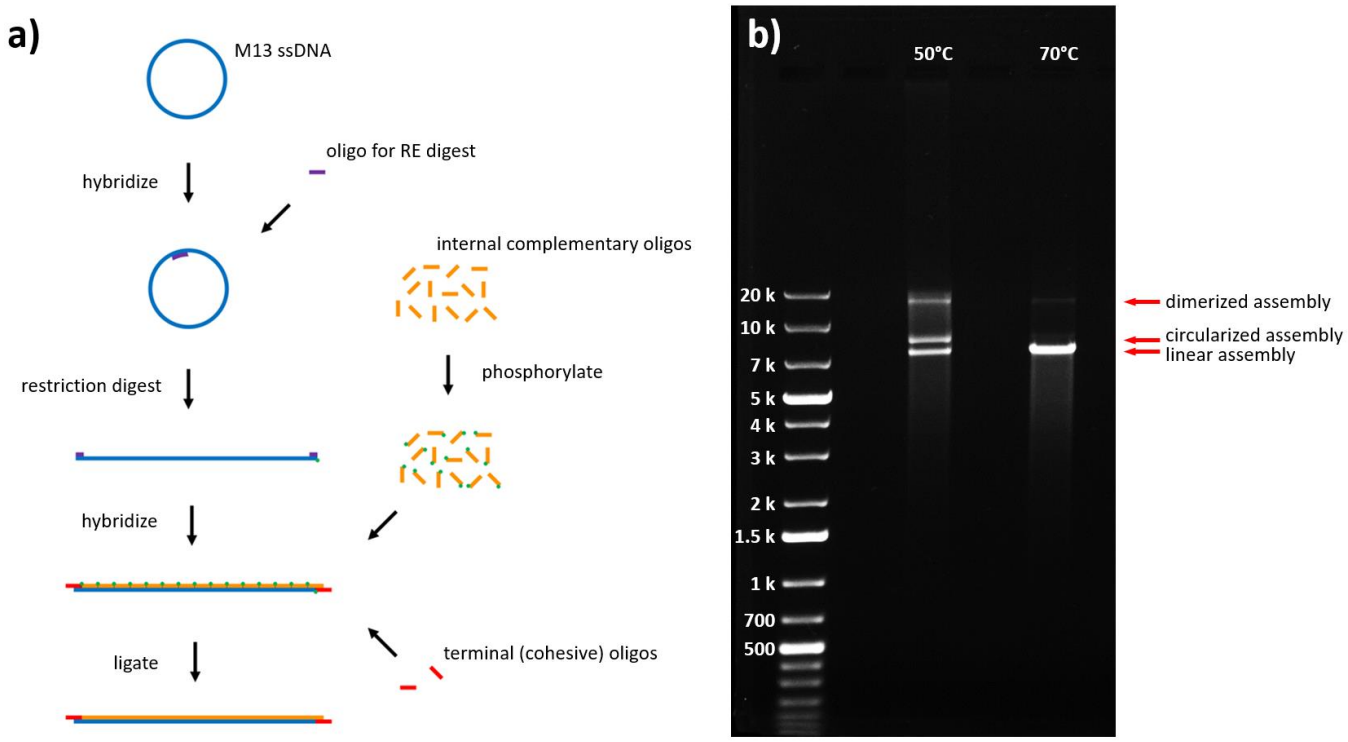


Figure 4.2: a) *Overview of a scheme for SE-DNA synthesis by oligonucleotide assembly onto a single-stranded DNA scaffold molecule. The internal oligos are phosphorylated (represented in the figure by green circles at their 5' ends), while the terminal oligos are left unphosphorylated to prevent the assembled molecules from being permanently locked into a circular or multimeric configuration after the final ligation step.* ***b)*** *Agarose gel (0.5%, 0.5× TAE, pre-stained with GelRed, 70 V) of 7.3 kbp SE-DNA before (lane 2) and after (lane 3) heat treatment for 10 min at 70 °C. The samples from both lanes were initially incubated for 1 hr at 50 °C in 0.25× PNK buffer (2.5 mM $MgCl_2$, 17.5 mM Tris-HCl, pH 7.6). The slower migrating bands, corresponding to circularized monomers and dimers, are greatly reduced in favour of the linearized monomer band when exposed to temperatures above the melting point of the 12-nt overhangs.*

Here we start with circular, single-stranded DNA sourced from a viral genome (7249 nt, M13mp18, NEB) as our proto-scaffold. We then linearize this with a restriction endonuclease (HincII, NEB) by first binding a short complementary oligo at the cleavage site for recognition by the enzyme (see Sections A1 and A2 in the Appendix for a full list of DNA sequences and detailed experimental protocols, respectively). For the purposes of assembly to the resulting scaffold, the staple strands (40-45 nt, IDT) are grouped into two categories: 1) "internal" oligos (×180) that will form the bulk of the linear complement to the scaffold and 2) "terminal" oligos (×2) that will form the cohesive, 12-nt ssDNA overhangs in the final structure (see Figure 4.2a). Before assembly, the internal oligos are phosphorylated (T4 polynucleotide kinase, NEB) to enable a subsequent ligation step to repair nicks in the backbone (T4 DNA ligase, NEB). This will result in a final molecule, featuring a continuous complement strand, that should be more stable during temperature cycling experiments (e.g. as in Figure 4.2b, detailed later)

as well as more directly comparable to the nick-free output of the second synthesis method investigated here (Figure 4.4). Conversely, the terminal oligos are left unphosphorylated in order to prevent unwanted permanent circularization/multimerization of the SE-DNA monomers during this ligation step.

After assembly, the final products are imaged by agarose gel electrophoresis (e.g. rightmost lane of Figure 4.2b), where we see a main product band that migrates at a speed that is consistent with a dsDNA fragment of the expected length (≳7 kbp) when compared to the DNA ladder. The gel migration was also compared to that of double-stranded versions of the original M13 genome in linear, relaxed circular, and supercoiled forms (see Section A3 in the Appendix), and the main band of the assembled product was indistinguishable from the linear form. In addition to the main band, the gel pattern of Fig. 4.2b is also characterized by a smear of faster migrating/lower molecular weight products below it, as well as at least two faint, slower bands above it. The spread of faster products might result from incomplete assemblies that are missing staples, or from partial assemblies that started with shorter ssDNA scaffolds. These truncated scaffolds can arise when some fraction of the circular ssDNA molecules have been prematurely sheared into linear form (a known issue with commercial M13 DNA[12]), resulting in a second, randomly placed cleavage after the restriction digest step that is located partway through the expected sequence.

Returning to the faint, slower-migrating bands that appear above the main population in the gel, these are presumably the result of interactions between the cohesive ends of properly-assembled molecules. For instance the two ends of a single assembly could hybridize to form a circular molecule with the same contour length as a linear monomer, or sticky ends from two separate molecules could combine to create a dimer with twice the original contour length. The migration patterns of the gel bands we observe are consistent with this interpretation, with the presumptive circular population migrating slightly slower than the linear monomers (mirroring the known behaviour of relaxed circle M13 dsDNA in these conditions, see Section A3), and the "dimer" band (which should contain 2 × 7.3 k = 14.6 k base pairs) migrating between the 10 kbp and 20 kbp bands of the ladder. Furthermore, because these second-order structures should not have been locked in during the ligation step of assembly (on account of the terminal oligos having unphosphorylated ends), they should be

capable of reversibly dissociating and re-forming simply by cycling the temperature around the melting point of the 12-nt cohesive ends. This is what is attempted in Figure 4.2b, which shows the distributions of populations in a sample of SE-DNA products above (70 °C, right lane) and below (50 °C, middle lane) this melting temperature. As the temperature is raised to 70 °C, the slower-migrating bands (corresponding to circles and dimers) are depleted while the faster (linear monomer) band is enhanced. This supports the idea that the intended single-stranded overhangs are present and active in a significant portion of the synthesized molecules.

4.2.2 – Varying SE-DNA Length with Method 1

Referring back to Figure 4.2a, we see that an oligo-assembly approach to SE-DNA production relies on having a single-stranded scaffold to start, and that the overall length of the final (double-stranded product) is directly determined by the length of this scaffold (in our case, the length of the M13 phage genome). Changing this final length could therefore begin with producing ssDNA scaffolds of alternate sizes. A strategy for this that was employed in this work is illustrated in Figure 4.3a. It uses PCR (Q5 Polymerase, NEB) to create a double-stranded product (whose length is defined by the distance between the chosen primers) and then one of the strands is selectively digested to leave a single-stranded scaffold of the same length.

The key to making this digestion selective to only one strand comes from modifying the 5' ends of the primers used during the PCR step[29]. Such a strategy emerges from the fact that the digest enzyme used here (lambda exonuclease, NEB) has a strong preference in activity for phosphorylated DNA substrates in particular[30]. This activity is greatly reduced on non-phosphorylated substrates and can be even further suppressed with the inclusion of chemically modified DNA bases into the sequence[31]. Following these principles, for each desired product length, two differentially 5'-terminated PCR primers are utilized (IDT, see A1 for sequences): a 5'-phosphorylated oligo that will prime the synthesis of the strand that will ultimately be digested, and a 5'-phosphorothioated oligo that will prime the synthesis of the scaffold strand and protect it from degradation. Phosphorothioted (PS) bases differ from native bases by the replacement of a particular oxygen atom in their phosphate group with a sulfur atom (forming a PS group, see Fig. 4.3a). Even small modifications like this to

bases placed at strategic locations within DNA substrates can have a drastic effect on the ability of many exonucleases to recognize and process these substrates, and the effect has been well-characterized for lambda exonuclease and 5′ PS bonds in particular[32].

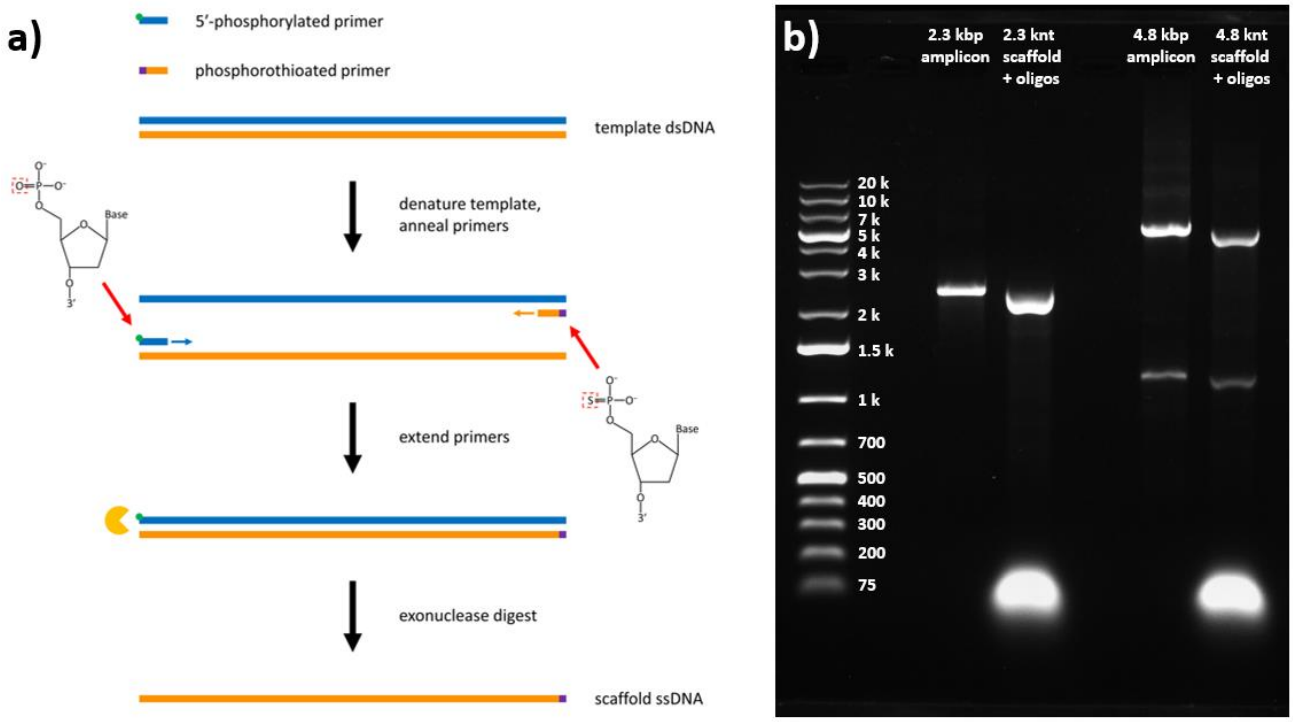


***Figure 4.3**: **a)** Schematic for creating ssDNA scaffolds of different lengths via PCR with modified primers. The phosphorothioated primer makes one of the two product strands resistant to subsequent digestion by lambda exonuclease. This resulting single-stranded product is then used downstream as a scaffold to create SE-DNA molecules via assembly of short oligonucleotides (see Fig. 4.2a). **b)** Agarose gel (0.7%, 0.5× TBE, pre-stained with GelRed, 70 V) documenting the synthesis step of staple strands assembling onto PCR-generated ssDNA scaffolds of two lengths. Lane 3: 2.3 kbp PCR product, Lane 4: λ-exo digest on 2.3 kbp PCR product incubated with full sequence of complimentary oligos to M13 ssDNA, Lane 6: 4.8 kbp PCR product, Lane 7: λ-exo digest on 4.8 kbp PCR product incubated with full sequence of complimentary oligos to M13 ssDNA.*

Figure 4.3b shows a gel image documenting representative stages of this process for two carrier lengths, 2.3 kbp and 4.8 kbp. The leftmost lanes in each case (lanes 3 & 6) show the initial PCR outputs for the two lengths. Both reactions shared M13 DNA as their template as well as a common phosphorothioted “left” primer, but differed in how many bases away (2.3 k or 4.8 k) their “right”, phosphorylated primers were placed. From the image, we see that both reactions resulted in a main product band that migrated to its expected position relative to the DNA ladder (in between 2 – 3 kbp, and 4 – 5 kbp, respectively), although the output of the 4.8 kbp reaction was

not quite as “clean”, featuring additional faint but noticeable populations between the 1 – 1.5 kbp ladder bands and smeared at slower migrating positions above 5 kbp (further optimization of that particular primer set may therefore be worth pursuing). The double-stranded PCR amplicons were then treated with lambda exonuclease, which resulted in a gel migration pattern of faster migrating, fainter bands (relative to the pre-digest pattern), consistent with single-stranded molecules that have become more flexible, and harder to stain with intercalating dyes (see Section A4 in the Appendix). Finally, the rightmost lanes (4 & 7) for each case in Fig. 4.3b show the outcomes of incubating the outputs of these two λ-exo digests with the full complement of staple strands to the M13 scaffold (i.e. the 180 “internal” oligos used to assemble the dsDNA product of Fig. 4.2). The result is the overall restoration of the original (pre-digest), double-stranded migration pattern plus a large, fast-migrating band at the bottom of the gel corresponding to residual unhybridized oligos (added in excess). This suggests that not only did one of the amplicon strands largely survive the exonuclease digest intact, but also that it was specifically the complement strand to the M13 staples (protected with the PS-primer), as intended.

Once confident in the scheme of Figure 4.3a to produce single-stranded scaffolds of alternate lengths, synthesis of sticky-ended assemblies was attempted as in Figure 4.2 by combining these scaffolds with staple strands (reduced from the full set for 7.3 kbp M13 DNA) that include the terminal oligos responsible for creating ssDNA overhangs. Subsequent gel shift assays (e.g. reversible circularization tests as in Fig. 4.2b) confirmed the presence and activity of sticky ends in the final products (see Section A5 in the Appendix). These results therefore represent a proof of concept for the method of Figures 4.2 & 4.3 to produce sticky-ended DNA fragments of variable length. As an alternative, a simpler approach to creating shorter scaffolds could involve cleaving the M13 genome at additional locations during the restriction digest step (currently used to linearize circular M13, see Fig. 4.2a). However, the method presented here of selectively digesting PCR amplicon strands provides at least a few key advantages over this simpler approach by 1) allowing scaffold lengths outside of just those from the available recognition sites on the M13 genome, 2) not requiring an additional gel purification step to isolate particular fragments from the RE digest (see Method 2, below, for a discussion of the problems

caused by gel extraction), and 3) being able to produce DNA substrates on demand via PCR amplification rather than continuously purchasing stocks of purified viral genomes.

4.2.3 – Method 2: Terminase Digest

The second SE-DNA synthesis method investigated in this work is illustrated in Figure 4.4a. The central idea here is to generate single-stranded overhangs on a synthetic DNA template by using the same enzyme that the bacteriophage λ uses to create the sticky ends of its own genome during its viral life cycle. This enzyme is "terminase", a heteromer built from two subunits (20 & 73 kDa), and it catalyzes this "terminus-generating" reaction by creating staggered nicks (offset by 12 bases) in the two strands of the λ genome[33]. Lambda terminase has been available in the past as a commercial product[34], where it was typically used in the process of generating genomic libraries through the use of cloning vectors such as λ-phage and cosmids, as well as in the physical mapping of these clones[35]. The terminase sample used in this work was generously supplied by Dr. Carlos Catalano (University of Colorado, Denver) and differs from the wild-type enzyme by the addition of six terminal histidine tags to aid in its purification by immobilized metal affinity chromatography[36].

To generate a suitable template for terminase to act on, we start with commercially-sourced λ-DNA (48.5 kbp, NEB). As mentioned, this phage genome ends with cohesive, single-stranded overhangs that play a role in the virus's reproductive cycle. During this process, the 12 base pairs that are to eventually form these overhangs are located within a relatively large ~200 bp region of the genome (the cohesive, or "*cos*" site) that is recognized by the terminase enzyme prior to the cleavage reaction[33]. The first step in our protocol (Figure 4.4a) is therefore to reconstitute this *cos* site by ligating the ends of our λ-DNA together (Taq DNA ligase, NEB). This is combined with a restriction digest step (BsrGI, NEB) to create more manageable fragment sizes that also make the progress of the reaction easier to track (see Section A6 in the Appendix for gel images from the various intermediate steps of Method 2). Sections of the central, *cos*-containing fragment of this ligation/digestion step can then be amplified by PCR (LongAmp polymerase, NEB) to create a renewable supply of terminase proto-template. Using terminase directly on this linear, *cos*-containing amplicon would result in single-stranded overhangs that are

split across two separate DNA fragments (see Figure 4.4a). Instead, another ligation step (T4 DNA ligase) is introduced to circularize the amplicons, which is facilitated by a preceding restriction digest (PciI, NEB) to generate shorter (4-nt) sticky ends.

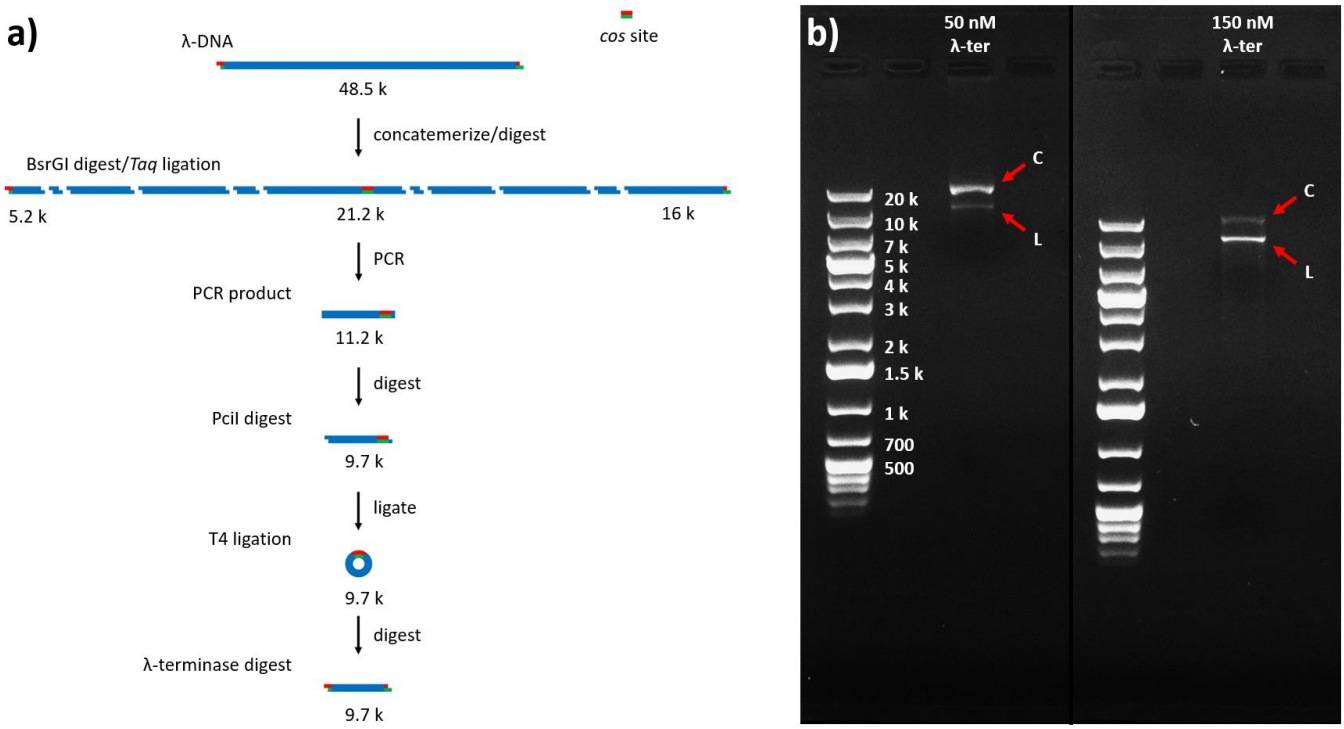


Figure 4.4: a) *Schematic of a method for in vitro SE-DNA production by PCR amplification of a cos-containing fragment and λ-terminase digestion.* ***b)*** *Left panel: Agarose gel (0.6%, 0.5× TAE, pre-stained with GelRed, 100 V) of a λ-terminase digest on a sample of 9.7 kbp circular DNA prepared by this method. Lane 1: DNA ladder. Lane 2: aliquot of sample after 1 hr incubation at 37 °C in the presence of 50 nM terminase. Right panel: Agarose gel of the same reaction as in the left panel but after an additional 1 hr incubation with more terminase added to a total concentration of 150 nM. Before and after the second incubation with additional enzyme, the fraction of linearized DNA circles rose from ~13% to ~75%, as estimated from the relative intensities of the slow-migrating circular band and the fast-migrating linear band.*

Circularized amplicon monomers were chosen as our enzyme template rather than multimers (repeats of the amplicon linked end-to-end) in order to maximize the efficiency of the terminase digest – only one cut per molecule is required to open up the circle into our final, sticky-ended product, compared to the two cuts (one on each side) needed to break out a product from within a multimer. This was important when working with a non-commercial enzyme due to both the finite total supply we had available as well as some uncertainty around its optimal reaction conditions (as compared to the precise formulations/instructions supplied with most commercial enzymes). Using circular monomers also sidesteps issues that arise from restriction enzymes

(including PciI) typically recognizing palindromic sequences[37], which makes it so that both of the freshly-cleaved ends of a restriction digest are identical and self-cohesive. If two monomers featuring these ends are ligated together in the "wrong" orientation (as they will be ~50% of the time), it will result in a fraction of final, post-terminase products that differ from the desired product in both their length and their overhang sequence (see Section A7 in the Appendix for details). Each end of a circularizing monomer having only one available ligation partner avoids this ambiguity.

In order to suppress unwanted multimers then during the circularization stage, the concentration of linear monomers during the ligation reaction becomes an important consideration. If this concentration is too high, an individual sticky end on one molecule will be more likely to encounter the end of a second, separate molecule than it will the other end of its same molecule. On the other hand, at very low DNA monomer concentration, the reaction volumes required to process a sufficient number of molecules (e.g. enough to generate data over several nanopore experiments) will become very large, as will the amounts of ligase used if constant enzyme concentrations/reaction times are maintained. An estimate of the relevant concentration ranges was made by modelling the amplicons as random walks of fixed-length (Kuhn) segments (see Section A8 in the Appendix). The takeaway is that for 1 – 10 kbp amplicons, concentrations within or below in the ~nM range (~a few ng/µL by mass) are required for circularization of monomers to be favoured (e.g. > 75%). Sample gels that document an amplicon circularization reaction under these conditions as well as a test for circularity via restriction digest are presented in the Appendix (Section A9).

Once a circular, *cos*-containing DNA template is produced, the final step is to re-linearize this template by λ-terminase digest at the *cos*-site, creating a molecule with extended, 12-nt ssDNA ends (Figure 4.4a). Attempts at optimizing the conditions of this reaction (e.g. by temperature, incubation time, enzyme concentration, and inclusion of co-factors) are documented in the Appendix (Section A10). Figure 4.4b shows the progress of an example terminase reaction involving the linearization of a 9.7 kbp *cos*-containing circle at two different enzyme concentrations. In the left panel, the DNA sample was incubated with 50 nM terminase (1 hr, 37°C), resulting in only a small fraction (~13%) of the total DNA shifting from the slower-migrating circular form to the faster-

migrating linear form, as estimated from the relative gel band intensities. In the right panel, an aliquot of this same sample was incubated with a higher, 150 nM enzyme concentration (1 hr, 37°C), and the populations were inverted – the majority of DNA (~75%) was now found in the linear band. Activity of the 12-nt sticky ends was tested by gel shift assay through the hybridization of DNA nanostructures (star junctions) featuring single-stranded extensions complementary to the overhang sequence (see Section A11 in the Appendix), and a noticeable shift was observed in the majority of SE-DNA molecules, signifying successful hybridization.

As a final note, in the course of designing this synthesis protocol we found that any purification steps that involved extracting DNA from agarose gels created problems for downstream nanopore sensing of these molecules. Such an agarose extraction step could be required for instance in between digesting the ends of our PCR amplicon and circularizing its central fragment (see Fig. 4.4a) – if the shorter terminal fragments are not removed before the ligation stage, they could re-attach to this central fragment and block it from being closed into a circle. However, we found that samples that had interacted with agarose reliably led to increased pore clogging (both reversible and irreversible with voltage inversion or solution flushing) and additional, low-amplitude blockages in the current signal (as if from contaminants), when compared to control samples that lacked agarose purification (see Section A12 in the Appendix). These issues persisted over all of the numerous extraction methods tested including gel dissolution (e.g. QIAquick, QIAGEN), centrifugal separation (e.g. Freeze 'N Squeeze, Bio-Rad), enzymatic digestion (e.g. agarase, NEB), and passive diffusion into supernatant out of submerged gel pieces. The protocol was therefore modified to avoid a gel extraction step by moving the location of the PCR primers to be only a short distance ($\lesssim$ 100 bp) away from the recognition site of the restriction enzyme (e.g. PciI) used to prepare the amplicon for circularization. In this way, very short terminal fragments are produced during the restriction digest that can be easily be separated from the larger (~kbp) central fragment by silica membrane-based spin columns (e.g. PureLink, Invitrogen) or alcohol precipitation, etc., thus negating the need for agarose purification. This alteration was successfully incorporated into the protocol and the remaining synthesis steps worked as before (see Section A13 in the Appendix).

4.2.4 – Varying SE-DNA Length with Method 2

Reviewing the synthesis protocol of Figure 4.4a, we see that the length of the final sticky-ended product here is determined during the restriction digest step (e.g. with PciI in Fig. 4.4) that occurs just prior to circularization of the λ-terminase substrate. An obvious way of altering this length is therefore to change the identity of the restriction endonuclease(s) in this step, which will serve to define the fragment size by the different sites they recognize on the molecule. There are several downsides to such an approach, however. For one, purchasing new restriction endonucleases for each new length can quickly become costly if more than just a few lengths are desired. Additionally, if a double digest is used to define the target fragment (or if variability exists in the recognition site of a single restriction enzyme), care must be taken to ensure that the resulting two ends of the fragment are compatible with each other (e.g. through adapter oligos) so as to enable the subsequent circularization step. Finally, as discussed in the analogous section of Method 1, above, the range of possible fragment sizes producible by this method is limited by the available restriction sites that are present in a given DNA template.

Rather than using the restriction digest step then to set the product length, a more flexible and consistent approach might be to instead use the distance between primers in the PCR step for this purpose, as was done in Method 1 (Fig. 4.3). The added complication in this case is the need for short cohesive overhangs at the ends of the amplicon in order to increase its efficiency at circularizing in the next stage (see Fig. 4.4). This can easily be achieved however by adding extra bases (not complementary to the template) to the start (5' end) of the PCR primers so as to include an arbitrary enzyme recognition site in this initial sequence[38]. To avoid overly long primers in this proof of concept (some enzymes also cannot cleave directly at the end of a fragment[38]), a related approach was employed in this work to *induce* a particular recognition site *within* the existing template sequence through the inclusion of point mismatches in the primers (see Figure 4.5a). In this scheme, a sequence in the DNA template is identified that is both: 1) close to being a recognition site for a particular restriction enzyme (i.e. sharing most of its bases with the "correct" sequence) and 2) located roughly the same distance away from a similar, recognition site-approximating sequence as the desired length of the final product. A

corresponding PCR primer is then designed that hybridizes to this sequence, except with the "pseudo-recognition" site replaced with the true version. As illustrated in Figure 4.5a, the result after PCR is a DNA fragment of the desired length that now contains recognition sites for a single enzyme at both its ends, and so can be cleaved in a single reaction to leave short, cohesive overhangs. Relying on pre-existing pseudo-recognition sites in the template sequence for this process is not actually very limiting to length selection. For instance, Section A14 in the Appendix shows a distribution of fragment sizes that are possible from selecting a single restriction enzyme (PciI), fixing the position of the first primer, and looking for sequences (inside the second primer) that differ from the 6-base recognition site of this enzyme by one base at most. Even in this highly restricted case, a full range of lengths (separated by ~220 bases on average) is achieved, all featuring compatible ends post-digestion. Selecting a particular product length for an experiment then involves modifying a core synthesis protocol by only a single, low-cost primer oligo.

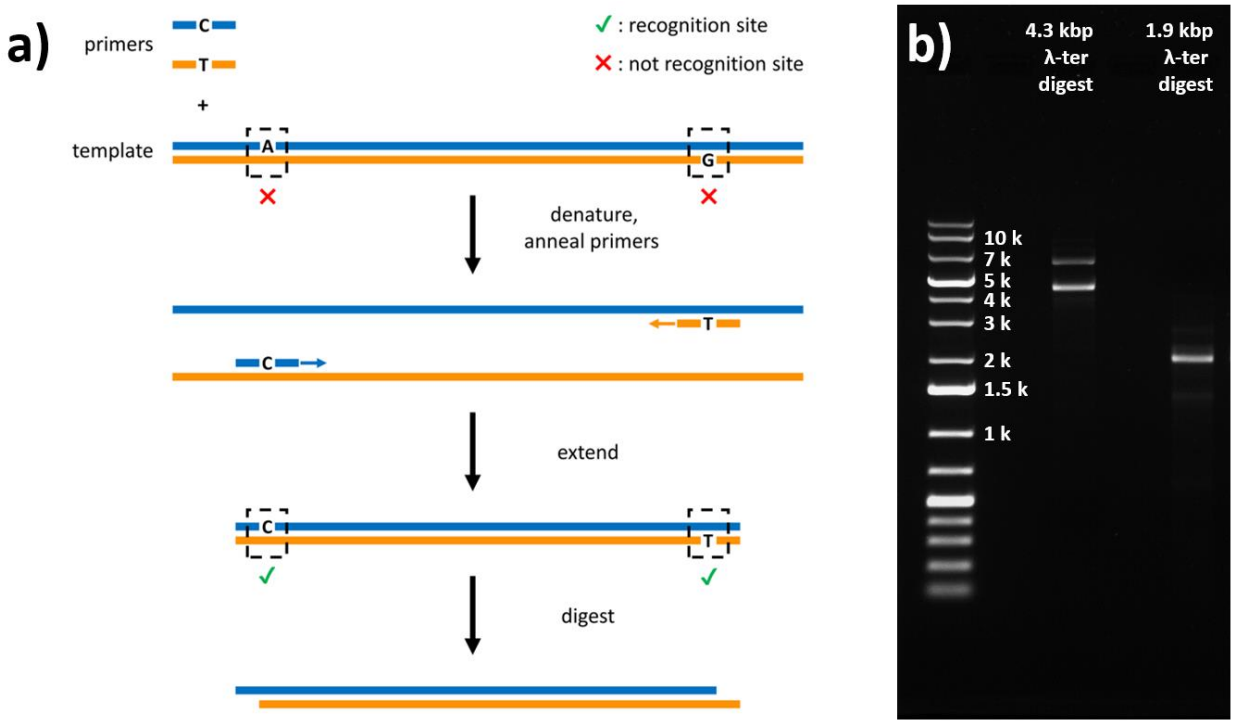


***Figure 4.5**: **a)** Schematic of introducing a recognition site for a restriction enzyme in the sequence of a DNA amplicon via partial primer-template mismatch during PCR. **b)** Agarose gel (0.7%, 0.5× TAE, pre-stained with GelRed, 70 V) of 4339 bp (Lane 3) and 1941 bp (Lane 5) sticky-ended DNA fragments produced by λ-terminase digest of circularized PCR template, where template length is determined through selection of primers containing PciI recognition sites during amplification. A more pronounced subpopulation of circular DNA is present in the ~4 kbp sample either from incomplete terminase digestion or re-hybridization of the sticky ends post-digest.*

Outputs of the synthesis protocols of Figures 4.4a/4.5a for two product lengths are presented in the gel image of Figure 4.5b. The middle and rightmost lanes show the λ-terminase digests of 4.3 kbp and 1.9 kbp circular, *cos*-containing substrates, respectively. As described in the previous paragraph, the production of these two molecules differed only in the use of alternate primer sets during the PCR amplification of a *cos*-containing template (see Section A1 for sequences). In this particular gel, the longer 4.3 kbp product (middle lane) features a more pronounced circular population (fainter, slower migrating band) when compared to the 1.9 kbp output, but this varies from sample to sample and may simply reflect less complete progress of the terminase digest in this instance (e.g. as in Fig. 4.4b). Finally, Figure 4.6 and Section A15 (in the Appendix) show the successful incorporation of these terminase-produced, sticky-ended molecules into nanopore sensing experiments involving the recognition of signals from DNA nanostructures attached to their sticky ends, and the separation of translocation events from two molecule lengths by equivalent charge deficit[20,28], respectively.

4.3 – Discussion/Conclusion

In summary, we explored two methods (oligonucleotide assembly and terminase digestion) for the *in vitro* production of sticky-ended DNA for use with nanopore sensing experiments. During the protocol design process, emphasis was placed on the capacity to synthesize products of variable lengths and the ability to carry out individual protocol steps with common molecular biology equipment and widely-available reagents (with one notable exception, discussed below). Both methods were successful in producing molecules that generate robust, easy-to-interpret nanopore signals by virtue of being linear, double-stranded DNA at their cores (whose translocations have been thoroughly studied in the past) and which feature cohesive, single-stranded ends through which application-specific functionality can be integrated. Moreover, the 12-nt length of these cohesive ends was observed to represent an good balance of creating stable hybridized constructs under standard nanopore sensing conditions (e.g. temperature, buffer composition) while also offering the ability to quickly dissociate these constructs and re-set the system simply by moderate cycling of the temperature (see Fig. 4.2b). Lastly, protocols to both methods were designed such that, after using a viral template for an initial run-though,

DNA substrates can thereafter be produced in-house by PCR amplification, eliminating the need to continuously purchase commercial stocks of DNA.

The benefits of choosing one of the synthesis methods presented here over the other may depend on the particular application being targeted. The oligo assembly approach (Method 1) is ultimately the more flexible of the two, with the ability to incorporate specifically-modified oligos at precise locations along the product and even to adjust the sequence of the terminal overhangs, if needed. This however comes at the cost, both financial and temporal, of purchasing and pipetting on the order of ~100 oligo species, whose mixed compositions will vary from molecule to molecule as different products are designed. The terminase digest approach (Method 2), for its part, is simpler in many ways, such as in the number of total reagents required and how many changes to these reagents are needed to make particular product modifications (e.g. replacing a single primer to create a range of product lengths, Section A14 in the Appendix), but this is associated with a loss of design flexibility. For instance, the terminase enzyme acts specifically on the *cos* region, defined above, which limits potential DNA templates to those that include this particular, relatively large sequence. The biggest downside to this method at present, however, may be in the availability of λ-terminase. Unlike the other enzymes encountered above (restriction endonucleases, ligases, polymerases, etc.), λ-terminase is not currently sold as a commercial product (as of 2023). This requires its custom production in smaller batches (e.g. by expression in a recombinant system[36]) which leads to increased cost (and experimental complexity if production is not outsourced), and creates uncertainty over enzyme purity and optimized reaction conditions from batch-to-batch. However, the past availability of λ-terminase for purchase as well as its use in previously-published techniques[34,35] suggest that these are addressable issues, if the enzyme proves particularly useful for an application.

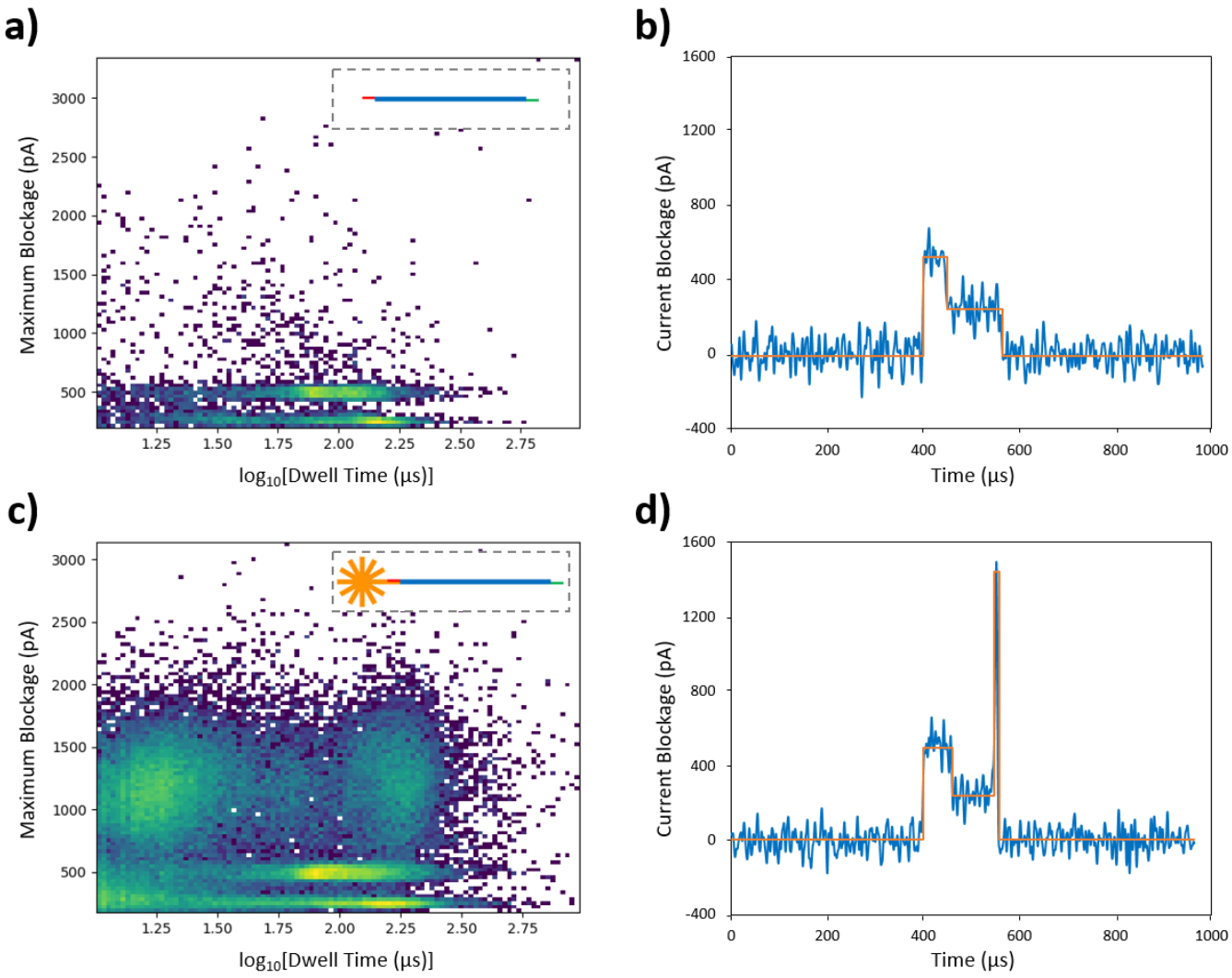


***Figure 4.6**: Nanopore signals of DNA nanostructures attached to synthesized SE-DNA. **a)** 2D histogram of maximum blockage vs. log[dwell time] for nanopore translocation events from a sample of 2-kbp SE-DNA produced by terminase digestion (~13 nm pore, 75 mV transmembrane potential, 3.6 M LiCl). Two main event populations are visible at ~250 pA and ~500 pA max blockage (representing unfolded and folded translocations, respectively). **b)** Representative current trace of a folded event from the dataset in (a) showing the fitted blockage states found in integer multiples of single-file dsDNA translocation. **c)** 2D histogram of max blockage vs. log[dwell time] for events from a sample of 2-kbp SE-DNA incubated with 12-arm DNA stars (25 bp arms) that feature complementary single-stranded extensions, detected with the same pore as in (a). Compared to the plot in (a), at least two additional populations are visible at deeper blockage depths (centred at ~5× single-file dsDNA): relatively short-lived events (dwell times ~20 µs) attributed to the passage of the free stars and relatively long-lived events (dwell times ~200 µs) attributed to the combination of a bulky star and a lengthy SE-DNA molecule into a single hybridized construct. **d)** Representative current trace of a deep, long-lived event from the dataset in (c) showing an event substructure that can be related to the assembly state and conformation of the translocating molecule. In this case, a hybridized SE-DNA + 12-arm star is initially captured by its free (linear) end in a folded state, continues translocation through its unfolded central section and terminates passage with its bulkier star-attached end, resulting in a deeper (and short-lived) final blockage state.*

In the future, several improvements to the main protocols presented here could be envisioned, beyond just the need for optimizing several of the reactions highlighted above (PCR of particular lengths, terminase digest, etc.). For instance, although Method 1 is characterized by the mixing of up to hundreds oligonucleotides to form the complementary strand of each product, strategies to minimize the pipetting effort involved in this could be implemented if the full set of desired products is known in advance. As an example, in the case of multiple product lengths, a "base mix" could be prepared that contained all of the internal oligos of the shortest product; for each additional length in the set, this is supplemented with an “extension mix” containing the internal oligos for that length that are not already contained in the base mix or in the extension mixes of shorter products. Assembly of a particular product length would then simply involve combining its single-stranded scaffold with aliquots of 1) the base mix, 2) extension mixes for that length and all shorter lengths, and 3) its terminal oligos. In this way, stocks of oligos common to multiple sticky-ended products in the set need only be pipetted once per experimental design.

Meanwhile in Method 2, of the four steps illustrated in Figure 4.4a preceding the terminase digest of a circular substrate, only the final three (PCR amplification, restriction digestion, ligase circularization) are necessary to perform for each new batch of products, since only a small (~ng) amount of *cos*-containing template from the first step is needed per amplification. This process could be simplified even further however (from three steps down to one) if a circularized amplicon could be produced directly – this is what has been previously demonstrated in a technique dubbed “ligation during amplification” (LDA)[39,40]. In this technique, a thermostable ligase is incorporated into a modified PCR protocol involving phosphorylated primers and a circular DNA template, where the ligase is used to close circular complements to this template produced in each temperature cycle. If LDA could be successfully performed on *cos*-containing circular substrates, it could therefore represent an improvement on a terminase digest method of SE-DNA production already characterized by its simplicity over Method 1.

Finally, as observed in the gel images of Figures 4.2-4.5, both synthesis methods tend to produce “extra” populations in their outputs (corresponding to molecules that differ in length or in structure from the intended

products) which could translate into unwanted background signals when these samples are run on a nanopore. Both methods may therefore benefit from a purification step prior to nanopore sensing in order to isolate the intended products exclusively. While such purification attempts have proven problematic when based on agarose gel electrophoresis (as detailed above), alternative separation techniques such as size exclusion chromatography[41] and density gradient centrifugation[42] exist and may be worth testing. Although some of these techniques require specialized, possibly institution-shared equipment (e.g. HPLC system, ultracentrifuge), if samples could be pooled such that large quantities of product lasting many nanopore experiments are processed in a single batch, the advantages to working with more purified samples may outweigh the added inconvenience.

With the synthesis methods presented in this chapter, sticky-ended DNA molecules have been prepared that are ready for use as molecular carriers in nanopore experiments. *Chapter 5* details an initial analysis approach to separating populations of these molecules with different lengths and attachment types by the translocation signals they generate. The core methods presented here can also easily be adapted through the incorporation of modified dNTPs[25,26] or oligos[9,12,14] during carrier synthesis to meet the needs of future applications that require specific functional groups along their length. The flexibility to integrate further functionality into these molecules post-assembly also exists through the ssDNA overhangs on both of their ends – these could, for instance, be used as attachment sites for molecular "barcodes" (as defined earlier) to more easily distinguish carriers probing for different targets and therefore expand the capability of the system to multiplex. Finally, although molecular carriers have been the focus of the current chapter, we note that sticky-ended DNA, as a basic construct, may have wider applications to nanopore sensing beyond this. For instance, individual DNA monomers could be chained together by their sticky ends (e.g. circularization[43], concatemerization[44]) to produce nanopore targets with lengths and structures suited for a particular application, starting from more easily synthesized (shorter, linear) building blocks. *Chapter 6* will also explore the use of sticky-ended DNA as probes of hybridization kinetics and temperature by using nanopore sensing to monitor the attachment states of their ends under different experimental conditions.

4.4 – References

(1) Raveendran, M.; Lee, A. J.; Sharma, R.; Wälti, C.; Actis, P. Rational Design of DNA Nanostructures for Single Molecule Biosensing. *Nat Commun* **2020**, *11* (1), 4384. https://doi.org/10.1038/s41467-020-18132-1.

(2) Plesa, C.; van Loo, N.; Ketterer, P.; Dietz, H.; Dekker, C. Velocity of DNA during Translocation through a Solid-State Nanopore. *Nano Lett* **2015**, *15* (1), 732–737. https://doi.org/10.1021/nl504375c.

(3) Xue, L.; Yamazaki, H.; Ren, R.; Wanunu, M.; Ivanov, A. P.; Edel, J. B. Solid-State Nanopore Sensors. *Nat Rev Mater* **2020**, *5* (12), 931–951. https://doi.org/10.1038/s41578-020-0229-6.

(4) Ding, T.; Yang, J.; Pan, V.; Zhao, N.; Lu, Z.; Ke, Y.; Zhang, C. DNA Nanotechnology Assisted Nanopore-Based Analysis. *Nucleic Acids Res* **2020**, *48* (6), 2791–2806. https://doi.org/10.1093/nar/gkaa095.

(5) Plesa, C.; Ruitenberg, J. W.; Witteveen, M. J.; Dekker, C. Detection of Individual Proteins Bound along DNA Using Solid-State Nanopores. *Nano Lett* **2015**, *15* (5), 3153–3158. https://doi.org/10.1021/acs.nanolett.5b00249.

(6) Plesa, C.; Kowalczyk, S. W.; Zinsmeester, R.; Grosberg, A. Y.; Rabin, Y.; Dekker, C. Fast Translocation of Proteins through Solid State Nanopores. *Nano Lett* **2013**, *13* (2), 658–663. https://doi.org/10.1021/nl3042678.

(7) Righetti, P. G.; Caravaggio, T. Isoelectric Points and Molecular Weights of Proteins. *J Chromatogr A* **1976**, *127* (1), 1–28. https://doi.org/10.1016/S0021-9673(00)98537-6.

(8) Sze, J. Y. Y.; Ivanov, A. P.; Cass, A. E. G.; Edel, J. B. Single Molecule Multiplexed Nanopore Protein Screening in Human Serum Using Aptamer Modified DNA Carriers. *Nat Commun* **2017**, *8* (1), 1552. https://doi.org/10.1038/s41467-017-01584-3.

(9) Bell, N. A. W.; Keyser, U. F. Digitally Encoded DNA Nanostructures for Multiplexed, Single-Molecule Protein Sensing with Nanopores. *Nat Nanotechnol* **2016**, *11* (7), 645–651. https://doi.org/10.1038/nnano.2016.50.

(10) Li, J.; Gershow, M.; Stein, D.; Brandin, E.; Golovchenko, J. A. DNA Molecules and Configurations in a Solid-State Nanopore Microscope. *Nat Mater* **2003**, *2* (9), 611–615. https://doi.org/10.1038/nmat965.

(11) Kowalczyk, S. W.; Tuijtel, M. W.; Donkers, S. P.; Dekker, C. Unraveling Single-Stranded DNA in a Solid-State Nanopore. *Nano Lett* **2010**, *10* (4), 1414–1420. https://doi.org/10.1021/nl100271c.

(12) Bell, N. A. W.; Keyser, U. F. Specific Protein Detection Using Designed DNA Carriers and Nanopores. *J Am Chem Soc* **2015**, *137* (5), 2035–2041. https://doi.org/10.1021/ja512521w.

(13) Rothemund, P. W. K. Folding DNA to Create Nanoscale Shapes and Patterns. *Nature* **2006**, *440* (7082), 297–302. https://doi.org/10.1038/nature04586.

(14) Kong, J.; Bell, N. A. W.; Keyser, U. F. Quantifying Nanomolar Protein Concentrations Using Designed DNA Carriers and Solid-State Nanopores. *Nano Lett* **2016**, *16* (6), 3557–3562. https://doi.org/10.1021/acs.nanolett.6b00627.

(15) Kong, J.; Zhu, J.; Keyser, U. F. Single Molecule Based SNP Detection Using Designed DNA Carriers and Solid-State Nanopores. *Chemical Communications* **2017**, *53* (2), 436–439. https://doi.org/10.1039/C6CC08621G.

(16) Chen, K.; Kong, J.; Zhu, J.; Ermann, N.; Predki, P.; Keyser, U. F. Digital Data Storage Using DNA Nanostructures and Solid-State Nanopores. *Nano Lett* **2019**, *19* (2), 1210–1215. https://doi.org/10.1021/acs.nanolett.8b04715.

(17) Chen, K.; Zhu, J.; Bošković, F.; Keyser, U. F. Nanopore-Based DNA Hard Drives for Rewritable and Secure Data Storage. *Nano Lett* **2020**, *20* (5), 3754–3760. https://doi.org/10.1021/acs.nanolett.0c00755.

(18) Zhu, J.; Ermann, N.; Chen, K.; Keyser, U. F. Image Encoding Using Multi-Level DNA Barcodes with Nanopore Readout. *Small* **2021**, *17* (28), 2100711. https://doi.org/10.1002/smll.202100711.

(19) Cai, S.; Sze, J. Y. Y.; Ivanov, A. P.; Edel, J. B. Small Molecule Electro-Optical Binding Assay Using Nanopores. *Nat Commun* **2019**, *10* (1), 1797. https://doi.org/10.1038/s41467-019-09476-4.

(20) Fologea, D.; Brandin, E.; Uplinger, J.; Branton, D.; Li, J. DNA Conformation and Base Number Simultaneously Determined in a Nanopore. *Electrophoresis* **2007**, *28* (18), 3186–3192. https://doi.org/10.1002/elps.200700047.

(21) Storm, A. J.; Storm, C.; Chen, J.; Zandbergen, H.; Joanny, J.-F.; Dekker, C. Fast DNA Translocation through a Solid-State Nanopore. *Nano Lett* **2005**, *5* (7), 1193–1197. https://doi.org/10.1021/nl048030d.

(22) Storm, A. J.; Chen, J. H.; Zandbergen, H. W.; Dekker, C. Translocation of Double-Strand DNA through a Silicon Oxide Nanopore. *Phys Rev E* **2005**, *71* (5), 051903. https://doi.org/10.1103/PhysRevE.71.051903.

(23) Ermann, N.; Hanikel, N.; Wang, V.; Chen, K.; Weckman, N. E.; Keyser, U. F. Promoting Single-File DNA Translocations through Nanopores Using Electro-Osmotic Flow. *J Chem Phys* **2018**, *149* (16), 163311. https://doi.org/10.1063/1.5031010.

(24) Roelen, Z.; Briggs, K.; Tabard-Cossa, V. Analysis of Nanopore Data: Classification Strategies for an Unbiased Curation of Single-Molecule Events from DNA Nanostructures. *ACS Sens* **2023**, *8* (7), 2809–2823. https://doi.org/10.1021/acssensors.3c00751.

(25) Soni, G. V; Singer, A.; Yu, Z.; Sun, Y.; McNally, B.; Meller, A. Synchronous Optical and Electrical Detection of Biomolecules Traversing through Solid-State Nanopores. *Review of Scientific Instruments* **2010**, *81* (1), 014301. https://doi.org/10.1063/1.3277116.

(26) Mai, D. J.; Marciel, A. B.; Sing, C. E.; Schroeder, C. M. Topology-Controlled Relaxation Dynamics of Single Branched Polymers. *ACS Macro Lett* **2015**, *4* (4), 446–452. https://doi.org/10.1021/acsmacrolett.5b00140.

(27) Fologea, D.; Gershow, M.; Ledden, B.; McNabb, D. S.; Golovchenko, J. A.; Li, J. Detecting Single Stranded DNA with a Solid State Nanopore. *Nano Lett* **2005**, *5* (10), 1905–1909. https://doi.org/10.1021/nl051199m.

(28) Bell, N. A. W.; Muthukumar, M.; Keyser, U. F. Translocation Frequency of Double-Stranded DNA through a Solid-State Nanopore. *Phys Rev E* **2016**, *93* (2), 022401. https://doi.org/10.1103/PhysRevE.93.022401.

(29) Zhang, H.; Chao, J.; Pan, D.; Liu, H.; Huang, Q.; Fan, C. Folding Super-Sized DNA Origami with Scaffold Strands from Long-Range PCR. *Chemical Communications* **2012**, *48* (51), 6405. https://doi.org/10.1039/c2cc32204h.

(30) Subramanian, K. The Enzymatic Basis of Processivity in Lambda Exonuclease. *Nucleic Acids Res* **2003**, *31* (6), 1585–1596. https://doi.org/10.1093/nar/gkg266.

(31) Nikiforov, T. T.; Rendle, R. B.; Kotewicz, M. L.; Rogers, Y. H. The Use of Phosphorothioate Primers and Exonuclease Hydrolysis for the Preparation of Single-Stranded PCR Products and Their Detection by Solid-Phase Hybridization. *Genome Res* **1994**, *3* (5), 285–291. https://doi.org/10.1101/gr.3.5.285.

(32) Civit, L.; Fragoso, A.; O'Sullivan, C. K. Evaluation of Techniques for Generation of Single-Stranded DNA for Quantitative Detection. *Anal Biochem* **2012**, *431* (2), 132–138. https://doi.org/10.1016/j.ab.2012.09.003.

(33) Catalano, C. E. The Terminase Enzyme from Bacteriophage Lambda: A DNA-Packaging Machine. *Cell Mol Life Sci* **2000**, *57* (1), 128–148. https://doi.org/10.1007/s000180050503.

(34) Wang, Y.; Wu, R. A New Method for Specific Cleavage of Megabase-Size Chromosomal DNA by λ-Terminase. *Nucleic Acids Res* **1993**, *21* (9), 2143–2147. https://doi.org/10.1093/nar/21.9.2143.

(35) Rackwitz, H. R.; Zehetner, G.; Murialdo, H.; Delius, H.; Chai, J. H.; Poustka, A.; Frischauf, A.; Lehrach, H. Analysis of Cosmids Using Linearization by Phage Lambda Terminase. *Gene* **1985**, *40* (2–3), 259–266. https://doi.org/10.1016/0378-1119(85)90048-4.

(36) Hang, Q.; Woods, L.; Feiss, M.; Catalano, C. E. Cloning, Expression, and Biochemical Characterization of Hexahistidine-Tagged Terminase Proteins. *Journal of Biological Chemistry* **1999**, *274* (22), 15305–15314. https://doi.org/10.1074/jbc.274.22.15305.

(37) Pingoud, A. Structure and Function of Type II Restriction Endonucleases. *Nucleic Acids Res* **2001**, *29* (18), 3705–3727. https://doi.org/10.1093/nar/29.18.3705.

(38) Delidow, B. C. Molecular Cloning of PCR Fragments with Cohesive Ends. *Mol Biotechnol* **1997**, *8* (1), 53–60. https://doi.org/10.1007/BF02762339.

(39) Chen, Z.; Ruffner, D. E. Amplification of Closed Circular DNA in Vitro. *Nucleic Acids Res* **1998**, *26* (23), v–v. https://doi.org/10.1093/nar/26.23.v.

(40) Le, Y.; Chen, H.; Zagursky, R.; Wu, J. H. D.; Shao, W. Thermostable DNA Ligase-Mediated PCR Production of Circular Plasmid (PPCP) and Its Application in Directed Evolution via In Situ Error-Prone PCR. *DNA Research* **2013**, *20* (4), 375–382. https://doi.org/10.1093/dnares/dst016.

(41) Xing, S.; Jiang, D.; Li, F.; Li, J.; Li, Q.; Huang, Q.; Guo, L.; Xia, J.; Shi, J.; Fan, C.; Zhang, L.; Wang, L. Constructing Higher-Order DNA Nanoarchitectures with Highly Purified DNA Nanocages. *ACS Appl Mater Interfaces* **2015**, *7* (24), 13174–13179. https://doi.org/10.1021/am505592e.

(42) Sambrook, J.; Russell, D. W. Purification of Bacteriophage λ Arms: Centrifugation through Sucrose Density Gradients. *Cold Spring Harb Protoc* **2006**, *2006* (1), pdb.prot3990. https://doi.org/10.1101/pdb.prot3990.

(43) Wang, L.; Meng, Z.; Martina, F.; Shao, H.; Shao, F. Fabrication of Circular Assemblies with DNA Tetrahedrons: From Static Structures to a Dynamic Rotary Motor. *Nucleic Acids Res* **2017**, *45* (21), 12090–12099. https://doi.org/10.1093/nar/gkx1045.

(44) Sun, L.; Åkerman, B. Characterization of Self-Assembled DNA Concatemers from Synthetic Oligonucleotides. *Comput Struct Biotechnol J* **2014**, *11* (18), 66–72. https://doi.org/10.1016/j.csbj.2014.08.011.

4.5 – APPENDIX

Section A1: Tables of primer/oligo sequences
("OA" = oligonucleotide assembly method, "TD" = terminase digest method)

OA: 7-kbp scaffold complement

(binds with linearized M13 ssDNA to create double-stranded SE-DNA, bases in red indicate the 12-nt ssDNA extensions, bases in green indicate 40-nt complementary regions across the two final oligos that are used to extend the scaffold)

#	sequence	length
1	GGGCGGCGACCTGACTCTAGAGGATCCCCGGGTACCGAGCTCGAATTCGTAATCATG	57
2	GTCATAGCTGTTTCCTGTGTGAAATTGTTATCCGCTCACA	40
3	ATTCCACACAACATACGAGCCGGAAGCATAAAGTGTAAAG	40
4	CCTGGGGTGCCTAATGAGTGAGCTAACTCACATTAATTGC	40
5	GTTGCGCTCACTGCCCGCTTTCCAGTCGGGAAACCTGTCG	40
6	TGCCAGCTGCATTAATGAATCGGCCAACGCGCGGGGAGAG	40
7	GCGGTTTGCGTATTGGGCGCCAGGGTGGTTTTTCTTTTCA	40
8	CCAGTGAGACGGGCAACAGCTGATTGCCCTTCACCGCCTG	40
9	GCCCTGAGAGAGTTGCAGCAAGCGGTCCACGCTGGTTTGC	40
10	CCCAGCAGGCGAAAATCCTGTTTGATGGTGGTTCCGAAAT	40
11	CGGCAAAATCCCTTATAAATCAAAAGAATAGCCCGAGATA	40
12	GGGTTGAGTGTTGTTCCAGTTTGGAACAAGAGTCCACTAT	40
13	TAAAGAACGTGGACTCCAACGTCAAAGGGCGAAAAACCGT	40
14	CTATCAGGGCGATGGCCCACTACGTGAACCATCACCCAAA	40
15	TCAAGTTTTTTGGGGTCGAGGTGCCGTAAAGCACTAAATC	40
16	GGAACCCTAAAGGGAGCCCCCGATTTAGAGCTTGACGGGG	40
17	AAAGCCGGCGAACGTGGCGAGAAAGGAAGGGAAGAAAGCG	40
18	AAAGGAGCGGGCGCTAGGGCGCTGGCAAGTGTAGCGGTCA	40
19	CGCTGCGCGTAACCACCACACCCGCCGCGCTTAATGCGCC	40
20	GCTACAGGGCGCGTACTATGGTTGCTTTGACGAGCACGTA	40
21	TAACGTGCTTTCCTCGTTAGAATCAGAGCGGGAGCTAAAC	40
22	AGGAGGCCGATTAAAGGGATTTTAGACAGGAACGGTACGC	40
23	CAGAATCCTGAGAAGTGTTTTTATAATCAGTGAGGCCACC	40
24	GAGTAAAAGAGTCTGTCCATCACGCAAATTAACCGTTGTA	40
25	GCAATACTTCTTTGATTAGTAATAACATCACTTGCCTGAG	40
26	TAGAAGAACTCAAACTATCGGCCTTGCTGGTAATATCCAG	40
27	AACAATATTACCGCCAGCCATTGCAACAGGAAAAACGCTC	40
28	ATGGAAATACCTACATTTTGACGCTCAATCGTCTGAAATG	40
29	GATTATTTACATTGGCAGATTCACCAGTCACACGACCAGT	40
30	AATAAAAGGGACATTCTGGCCAACAGAGATAGAACCCTTC	40
31	TGACCTGAAAGCGTAAGAATACGTGGCACAGACAATATTT	40
32	TTGAATGGCTATTAGTCTTTAATGCGCGAACTGATAGCCC	40
33	TAAAACATCGCCATTAAAAATACCGAACGAACCACCAGCA	40
34	GAAGATAAAACAGAGGTGAGGCGGTCAGTATTAACACCGC	40

35	CTGCAACAGTGCCACGCTGAGAGCCAGCAGCAAATGAAAA	40
36	ATCTAAAGCATCACCTTGCTGAACCTCAAATATCAAACCC	40
37	TCAATCAATATCTGGTCAGTTGGCAAATCAACAGTTGAAA	40
38	GGAATTGAGGAAGGTTATCTAAAATATCTTTAGGAGCACT	40
39	AACAACTAATAGATTAGAGCCGTCAATAGATAATACATTT	40
40	GAGGATTTAGAAGTATTAGACTTTACAAACAATTCGACAA	40
41	CTCGTATTAAATCCTTTGCCCGAACGTTATTAATTTTAAA	40
42	AGTTTGAGTAACATTATCATTTTGCGGAACAAAGAAACCA	40
43	CCAGAAGGAGCGGAATTATCATCATATTCCTGATTATCAG	40
44	ATGATGGCAATTCATCAATATAATCCTGATTGTTTGGATT	40
45	ATACTTCTGAATAATGGAAGGGTTAGAACCTACCATATCA	40
46	AAATTATTTGCACGTAAAACAGAAATAAAGAAATTGCGTA	40
47	GATTTTCAGGTTTAACGTCAGATGAATATACAGTAACAGT	40
48	ACCTTTTACATCGGGAGAAACAATAACGGATTCGCCTGAT	40
49	TGCTTTGAATACCAAGTTACAAAATCGCGCAGAGGCGAAT	40
50	TATTCATTTCAATTACCTGAGCAAAAGAAGATGATGAAAC	40
51	AAACATCAAGAAAACAAAATTAATTACATTTAACAATTTC	40
52	ATTTGAATTACCTTTTTTAATGGAAACAGTACATAAATCA	40
53	ATATATGTGAGTGAATAACCTTGCTTCTGTAAATCGTCGC	40
54	TATTAATTAATTTTCCCTTAGAATCCTTGAAAACATAGCG	40
55	ATAGCTTAGATTAAGACGCTGAGAAGAGTCAATAGTGAAT	40
56	TTATCAAAATCATAGGTCTGAGAGACTACCTTTTTAACCT	40
57	CCGGCTTAGGTTGGGTTATATAACTATATGTAAATGCTGA	40
58	TGCAAATCCAATCGCAAGACAAAGAACGCGAGAAAACTTT	40
59	TTCAAATATATTTTAGTTAATTTCATCTTCTGACCTAAAT	40
60	TTAATGGTTTGAAATACCGACCGTGTGATAAATAAGGCGT	40
61	TAAATAAGAATAAACACCGGAATCATAATTACTAGAAAAA	40
62	GCCTGTTTAGTATCATATGCGTTATACAAATTCTTACCAG	40
63	TATAAAGCCAACGCTCAACAGTAGGGCTTAATTGAGAATC	40
64	GCCATATTTAACAACGCCAACATGTAATTTAGGCAGAGGC	40
65	ATTTTCGAGCCAGTAATAAGAGAATATAAAGTACCGACAA	40
66	AAGGTAAAGTAATTCTGTCCAGACGACGACAATAAACAAC	40
67	ATGTTCAGCTAATGCAGAACGCGCCTGTTTATCAACAATA	40
68	GATAAGTCCTGAACAAGAAAAATAATATCCCATCCTAATT	40
69	TACGAGCATGTAGAAACCAATCAATAATCGGCTGTCTTTC	40
70	CTTATCATTCCAAGAACGGGTATTAAACCAAGTACCGCAC	40
71	TCATCGAGAACAAGCAAGCCGTTTTTATTTTCATCGTAGG	40
72	AATCATTACCGCGCCCAATAGCAAGCAAATCAGATATAGA	40
73	AGGCTTATCCGGTATTCTAAGAACGCGAGGCGTTTTAGCG	40
74	AACCTCCCGACTTGCGGGAGGTTTTGAAGCCTTAAATCAA	40
75	GATTAGTTGCTATTTTGCACCCAGCTACAATTTTATCCTG	40
76	AATCTTACCAACGCTAACGAGCGTCTTTCCAGAGCCTAAT	40
77	TTGCCAGTTACAAAATAAACAGCCATATTATTTATCCCAA	40
78	TCCAAATAAGAAACGATTTTTTGTTTAACGTCAAAAATGA	40

79	AAATAGCAGCCTTTACAGAGAGAATAACATAAAAACAGGG	40
80	AAGCGCATTAGACGGGAGAATTAACTGAACACCCTGAACA	40
81	AAGTCAGAGGGTAATTGAGCGCTAATATCAGAGAGATAAC	40
82	CCACAAGAATTGAGTTAAGCCCAATAATAAGAGCAAGAAA	40
83	CAATGAAATAGCAATAGCTATCTTACCGAAGCCCTTTTTA	40
84	AGAAAAGTAAGCAGATAGCCGAACAAAGTTACCAGAAGGA	40
85	AACCGAGGAAACGCAATAATAACGGAATACCCAAAAGAAC	40
86	TGGCATGATTAAGACTCCTTATTACGCAGTATGTTAGCAA	40
87	ACGTAGAAAATACATACATAAAGGTGGCAACATATAAAAG	40
88	AAACGCAAAGACACCACGGAATAAGTTTATTTTGTCACAA	40
89	TCAATAGAAAATTCATATGGTTTACCAGCGCCAAAGACAA	40
90	AAGGGCGACATTCAACCGATTGAGGGAGGGAAGGTAAATA	40
91	TTGACGGAAATTATTCATTAAAGGTGAATTATCACCGTCA	40
92	CCGACTTGAGCCATTTGGGAATTAGAGCCAGCAAAATCAC	40
93	CAGTAGCACCATTACCATTAGCAAGGCCGGAAACGTCACC	40
94	AATGAAACCATCGATAGCAGCACCGTAATCAGTAGCGACA	40
95	GAATCAAGTTTGCCTTTAGCGTCAGACTGTAGCGCGTTTT	40
96	CATCGGCATTTTCGGTCATAGCCCCCTTATTAGCGTTTGC	40
97	CATCTTTTCATAATCAAAATCACCGGAACCAGAGCCACCA	40
98	CCGGAACCGCCTCCCTCAGAGCCGCCACCCTCAGAACCGC	40
99	CACCCTCAGAGCCACCACCCTCAGAGCCGCCACCAGAACC	40
100	ACCACCAGAGCCGCCGCCAGCATTGACAGGAGGTTGAGGC	40
101	AGGTCAGACGATTGGCCTTGATATTCACAAACAAATAAAT	40
102	CCTCATTAAAGCCAGAATGGAAAGCGCAGTCTCTGAATTT	40
103	ACCGTTCCAGTAAGCGTCATACATGGCTTTTGATGATACA	40
104	GGAGTGTACTGGTAATAAGTTTTAACGGGGTCAGTGCCTT	40
105	GAGTAACAGTGCCCGTATAAACAGTTAATGCCCCCTGCCT	40
106	ATTTCGGAACCTATTATTCTGAAACATGAAAGTATTAAGA	40
107	GGCTGAGACTCCTCAAGAGAAGGATTAGGATTAGCGGGGT	40
108	TTTGCTCAGTACCAGGCGGATAAGTGCCGTCGAGAGGGTT	40
109	GATATAAGTATAGCCCGGAATAGGTGTATCACCGTACTCA	40
110	GGAGGTTTAGTACCGCCACCCTCAGAACCGCCACCCTCAG	40
111	AACCGCCACCCTCAGAGCCACCACCCTCATTTTCAGGGAT	40
112	AGCAAGCCCAATAGGAACCCATGTACCGTAACACTGAGTT	40
113	TCGTCACCAGTACAAACTACAACGCCTGTAGCATTCCACA	40
114	GACAGCCCTCATAGTTAGCGTAACGATCTAAAGTTTTGTC	40
115	GTCTTTCCAGACGTTAGTAAATGAATTTTCTGTATGGGAT	40
116	TTTGCTAAACAACTTTCAACAGTTTCAGCGGAGTGAGAAT	40
117	AGAAAGGAACAACTAAAGGAATTGCGAATAATAATTTTTT	40
118	CACGTTGAAAATCTCCAAAAAAAAGGCTCCAAAAGGAGCC	40
119	TTTAATTGTATCGGTTTATCAGCTTGCTTTCGAGGTGAAT	40
120	TTCTTAAACAGCTTGATACCGATAGTTGCGCCGACAATGA	40
121	CAACAACCATCGCCCACGCATAACCGATATATTCGGTCGC	40
122	TGAGGCTTGCAGGGAGTTAAAGGCCGCTTTTGCGGGATCG	40

123	TCACCCTCAGCAGCGAAAGACAGCATCGGAACGAGGGTAG	40
124	CAACGGCTACAGAGGCTTTGAGGACTAAAGACTTTTTCAT	40
125	GAGGAAGTTTCCATTAAACGGGTAAAATACGTAATGCCAC	40
126	TACGAAGGCACCAACCTAAAACGAAAGAGGCAAAAGAATA	40
127	CACTAAAACACTCATCTTTGACCCCCAGCGATTATACCAA	40
128	GCGCGAAACAAAGTACAACGGAGATTTGTATCATCGCCTG	40
129	ATAAATTGTGTCGAAATCCGCGACCTGCTCCATGTTACTT	40
130	AGCCGGAACGAGGCGCAGACGGTCAATCATAAGGGAACCG	40
131	AACTGACCAACTTTGAAAGAGGACAGATGAACGGTGTACA	40
132	GACCAGGCGCATAGGCTGGCTGACCTTCATCAAGAGTAAT	40
133	CTTGACAAGAACCGGATATTCATTACCCAAATCAACGTAA	40
134	CAAAGCTGCTCATTCAGTGAATAAGGCTTGCCCTGACGAG	40
135	AAACACCAGAACGAGTAGTAAATTGGGCTTGAGATGGTTT	40
136	AATTTCAACTTTAATCATTGTGAATTACCTTATGCGATTT	40
137	TAAGAACTGGCTCATTATACCAGTCAGGACGTTGGGAAGA	40
138	AAAATCTACGTTAATAAAACGAACTAACGGAACAACATTA	40
139	TTACAGGTAGAAAGATTCATCAGTTGAGATTTAGGAATAC	40
140	CACATTCAACTAATGCAGATACATAACGCCAAAAGGAATT	40
141	ACGAGGCATAGTAAGAGCAACACTATCATAACCCTCGTTT	40
142	ACCAGACGACGATAAAAACCAAAATAGCGAGAGGCTTTTG	40
143	CAAAAGAAGTTTTGCCAGAGGGGGTAATAGTAAAATGTTT	40
144	AGACTGGATAGCGTCCAATACTGCGGAATCGTCATAAATA	40
145	TTCATTGAATCCCCCTCAAATGCTTTAAACAGTTCAGAAA	40
146	ACGAGAATGACCATAAATCAAAAATCAGGTCTTTACCCTG	40
147	ACTATTATAGTCAGAAGCAAAGCGGATTGCATCAAAAAGA	40
148	TTAAGAGGAAGCCCGAAAGACTTCAAATATCGCGTTTTAA	40
149	TTCGAGCTTCAAAGCGAACCAGACCGGAAGCAAACTCCAA	40
150	CAGGTCAGGATTAGAGAGTACCTTTAATTGCTCCTTTTGA	40
151	TAAGAGGTCATTTTTGCGGATGGCTTAGAGCTTAATTGCT	40
152	GAATATAATGCTGTAGCTCAACATGTTTTAAATATGCAAC	40
153	TAAAGTACGGTGTCTGGAAGTTTCATTCCATATAACAGTT	40
154	GATTCCCAATTCTGCGAACGAGTAGATTTAGTTTGACCAT	40
155	TAGATACATTTCGCAAATGGTCAATAACCTGTTTAGCTAT	40
156	ATTTTCATTTGGGGCGCGAGCTGAAAAGGTGGCATCAATT	40
157	CTACTAATAGTAGTAGCATTAACATCCAATAAATCATACA	40
158	GGCAAGGCAAAGAATTAGCAAAATTAAGCAATAAAGCCTC	40
159	AGAGCATAAAGCTAAATCGGTTGTACCAAAAACATTATGA	40
160	CCCTGTAATACTTTTGCGGGAGAAGCCTTTATTTCAACGC	40
161	AAGGATAAAAATTTTTAGAACCCTCATATATTTTAAATGC	40
162	AATGCCTGAGTAATGTGTAGGTAAAGATTCAAAAGGGTGA	40
163	GAAAGGCCGGAGACAGTCAAATCACCATCAATATGATATT	40
164	CAACCGTTCTAGCTGATAAATTAATGCCGGAGAGGGTAGC	40
165	TATTTTTGAGAGATCTACAAAGGCTATCAGGTCATTGCCT	40
166	GAGAGTCTGGAGCAAACAAGAGAATCGATGAACGGTAATC	40

167	GTAAAACTAGCATGTCAATCATATGTACCCCGGTTGATAA	40
168	TCAGAAAAGCCCCAAAAACAGGAAGATTGTATAAGCAAAT	40
169	ATTTAAATTGTAAACGTTAATATTTTGTTAAAATTCGCAT	40
170	TAAATTTTTGTTAAATCAGCTCATTTTTTAACCAATAGGA	40
171	ACGCCATCAAAAATAATTCGCGTCTGGCCTTCCTGTAGCC	40
172	AGCTTTCATCAACATTAAATGTGAGCGAGTAACAACCCGT	40
173	CGGATTCTCCGTGGGAACAAACGGCGGATTGACCGTAATG	40
174	GGATAGGTCACGTTGGTGTAGATGGGCGCATCGTAACCGT	40
175	GCATCTGCCAGTTTGAGGGGACGACGACAGTATCGGCCTC	40
176	AGGAAGATCGCACTCCAGCCAGCTTTCCGGCACCGCTTCT	40
177	GGTGCCGGAAACCAGGCAAAGCGCCATTCGCCATTCAGGC	40
178	TGCGCAACTGTTGGGAAGGGCGATCGGTGCGGGCCTCTTC	40
179	GCTATTACGCCAGCTGGCGAAAGGGGGATGTGCTGCAAGG	40
180	CGATTAAGTTGGGTAACGCCAGGGTTTTCCCAGTCACGACGTTG	44
181	TAAAACGACGGCCAGTGCCAAGCTTGCATGCCTGCAGGTCATCACTTTACGGGTCCTTTCCGGTGATCCGACAGGTTACG	80
182	AGGTCGCCGCCCCGTAACCTGTCGGATCACCGGAAAGGACCCGTAAAGTGAT	52

OA: M13-HincII oligo

(binds to region of M13 ssDNA to create a double-stranded HincII recognition sequence – in blue)

CTGCAGGTCGACTCTAGAGG

OA: 2-kbp scaffold complement

(use oligos #106-159 & #182 of the 7-kbp set, plus the following four)

'PS24_P105'

GGGCGGCGACCTGCCTTGAGTAACAGTGCCCGTATAAACAGTTAATGCCCCCTGCCT

'PS24_P160'

CCCTGTAATACTTTTGCGGGAGAAGCCTTTATTTCAACGCAAGG

'PS24_P161'

ATAAAAATTTTTAGAACCCTCATATATTTTAAATGCAATGCCTGA

'PS24_P162'

GTAATGTGTAGGTAAAGATTCAAAAGGGTGAGAAAGGCCGATCACTTTACGGGTCCTTTCCGGTGATCCGACAGGTTACG

OA: 2-kbp scaffold primers

(produce 2-kbp PCR amplicon to be digested with λ exonuclease to create a ssDNA scaffold, "*Phos*" indicates 5'-phosphrylation, asterisks "*" indicate replacement of phosphodiester bonds with phosphorothiote bonds)

'OA-PS-S5-L'

C*G*G*C*C*TTTCTCACCCTTTTGAATC

'5P-OA-PS-S5-2.4-R'

Phos-GCCTTGAGTAACAGTGCCCGTATA

OA: 5 kbp scaffold primers

(use ‘OA-PS-S5-L’ as left primer plus following oligo as right primer to produce 5-kbp PCR amplicon for later λ exonuclease digest)

‘5P-OA-PS-S5-4.8-R’

Phos-GCGGAACAAAGAAACCACCAGAAG

TD: 10-kbp template primers

(produce 11-kbp PCR amplicon to be digested with PciI to create 10-kbp sticky ended template for subsequent circularization and λ terminase digest)

‘λC-PF2-L’

CAAGTTGCTGCGATTCTCACCAAT

‘λC-PF2-R’

TTCAATACGCTTGTCACCCAGGAA

TD: 10-kbp short end template primers

(produce 10-kbp amplicon with short gaps < 100 bp between ends of amplicon and PciI recognition sites)

‘λC-SE1-L’

CACCATCAGCCAGAAAACCGAATT

‘λC-SE1-R’

TATTCACTCAGCAACCCCGGTATC

TD: 4-kbp template primers

(use ‘λC-SE1-R’ as right primer plus following oligo as left primer to produce 4-kbp amplicon for subsequent circularization and λ terminase digest, underlined base indicates point mutation in primer to introduce PciI recognition site in orange into amplicon sequence)

‘λC-M4-SE-L’

TGCACAACATGTAAGAGCATTGAGT

TD: 2-kbp template primers

(use ‘λC-SE1-R’ as right primer plus following oligo as left primer to produce 2-kbp amplicon for subsequent circularization and λ terminase digest, underlined base indicates point mutation in primer to introduce PciI recognition site in orange into amplicon sequence)

‘λC-M2-SE-L’

CCCTGTATTGCTGACATGTGATTTCTC

Section A2: Detailed synthesis protocols
("OA" = oligonucleotide assembly method, "TD" = terminase digest method)

A) OA: M13 Linearization

1. Combine in tube:
    - 24 µL M13mp18 ssDNA (~112 nM, New England Biolabs)
    - 20 µL NEBuffer 3.1 (10×, NEB)
    - 10 µL M13-HincII oligo (50 µM, Integrated DNA technologies)
    - 142 µL nuclease-free water (Invitrogen)
2. Incubate at 65 °C for 10 min, slowly cool to RT (~1 °C/min)
3. Add 4 µL HincII (10 U/µL, NEB) restriction enzyme to tube
4. Incubate at 37 °C for 1+ hrs, heat inactivate enzyme at 65 °C for 10 min
5. Recover DNA scaffold with spin columns (e.g. Monarch PCR & DNA Cleanup Kit, NEB) per manufacturer's instructions

B) OA: Phosphorylation of Internal Oligos

1. Combine in tube on ice:
    - 24 µL internal oligo mix (100 µM total, IDT)
    - 40 µL T4 DNA ligase buffer (10×, NEB)
    - 8 µL T4 polynucleotide kinase (10 U/µL, NEB)
    - 328 µL nuclease-free water (Invitrogen)
2. Incubate at 37 °C for 1+ hrs, heat inactivate enzyme at 65 °C for 20 min
3. Recover phosphorylated oligos with spin columns (e.g. Monarch PCR & DNA Cleanup Kit, NEB) per manufacturer's instructions

C) OA: Internal Oligo Hybridization and Ligation

1. Combine in tube:
    - 10 µL linearized M13 ssDNA (≳ 100 nM, from reaction A)
    - 30 µL phosphorlyated oligos (≳ 200 nM each, from reaction B)
    - 8 µL Taq DNA ligase buffer (10×, NEB)
    - 32 µL nuclease-free water (Invitrogen)
2. Incubate at 70 °C for 10 min, slowly cool to RT (~1 °C/min)
3. Add to tube:
    - 2 µL Taq DNA ligase buffer (10×, NEB)
    - 4 µL Taq DNA ligase (40 U/µL, NEB)
    - 14 µL nuclease-free water (Invitrogen)
4. Incubate at 45 °C overnight
5. Recover DNA product with spin columns (e.g. PureLink PCR Purification Kit, Invitrogen) per manufacturer's instructions

D) OA: Terminal Oligo Hybridization and Ligation

1. Combine in tube:
   - 50 µL M13 assembly (≳ 10 nM, from reaction C)
   - 4 µL internal oligo mix (4 µM total, IDT)
   - 8 µL Taq DNA ligase buffer (10×, NEB)
   - 18 µL nuclease-free water (Invitrogen)
2. Incubate at 70 °C for 10 min, slowly cool to RT (~1 °C/min)
3. Add to tube:
   - 2 µL Taq DNA ligase buffer (10×, NEB)
   - 4 µL Taq DNA ligase (40 U/µL, NEB)
   - 14 µL nuclease-free water (Invitrogen)
4. Incubate at 45 °C for 1+ hrs
5. Recover sticky-ended product with spin columns (e.g. PureLink PCR Purification Kit, Invitrogen) per manufacturer's instructions

E) OA: Scaffold Synthesis – Polymerase Chain Reaction

1. Combine in tube on ice:
   - 2 µL dNTPs (10 mM, NEB)
   - 5 µL phosphorothioated forward primer (10 µM, IDT)
   - 5 µL phosphorylated reverse primer (10 µM, IDT)
   - 2 µL M13-DNA template (10 ng/µL, NEB)
   - 1 µL Q5 DNA polymerase (2 U/µL, NEB)
   - 20 µL Q5 reaction buffer (5×, NEB)
   - 65 µL nuclease-free water (Invitrogen)
2. Place in thermocycler:

Time	Temperature	Cycles
30 s	98 °C	×1
10 s 30 s 1 min 33 s	98 °C 69 °C 72 °C	×30
2 min	72 °C	×1

3. Recover PCR product with spin columns (e.g. PureLink PCR Purification Kit, Invitrogen) per manufacturer's instructions

F) OA: Scaffold Synthesis – Digest of Phosphorylated Strand

1. Combine in tube on ice:
   - 35 µL PCR amplicon (≳ 100 ng/µL, from reaction E)
   - 1 µL λ-exonuclease (5 U/µL, NEB)
   - 5 µL λ-exonuclease buffer (5×, NEB)
   - 9 µL nuclease-free water (Invitrogen)
2. Incubate at 37 °C for 1+ hrs, add EDTA to 10 mM, heat inactivate enzyme at 75 °C for 10 min

3. Recover ssDNA product with spin columns (e.g. Monarch PCR & DNA Cleanup Kit, NEB) per manufacturer's instructions

G) TD: λ Digestion and *cos* Ligation

1. Combine in tube on ice:
   - 2 µL λ-DNA (500 ng/µL, NEB)
   - 1 µL NAD^+ (50 mM, NEB)
   - 1 µL Bsp1407I restriction enzyme (10 U/µL Thermo Scientific)
   - 2 µL Taq DNA ligase (40 U/µL, NEB)
   - 5 µL Tango buffer (10×, Thermo Scientific)
   - 39 µL nuclease-free water (Invitrogen)
2. Incubate at 40 °C for 1+ hrs
3. Recover digest fragments with spin columns (e.g. PureLink PCR Purification Kit, Invitrogen) per manufacturer's instructions

H) TD: Amplification of *cos*-containing Fragment

1. Combine in tube on ice:
   - 3 µL dNTPs (10 mM, New England Biolabs)
   - 4 µL forward primer (10 µM, IDT)
   - 4 µL reverse primer (10 µM, IDT)
   - 2 µL template DNA (~10 ng/µL, from reaction G)
   - 4 µL LongAmp Taq polymerase (2.5 U/µL, NEB)
     20 µL LongAmp reaction buffer (5×, NEB)
   - 63 µL nuclease-free water (Invitrogen)
2. Place in thermocycler:

Time	Temperature	Cycles
30 s	94 °C	×1
15 s 30 s 10 min	94 °C 60 °C 65 °C	×30
10 min	65 °C	×1

3. Recover DNA product with spin column (e.g. PureLink PCR Purification Kit, Invitrogen) per manufacturer's instructions

I) TD: Fragment Circularization – Restriction Digest

1. Combine in tube on ice:
   - 20 µL *cos*-containing amplicon (≳ 50 ng/µL, from reaction H)
   - 1 µL PciI restriction enzyme (10 U/µL, NEB)
   - 5 µL NEBuffer 3.1 (10×, NEB)
   - 24 µL nuclease-free water (Invitrogen)
2. Incubate at 37 °C for 1+ hrs

3. Recover digest fragment with spin columns (e.g. PureLink PCR Purification Kit, Invitrogen) per manufacturer's instructions

J) TD: Fragment Circularization – Ligation

1. Combine in tube on ice:
   - 30 µL *cos*-containing digest fragment (~15 ng/µL, from reaction I)
   - 5 µL T4 DNA ligase (5 U/µL, Thermo Scientific)
   - 40 µL T4 ligase buffer (10×, NEB)
   - 325 µL nuclease-free water (Invitrogen)
2. Incubate at 16 °C overnight
3. Purify DNA by phenol-chloroform extraction:
   - Mix ligation reaction with equal volume phenol-chloroform (25:24:1 phenol/chloroform/ isoamyl alcohol, Acros Organics)
   - Centrifuge at 15,000 × *g* for 15 min, transfer aqueous phase to fresh tube
   - Mix aqueous phase with equal volume chloroform ()
   - Centrifuge at 15,000 × *g* for 15 min, transfer aqueous phase to fresh tube
4. Concentrate DNA by ethanol precipitation:
   - Mix phenol-chloroform extraction with 2.5× volume ethanol and 0.1× volume sodium acetate (3 M, pH 5.2, Thermo Scientific)
   - Incubate at -20 °C for 1 hr
   - Centrifuge at 15,000 × *g* for 15 min, remove supernatant
   - Add 1 mL 70% ethanol, centrifuge at 15,000 × *g* for 15 min, remove supernatant
   - Air dry resulting pellet, re-suspend in 20 µL TE buffer (Invitrogen)

K) TD: λ-Terminase Digestion

1. Combine in tube on ice:
   - 30 µL circularized template DNA (≳ 10 nM, from reaction J)
   - 2 µL integration host factor (3 µM, Excellgen)
   - 2 µL λ-terminase (1.5 µM, Catalano Lab – University of Colorado)
   - 6 µL T4 DNA ligase buffer (10×, NEB)
   - 20 µL nuclease-free water (Invitrogen)
2. Incubate at 37 °C for 1 hr
3. Analyze aliquot of reaction (e.g. 4 µL) by gel electrophoresis (0.7% agarose, 0.5× TBE, 100 V) – assess reaction progress by comparing intensities of circular and linear gel bands
4. If reaction progress is unsatisfactory (e.g. < 70% digested), supplement reaction mixture with additional terminase (to 50 mM increment in concentration), repeat steps 2-4 as necessary
5. Recover sticky-ended product with spin columns (e.g. PureLink PCR Purification Kit, Invitrogen) per manufacturer's instructions

Section A3: Comparison of oligo assembly product with linear, open circle, and supercoiled M13 on gel

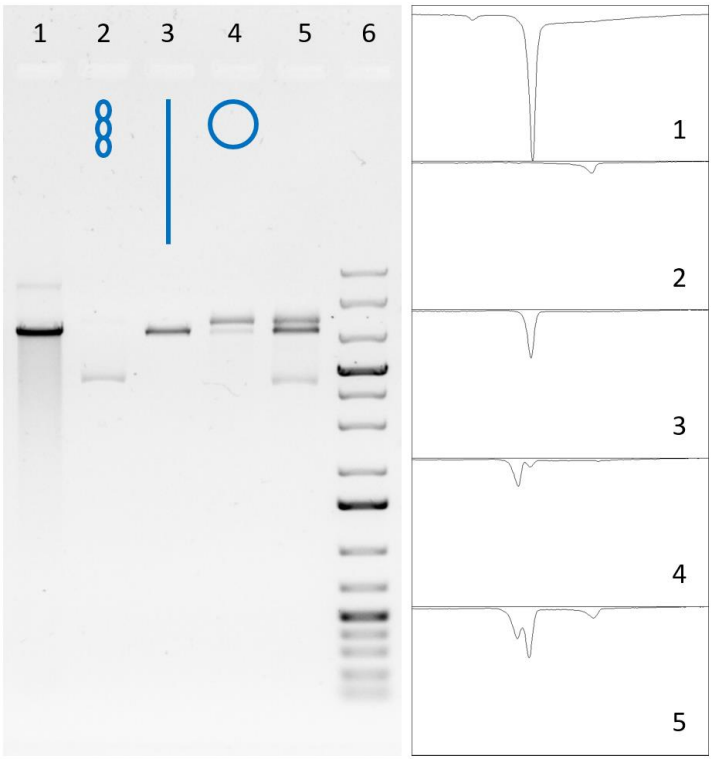


Figure 4A.1: *Agarose gel (0.5%, 0.5× TAE, pre-stained with GelRed, 70 V) of SE-DNA synthesized via oligo assembly onto a 7.25 knt ssDNA scaffold. Lane 1: Annealed, ligated mix of linearized M13mp18 + internal complementary oligos (×180) + terminal oligo extensions (×2), subsequently spin-column purified (PureLink Quick PCR Purification Kit, Invitrogen). Lane 2: M13mp18-RFI (supercoiled circular dsDNA). Lane 3: HincII digest on M13mp18-RFI (linear dsDNA). Lane 4: Nb.BsmI digest on M13mp18-RFI (relaxed circular dsDNA). Lane 5: Equal volume mix of Lanes 2-4. Lane 6: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). The final oligo assembly product is characterized by i) a dominant band migrating in line with linear, double stranded M13, ii) a faint, slow migrating band consistent with molecular weight of ~2× that of linear monomer (dimer), iii) a diffuse smear of faster migrating populations possibly indicative of truncated or incompletely assembled products. The panel to the right represents the integrated band intensities along each numbered lane.*

Section A4: Gel of ssDNA scaffold production from dsDNA template (λ-exo digest)

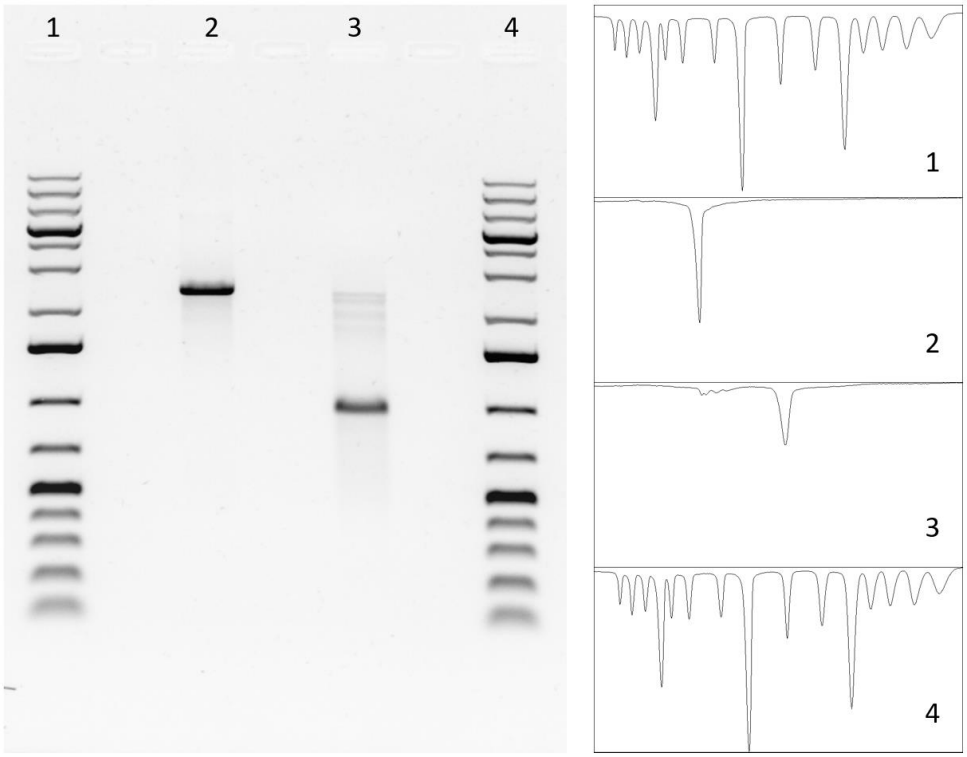


Figure 4A.2*: Agarose gel (0.7%, 0.5× TBE, pre-stained with GelRed, 70 V) documenting the digestion of a single strand of PCR-generated dsDNA template to produce a ssDNA scaffold for oligonucleotide assembly. Lanes 1 & 4: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). Lane 2: PCR amplicon (2.3-kbp) produced using a 5'-phosphorothioated forward primer and a 5'-phosphorylated reverse primer (M13 DNA template). Lane 3: λ-exonuclease digest on the 2.3 kbp PCR amplicon. The panel to the right represents the integrated band intensities along each numbered lane. Comparing pre- and post-digestion gel patterns, the main population shifts from migrating between the 2- and 3-kbp bands of the ladder to a faster band near ~1 kbp, consistent with a transition from relatively rigid dsDNA to an equivalent length of relatively flexible ssDNA. Incubating the output of the λ-exo digest with a full suite of oligos complementary to the 5'-phosphorothioated strand results in the restoration of the pre-digest gel pattern (see Figure 4.3b in the main text), indicating that the 5'-phosphorylated strand was indeed preferentially digested.*

Section A5: Test of end activity from oligo assembled SE-DNA (scaffold produced with λ-exo digest)

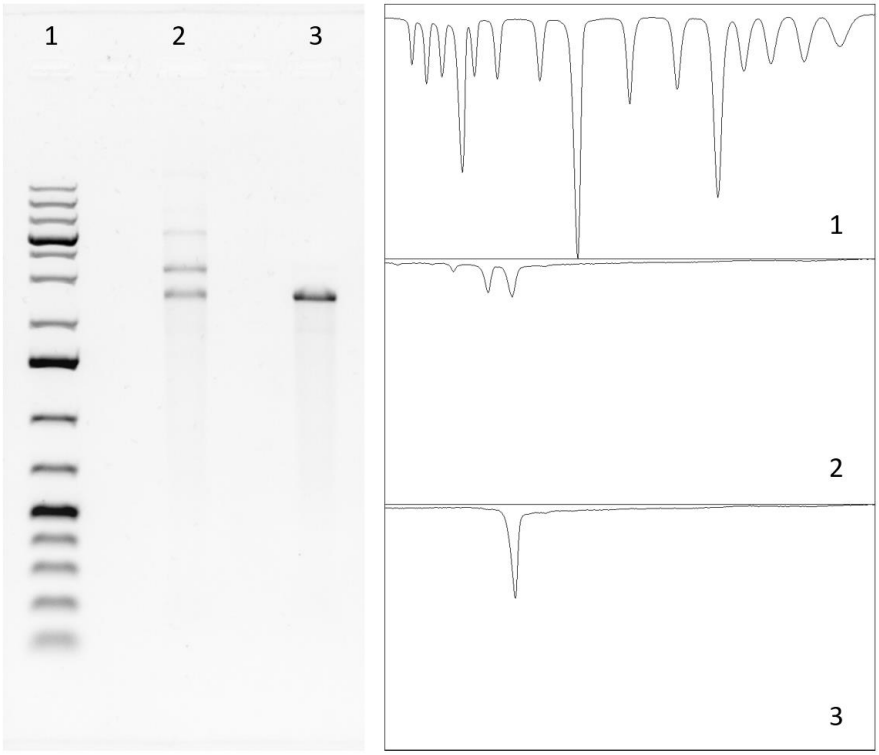


Figure 4A.3: *Agarose gel (0.7%, 0.5× TBE, pre-stained with GelRed, 70 V) of 2.3 kbp SE-DNA resulting from staple strand assembly to a ssDNA scaffold produced by PCR and lambda exonuclease digestion. Lane 1: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). Lane 2: incubation of SE-DNA product at 50 °C (2 hrs) in 10 mM $MgCl_2$, Lane 3: incubation of SE-DNA product at 50 °C (2 hrs) in 10 mM $MgCl_2$ followed by a 10 min. incubation at 70 °C and flash cooling. Higher-order combinations of the sticky ends (circularized monomers, dimers, etc.) are reversibly melted by this 70 °C treatment. The panel to the right represents the integrated band intensities along each numbered lane.*

Section A6: Gels of intermediate synthesis steps (terminase digest method)

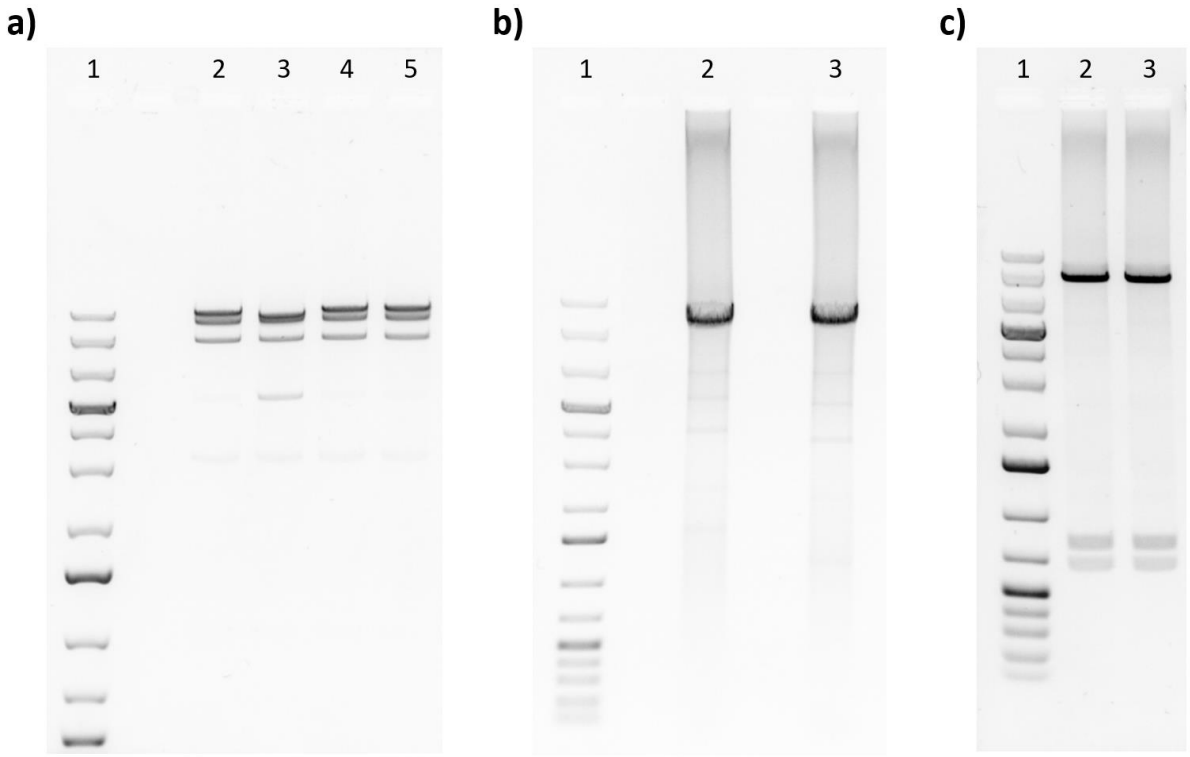


Figure 4A.4: a) *Agarose gel (0.7%, 0.5× TAE, pre-stained with GelRed, 70 V) of restriction endonuclease digest (BsrGI) on λ-DNA, confirming the effect of adding a ligase (Taq) to the reaction mixture to repair nicks in the annealed cohesive (cos) ends. Lane 1: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). Lane 2: BsrGI digest on λ-DNA. Lane 3: BsrGI digest on λ-DNA, heated to 70°C and rapidly cooled. Lane 4: BsrGI digest + Taq ligation on λ-DNA. Lane 5: BsrGI digest + Taq ligation on λ-DNA, heated to 70°C and rapidly cooled. In the presence of Taq ligase, the ~21 kbp fragment containing the cos sequence is stable against heating past the melting point of the 12-nt overhangs – without it (Lane 4) the fragment melts apart into smaller 16 kbp and 5 kbp pieces.* ***b)*** *Agarose gel (0.5%, 0.5× TAE, pre-stained with GelRed, 70 V) of polymerase chain reactions that use the ligated, cos-containing fragment from (a) as a template to produce ~11 kbp products. Lanes 2 & 3 correspond to two different choices of primer set, both of which appeared successful in creating a single dominant PCR product (NEB LongAmp Taq polymerase, 30 cycles) that migrates just above the 10-kbp band of the DNA ladder.* ***c)*** *Agarose gel (0.7%, 0.5× TAE, pre-stained with GelRed, 70 V) of a PciI restriction digest on the PCR product of Lane 3 in (b) – Lanes 2 & 3 in (c) are duplicates. Two new bands appear, corresponding to the cleaved 688 bp and 820 bp ends, leaving cohesive 4-nt sticky ends on the remaining ~9.7 kbp fragment (to be used in the subsequent circularization step).*

Section A7: Complications of ligating *cos*-containing DNA multimers (palindromic restriction enzymes)

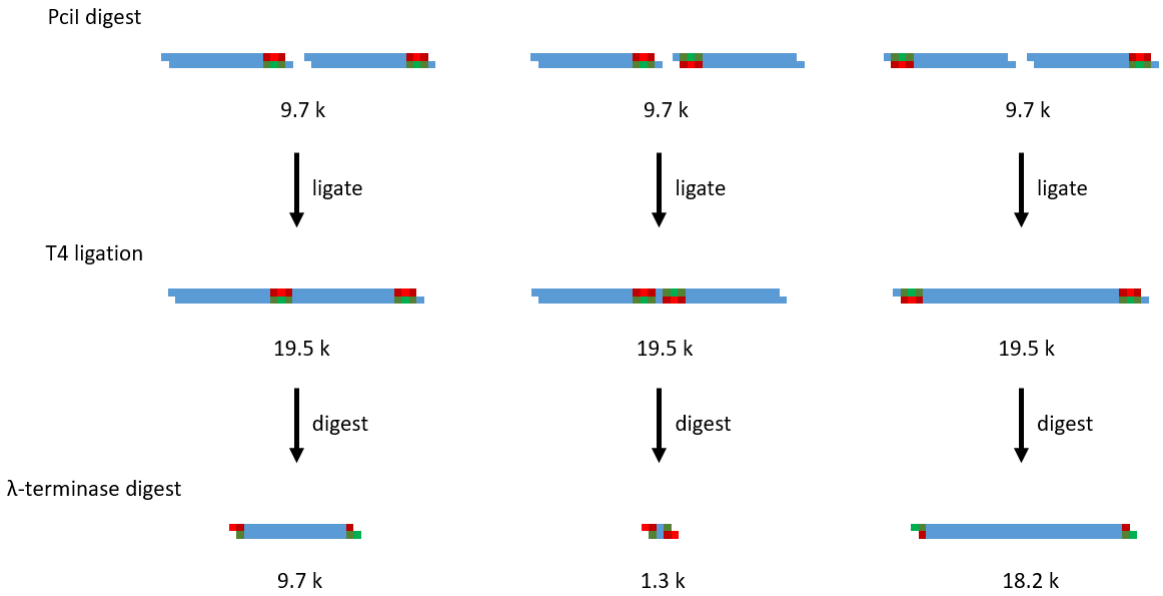


Figure 4A.5: *An illustration of a complication that arises when ligating multimers together of cos-containing DNA fragments cleaved at both ends with a palindromic restriction endonuclease. Each ligation has only a 50% chance of attaching two fragments together in the desired orientation (leftmost column in figure), leading to a variety of possible fragment lengths after the subsequent λ-terminase digestion step. These final products (of different lengths) are also expected to differ in the sequences of their 12-nt sticky ends – fragments from "properly oriented" ligations will have two different (cohesive) ends while "misoriented" ligations will lead to fragments with identical 5' extensions on both sides.*

Section A8: Estimating DNA concentrations to circularize linear monomers

To increase the proportion of circular monomers formed during the T4 ligation step of Method 2 (see Fig. 4.4a), the initial concentration of (linear) sticky-ended DNA templates in the reaction can be lowered such that the rate at which two ends from different individual molecules encounter each other becomes smaller and smaller in comparison to the encounter rate of two ends from the same molecule.

One way of estimating the relevant concentration range of this situation is to model the conformations of the DNA molecules as random walks ("freely-jointed chains") of *N* segments of Kuhn length *b* (for a total contour length *L* = *Nb*). The probabilities of end-to-end distances *R* for the molecules are then given by the following Gaussian distribution[1]:

$$p(\boldsymbol{R}) = \left(\frac{3}{2\pi Lb}\right)^{3/2} e^{-3R^2/2Lb} \tag{A1}$$

In order to ligate a DNA molecule into a circle, its two ends need to find themselves within some small distance $\delta << \sqrt{N}b$ of each other. This will occur with probability:

$$p_C = \int_{R=0}^{R=\delta} p(\boldsymbol{R}) \cdot dV = \int_0^{\delta} \left(\frac{3}{2\pi Lb}\right)^{3/2} e^{-3R^2/2Lb} \cdot (4\pi R^2 dR) \approx \int_0^{\delta} \left(\frac{3}{2\pi Lb}\right)^{3/2} \cdot (4\pi R^2 dR) \tag{A2}$$

Where $e^{-3R^2/2Lb} \approx 1$ over the range [0, δ] since $\delta << \sqrt{N}b = \sqrt{Lb}$.

The factor $\left(\frac{3}{2\pi Lb}\right)^{3/2}$ then (with units of inverse volume) represents the effective concentration of one end in the vicinity in the other – it is this quantity that is being integrated within small volumes centred on one end to find the probability of the other end being nearby. It is frequently called the Jacobson-Stockmayer factor, *j*, after an early paper on polymer cyclization[2].

$$j = \left(\frac{3}{2\pi Lb}\right)^{3/2} \tag{A3}$$

If during a ligation reaction this effective concentration of ends from the same molecule, *j,* exceeds the bulk concentration of free ends in solution, *i*, then the closing of a single molecule into a circle is predicted to predominate over connecting ends from separate molecules. Conversely if *i* > *j*, then multimerization would be favoured. Some authors[3] have estimated the proportion of monomeric DNA circles expected in a ligation reaction by considering the ratio of *j* / *i* to be equal to the expected ratio of ligated monomers vs. multimers. This leads to a prediction of[3]:

$$\%\ circles \stackrel{?}{=} \frac{j}{i+j} \times 100 \tag{A4}$$

However, this seems to underestimate the magnitude of this quantity (based on concentration considerations alone), as we would not expect the bulk concentration of sticky ends *i*, to remain fixed over the course of a ligation reaction (since each successful ligation removes two ends from the mix). Circles should therefore become and more and more favoured as the reaction proceeds and *i*(*t*) decreases. For instance, in the extreme limit of having only two linear monomers left at the end of the reaction, it is extremely unlikely they would be able to find each other within a macroscopic volume before circularizing, assuming undamaged, active ends.

The value defined on the right hand side of Equation A4, however, could map well to the *initial* proportion of circular monomers formed (when *i* takes on its initial value) at least, so it is here defined as a proxy measure of monomer circularization that we will call the "circularization factor", f_C:

$$f_C \equiv \frac{j}{i+j} \tag{A5}$$

For a given *minimum* desired value of f_C (e.g. 95%) during a ligation reaction, there is then a corresponding *maximum* allowed bulk concentration of initial ends, i_{max}. This maximum *i* also depends on the molecule-specific value of *j*. As the DNA contour length *L* gets longer for instance, the two ends of a molecule have a harder time finding each other and *j* decreases ($\propto L^{-3/2}$, per Eq. A3). Similarly, by working in conditions that decrease polymer stiffness (e.g. by charge screening in high salt), the Kuhn length *b* decreases and *j* increases as the polymer adopts a more compact shape (with its ends found close together more often).

Putting it all together and assuming a Kuhn length of ~100 nm for dsDNA[4], we can plot the maximum allowed concentrations of DNA ends for different values of f_C as a function of DNA length (see Figure 4A.6). From this plot, several general trends can be observed. Of primary importance, the "correct" upper limit of DNA concentration for ligation reactions that favour circles is predicted to be in the ~nM range for 1–10 kbp templates (corresponding to mass concentrations of the order of ~a few ng/µL). Figure 4A.7 in the next section shows the gel pattern resulting from a ligation reaction on 9.7-kbp template at a concentration of ~1 ng/µL (~0.2 nM). There we see a majority of ligated products migrating slightly slower than the unligated/linear template (consistent with circularized monomers), while a much smaller subset ran closer to the 20-kbp ladder band (consistent with a dimerized product of 2 × 9.7 = 19.4 kbp).

Returning to Figure 4A.6, we also observe that relaxing the strictness of the f_C condition rapidly increases the allowable concentration of linear DNA template during the reaction. Using 2 kbp DNA as an example, f_C = 95% corresponds to an i_{max} of 1.63 nM, while reducing f_C to 75% (a 1.27× reduction) raises i_{max} all the way to 10.3 nM (a 6.33× increase). The main takeaway is that if a DNA concentration (e.g. in the nM range) is found that works (i.e. makes almost entirely circular monomers) for a given molecule, accidently doubling the concentration next time, for instance, is not likely to lead to a complete inversion in the population (i.e. now mostly multimers). This means that careful DNA concentration measurements (e.g. by plate reader) followed by nanolitre-level volume adjustments should not really be necessary to achieve relatively consistent reaction outcomes.

Finally, it is clear from Figure 4A.6 that as the length of the DNA template increases, the maximum concentration of ends for a given value of f_C decays rapidly. This of course could be predicted from the strong $L^{-3/2}$-dependence of the effective intramolecular end concentration *j*. At f_C = 90%, for example, i_{max} goes from 9.7 nM at 1 kbp to 0.31 nM at 10 kbp ($9.7/0.31 = 31.6 = 10^{3.2}$). In practice, in order to work with similar molar amounts of DNA during an experiment, much larger total reaction volumes will be required for longer DNA fragments, as well as more total enzyme if the concentration of ligase is meant to remain fixed across samples.

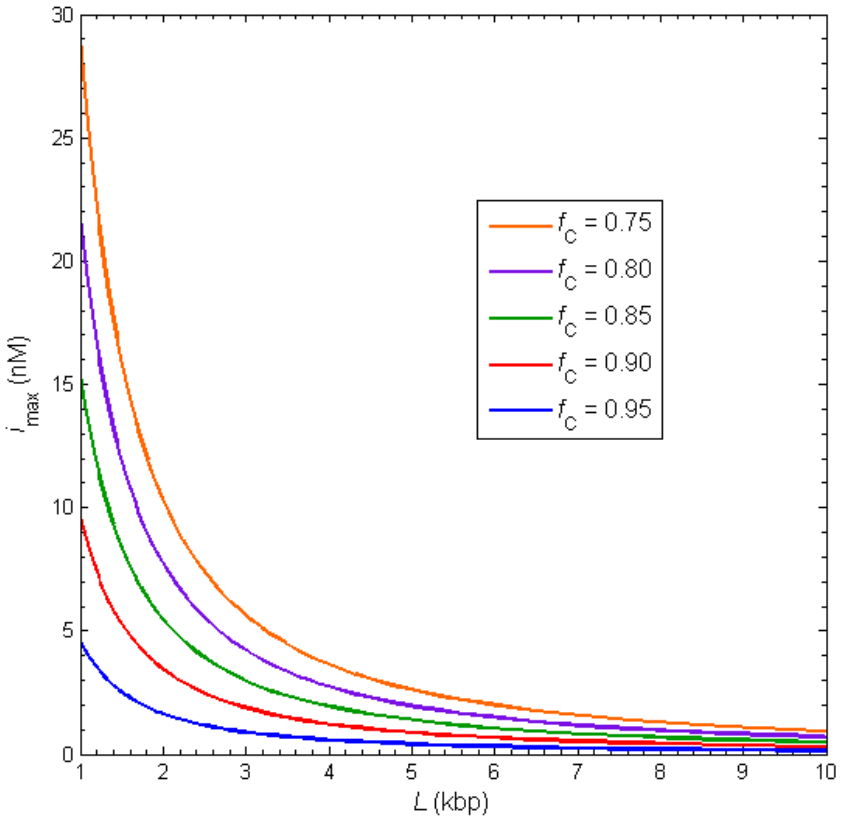


Figure 4A.6: *Maximum concentrations of DNA ends during a ligation reaction to achieve a given "circularization factor" f_C (defined via Eq. A5) plotted vs. DNA length. The circularization factor should give a rough estimate of the fraction of DNA molecules that will be ligated into circular monomers, the desired product in the T4 ligation reaction of Method 2 (see Figure 4.4a in the main text).*

Section A9: Restriction enzyme digest of ligated *cos*-containing fragments (circularity test)

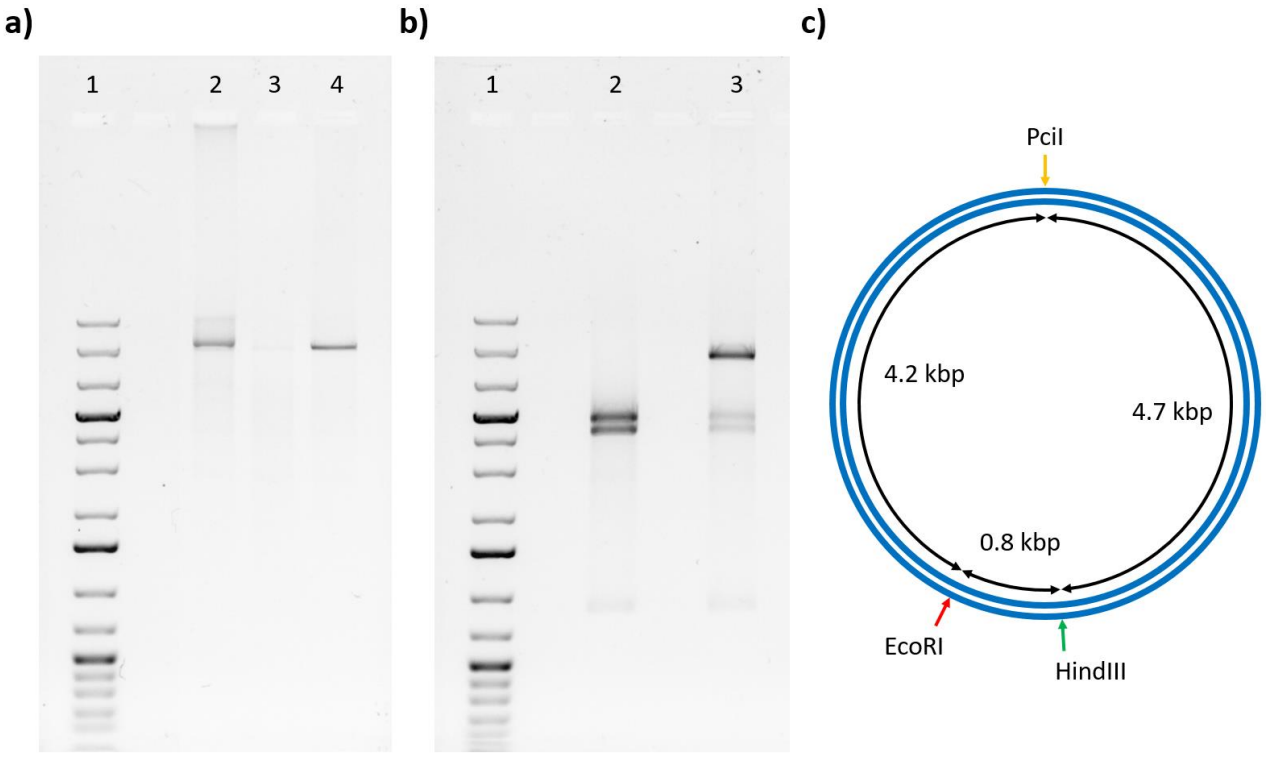


Figure 4A.7: a) *Agarose gel (0.5%, 0.5× TAE, pre-stained with Gel Red, 70 V) of a T4 ligase reaction on 9.7 kbp PciI-digested, cos-containing, linear DNA. Lane 1: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). Lane 2: products of the ligation reaction purified by phenol-chloroform extraction. Lane 3: products of the ligation reaction purified by silica membrane-based spin columns (Monarch PCR Cleanup Kit, NEB). Lane 4: original 9.7 kbp linear DNA template (PciI- digested). We found phenol-chloroform extraction to be noticeably more efficient at recovering the ligation products when compared to a spin column approach.* ***b)*** *Agarose gel (0.5% agarose, 0.5× TAE, pre-stained with Gel Red, 70 V) of an EcoRI-HindIII double digest on the linear 9.7 kbp template fragment (lane 2) and the T4 ligation products (lane 3). The shift from two dominant bands at 4.7 kbp and 4.2 kbp corresponding to the end fragments of the linear DNA to a single dominant band at 8.9 kbp in the case of the ligation products suggests that the majority of DNA had been successfully circularized.* ***c)*** *Restriction map of the presumed 9.7 kbp circular output of the ligation reaction of (a).*

Section A10: Optimizing terminase reaction conditions

To investigate the activity of our non-commercial sample of λ-terminase enzyme (Catalano Lab, UC Denver) and assess its efficiency under various reaction conditions, we employed a "test molecule" of template DNA. This consisted of a 14 kbp long, *cos*-containing fragment generated from a BstEII digest on λ-DNA that had been ligated together at its single-stranded ends. Upon cleavage by terminase, it is expected to split open at its *cos*-site into 6-kbp and 8-kbp fragments that should be easily separable from each other (as well as from their 14-kbp parent) on a gel (see Figure 4A.8).

As background, the terminase sample used here is modified with extra histidine residues (×6) on the C-terminus of its large subunit (gpA) in order to facilitate its purification by immobilized metal affinity chromatography. However, these purification tags may also reduce its efficiency as a *cos*-cleaving enzyme[5]. To help counteract this loss in efficiency, we tried including IHF in the reaction mix. IHF (**i**ntegration **h**ost **f**actor) is a protein produced by the *E. coli* bacterial host of λ phage that binds to specific sequences of the viral genome during infection and helps assemble terminase into an active nucleoprotein complex (viral DNA + terminase + IHF + cofactors) that will carry out the some of the DNA packaging reactions (e.g. *cos*-cleavage) *in vivo* [6]. Past evidence also exists of IHF improving the yield of *cos*-cleavage *in vitro*, particularly when using these reduced-efficiency His-tagged enzymes[5].

Figure 4A.8 shows an agarose gel documenting the effect of IHF on the terminase digest of a sample of 14-kbp test DNA. The reactions with IHF indeed saw improved yields, most notably for reactions featuring lower enzyme concentrations (compare the lanes with 100 nM and 200 nM terminase). The inclusion of IHF could therefore help extend our limited supply of terminase by enabling the use of less enzyme per reaction while maintaining yields similar to those with higher enzyme levels in the absence of IHF.

Attempts to optimize other reaction conditions are shown in Figure 4A.9. A sweep of incubation time for the reaction (0.5, 1, 2, & 4 hrs) did not result in a drastic change in reaction progress over the range tested (see Figure 4A.9a). On the other hand, changing the incubation temperature from 20 °C to 37 °C (Figure 4A.9b) noticeably shifted the reaction toward completion. Based on these results, unless otherwise stated, 1 hr incubations at 37 °C were used going forward.

Another effect that was observed occasionally was the non-specific digestion of the DNA template at high enzyme concentration. As enzyme concentration is increased, the reaction can sometimes quickly transition from incomplete specific digestion → mostly complete specific digestion → non-specific digestion. On a gel, this is visible as a transition from the discrete bands of the template (14 kbp) and the intended product (8 & 6 kbp) fragments to a diffuse smear of faster-migrating products (see Figure 4A.9c). A full understanding of this phenomenon was not achieved of the course of the present work – parallels could perhaps exist with the "star activity" commonly described in commercial restriction endonucleases that can occur at elevated ratios of enzyme to DNA template[7]. Terminase is a complex enzyme, responsible for catalyzing several reactions during the multi-stage process of viral DNA packaging[6]. As such, having it interact optimally and exclusively with *cos*-containing DNA in a single way (*cos*-cleavage) is likely to require some fine tuning of the reaction conditions. As a practical matter, in order to reduce the impact of this non-specific digestion effect, the goal in future reactions was to use as little enzyme as needed to achieve the desired level of (specific) digestion on a given DNA template.

Finally, the use of alternative enzyme co-factors during terminase digestion was investigated. Studies with wild-type terminase have previously shown ATP to stimulate its *cos*-cleavage activity[8]. Specifically, ATP seems to act as an "allosteric effector" for the nicking of the DNA backbone (binds to a secondary site on the enzyme to enhance activity) as well as being actively hydrolyzed by the enzyme to power the separation of the otherwise

stable resulting 12-nt cohesive ends. Here we sought to verify that the behaviour of our current terminase system was similar in the absence/presence of this co-factor. Samples of 14-kbp test DNA were digested under three different conditions: 1) with ATP, 2) with a non-hydrolyzable ATP analogue (AMP-PNP, see Figure 4A.10a), and 2) in the absence of co-factor (control). To distinguish between DNA nicking and active strand separation for each condition, after an initial 1 hr incubation at 37 °C, EDTA was first added to stop the reaction and then aliquots of each mixture were further incubated for 10 minutes at either 70 °C (above melting temperature of ends) or room temperature.

An example of such an experiment is shown in Figure 4A.10. From the relative band intensities, AMP-PNP (the non-hydrolyzable ATP analogue) appears to lead to enhanced levels of nicking compared to the co-factor-free control (though lower than with ATP proper) in addition to having the largest proportion of its total strand separation resulting from the melting step at 70 °C. This overall trend is in line with the idea that ATP binding during the nicking stage stimulates the action of the enzyme, while ATP hydrolysis is necessary to maximize the degree of strand separation below the melting temperature of the 12-nt ends. Also, the fact that there is a noticeable increase of strand separation above the melting point of the ends (but after the enzyme has been deactivated with EDTA) is a positive sign that terminase is nicking both strands at the intended sites 12 nt apart (at least for a noticeable fraction of the DNA templates). Nicks separated by only ~4 nt, such as those created by common restriction endonucleases, or even λ-terminase itself under certain unfavourable conditions[8], would create double-stranded segments that are too short to remain stably annealed at these moderate temperatures (25 - 37°C).

However, our observed results are not fully consistent with those documented by Higgins et al[8]. For instance, there exists a substantial amount of strand separation in ATP-free conditions (AMP-PNP & no co-factor) at temperatures below the melting point of 12 nt overhangs, perhaps an indication that some portion of the products feature shorter, “incomplete” single-stranded segments. The non-hydrolyzable ATP analogue AMP-PNP was also noticeably less effective at stimulating *cos*-nicking than ATP (40-60% of cutting efficiency vs. ATP – see Figure 4A.10b – compared to ~80% of ATP efficiency observed by Higgins et al.[8]).

Together with the general non-specific digestion observed above (Fig. 4A.9c) and the overall low/variable efficiency of the enzyme (100s of nM enzyme needed for ~1 nM of DNA) these results may suggest that some of the current digest conditions (e.g. enzyme sample, reaction buffer, incubation times/temperatures, etc.) are in some ways unoptimized or non-ideal. However, pinning down the exact biochemistry behind the causes of this is beyond the scope of this initial proof of principle of a general SE-DNA synthesis scheme. The fact however that λ-terminase has, in the past, been extensively isolated and characterized for use in biotechnology[9,10] as well as previously existing as a commercial product complete with defined digest protocols and supplied reaction buffers suggest that optimized reactions exist and could be used in future studies, should the need arise.

For the integration of a terminase digest step into the SE-DNA protocol of Method 2 (acting on DNA circles with an internal *cos* site, see Figure 4.4a) then, lessons learned with the 14-kbp linear test molecule were applied (e.g. include IHF in the reaction, incubate at ~1 hr at 37 °C, use ATP as a co-factor for higher yields). We also adopted an enzyme titration approach whereby we initially add a “small” amount of enzyme (50-100 nM), check the progress of the reaction on a gel (aliquotted from the main tube kept on ice), and continue adding additional enzyme in this manner as needed until the desired level of digestion is achieved. This has two benefits: it minimizes the amount of enzyme used per reaction from our limited supply and it reduces the danger of entering a regime of “excess” enzyme where overdigestion (non-specific) may occur.

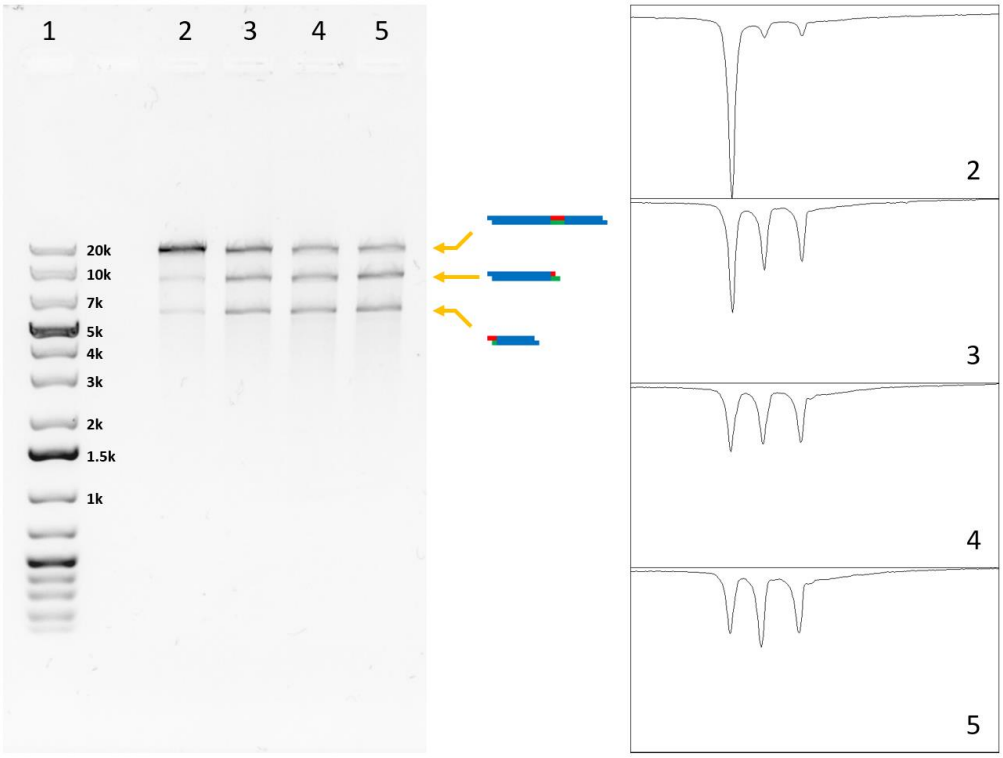


Figure 4A.8: *Agarose gel (0.6%, 0.5× TAE, pre-stained with Gel Red, 70 V) demonstrating the effect of IHF on cos-cleavage by His-tagged λ-terminase (Catalano Lab, UC Denver) in vitro. The DNA template was a 14-kbp fragment of the λ-phage genome featuring an internal, ligated cos sequence and all digests were carried out at 37°C for 1 hr (~1 nM DNA, 10 mM $MgCl_2$, 1 mM ATP, 50 mM Tris-HCl pH 7.5). Lane 1: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). Lane 2: DNA template + 100 nM terminase, Lane 3: DNA template + 100 nM terminase + 50 nM IHF, Lane 4: DNA template + 200 nM terminase, Lane 5: DNA template + 200 nM terminase + 50 nM IHF. The improvement in reaction progress in the presence of IHF is more drastic at the lower enzyme concentration (100 nM terminase: ~10% → ~50%) than at the higher enzyme concentration (200 nM terminase: ~60% → ~70%). The panel to the right represents the integrated band intensities along the designated numbered lane.*

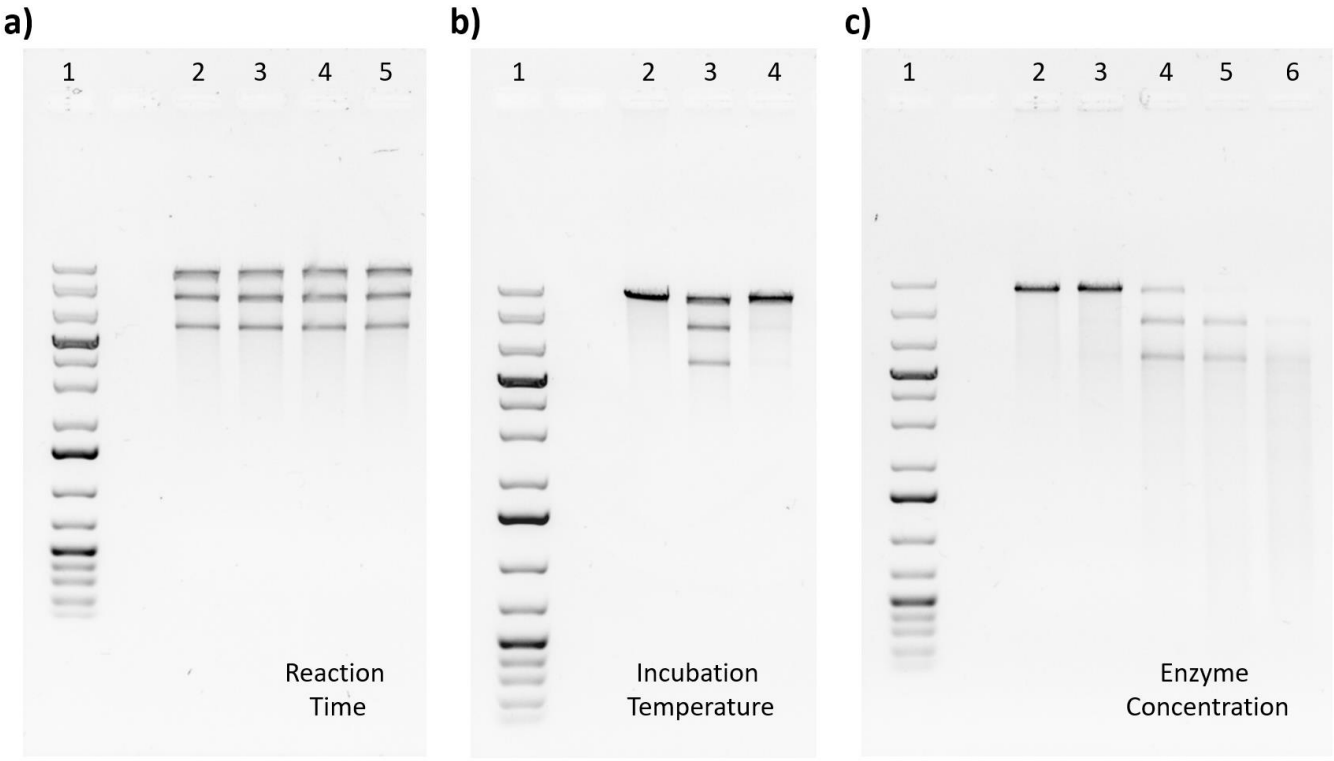


Figure 4A.9: *Agarose gels (0.6%, 0.5× TAE, pre-stained with Gel Red, 70 V) exploring ideal reaction conditions for the λ-terminase digest of a 14-kbp, cos-containing DNA fragment.* ***a)*** *Tracking the reaction progress over time (200 nM terminase, 50 nM IHF, ~1 nM DNA, 37° C). Lane 1: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). Lane 2: 0.5 hr incubation, Lane 3: 1 hr incubation, Lane 4: 2 hr incubation, Lane 5: 4 hr incubation. From the relative intensities of the product (6 & 8 kbp) vs. the reactant (14 kbp) bands, the reaction seemingly saturated early on, with no benefits to prolonged (2+ hr) incubations.* ***b)*** *Effect of incubation temperature on the reaction (100 nM terminase, 100 nM IHF, ~1 nM DNA). Lane 2: DNA template without enzyme added (control). Lane 3: 1 hr incubation at 37 °C. Lane 4: 18 hr incubation at 22 °C. On account of the improved yield, 37 °C was used consistently in future experiments.* ***c)*** *Example of non-specific digestion observed at elevated enzyme concentrations (100 nM IHF, < 1 nM DNA, 1 hr at 37° C). Lane 2: 0 nM terminase (control). Lane 3: 100 nM terminase. Lane 4: 200 nM terminase. Lane 5: 300 nM terminase. Lane 6: 400 nM terminase. Between 200 nM and 400 nM enzyme concentrations, the reaction transitioned from proceeding well at the desired cleavage site (~70% cleaved) to a regime of non-specific digestion.*

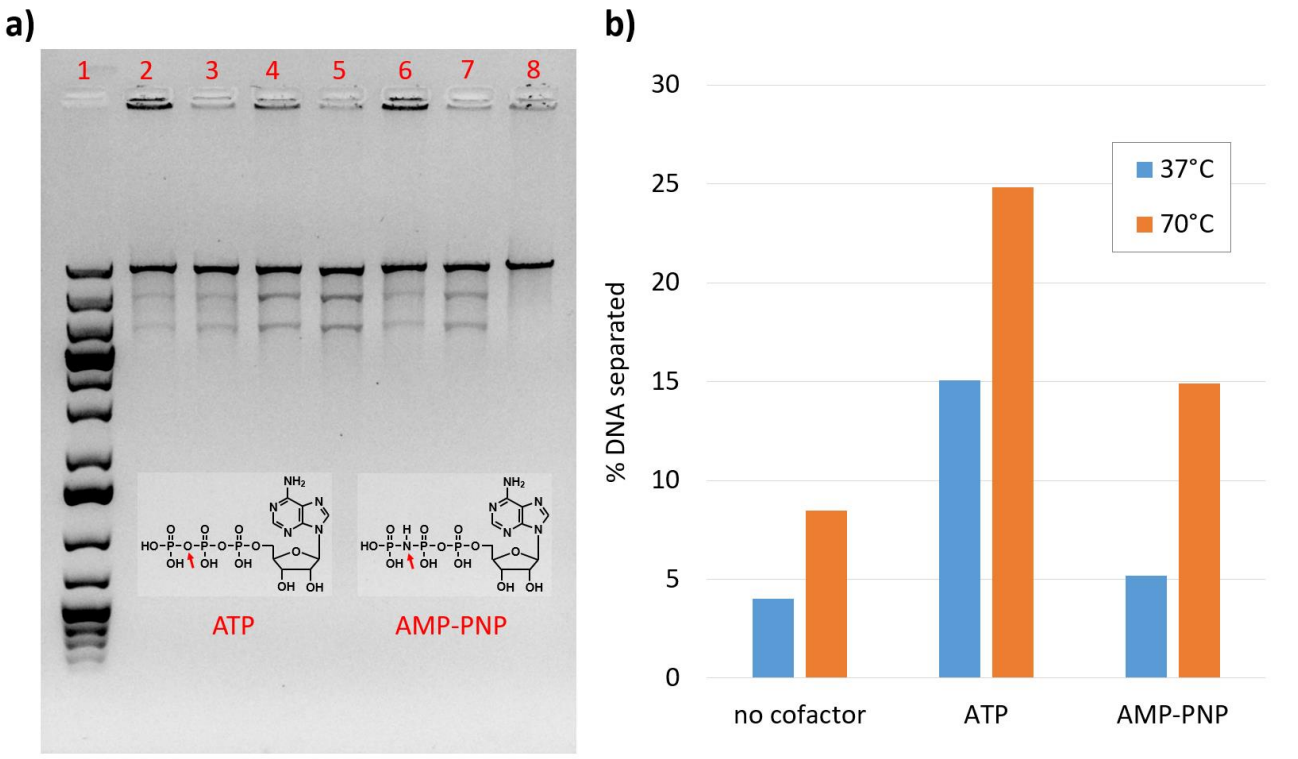


Figure 4A.10: a) *Agarose gel (0.6%, 0.5× TAE, pre-stained with Gel Red, 70 V) showing the effect of the choice of co-factor on the cos-nicking/strand-separation activity of a sample of λ-terminase. General conditions: 300 nM terminase, 100 nM IHF, ~1 nM DNA template, 1 hr incubation at 37 °C followed by injection of EDTA-containing "stop" solution and further 10 min incubation at either room temperature (RT) or 70 °C. Lane 1: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). Lane 2: no co-factor (control), RT secondary incubation. Lane 3: no co-factor (control), 70 °C secondary incubation. Lane 4: 1 mM ATP, RT secondary incubation. Lane 5: 1 mM ATP, 70 °C secondary incubation. Lane 6: 1 mM AMP-PNP, RT secondary incubation. Lane 7: 1 mM AMP-PNP, 70 °C secondary incubation. Lane 8: DNA template (no enzyme, IHF, or co-factor).* ***b)*** *Proportion of DNA mass in each condition separated into 6- & 8-kbp products, as estimated from the relative integrated band intensities in each respective lane. The labels 37 °C and 70 °C refer to the highest temperatures experienced by the sample in question.*

Section A11: Gel-shift assay with 12-arm stars + sticky-ended DNA (produced by λ-terminase method)

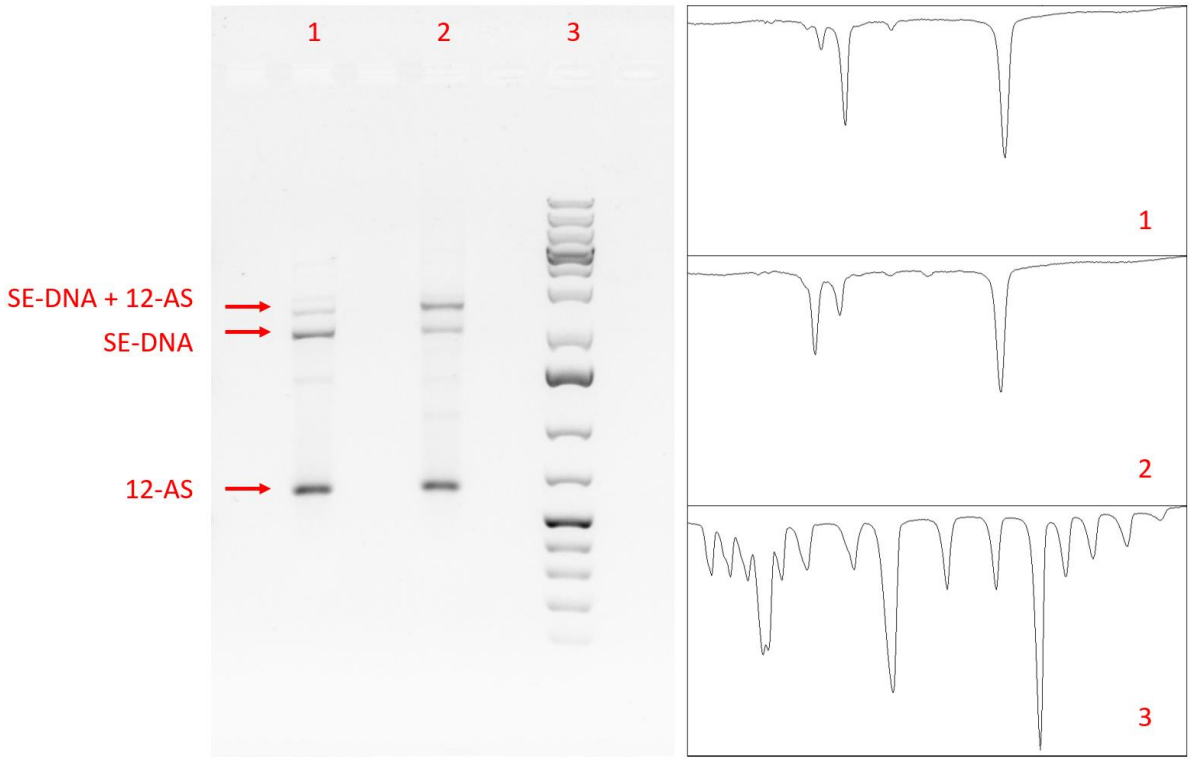


Figure 4A.11: *Agarose gel (0.7% agarose, 0.5× TBE, pre-stained with Gel Red, 70 V) showing the effect of incubating SE-DNA (2 kbp, produced by λ-terminase digest) together with DNA nanoparticles (12-arm stars) and a linker oligonucleotide complementary to single-stranded extensions on the two pieces (mixture incubated overnight at 50°C, in 10 mM $MgCl_2$). Lane 1: SE-DNA (~2 nM) + 12-arm stars (20 nM). Lane 2: SE-DNA (~2 nM) + 12-arm stars (20 nM) + linker oligo (20 nM). Lane 3: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). In the presence of the linker oligo that can bridge across the SE-DNA and DNA stars (Lane 2), the main SE-DNA population shifts to a slower-migrating band (relative to Lane 1), signifying the successful attachment of the DNA stars .*

Section A12: Complications from agarose-extracted DNA fragments

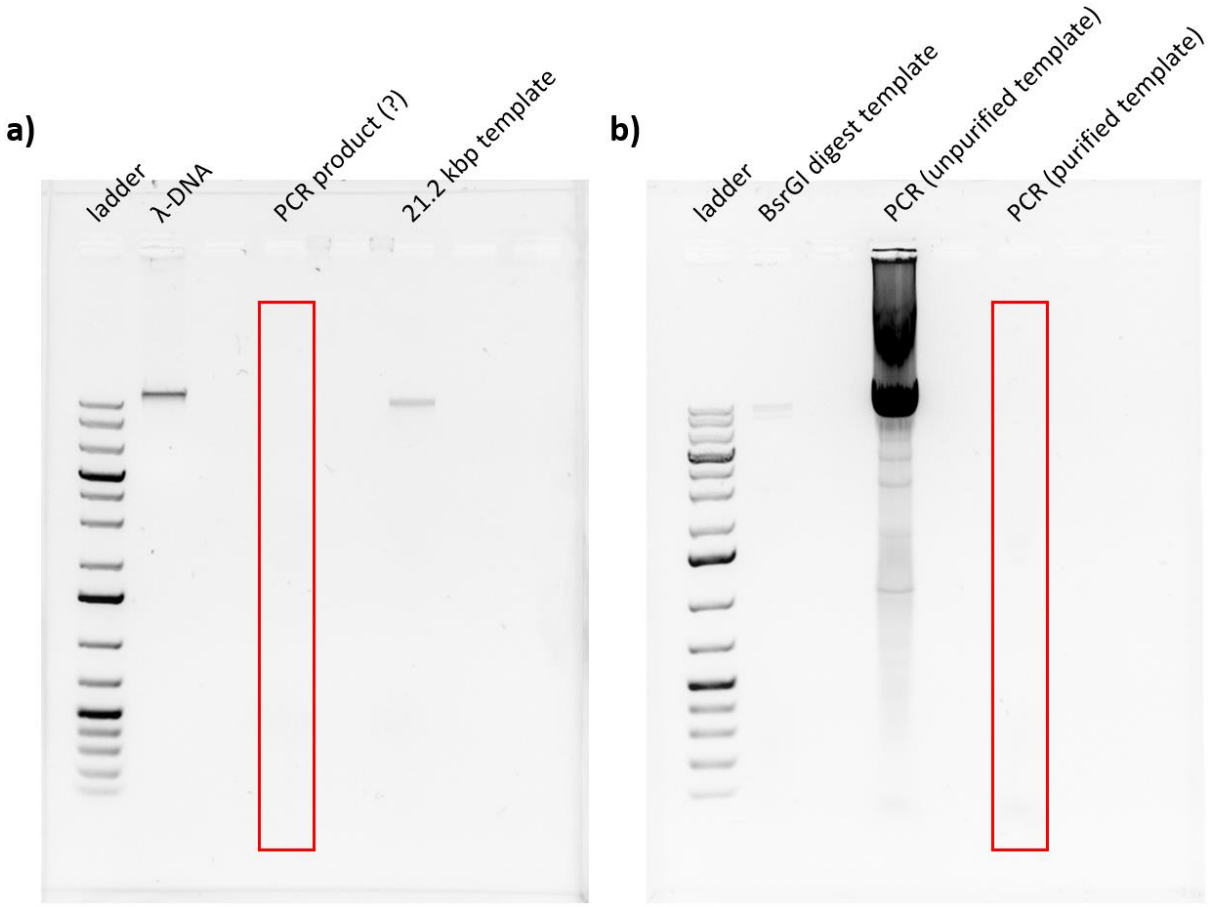


Figure 4A.12: a) *Agarose gel (0.7%, 0.5× TAE, pre-stained with GelRed, 70 V) showing the lack of a visible 11.2-kbp PCR product (LongAmp polymerase, NEB) when an agarose-purified (Freeze-'N-Squeeze) 21.2 kbp DNA fragment was used as the DNA template in the reaction.* ***b)*** *Agarose gel (0.7%, 0.5× TBE, pre-stained with GelRed, 70 V) showing the experiment repeated by aliquoting two identical mixtures of PCR reagents from the same stock (minus the DNA template) and then adding the purified 21.2 kbp fragment to one and the raw restriction digest that generated it to the other (such that the mass of 21.2 kbp fragment was similar in both instances). After temperature cycling, the PCR with agarose-purified template again showed no visible product while the raw template worked well.*

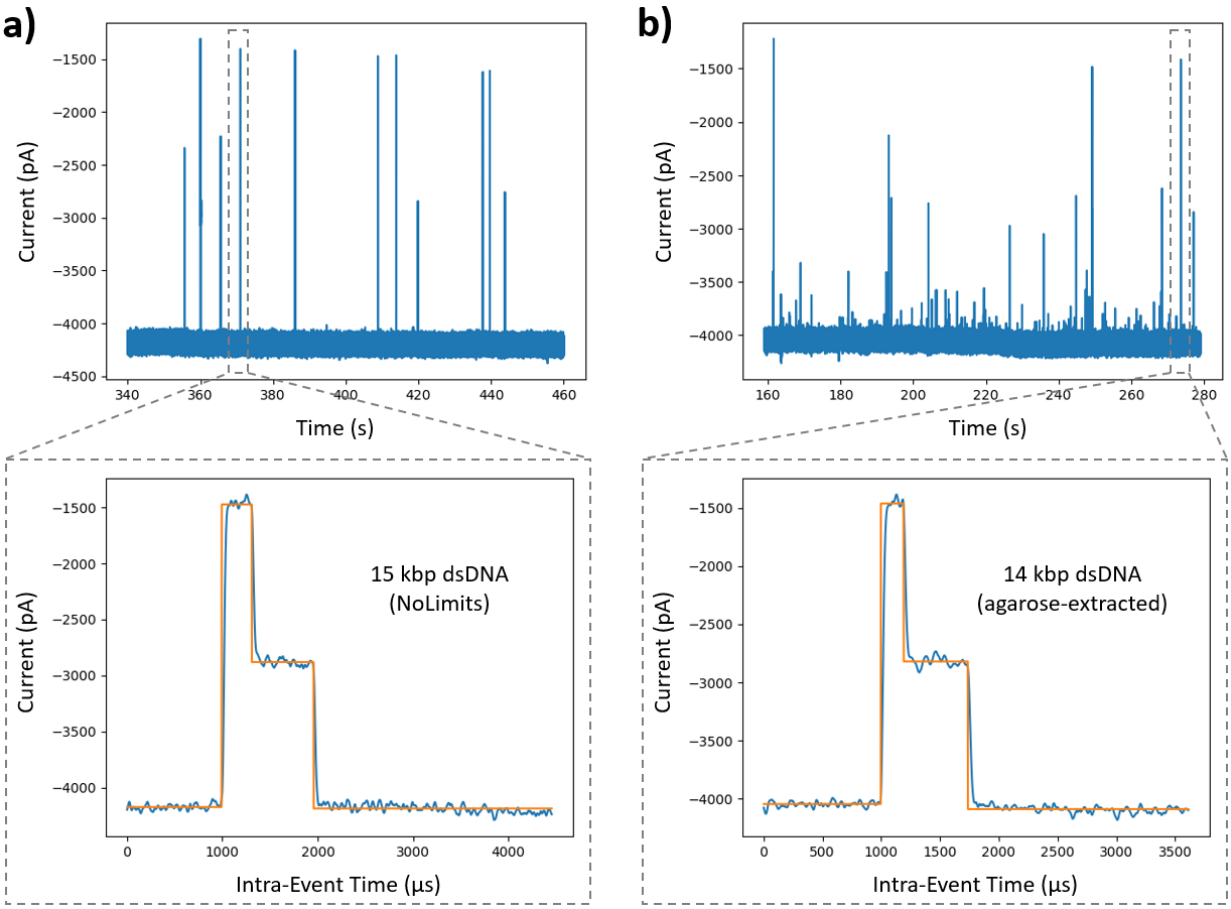


Figure 4A.13: *Comparing nanopore signals (~4.9-nm pore diameter, 200 mV, 3.6 M LiCl pH 8) produced by commercial and agarose-extracted DNA samples of a similar length.* ***a)*** *Representative (2 minute) current trace of a 15 kbp commercial dsDNA sample (NoLimits, Thermo Scientific) passing through the pore.* ***b)*** *Representative (2 minute) current trace of a 14 kbp dsDNA sample (BstEII digest fragment of λ-DNA) extracted from an agarose gel (GeneJET Gel Extraction Kit, Thermo Scientific) passing through the same pore as in (a). While both samples produced recognizable dsDNA signals (example traces expanded as insets), agarose-extracted DNA reliably led to the appearance of additional, lower amplitude signals (see b) as if from contaminants, as well as increased pore clogging (both reversible and irreversible with voltage inversion or solution flushing).*

Section A13: Avoiding agarose purification with short-end PCR

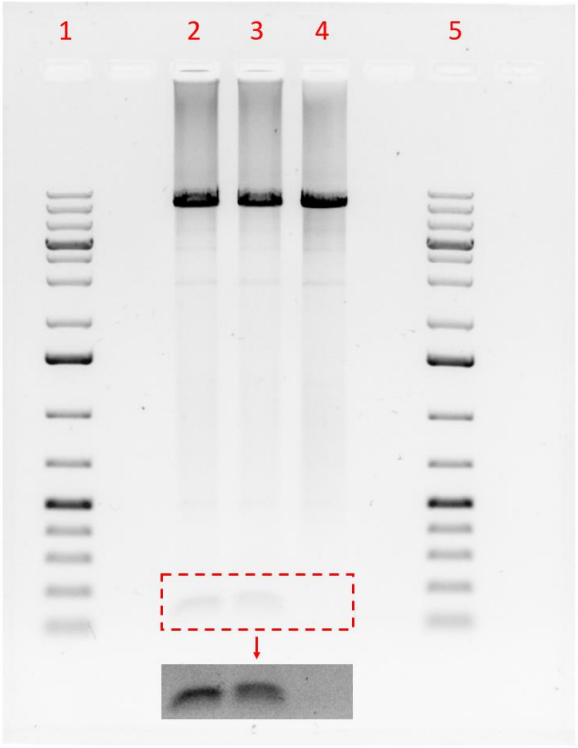


Figure 4A.14: *Agarose gel (0.7%, 0.5× TBE, pre-stained with GelRed, 70 V) of restriction digests (PciI) on PCR products generated by primers placed relatively close to the recognition sites of the restriction enzyme. The short ends cleaved off of the main 9.7 kbp product are then removed by silica membrane-based spin column (PureLink Quick PCR Purification Kit, Invitrogen). Lanes 1 & 5: GeneRuler 1 kb Plus DNA Ladder (Thermo Scientific). Lane 2: PciI digest on PCR product from one primer set (generates two 84-bp end fragments per template molecule). Lane 3: PciI digest on PCR product from a second primer set (generates a 62-bp and a 110-bp end fragment per template molecule). Lane 4: spin column clean-up on a mix of the restriction digest reactions in Lanes 2 & 3, showing the removal of these short ends (see inset for higher-contrast).*

Section A14: Distribution of fragment sizes with single restriction enzyme + single mismatch in primer

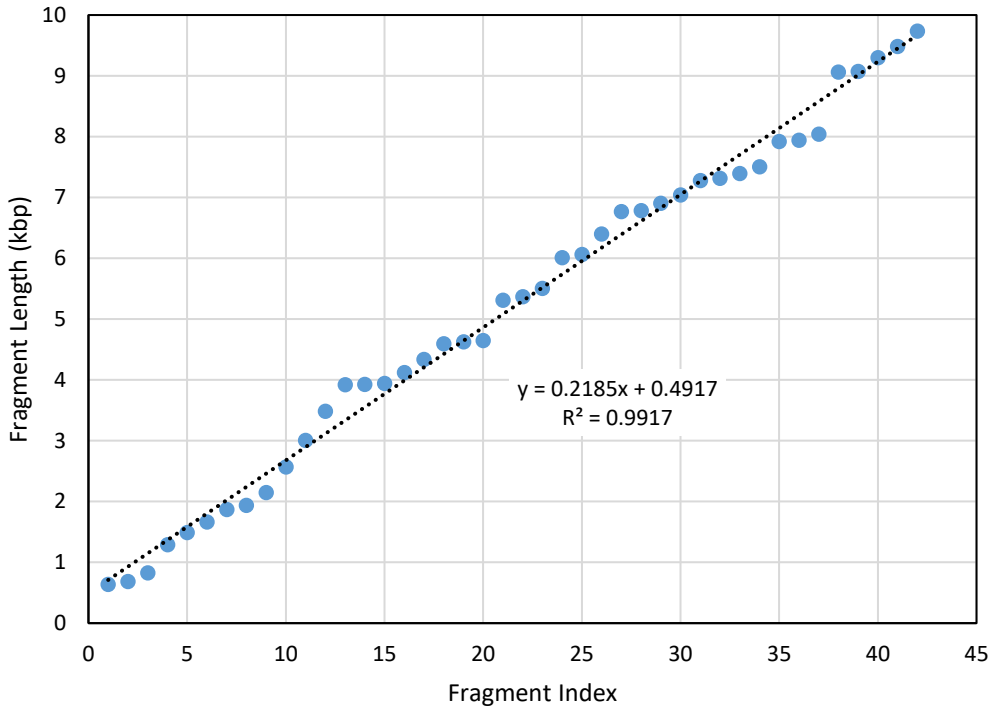


Figure 4A.15: *Distribution of possible sizes of cos-containing fragments, ordered by length, from digesting a ~10-kbp region of circularized λ-DNA (bases 39395 – 633) with PciI, where: 1) the position of the "right" cut is fixed (at bases 628-633), and 2) the position of the "left" cut is allowed to move by introducing a single point mutation in the native λ sequence to create a novel PciI recognition site within this region. An average gap of ~220 bp exists between adjacent fragment sizes.*

Section A15: Nanopore data from terminase-produced SE-DNA (separating lengths by ECD)

DNA molecules of different lengths have previously been separated by the "equivalent charge deficit" (ECD) of their nanopore translocation signals [11–13]. This quantity is defined as the integrated area underneath a trace of current vs. time for each event, in between the open-pore baseline current level and the levels of blocked current where the molecule is in the pore obstructing the background flow of electrolyte ions (e.g. Li^+ & Cl^-).

Under conditions (e.g. large pore diameters) that permit the folded passage of target molecules through the pore, ECD can be a more useful metric for characterizing molecular length than overall passage time since the more compact shape of folded molecules can lead to them translocating as quickly as unfolded molecules of a shorter length (in the absence of pore wall interactions). On the other hand, folded passage also tends to produce deeper average current blockages during translocation due to the presence of multiple molecular segments inside the pore volume simultaneously. If the two effects (shorter dwell times & deeper blockages) have compensating impacts on the total blocked charge (area) of the event, then ECD can be conserved among linear molecules of a fixed length, independent of their conformation through the pore.

When all the events of a single nanopore experiment are plotted in a graph of average blockage ($\bar{I}_b$) vs. dwell time (τ_D), populations of conserved ECD (A_{ECD}) can be easily identified as falling along either hyperbolas when using linearly-scaled axes or along straight lines with logarithmically-scaled axes:

$$\bar{I}_b = \frac{A_{ECD}}{\tau_D} \Rightarrow y \equiv \log \bar{I}_b = -\log \tau_D + \log A_{ECD} \equiv -x + b \tag{A6}$$

In particular, straight lines corresponding to different ECD values (different molecule lengths) are expected to run parallel to each other (sharing a slope of -1, see eq. A6) and be bound by the same y-limits (blockage depths) corresponding to events that are either fully unfolded [$(\bar{I}_b)_1$] and fully folded [$(\bar{I}_b)_2 \approx 2(\bar{I}_b)_1$]. Their horizontal separation, meanwhile, is dictated by the ratio of these different ECD values:

$$\text{for } y_1 = y_2 \Rightarrow -x_1 + b_1 = -x_2 + b_2 \Rightarrow \Delta x \equiv x_2 - x_1 = b_2 - b_1 = \log \frac{(A_{ECD})_2}{(A_{ECD})_1} \tag{A7}$$

Figure 4A.16a shows such a plot for a mixed sample of sticky-ended DNA molecules prepared by terminase digest ("Method 2" in the main text) at two different lengths – 2 kbp and 4 kbp (~10 nm pore diameter, 200 mV, 3.6 M LiCl). Two populations are indeed visible as a set of roughly parallel lines running from unfolded (~1×) to folded (~2×) blockages in both cases ($\Delta y \approx 0.3 \approx \log_{10}2$). The increased event density observed for fully folded blockages in particular is likely due to subpopulations of circular molecules, which, due to their connected ends can only pass through the pore in pairs of dsDNA segments. These may be molecules that were never fully digested with terminsase to begin with (e.g. see Figure 4.5b in the main text) or that have subsequently had their complementary sticky ends re-anneal to each other.

Finally, Figure 4A.16b shows the distribution of ECD specifically for this same set of events (2- & 4-kbp SE-DNA). Two main peaks are again visible, corresponding to the two DNA species in the mix, and are separated roughly in correspondence with the ratio of their lengths ($\Delta x \approx 0.4$ vs. $\log_{10}$[4339 bp/1941 bp] = 0.35).

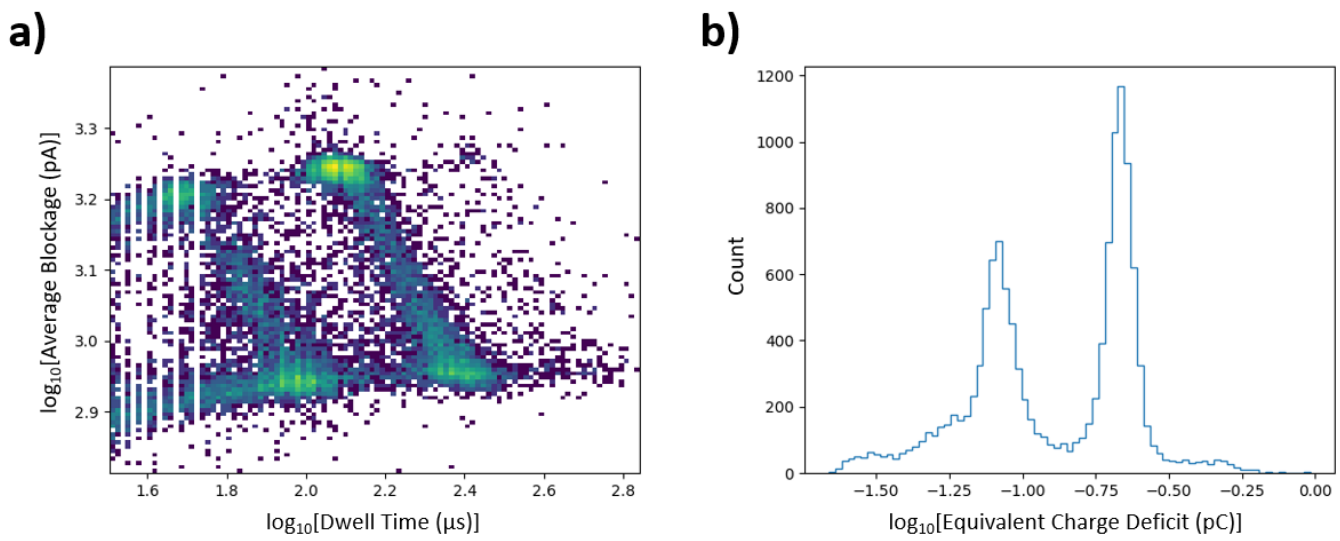


Figure 4A.16: a) *2D histogram (log-log) of average blockage per event vs. event duration from a nanopore experiment (~10 nm pore diameter, 3.6 M LiCl, 200 mV transmembrane potential) on a sample containing a mixture of 2- and 4-kbp SE-DNA produced by λ-terminase digest of circularized, cos-containing templates ("Method 2" in the main text). The main populations of full-length products for both DNA sizes are visible as two roughly parallel lines running from fully unfolded (~920 pA) to fully folded (~1855 pA) average blockages, consistent with the conservation of equivalent charge deficit (ECD) under these conditions – see equations A6 and A7.* ***b)*** *1D log-histogram of ECD for the events. A factor of ~$10^{0.4}$ = 2.5 separates the two main peaks, in reasonable agreement with the ratio of DNA lengths (4339 / 1941 ≈ 2.24).*

References

(1) Wang, J. C.; Davidson, N. On the Probability of Ring Closure of Lambda DNA. *J Mol Biol* **1966**, *19* (2), 469–482. https://doi.org/10.1016/S0022-2836(66)80017-7.

(2) Jacobson, H.; Stockmayer, W. H. Intramolecular Reaction in Polycondensations. I. The Theory of Linear Systems. *J Chem Phys* **1950**, *18* (12), 1600–1606. https://doi.org/10.1063/1.1747547.

(3) Collins, F. S.; Weissman, S. M. Directional Cloning of DNA Fragments at a Large Distance from an Initial Probe: A Circularization Method. *Proceedings of the National Academy of Sciences* **1984**, *81* (21), 6812–6816. https://doi.org/10.1073/pnas.81.21.6812.

(4) Rizzo, V.; Schellman, J. Flow Dichroism of T7 DNA as a Function of Salt Concentration. *Biopolymers* **1981**, *20* (10), 2143–2163. https://doi.org/10.1002/bip.1981.360201009.

(5) Hang, Q.; Woods, L.; Feiss, M.; Catalano, C. E. Cloning, Expression, and Biochemical Characterization of Hexahistidine-Tagged Terminase Proteins. *Journal of Biological Chemistry* **1999**, *274* (22), 15305–15314. https://doi.org/10.1074/jbc.274.22.15305.

(6) Catalano, C. E. The Terminase Enzyme from Bacteriophage Lambda: A DNA-Packaging Machine. *Cell Mol Life Sci* **2000**, *57* (1), 128–148. https://doi.org/10.1007/s000180050503.

(7) Nasri, M.; Thomas, D. Relaxation of Recognition Sequence of Specific Endonuclease HlndIII. *Nucleic Acids Res* **1986**, *14* (2), 811–821. https://doi.org/10.1093/nar/14.2.811.

(8) Higgins, R. R.; Lucko, H. J.; Becker, A. Mechanism of Cos DNA Cleavage by Bacteriophage λ Terminase: Multiple Roles of ATP. *Cell* **1988**, *54* (6), 765–775. https://doi.org/10.1016/S0092-8674(88)91021-5.

(9) Wang, Y.; Wu, R. A New Method for Specific Cleavage of Megabase-Size Chromosomal DNA by λ-Terminase. *Nucleic Acids Res* **1993**, *21* (9), 2143–2147. https://doi.org/10.1093/nar/21.9.2143.

(10) Rackwitz, H. R.; Zehetner, G.; Murialdo, H.; Delius, H.; Chai, J. H.; Poustka, A.; Frischauf, A.; Lehrach, H. Analysis of Cosmids Using Linearization by Phage Lambda Terminase. *Gene* **1985**, *40* (2–3), 259–266. https://doi.org/10.1016/0378-1119(85)90048-4.

(11) Fologea, D.; Gershow, M.; Ledden, B.; McNabb, D. S.; Golovchenko, J. A.; Li, J. Detecting Single Stranded DNA with a Solid State Nanopore. *Nano Lett* **2005**, *5* (10), 1905–1909. https://doi.org/10.1021/nl051199m.

(12) Fologea, D.; Brandin, E.; Uplinger, J.; Branton, D.; Li, J. DNA Conformation and Base Number Simultaneously Determined in a Nanopore. *Electrophoresis* **2007**, *28* (18), 3186–3192. https://doi.org/10.1002/elps.200700047.

(13) Bell, N. A. W.; Muthukumar, M.; Keyser, U. F. Translocation Frequency of Double-Stranded DNA through a Solid-State Nanopore. *Phys Rev E* **2016**, *93* (2), 022401. https://doi.org/10.1103/PhysRevE.93.022401.

CHAPTER 5 – ANALYSIS OF SINGLE-MOLECULE EVENTS FROM DNA NANOSTRUCTURES

This chapter is published in *ACS Sensors* as:

"Analysis of Nanopore Data: Classification Strategies for an Unbiased Curation of Single-Molecule Events from DNA Nanostructures" (2023)

by Zachary Roelen, Kyle Briggs, and Vincent Tabard-Cossa

5.1 – Abstract

Nanopores are versatile single-molecule sensors that are being used to sense increasingly complex mixtures of structured molecules, with applications in molecular data storage and disease biomarker detection. However, increased molecular complexity presents additional challenges to the analysis of nanopore data including more translocation events being rejected for not matching an expected signal structure and a greater risk of selection bias entering this event curation process. To highlight these challenges, here we present the analysis of a model molecular system consisting of a nanostructured DNA molecule attached to a linear DNA carrier. We make use of recent advances in the event segmentation capabilities of Nanolyzer, a graphical analysis tool provided for nanopore event fitting, and describe approaches to event substructure analysis. In the process, we identify and discuss important sources of selection bias in that emerge in the analysis of this molecular system and consider the complicating effects of molecular conformation and variable experimental conditions (e.g. pore diameter). We then present additional refinements to existing analysis techniques, allowing for improved separation of multiplexed samples, fewer translocation events rejected as false negatives, and a wider range of experimental conditions for which accurate molecular information can be extracted. Increasing the coverage of analyzed events within nanopore data is not only important for characterizing complex molecular samples with high fidelity, but is also becoming essential to the generation of accurate, unbiased training data as machine learning approaches to data analysis and event identification continue to increase in prevalence.

5.2 – Introduction

Solid state nanopores are a promising detection platform for a range of single-molecule biosensing applications.[1] In their basic form, they consist of a single nanoscale hole through a thin dielectric membrane separating two electrolyte-filled reservoirs. When an external driving force (such as a voltage or pressure gradient) is applied across the membrane, a flow of ions through the pore is produced that can be measured as ionic current.[2] Passage of a single molecule of interest through the pore can then be detected as a transient blockage in this ionic current, with the magnitude of the deviation from the open pore baseline at any moment relating to the local properties (size, shape, charge, etc.) of the part of the molecule that is in or near the pore at that time.[3] In this way, the presence of specific biomolecules (e.g. nucleic acids, proteins) in a sample can be deduced from the appearance of their signatures in the electrical signal.[1,4]

Typically, to analyze nanopore data, a (small) set of descriptive statistics is extracted from fits to the ionic current signal of each translocation event (such as the total event duration and the maximum blockage depth) and the values of these parameters are plotted for all events in each experiment. Sub-populations of events are then assigned to target molecules based on clustering of these statistics. In recent years a handful of analysis tools[5–8] (each with their own unique limitations on event fitting) have been proposed to facilitate such processing of raw nanopore data at a level that can suit most analysis needs. In practice, despite the availability of analysis tools that support deeper functionalities such as OpenNanopore,[6] MOSAIC,[5] Transalyzer,[7] AutoNanopore,[8] and others, analyte characterization is often limited to basic, single-valued metrics of their translocation signals and ignores information encoded in the internal structure of the events.

Distinguishing molecules with similar physical and chemical characteristics under this scheme can be challenging, however, as significant overlap can exist in the distributions of their extracted parameters. For example, the stochasticity in the transport of DNA can limit the ability of the nanopore to accurately separate DNA fragment sizes by passage time[9–13] and, likewise, the different possible conformations molecules can take during translocation can limit the ability of the nanopore to identify molecules based on shape.[14–16] For this

reason, structured DNA polymers have been used to create barcode-like signals or more unique electrical signatures to improve the capability of the nanopore to distinguish molecules in a mixture or to multiplex detection during sensing. Applications in molecular data storage[17–21] and target biomarker detection,[22–27] for instance, have begun to use these structured polymers to encode nanopore-readable information about a sample.

When working with structured molecules of increasing complexity, a simple analysis workflow based on whole-event metrics (e.g. total event passage time or maximum blockage level) may not be sufficient to fully capture the information content of the signal, which is instead contained in localized features of each event, (hereafter referred to as "event substructure"), arising from specific local details of the molecule being sensed serially during passage through the pore. Furthermore, when selecting the subset of events that satisfy the expected event substructure associated with a given target molecule, many events (often >50%[19]) must be discarded as a result of conformational effects,[28] fitting errors,[5,29] or bandwidth limitations.[30] This necessitates the use of longer sensing times and/or higher target concentrations in order to generate sufficient statistics for each structure variant in a given sample.

Focusing on a narrow window of total events in this way may also introduce undesirable biases into downstream analysis since there are often strong statistical correlations between the types of events that are discarded. This can significantly alter the statistical distributions of key parameters of interest from the ground truth when considering only the subset of events that pass these filters. For example, a common source of error in nanopore measurements is bandwidth limitation, which results in the distortion of events/subevents with short duration and translates to a higher probability of failed fits as sublevel durations are reduced.[5] Removing badly fitted events in this case will cause any downstream analysis of the passage times of the molecules (or substructures of the molecules), for instance, to only use events with long sublevel durations, potentially leading to erroneous conclusions as this bias propagates into the overall distributions of these quantities.

Recently, the field of nanopore analysis has seen a significant increase in the use of machine learning models for event recognition.[31–35] While these developments are exciting, the challenges involved in assembling appropriate training datasets are not often considered and are critical to sound model design. Most importantly, as machine learning approaches to data analysis and event identification become more common,[31] any bias that exists in selecting events as the training data for the machine learning algorithms may ultimately be encoded in the model itself, propagating bias-induced errors forward anywhere that model is used. It is therefore critical that tools for event segmentation into model training categories be developed to assist in avoiding these biases at the level of training data curation.

To illustrate these analysis challenges when working with complex molecules, as well as to develop strategies and tools to address them, in this work we consider translocation events from a basic system of structured DNA molecules. The system is composed of two parts: a long, double-stranded piece of linear DNA of variable length ("carrier") and a relatively short, bulky DNA nanostructure of variable size ("cargo") that is reversibly attached to one end of the linear DNA through the hybridization of complementary DNA sequences. Here we will apply different analysis frameworks to segment our translocation events by both the carrier length and the type of attached nanostructure, and carefully identify and discuss errors that arise in this dataset curation process and how they can be mitigated.

As a practical application, such a structured molecular system could be used to translate a biomarker of interest into a more specific nanopore signal, for instance using toehold-mediated strand displacement reactions.[36,37] In this scheme, a DNA oligonucleotide ("displacement strand") that acts as a label for a particular biomarker is generated by an upstream assay step[19,20,24,38,39] and then releases a cargo molecule from its attached carrier by binding to a complementary single-stranded extension ("toehold") on one of the two pieces. If separate carrier length / cargo species pairs were each associated with a unique displacement strand sequence, then multiple biomarkers could be detected simultaneously in a single nanopore experiment, provided that the translocation events from these separate carrier-cargo pairs were distinguishable by their current signatures.

Importantly, the translocation events of these molecules contain essential features shared by many of those from other structured nanopore targets: all of the identifying information is encoded in the event substructure, in this case, local current spikes generated by the bulky nanostructure on top of an extended underlying signal from the DNA carrier, with variability in this signal arising from the multiple possible orientations/conformations the molecule can adopt prior to entering the pore. As such, this model system is representative of many nanopore sensing schemes currently in use in the field, and the analysis challenges we discuss are readily applicable to other, similar systems.[19,27,40]

In this paper we describe a process to analyze the nanopore signals of structured polymers and introduce refinements of this basic approach to better realize our goal of separating events from mixed samples. This process is centred around the use of Nanolyzer, a graphical analysis tool provided by Northern Nanopore Instruments for nanopore event fitting, segmentation, and event substructure analysis. We note, however, that the general approach to complex event analysis outlined in this work is compatible with any alternative software packages that support the core functions of: fitting event substructure, adding labels to subevents identified by the fits, and segmenting populations of events based on this labelled substructure, and therefore aims to be broadly integrable with existing experimental tools and applications in the field.

Initially, we sort events from a dataset with a single type of carrier-cargo molecule into categories that reflect the conformation of their parent molecules as they passed through the pore, using database query tools to segment events based on the fits. In the process, we discuss what additional information can be gained by breaking down subpopulations by the event substructure, as well as how variation in molecular conformation can impact both the spread of the total population in parameter-space and the rate of generating false negatives that are rejected during analysis due to fitting errors or unrecognized signal shapes. Following on this, we develop refinements of our analysis frameworks that allow for 1) improved separation of event populations by their carrier length and attachment type, 2) fewer events rejected as false negatives / less introduced selection bias, and 3) a wider range of compatible experimental conditions used to generate data with these

molecules. Datasets from mixed samples featuring multiple carrier lengths or attachment types are used to compare the results of analyses performed with and without these analysis refinements in place.

5.3 – Results and Discussion

Figure 5.1a shows a schematic of the molecular system used here. It consists of a variable length of linear double-stranded DNA (2-7 kbp, referred to as "carrier") attached to a DNA multi-way junction (stars with 6 or 12 arms, 24-bp arm length, referred to as "cargo") via complementary single-stranded extensions (12 nt) on both pieces – see Methods section for synthesis details. This cargo has been well characterized in the past as a nanopore sensing target – mixes of stars with differing numbers of arms run on the same pore have previously been separated by the depth of the current blockages they generate.[14,20] A typical current trace of translocating carrier + cargo molecules is presented in Figure 5.1b, as well as a magnified view of the current signature from the passage of a single molecule.

Once the ionic current data is recorded as a continuous timeseries, the next step is to use Nanolyzer to detect the translocation events within this data as statistically significant departures from the open-pore baseline current, as well as to extract numerical values of key descriptive statistics ("metadata") for each event. For instance, Figure 5.1c shows a 2D histogram over two such statistics (using the analyzed events from same dataset as Figure 5.1b): "dwell time" and "maximum deviation". Dwell time is simply the difference between the start and end times of the fitted event (signalled by the departure from, and return to the baseline current, respectively) as determined by the event-fitting algorithm. It characterizes the duration of an interaction between an analyte molecule and the nanopore. Maximum deviation corresponds to the single data point within the event that is furthest away from the open-pore baseline current and is calculated as the absolute difference between that point and the average value of the fitted baseline immediately preceding/following the event. For data featuring very short-lived current blockages within events, (as due to translocations of the small cargo molecules in our case, see sample event in Figure 5.1b), max deviation can often be a more useful measure of the magnitude of these blockages when compared to approaches that attempt to fit these fast

current spikes to step functions, since this fitting tends to fail with increased frequency as the blockage durations get shorter[5] (see also the discussion surrounding Figure 5.5, below).

From the distributions in dwell time and max deviation of Figure 5.1c, at least three distinct populations of events are visible: i) short-lived, deep events (top left), likely due to the passage of cargo molecules (i.e. short but bulky star junctions) that are unattached to carriers, ii) longer events with two shallower blockage levels (bottom) likely corresponding to folded and unfolded conformations of free carrier molecules, and iii) events that are both long-lived *and* deep (top right) from the fully-assembled carrier-cargo molecule.

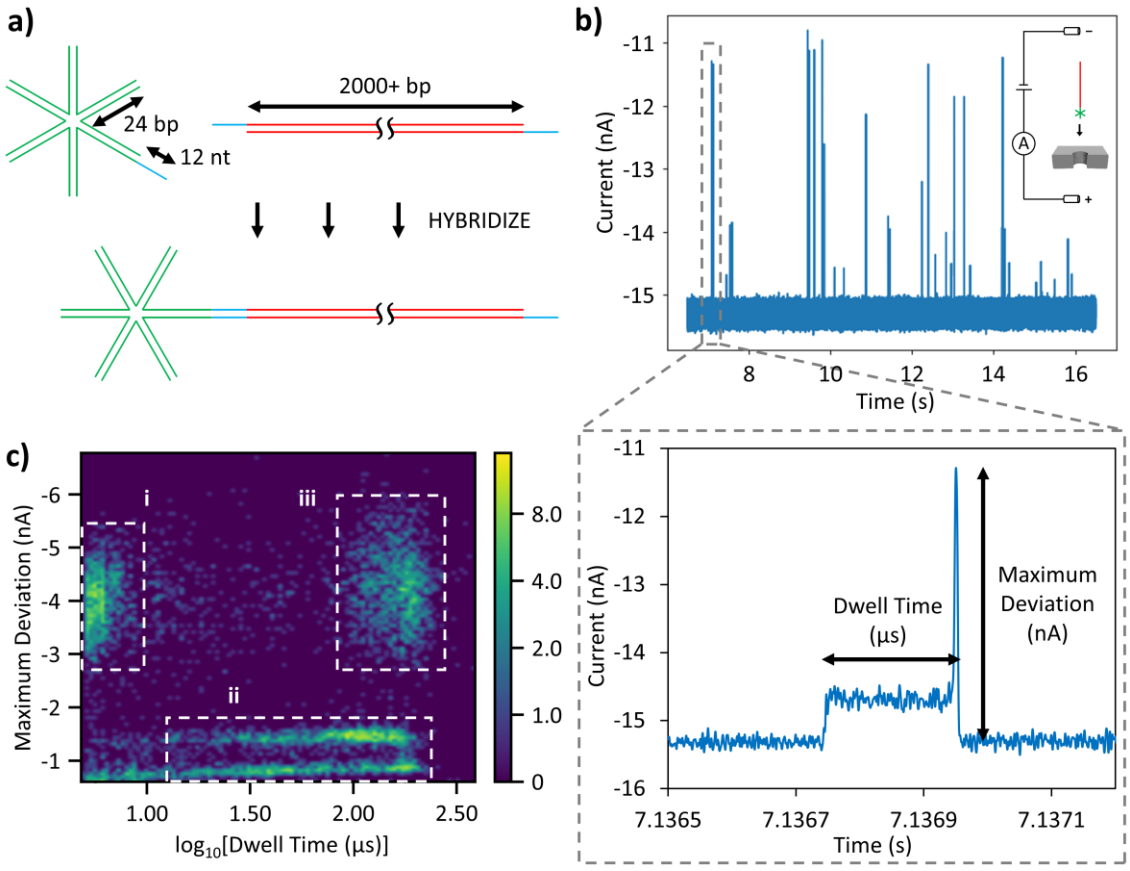


Figure 5.1: a) *Overview of DNA structures in "carrier" and "cargo" system. Illustrated here as a cargo example is a 6-arm star junction (24-bp arms) with a 12-nt ssDNA extension that can be hybridized to its complementary extension on a linear carrier molecule.* ***b)*** *Representative current trace of a sample of DNA carriers + cargo translocating though a nanopore under an applied electric field (2-kbp carrier DNA + 12-arm stars, ~11-nm pore, 3.6 M LiCl, 175 mV). Inset: Zoomed-in view of an individual event.* ***c)*** *2D histogram of the maximum deviation from the baseline current for each event vs. its translocation time (log scale), using the same dataset as in (b). Population i: free cargos (12-arm DNA stars). Population ii: free carriers (linear 2 kbp dsDNA). Population iii: hybridized cargo-carrier molecules.*

5.3.1 – Rigid Event Classification

A proposed categorization scheme for the structured ionic current signals generated from hybridized carrier-cargo targets passing through a nanopore is presented in Figure 5.2. It is based on simultaneously grouping the events on two criteria that characterize the conformational state of the molecule as it translocates: 1) the folding state of the dsDNA carrier and 2) in what order the two ends of the carrier (free end and target-attached end) enter the pore. In this scheme, the carrier can be unfolded ("Type U"), partially folded ("Type P") or fully folded ("Type F"), while the attachment can enter the pore first ("Type 1"), second ("Type 2") or not at all ("Type 0"). Type 0 events are expected when a sample contains at least some carrier molecules with two free ends, that is, without an attached cargo. More complex event shapes (e.g. involving a doubly-folded or knotted carrier translocating the pore) have not been included in this categorization scheme and thus any events characterized by such shapes will be "rejected" by filters implementing it. However, based on previous results, simple DNA conformations featuring at most an initial single fold should represent the majority of translocation events under the experimental conditions tested here[12,41,42] (double-stranded DNA, pore diameters < 20 nm), and no significant bias is expected to be introduced by ignoring events with more than one carrier fold.

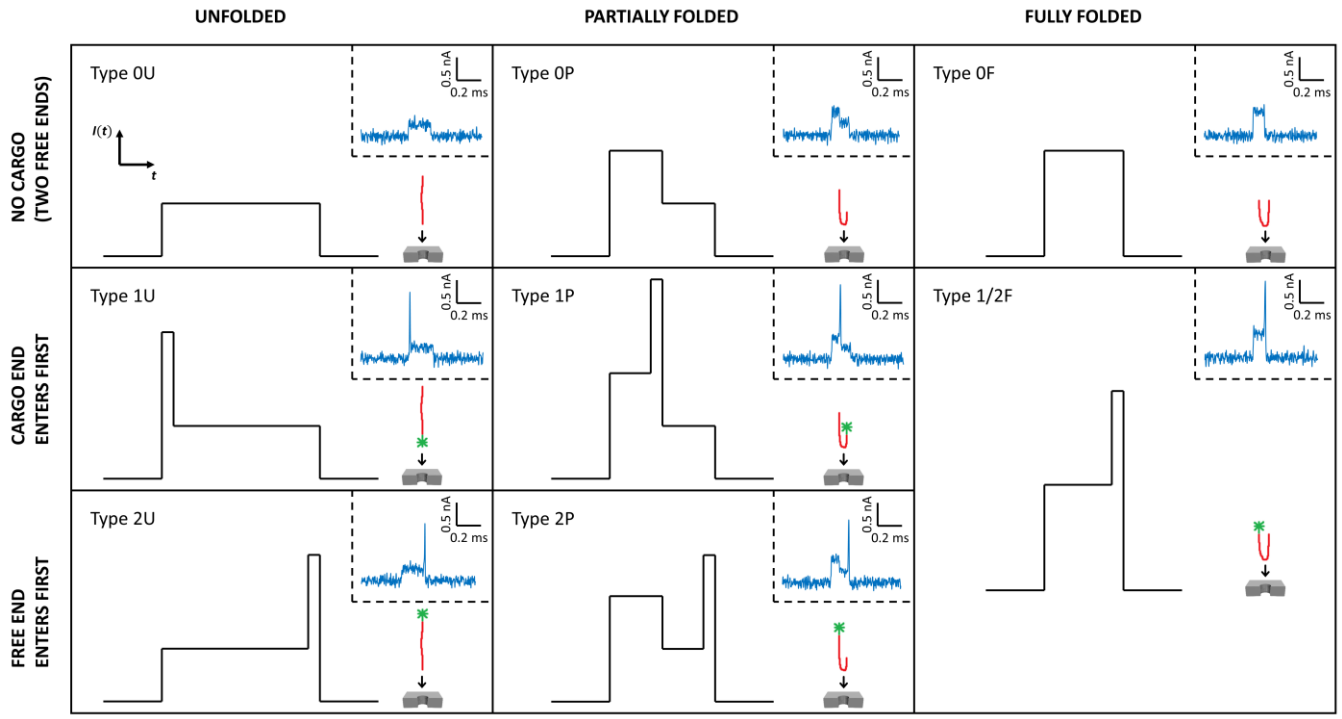


Figure 5.2: *Schematic of an event categorization scheme for nanopore translocations of a bulky cargo attached to a linear carrier. Ionic current events are grouped as entries of a 2D matrix related to the underlying conformation of the translocating molecules that generated them (see main text for a description of the two criteria that create the axes of this grid). A visualization of the sublevel sequence of each event type is presented, with the longer, shallower levels representing fitted states of the linear carrier as it passes through the pore and the shortest, deepest levels representing rapid translocations of the bulkier cargo. Insets for each event type show an illustration of the molecular conformation right before translocation through the pore (bottom right), as well as a representative current trace from the dataset of Table 5.1 (top right) that was sorted into the given category with the sublevel labelling approach outlined in Figure 5.3.*

The process for performing this categorization on a dataset of nanopore events is outlined in Figure 5.3. It starts with passing the raw or digitally low-pass filtered nanopore data to Nanolyzer to first locate the events within the ionic current timeseries and then to fit sublevels to the event substructure (relating to different possible *physical* substructures or configurations of the translocating molecule), based on criteria that can be adjusted in the software settings. The full configuration details of the analysis are provided in Supporting Information Section S1. The information from the output of this step on a single event is illustrated in Figure 5.3a and consists of sublevel duration, blockage depth, and time ordering, as well as approximately 30 additional metadata (see Supporting Information Section S2 for a full list of metadata used in this work).

From there, all the sublevels found in the first step (across all translocation events in the dataset) can be grouped into clusters based on any combination of these metadata using the clustering module in Nanolyzer. For the purposes of this work, we use sublevel duration and sublevel blockage. The result is a numerical label assigned to each fitted sublevel that maps it to the cluster to which it was assigned, as seen in the example scatterplot in Figure 5.3b. This associates each event with an ordered series of sublevel labels that encode the event substructure.

Ideally, for the types of ionic current signatures being investigated here (generated by carrier-cargo molecules), separate clusters will emerge from this process corresponding to: 1) unfolded carrier blockages, 2) folded carrier blockages, and 3) cargo blockages. The carrier blockages, corresponding to folded and unfolded segments of linear dsDNA, can have a broad spread in sublevel durations ($\gtrsim$ 1 order of magnitude)[10,42,43] and are typically observed with current blockage ratios ~2:1 when comparing folded (two DNA fragments in pore) and unfolded (one DNA fragment) blockages. Conversely, the DNA nanostructures comprising the cargo attachments are both bulkier (6+ dsDNA fragments) and shorter (~100×) than the carriers and so are expected to result in sublevels that are substantially smaller in duration and larger in blockage magnitude (see Figure 5.3b).

Once these cluster labels are defined, database filters based on SQL queries ("selection filters") are written in Nanolyzer to extract only the subset of events that follow a specific sequence of sublevel labels. For instance, Type 2P events feature a sublevel sequence of 'folded carrier' → 'unfolded carrier' → 'cargo'. This translates to a *label* sequence of '2' → '1' → '3' in the example of Figure 5.3 (see Figure 5.3c). Selection filters corresponding to each of the event types of Figure 5.2 are thus written based on their associated sequence of sublevel labels and are applied to the dataset to separate events into categories to be analyzed separately. A full list of selection filters is provided in the Supporting Information Section S3.

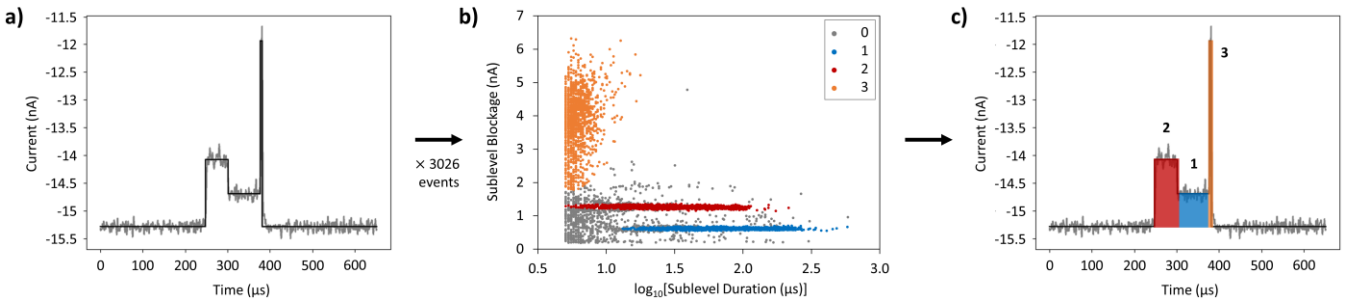


Figure 5.3: *Outline of the process to assign labels to event sublevels, using an example dataset of a 12-arm star cargo attached to a 2-kbp dsDNA carrier and passed through a ~11-nm pore (3.6 M LiCl, 175 mV).* ***a)*** *Sample event showing five fitted sublevels (including the two at the beginning and end of the event, respectively, associated with the open baseline).* ***b)*** *Scatterplot of blockage depth and dwell time for all fitted sublevels in the dataset (across all events), clustered into four groups with Nanolyzer. Group 1 will be associated with unfolded sublevels in the carrier, Group 2 with folded carrier sublevels, and Group 3 (short, deep sublevels) with blockages from the attached cargo. Group 0, which includes short, shallow fitted noise spikes, 3× multiples of the unfolded level, etc. will not feature in the analysis of this sample and so events containing these sublevels will be discarded by a rigid event sorting approach.* ***c)*** *Original event in (a) with newly-assigned sublevel labels superimposed.*

We performed this categorization process on a large dataset (>10,000 events) generated from passing 12-arm star nanostructures attached to a 2-kbp dsDNA carrier through a ~13-nm pore (3.6 M LiCl, 75 mV) with a measurement bandwidth of 1 MHz (sampled at 4.17 MHz) and digitally low-pass filtered for fitting at 250 kHz. As an initial step, a selection filter was applied on the fitted events to remove those without a "large" max deviation (>800 pA), to look at only events featuring an attached star (remove Type 0 events as labelled in Figure 5.2, see Supporting Information Section S4). The remaining events were then sorted as outlined in Figure 5.3, with the resulting statistics presented in Table 5.1. The majority (57.7%) of total events were successfully binned into the categories of Figure 5.2, and manually inspecting the current traces of a sample of events in each category confirmed that this assignment was accurate (i.e. had a low false positive rate) (See Supporting Information Section S5 for the first ten events from each category). This also implies that 42.3% of deep-blockage events did not match an expected event substructure ("Rejected" entry of Table 5.1), and while this is a significant fraction, it is regretfully typical for experiments involving structured molecules (i.e. rigid selection filters can have a high false negative rate). A later section, "*Rejected Events and Induced Biases*", will examine this subset of events in more detail.

Interestingly, there is a roughly equal distribution of the accepted events into each of the five categories, with the exception of Type 1/2F events which are sparsely distributed. This latter fact is perhaps to be expected, as it is unlikely that folded events will be folded exactly in the middle of the molecule within the time resolution of an experiment. In that sense, Type 1/2F events can be thought of as a mix of Type 1P and Type 2P events with exceedingly short, unfolded subevents that in the limit of infinite resolution could be properly sorted amongst those two categories. The roughly equal proportions of the remaining event types (Types 1U, 1P, 2U, and 2P) imply that there was no strong preference for these molecules to enter the pore in an unfolded or folded state, or with the star-end captured before the free end (or vice versa). Such trends are expected to depend strongly on the exact conditions of any particular experiment, however. For instance, shorter DNA carriers or smaller pores may reduce the degree of folding (promote Type U events over Type P) as the number of persistence lengths per molecule decreases or as the physical volume to accommodate multiple molecular fragments in the pore is reduced.[13] Similarly, particularly bulky attached cargos (relative to the size of the pore) may have to be forced through by prior capture of the linear carrier, imparting a preference for Type 2 events over Type 1 events.[25]

Table 5.1*: Statistics of analyzed translocations grouped by event type for the dataset of Figure 5.4 (12-arm stars bound to 2-kbp dsDNA carrier, ~13 nm pore, 75 mV, 3.6 M LiCl). The tabulated entries of dwell time, ECD, and max deviation for each event type represent the peak values of Gaussian fits to histograms of these quantities, while the range above and below these peak values is taken from the standard deviation of the fits. Since the widely-spread dwell time and ECD distributions are fitted to log-normal functions, their range extends slightly farther for positive deviations than for negative deviations when mapped back to a linear scale (see Figure 5.4). Recall that Type U refers to unfolded carriers, Type P to partially folded, Type F to fully folded (i.e. capture in the middle of the carrier), while Type 1 refers to the attached cargo entering the pore first, and Type 2 after the carrier. Here Type 0 events (missing cargo) are filtered out.*

Event Category	Count	Percentage (%)	Dwell Time (μs)	ECD (fC)	Max Deviation (nA)
Type 1U	1436	14.6	187^{+61}_{-46}	50^{+19}_{-14}	1.27 ± 0.43
Type 1P	1432	14.5	146^{+85}_{-54}	55^{+21}_{-15}	1.54 ± 0.44
Type 2U	1379	14.0	189^{+75}_{-54}	52^{+27}_{-18}	1.24 ± 0.37
Type 2P	1233	12.5	148^{+87}_{-55}	55^{+23}_{-16}	1.26 ± 0.44
Type 1/2F	209	2.1	100^{+34}_{-25}	54^{+16}_{-12}	1.41 ± 0.55
Rejected	4169	42.3			
Total	9858	100			

5.3.2 – Classifying by Dwell Time

Figure 5.4 shows how events from each subset are distributed over key metadata that will be relevant to our eventual goal of sorting mixtures of carrier lengths and attachment species by their current signatures. A scatter plot of max deviation over log[dwell time] for all events (colour-coded by event category) is presented in Figure 5.4a. From the dwell time distributions (Figure 5.4b), we see that fully unfolded events (Type U) pass through the pore in the longest amount of time on average, followed by partially folded events (Type P) and finally fully folded events (Type F). This makes sense if the impact of folding state on the electrophoretic velocity (e.g. due to interactions between the pore walls and the DNA as it passes through) is minimal here. Indeed, the more folded a molecule is, the shorter the distance between the segments of the molecule that first enter and last leave the pore, up to a factor of 2 when comparing fully folded and fully unfolded molecules, and we observed a factor of ~1.9 between the peak dwell times of Type U and Type F events of the current dataset (see Table 5.1).

Additionally, within a given folding state of the carrier, the two subtypes (Types 1 and 2) – corresponding to the sequence of molecule ends that enter the pore – remain consistent in dwell time (see Table 5.1). Together with the tight correlation of carrier folding state and peak dwell time, this is another indication that the relative sizes of the pore and the target molecule used here were in a regime where interactions between the pore walls and passing DNA were negligible. For instance, comparing the shapes of Type 1P and Type 2P events in Figure 5.2, we see that the cargo sublevel is directly preceded by a folded carrier sublevel in the case of Type 1P events and an unfolded carrier sublevel for Type 2P events. This means that there is an additional fragment of dsDNA in the pore during the translocation of the (already bulky) 12-arm star cargo of Type 1P events vs. those of Type 2P events. However, this is not associated with a significant difference in dwell times between the two subpopulations that might result from extra friction during Type 1P translocations (see the discussion surrounding Figure 5.7, below, for examples with smaller pore sizes where pore wall/cargo interactions cannot be ignored).

One issue with the wide distribution of dwell times observed above (due at least partly to variation in the folding state of the carrier), is that it limits the resolving power of this parameter in our original application of separating target molecules on carrier length. For instance, we would expect a carrier molecule that was 1.9× longer in length to, on average, take a longer time to pass through the pore by roughly that factor. But as we saw with the data of Table 5.1/Figure 5.4 this effect might be exactly undone if comparing fully folded instances of the longer molecule with unfolded instances of the shorter molecule (compounded further by the fact that a greater proportion of folding is typically observed with longer molecules[12]). One solution might be to discard all of the folded events (~50% of events for the selected data above, and ~29% of the entire dataset) in each population but 1) this may not be practical for already small datasets and 2) the proportion of folded events is again not expected to remain fixed across carrier lengths – the data from longer carriers in a sample will likely be depleted to a greater extent. Rejecting folded events could therefore severely limit our ability to compare experiments and extract meaningful conclusions when using a variety of molecular lengths.

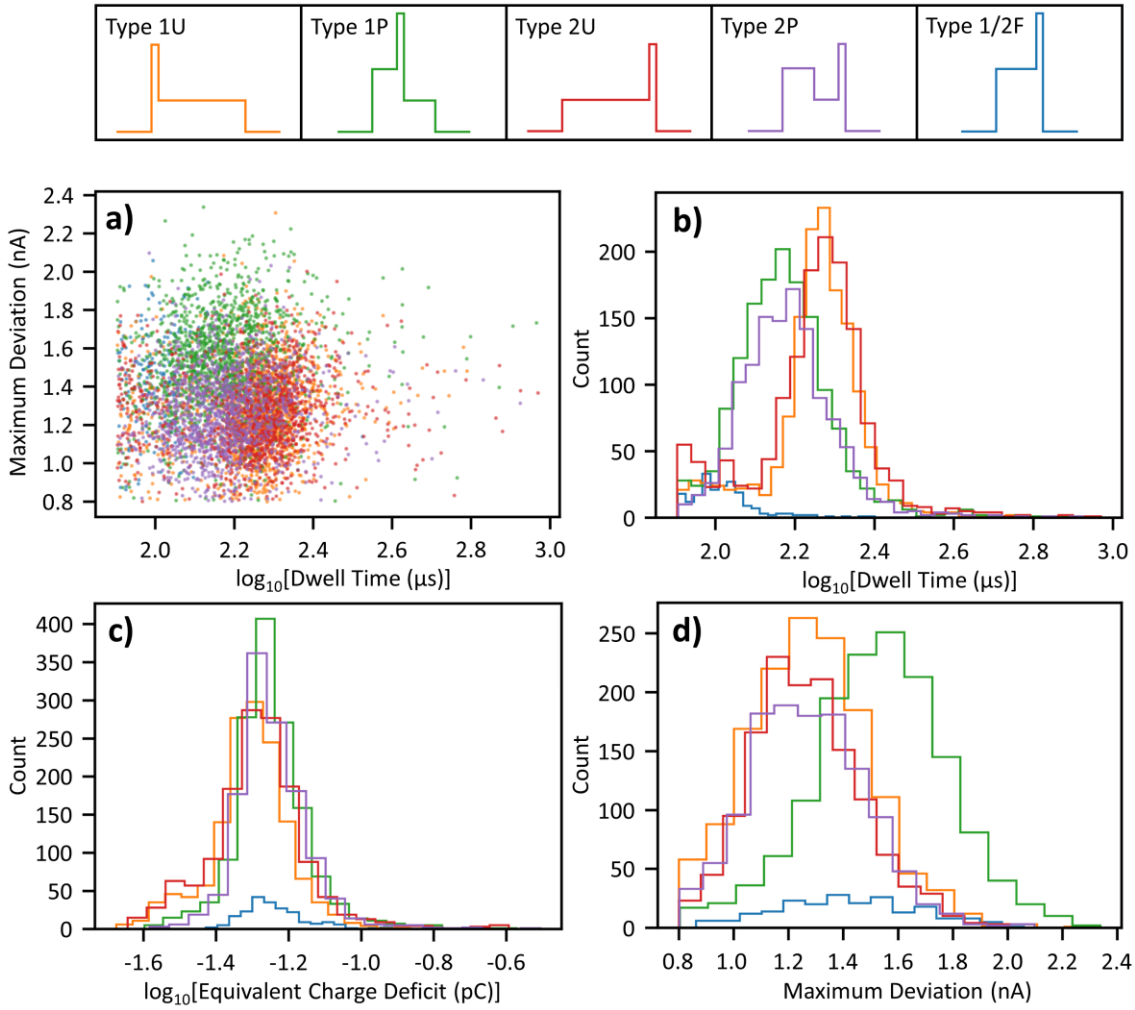


Figure 5.4: *Distributions of extracted translocation parameters grouped by event type (see Figure 5.2 for definitions) for the dataset of Table 5.1 (12-arm star cargo bound to 2-kbp dsDNA carrier, ~13 nm pore, 75 mV, 3.6 M LiCl).* ***a)*** *Scatter plot of max deviation vs. log[dwell time] for all events, colour-coded by subtype (see legend overhead).* ***b)*** *Histograms of log[dwell time] for each subtype.* ***c)*** *Histograms of log[ECD] for each subtype.* ***d)*** *Histograms of max deviation for each subtype.*

5.3.3 – Classifying by Equivalent Charge Deficit

A potentially more flexible solution is to instead bin the events on "equivalent charge deficit" (ECD). Equivalent charge deficit is the integrated area under each event in a current vs. time plot, in between the blocked current values comprising the event and the open-pore baseline (Figure 5.6a later illustrates a sample calculation). It is the deficit of electric charge that would have otherwise passed through the pore as ionic current had the translocating molecule not blocked it. Folded segments of linear molecules thus contribute *more* to an ECD

calculation compared to an equivalent unfolded segment through their higher current blockages but also *less* though their shorter dwell times (assuming similar passage speeds). If ECD is conserved for molecules of the same length but different folding states, the two effects perfectly cancel, within statistical variation.[11,13,44]

Figure 5.4c shows the ECD distribution for each subset of Table 5.1. The events from different subsets, corresponding to all the possible carrier folding states, are much more consistent in ECD, collapsing the separate dwell time distributions in Figure 5.4b to a shared ECD distribution in Figure 5.4c (compare also the two corresponding columns in Table 5.1). Now this collection of translocation events resulting from molecules of a single carrier length can be characterized by a single peak value relating to this length, making comparisons to other carrier lengths much easier. One thing to note is that this common peak ECD value for Types 1 and 2 events in the sample is slightly larger than that of Type 0 events in the same sample (analyzed separately, e.g. 55^{+21}_{-15} fC vs. 40^{+21}_{-14} fC for folded events with and without stars, respectively, see Supporting Information Section S6), which is related to the deep extra blockage present in Types 1 and 2 events from the attached cargo. The presence of a cargo can therefore influence the ECD value of a translocation event, which may impact the accuracy of this classification scheme. An approach to removing the influence of the cargo on carrier-specific length metrics will be discussed in a later section "*Classifying by Truncated ECD*", below.

5.3.4 – Classifying by Max Deviation

Finally, Figure 5.4d shows the distribution in maximum deviation for each subset (see also the corresponding column in Table 5.1). As a reminder, the data point corresponding to the max deviation in each event is expected to appear during the subevent associated with the translocation of the bulky attached cargo (assuming it is present, see Figure 5.2) and so indirectly characterizes the size of the attachment through how much current is blocked. Interestingly, three of the five event types (Types 1U, 2U, and 2P) share an approximately equal max deviation distribution, while that of Type 1P events is noticeably shifted toward deeper blockages. This can be taken as evidence of the extra carrier fragment alluded to earlier that is present in the pore during the passage of the cargo, blocking extra current. Type 1/2F events, which can be seen as a mixture of Type 1P

and Type 2P events with unfolded subevents too fast to be resolved, have a peak max deviation that is intermediate between the two other subpopulations (compare the relevant entries for max deviation in Table 5.1), as might be expected. In the context of our original goal of separating molecular mixtures of carrier-attached cargos of different sizes by their current signatures, we see how the presence of Type 1P and Type 1/2F events could complicate this process by "artificially" broadening distributions that characterize their blockage depths, creating more overlap between distributions of different cargos. An approach of first grouping translocation events by molecular conformation (rather than just considering metrics of the entire global population) therefore has the advantage of separating some of the carrier versus cargo effects and provides a clearer picture of the true magnitude of the attachment blockage.

5.3.5 – Rejected Events and Induced Biases

Current traces and their fitted sublevels for a few representative examples of the events that were *not* binned into the categories of Figure 5.2 (42.3%) are shown in Figure 5.5. Manually inspecting a sample of these rejected events, we note that almost all of them fall into rejection modes corresponding to recognizable cargo-carrier events (Figs 5.5a – 5.5f), with a majority of these involving unfitted (Fig. 5.5a) or *over*fitted (Fig. 5.5b) cargo sublevels. These are false negatives – easily human-recognizable as representing our molecule of interest, with subevents containing usable information that is currently being discarded with this rigid sorting approach due to fitting artefacts. This contrasts with events as in Figures 5.5g and 5.5h, which cannot easily be associated with a particular conformation of our intended target molecule. An important point to note is that if events from each of the categories of Figure 5.2 differ in their rates of generating false negatives, it will introduce biases into the statistics of the population (such as in Table 5.1). For instance, as discussed previously, the supposed "cargo" sublevel of Type 1P events (and some of Type 1/2F events) is actually the convolution of subevents from both the cargo and an extra carrier fragment translocating together. If the additional blockage from the carrier fragment results in a deeper or longer-lived "cargo" subevent, it may stand a better chance of being fitted to a sublevel and thus not falling into the rejection mode of Figure 5.5a when compared to its Type 2P analogue.

Alternatively, deeper or longer-lived cargo subevents may be fitted with extra structure more frequently (beyond the single level expected by the selection filters of Figure 5.2), which is the rejection mode of Figure 5.5b. More generally, even among the events of a single type, if there is a correlation between the value of a parameter and the rate at which its parent event is rejected, the extracted values for the population will be biased in that direction rather than reflecting the entirety of the data. Understanding and limiting these biases is critical for nanopore sensing applications, especially those involving low numbers of total events and/or precisely quantifying the statistics of a parameter of interest (e.g. as in an assay for a low copy number target).

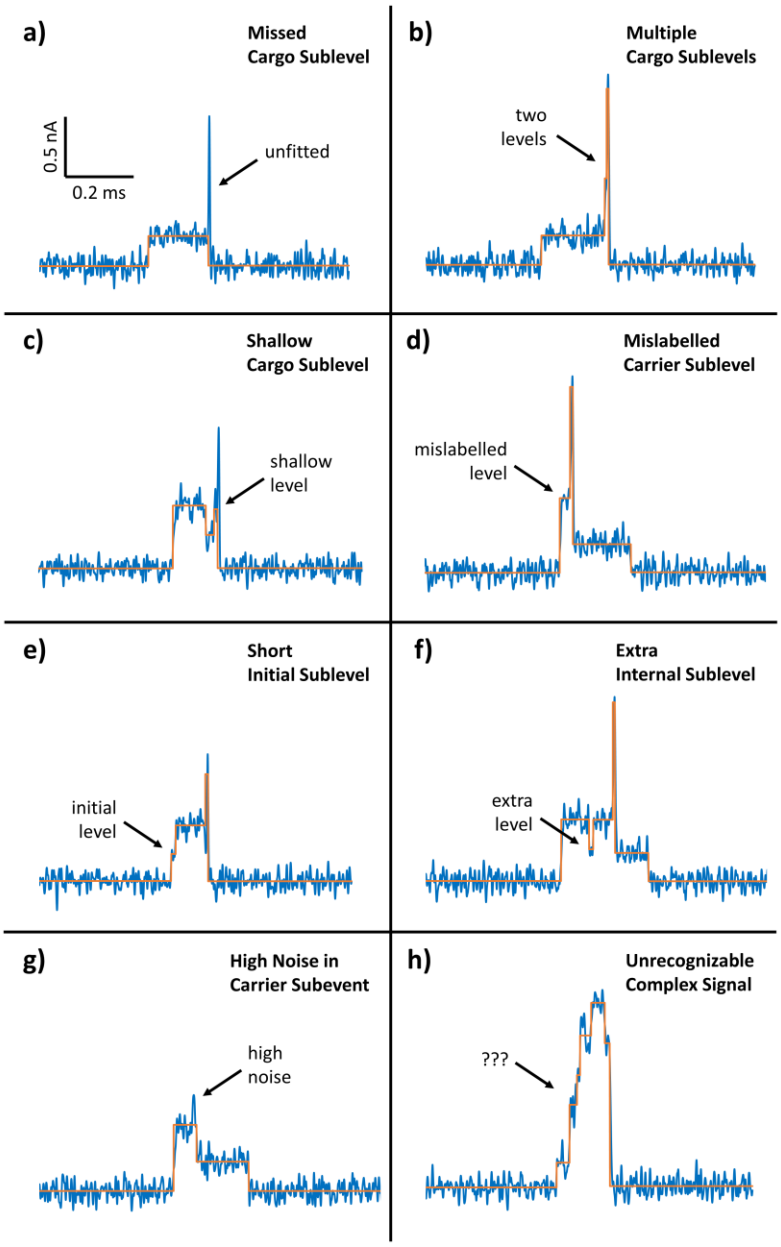


Figure 5.5: *Gallery of several events from the dataset of Figure 5.4 that failed to be sorted into the categories of Figure 5.2 ("Rejected" row in Table 5.1) illustrating some of the reasons for this assignment.* ***a)*** *Cargo subevent missed.* ***b)*** *Cargo subevent fitted with multiple levels.* ***c)*** *Cargo subevent fitted with shallow level that will not be separable from folded carrier level.* ***d)*** *One of carrier levels not assigned into a carrier label (e.g. slightly deeper/shorter/etc. than bulk of other like sublevels).* ***e)*** *Event features short fitted initial level, perhaps due to access resistance blockage.* ***f)*** *One of carrier subevents features extra sublevel (arising perhaps from noise in signal or additional structure in physical conformation, e.g. knot/kink in backbone).* ***g)*** *Events lacking obvious cargo subevent (likely Type 0 events that were not removed in the initial max deviation threshold due to noise that exceeded this threshold in their deepest/folded subevent).* ***h)*** *Events that are not recognizable as resulting from simple conformations of our molecule of interest (may be due to contamination in the DNA sample, aggregates of multiple molecules, highly contorted individual molecules, etc.) Broadly speaking, (a) – (f) represent false negatives while (g) – (h) are examples of true negatives.*

This issue of overzealous event rejection resulting from a rigid classification approach is expected to only become more problematic when sensing mixtures of target molecules with several cargo sizes. The dataset of Table 5.1/Figure 5.4 is one that is particularly amenable to such an analysis approach and is not necessarily representative of all datasets featuring combinations of DNA carriers and cargo molecules. First of all, it is large, containing $\gtrsim$ 10,000 events for a single carrier length/cargo size pair. In other cases, such as when a pore features a low capture rate, or is rendered inoperable early on from a permanent clog, or is sampling many carrier + cargo pairs each at a reduced concentration, it may not be possible to reject a large fraction of the total events (~42% in the example above) and still be left with statistically valid counts in each category with which to draw conclusions.

More importantly, this dataset results from experimental conditions (pore size, cargo type, applied voltage, RMS noise vs. amplifier bandwidth, etc.) in which clustering sublevels and filtering events on sublevel sequence is successful in its simplest form. For instance, if a nanopore is "too big" relative to a particular cargo, the event substructure resulting from translocations of this attachment on the carrier may be too fast to reliably fit[45], or feature attenuated blockage depths that are not easily separable from one another. On the other hand, if a nanopore is "too small", passing cargos may interact strongly with its walls, resulting in long events with complex substructure that reflects cargo-pore interactions rather than molecular identity, dominating the event metadata[46] (e.g. dwell time, ECD) and obscuring the subevent sequence that identifies the carrier conformation – see Supporting Information Section S9 for example current traces of carrier-cargo hybrids passing through a smaller (~8 nm) pore. Moreover, for a sample containing a mixture of cargo molecules, the size of the pore may represent one interaction regime for the smallest of the cargos (e.g. fast, shallow translocations) and a different one for the largest (e.g. slow, complex translocations). This means that if two carrier + cargo species are sensed by the same pore, it is very unlikely that similar rigid event filters applied to the translocation signals will give equal overall rejection rates (or equal weights to each rejection mode) in both cases, resulting in different imparted biases to the data from each species, and making comparison between experimental replicates difficult.

For these reasons, we next explore a refinement of the analysis approach that relies less on (inconsistent) sublevel fitting to the cargo subevents and develop additional event metrics that allow for the accurate characterization of a wider range of molecules with a given nanopore.

5.3.6 – Refinements to Cargo-Carrier Event Characterization

As observed in the earlier ECD classification section, it is tempting to choose ECD as a metric to classify target molecules by carrier length, as this quantity is better conserved across folding states of a given carrier compared to dwell time (when carrier-pore wall interactions are negligible). One issue however, is that ECD is a measure of the entire event, including contributions from the attached cargo molecules in our case. Evidence of this was also observed in that earlier section (discussing the data of Figure 5.4) where the peak ECD value of Type 1 & 2 events in that dataset was noticeably larger than that of Type 0 events, preventing ECD from being distinctively a measure of the carrier type used there. One consequence of this is that a spread in the magnitude of current blockages of the cargo contributes to a spread in the ECD of the entire event, reducing our ability to resolve two or more populations centred on different ECD values. Figure 5.6c shows an example of this effect with a 2D histogram (max deviation vs. log[ECD]) of events from a mixture of 2-kbp and 7-kbp carriers with attached 12-arm star cargos (~11-nm pore, 3.6 M LiCl, 200 mV). Two populations are visible at smaller (2 kbp) and larger (7 kbp) ranges of ECD values but are not particularly well separated despite the ~3.5× difference in carrier length of the molecules that produced them. Part of the problem is the correlation between max deviation and ECD for these events. When plotted on this logarithmic scale, the effect is more drastic for the shorter molecules, since the ECD contribution from the cargo represents a larger fraction of the total ECD in that case and so results in a steeper slope of log[ECD] vs. max deviation. Clearly more carrier-focused metrics of the event signatures are needed to better classify events.

5.3.7 – Classifying by Truncated ECD

To counteract the effect of the cargo on the total ECD, a parameter is defined that we call "truncated ECD" (TrECD). TrECD is defined such that only the sublevels that were associated with current blockages from the carrier during the clustering stage (clusters '1' & '2' in the example of Figure 5.3) enter the ECD calculation. Each such sublevel adds a contribution equal to its fitted blockage level multiplied by its duration (see Figure 5.6b). Using the fitted levels serves to further insulate this carrier-specific metric from any large but fast current deviations (e.g. from unfitted cargo blockages, extra noise, etc.) that could lead to less consistent calculated values among events from a single carrier-length. In this way, Figure 5.6d includes the same events as in Figure 5.6c but with its *x*-axis binned on TrECD instead. The two populations are now better separated in *truncated* ECD-space – especially the largest max deviation 2-kbp events from the smallest max deviation 7-kbp events – due to the greater consistency in TrECD value across the range of current blockages from the star junction. Two peaks exist in the TrECD distribution at ~134 fC and ~493 fC, separated by a factor of ~3.7×, in reasonable agreement with ratio of carrier lengths. (See Supporting Information Section S7 for the corresponding 1D histograms on ECD and TrECD).

Figure 5.6e shows another 2D histogram of max deviation vs. log[ECD], this time for a sample with *multiple* star cargo types (12-arm and 6-arm) attached to the *same* carrier length of 2 kbp (~8-nm pore, 3.6 M LiCl, 200 mV), and serves to illustrate the shortcomings of ECD as a carrier-focused metric under these circumstances. At least three populations are visible in the plot, from lowest to highest max deviation: i) a compact cluster at the bottom corresponding to folded carriers without an attached DNA star (unfolded Type 0 events have been removed prior to plotting via a threshold on max deviation to reduce clutter), ii) an intermediate population with a narrow spread in ECD but a large spread in max deviation corresponding to carriers attached to 6-arm stars, and iii) events with the largest blockages and a wide spread in ECD corresponding to carriers attached to 12-arm stars. (See Supporting Information Section S8 for a control of 6- and 12-arm star-attached carriers run separately on a single pore). The folded Type 0 events, lacking a star attachment to interfere with an ECD calculation, are presumably peaked around the "true" ECD value of the 2-kbp carrier in these conditions.

Above that, the 6-arm star + carrier events start with their lowest max deviations at ECD values consistent with the peak of the Type 0 events, but then slant off toward larger ECDs as max deviation increases, just as observed for the samples in Figure 5.6c. Finally, the 12-arm star cargo + carrier events have both the largest ECD values and a very broad spread in these values.

The reason for this disparity between molecules with different cargos in the widths of their ECD distributions is that with the smaller nanopore diameter used here (~8 nm), the larger 12-arm stars interact strongly with the pore as they translocate, slowing them down considerably and imparting a significant variance to their dwell times, while the smaller 6-arm stars pass relatively freely (see Supporting Information Section S9 for sample current traces of both cases). Applying truncated ECD to this distribution again (Figure 5.6f), we now see both populations aligned above the peak value of the Type 0 events. Despite the chaotic cargo subevents that appear in the current signals of the translocating 2-kbp + 12-arm star hybrids, TrECD removes their influence on each event by looking only at the well-behaved, easily-fitted sublevels of the carrier (while relying on the presence/magnitude of the star subevent to further classify them). Comparing Figures 5.6e and 5.6f, in the latter case there now exists clear room to the “right” of these 2-kbp populations in TrECD-space where events from longer carrier lengths could reside and be well separated.

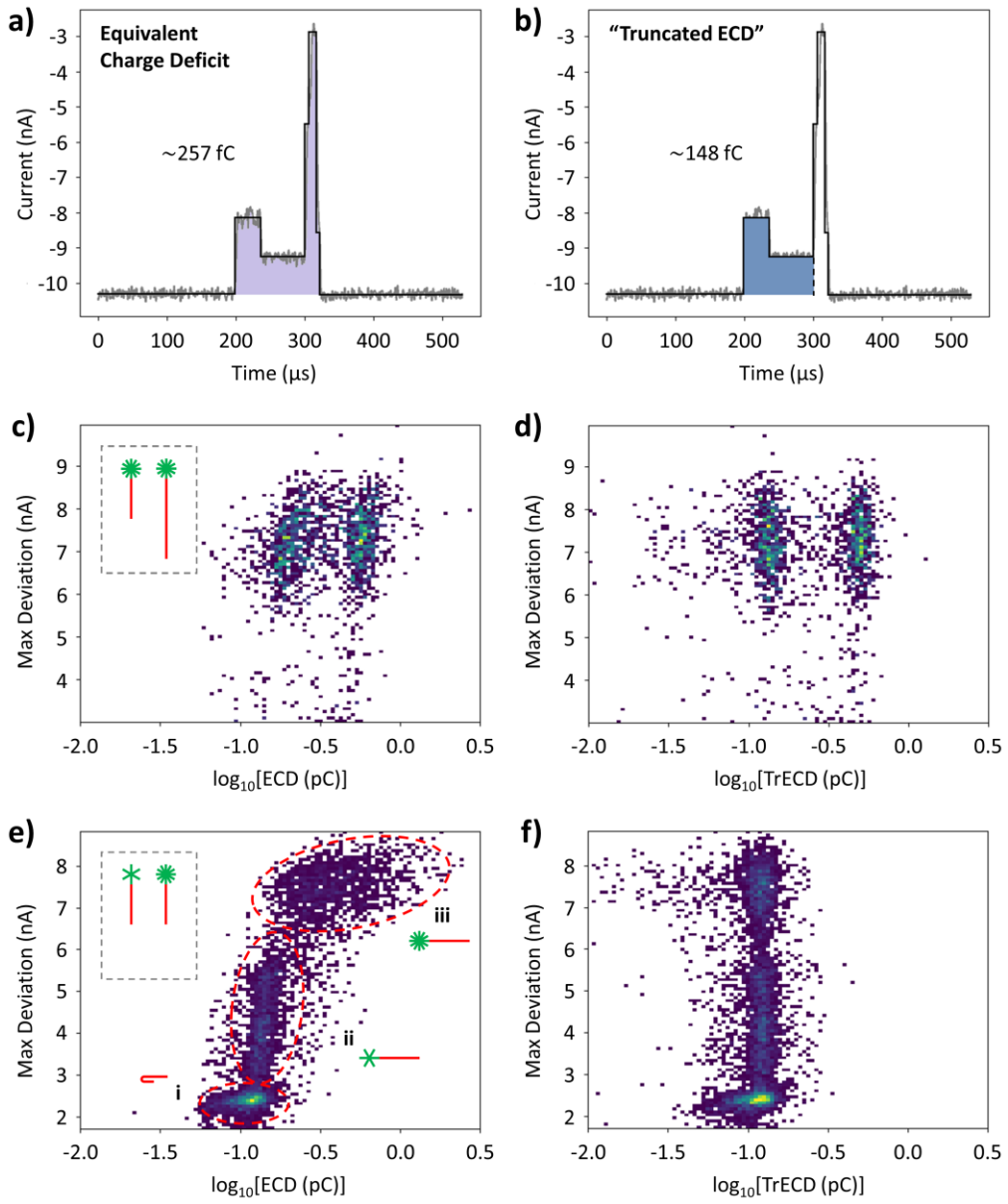


Figure 5.6: *Comparison of total equivalent charge deficit (ECD) and an optimized version of this quantity, truncated ECD ("TrECD"), for classifying populations of DNA carrier + cargo translocation events by carrier length.* ***a)*** *Illustration of an ECD calculation for a sample translocation event (2 kbp carrier + 12-arm star cargo, ~8-nm pore, 3.6 M LiCl, 200 mV), visualized as the shaded area between the data points comprising the blockages and the fitted baseline (open pore) current.* ***b)*** *Illustration of a TrECD calculation for the same event as in (a), incorporating only the sublevels that were associated to a blockage from the carrier during a clustering step (see Figure 5.3). Note that these sublevels enter the TrECD area calculation only through their single-valued, fitted current blockages (multiplied by their durations in time), rather than by the distance of every raw current point to the baseline as in (a).* ***c)*** *2D histogram of max deviation vs. log[ECD] for translocation events from 12-arm star junctions attached to a mixture of carrier lengths (2- & 7-kbp, ~11-nm pore, 3.6 M LiCl, 200 mV).* ***d)*** *2D histogram of max deviation vs. log[TrECD] for the same events as in (c).* ***e)*** *2D histogram of max deviation vs. log[ECD] for translocation events from 2-kbp carriers attached to a mixture of star types (6- & 12-arm, ~8-nm pore, 3.6 M LiCl, 200 mV).* ***f)*** *2D histogram of max deviation vs. log[TrECD] for the same events as in (e).*

5.3.8 – Extracting Events with Relaxed Selection Filters

Another step of the analysis process with room for improvement is in isolating the subsets of events that are the most useful in identifying the carrier lengths/cargo types of a sample. As examined in earlier sections, an approach to this that is based on recognizing an exact sequence of expected sublevels is optimized only for a narrow range of experimental conditions such as pore size: large pores (relative to the cargo dimensions) lead to short, poorly-fitted cargo subevents, while small pores can produce long, complex subevents that generate a variable number of fitted sublevels. Even with a molecular sample featuring a single cargo type that is well matched in size to the pore diameter, there can exist enough variability in how the events are fitted that a substantial portion of them is rejected as false negatives under this scheme, as evidenced by the statistics of Table 5.1 and rejection modes of Figure 5.5, above. While Nanolyzer settings can be tuned to optimize fitting, a certain amount of variability will always exist and will usually be correlated with particular subpopulations of events (e.g. those with subevents that are short compared to the rise time of the system are routinely underfitted).

As discussed, much of this variability in event substructure can be linked to the fitting of subevents from the cargo signals. A more versatile approach then would involve “relaxed” event filters based only on the (well-fitted) sublevels of the carriers as well as individual points of “large” current deviation that could be associated with the presence of the bulky attachments, regardless of how they were fitted.

5.3.9 – Selection Filters Based on Local Max Deviation

An example of a simple set of two such selection filters that follow these principles is presented in Table 5.2, where each selection filter is designed to capture events in a separate regime of cargo subevent fitting. The first (“Filter A”) looks for Type 2 events (“free end first”, see Figure 5.2) where the cargo subevent has been missed by sublevel fitting, and so the last fitted sublevel of the event is instead that of an unfolded carrier (see Figures 5.5a and 5.7a for representative examples). This condition alone (#1 in Table 5.2) could be satisfied by an event from any of the unfolded or partially folded event categories (Types U and P), regardless of which end

(if any) the cargo was attached to, and so it is supplemented with a second condition (#2) that checks if the event is also characterized by a large terminal spike in current, as would only be the case with Type 2 events. The test for this spike is if the maximum deviation *within only* the final baseline sublevel ("local max deviation") exceeds a set threshold value (see Fig. 5.7a). Some strategies for choosing this threshold value are presented in the discussion surrounding the specific dataset of Figure 5.7, below.

A second filter ("Filter B") looks for the remainder of Type 2 events – those with at least one fitted cargo sublevel. It starts with the condition (#1) that the event ends on a sublevel *other* than that of an unfolded carrier. Note that this makes Filters A and B mutually exclusive from their respective first conditions alone – they will extract out different sets of events that can later be combined without any repeats. Filter B also contains a condition (#3) that checks for local max deviations above a threshold value (analogous to Condition #2 of Filter A), but this time the threshold may be crossed in one or both of the last two sublevels, to account for Filter B events having at least one extra sublevel before the baseline through Condition #1 (see Figure 5.7b). Finally, an extra condition (#2) is added to Filter B that its events contain at least one unfolded carrier level somewhere (though not at their very ends, as this is precluded by Condition #1). This is aimed at removing any fully folded (Type F) events as well as any true negatives that are not recognizable as simple carrier-cargo events, as in the example of Figure 5.5h.

Together then, Filters A and B are capable of extracting Type 2 ("free end first") carrier-cargo events with relative indifference to how well their cargo subevent signals were fitted to sublevels. By focusing only on the end of these events, they also lead to fewer false negatives arising from extra/alternative sublevels being fitted to earlier carrier subevents, as in the examples of Figures 5.5d-f. Note that this example set of selection filters ignores Type 1 events, since these may be problematic to our application of sorting cargo-carrier types due to the deeper blockages of Type 1P events (see Figure 5.4d) and the potential for confusion between the blockages of small cargo attachments and those of folded free carriers (Type 1U vs. Type 0P events). However, very similar filters that select *for* Type 1 events could be constructed by focusing instead on the sublevels at the start of an event (fitted carrier sublevels and sublevel max deviations), should the need arise.

***Table 5.2**: Example set of "relaxed" selection filters that select for Type 2 events using only the sequence of fitted sublevels from the carrier and the "local max deviation" of individual sublevels. All numbered conditions must be satisfied simultaneously for an event to be classified under a given filter.*

Event Filter	Conditions	Target
Filter A	1. event ends in unfolded carrier sublevel 2. max deviation of last (baseline) sublevel > threshold value	Type 2 events without fitted cargo sublevel(s)
Filter B	1. event ends not in unfolded carrier sublevel 2. event contains unfolded carrier sublevel (elsewhere) 3. max deviation of last (baseline) sublevel > threshold value or max deviation of 2nd last sublevel > threshold value	Type 2 events with fitted cargo sublevel(s)

Figure 5.7 shows the result of applying the selection filters of Table 5.2 to some carrier-cargo translocation events. The dataset chosen is the same as in Figures 5.6e and 5.6f and was generated from a sample of 6- and 12-arm stars attached to 2-kbp carriers and passed through a relatively small, ~8-nm pore (3.6 M LiCl, 200 mV). As discussed above, this combination of pore size and attachment types resulted in distinct regimes of cargo subevent fitting being represented: fast, poorly-fitted subevents (from "small" 6-arm stars) and slow, multi-level subevents (from "large" 12-arm stars) – see also Supporting Information Section S10 for an application of these filters with a larger, ~12-nm pore where both attachments fall in the "fast" regime. For the local max deviation threshold required by the filters, a value was chosen (1500 pA) that was intermediate between the average values of the unfolded (1040 pA) and folded (2142 pA) carrier levels. This represents a reasonable balance between the need to clear noise above an unfolded level while still capturing attenuated spikes from fast, folded subevents; an alternative criterion might be to use a fixed number of standard deviations above the unfolded carrier level (~4σ above 1040 pA here with the use of a 1500 pA threshold and a 500 kHz low-pass Bessel filter).

The output of Filter A on this dataset is presented in Figure 5.7c as a 2D-histogram of max deviation vs. log[ECD]. Most events captured here are from molecules with the smaller 6-arm star attachments, specifically those without a fitted cargo sublevel (see Figure 5.7a for the current trace and fitted sublevels of a typical event). In contrast, applying Filter B to the same dataset mainly extracts events from molecules with 12-arm star attachments, which feature wide spreads in dwell time/ECD due to interactions between this larger structure

and the pore walls (top right population in Figure 5.7d, sample event in Figure 5.7b). Also collected by Filter B is a smaller subset with shallower blockages and shorter dwell times (middle population in Figure 5.7d) that represents some of the deepest of the 6-arm star events where a cargo sublevel was successfully fitted.

Figure 5.7e shows the union of the outputs from Filters A and B plotted on max deviation vs. log[ECD] while Figure 5.7f plots the same data on max deviation vs. log[TrECD], as defined earlier. This last plot illustrates the success of these filters – they result in two distinct populations (corresponding to the two attachment types) that are separable in max deviation and characterized by a single peak in TrECD, reflecting their shared 2-kbp carrier lengths. Note that part of the reason for this degree of separability came from targeting only Type 2 events, which had the effect of distinguishing the shallowest of the 6-arm star events from those of folded free carriers (no Type 1U events to be confused with Type 0P) as well as distinguishing the shallowest of the 12-arm star events from the deepest of the 6-arm star events (Type 1P, removed) – compare Figure 5.7f with Figure 5.6f, above.

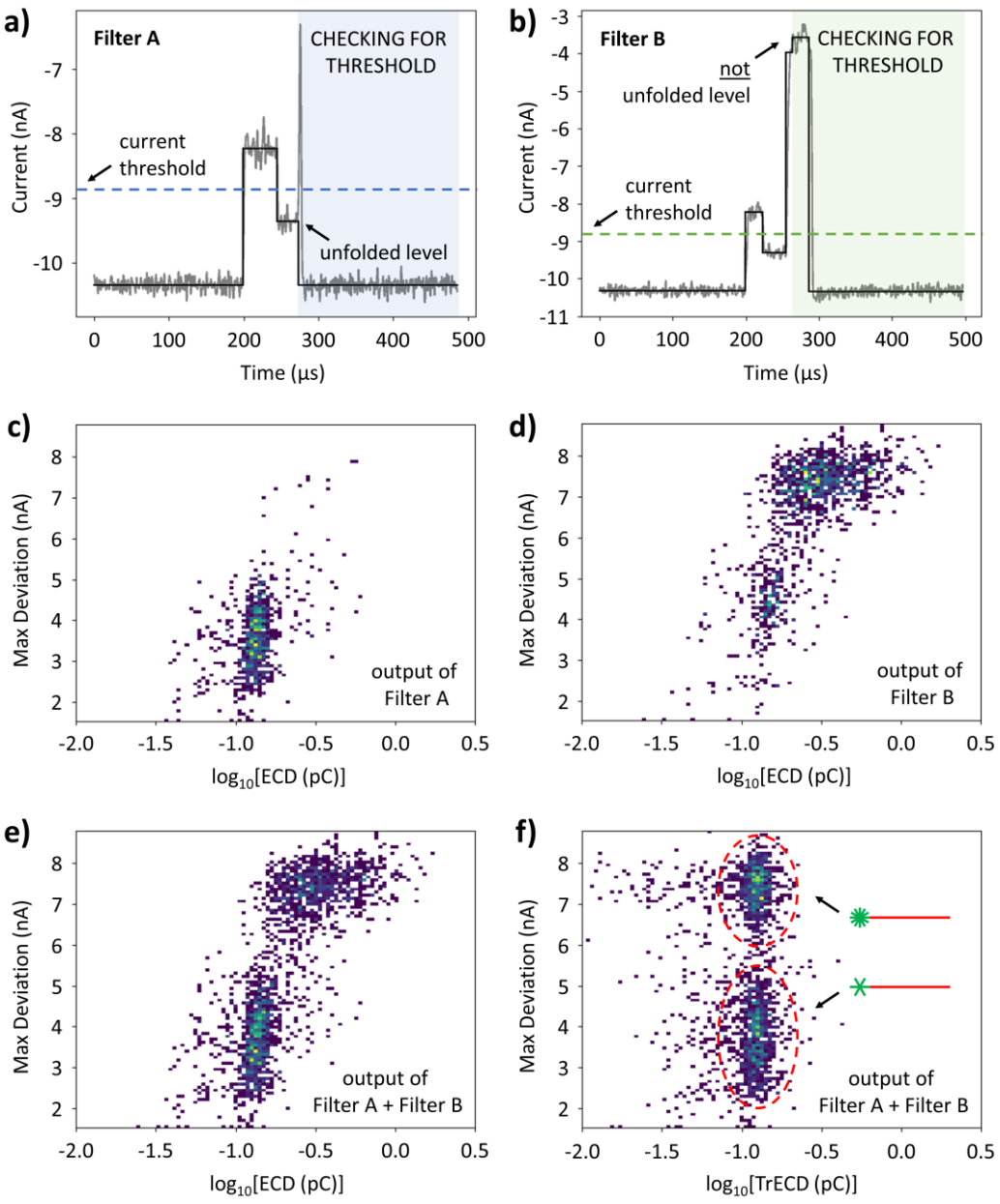


Figure 5.7: *Using simple selection filters based on fitted carrier sublevels and the "local max deviation" of individual sublevels exceeding thresholds to extract Type 2 events from a complex dataset (6- or 12-arm star cargos attached to 2-kbp carriers, ~8-nm pore, 3.6 M LiCl, 200 mV).* ***a)*** *Schematic of the action of "Filter A" (see Table 5.2 for definition) on a typical event captured with this filter. Illustrated are the fitted sublevels of the event, the current threshold for the local max deviation to cross, and the sublevels over which this threshold is active.* ***b)*** *Schematic of the action of "Filter B" (see Table 5.2 for definition) on a typical event captured with this filter.* ***c)*** *2D histogram of max deviation vs. log[ECD] for the output of Filter A.* ***d)*** *2D histogram of max deviation vs. log[ECD] for the output of Filter B.* ***e)*** *2D histogram of max deviation vs. log[ECD] for the output of the union of Filter A and Filter B.* ***f)*** *2D histogram of max deviation vs. log[TrECD] for the output of the union of Filter A and Filter B. (See the discussion surrounding Figure 5.6 for a definition of "truncated ECD").*

In total, 1976 Type 2 events were extracted from the data of Figure 5.7 using the selection filters of Table 5.2 – 795 events with Filter A and 1181 events with Filter B – and importantly, these events spanned a range of cargo subevent fitting behaviors (0, 1, or 2+ fitted sublevels) and cargo subevent blockage depths resulting from the two different-sized star cargos. In comparison, when the rigid sublevel sequence approach of earlier sections was applied to the same data (see Supporting Information Section S11) only 540 total events, including 217 Type 2 events (an 89% reduction), were recovered, due primarily to the requirement that exactly one cargo sublevel be present per event and that this cargo sublevel be substantially deeper than that of a folded carrier in order to be labelled as such. Moreover, when examining how these 540 events are distributed over max deviation, a strong bias selecting specifically for Type 1P events from molecules with 6-arm star attachments becomes apparent (Figure 5S.8d). As was postulated in the discussion of Figure 5.5, this likely corresponds to the type of event that is most easily fitted with a single cargo sublevel under these experimental conditions (e.g. deeper than other 6-arm star events due to the extra carrier fragment but not as prone to being fitted to multiple sublevels as 12-arm star events).

In summary, an approach to extracting useful events that is based on minimum sets of shared characteristics (e.g. carrier sublevels, threshold crossings) of their current signals rather than exact sublevel sequences can provide a dramatic improvement to the data curation and event segmentation process by being applicable to molecules in different size regimes simultaneously and by introducing fewer selection biases on the output (arising from bandwidth limitations, event fitting issues, or otherwise).

5.4 – Conclusion

In this work, we presented the analysis of nanopore data generated by a basic structured polymer: an extended length of linear, double-stranded DNA (carrier) attached at one end to a DNA nanostructure (cargo). First, we organized the translocation events of a dataset from a single carrier-cargo pair by the folding state of their parent molecules to assess the contribution of molecular conformation to the distribution of several key metadata such as event duration and maximum (current) deviation. In the process, it was also observed how

inconsistency in the fitting of discrete levels to the current signals can lead to many events being rejected by a categorization scheme that relies on recognizing exact sequences of these fitted sublevels. With this knowledge in hand, refinements of existing analysis tools/metrics were developed (e.g. truncated ECD, local max deviation) which, when combined with basic filtering and applied to datasets from mixes of molecular species (carrier length and cargo type), led to improved separation of different populations, lower rates of false negatives, and selection filters suited for data from a wider range of experimental conditions.

Over the course of analyzing data from this model system, a general approach emerged that should be applicable to most nanopore experiments on structured polymers: focusing on the most reduced subset of features from the full event substructure that, in a given application, groups populations of like molecular signals together. This allows for more events to be retained that might otherwise be rejected (e.g. folded events) on the basis of variation that emerges from different molecular conformations or different levels of fitting success. In the future, this approach could be further improved by incorporating a peak-finding or threshold-crossing algorithm to locate the short current spikes in the signal and therefore completely decouple the information found in them from fitted levels. Each event type could then be identified by the ordered list of carrier sublevels and local maxima/threshold crossings, which should provide rich information on which to base more universal event filters.

Another avenue for future investigation is in characterizing the efficiency of particular selection filters by quantifying the numbers of false negatives removed and the numbers of false positives that remain in the subsets they create. This would require the development of further metrics to distinguish accurately- from inaccurately-labelled events in both cases, something that may be challenging depending on how complex these distinctions are in a particular application. This is an area where approaches that incorporate a machine learning model may prove useful by only having to manually identify a limited number of representative examples from each set (or in simulated data with known ground truth) as training data; from there, the remainder of the false positives/negatives could be identified by progressively more accurate automated classification. Regardless of the method used, reducing the numbers of mislabelled signals will remain an important consideration for

minimizing statistical bias as nanopore sensing of complex analytes continues to be established as a powerful bioanalysis technique into the future.

Going forward, the somewhat laborious event characterization process presented here (of clustering sublevels and defining selection filters) would ideally not be something to be performed after each nanopore experiment. Automation of this process will be a key component of incorporating nanopore-based molecular assays into real-world diagnostic technologies. To that end, we hope that some of the tools and approaches presented here will also find utility in the development of more complex AI/ML event classification models, by providing unbiased training data curated through these techniques. This way, event classification will only need to be performed once when a new experimental system is explored, and subsequent analysis can happen quickly based on the resulting trained model.

5.5 – Methods

5.5.1 – DNA Sample Preparation

DNA carriers used in this study were synthesized by one of two methods. In the first method, a polymerase-chain reaction (Q5 DNA Polymerase, New England Biolabs) was carried out on a M13mp18 template (NEB) to create a double-stranded fragment of variable length using a phosphorylated forward primer and a phosphorothioated reverse primer (Integrated DNA Technologies). The phosphorylated strand was then selectively digested with λ-exonuclease (NEB) to leave a single-stranded scaffold. Finally, short (~40 nt) complementary oligonucleotides (IDT) were hybridized to this scaffold, creating a dsDNA fragment with many regularly-spaced nicks on one strand[47] (subsequently repaired with T4 DNA ligase, NEB). The terminal oligonucleotides were chosen to be longer than the ends of the scaffold by 12 bases, resulting in single-stranded overhangs through which cargo molecules could be attached via a hybridization reaction.

In the second method of carrier synthesis, a variable length PCR product was again created, this time using lambda DNA (NEB) as a template. With this method, the PCR product is then digested with a restriction enzyme

(e.g. PciI, NEB) to create short (e.g. 4 nt) sticky ends which are used to circularize the fragment (T4 DNA ligase, NEB). Finally, a unique sequence (*cos*) of lambda DNA[48] present in the circularized fragments is used as the target of a λ-terminase enzyme (Catalano Lab, University of Colorado) which acts to create staggered nicks in the *cos* sequence,[48] opening the DNA circles into linear fragments with 12-nt single-stranded overhangs. Carrier molecules generated by each of the two methods (oligo assembly and λ-terminase digestion) were found to perform comparably and were used interchangeably throughout the study.

For cargo molecules, DNA star junctions were synthesized as previously described.[14,27] Briefly, short (48-nt) DNA oligonucleotides, each spanning two of the star junction arms, were mixed in equimolar ratios, heated to 95°C and allowed to cool slowly back to room temperature. In this way, each strand is hybridized to its two neighbours within the completed structure to form a multi-arm junction (see Fig. 5.1a). For this application, one of the oligonucleotides was extended by 12 bases to create a single-stranded overhang that is complementary to the analogous overhang on the carriers. To purify the 6- and 12-arm junctions used here, crude DNA star samples were imaged by polyacrylamide gel electrophoresis, the main gel bands were excised, and the contained nanostructures re-suspended by electroelution using dialysis tubes (D-Tube Dialyzer Maxi, EMD Milipore).

5.5.2 – Nanopore Fabrication & Sensing

The nanopores used in this study were fabricated by controlled dielectric breakdown, as described previously.[49,50] In short, thin silicon nitride membranes (40×40 μm^2, 12 nm thick) were purchased from Norcada Inc. as TEM windows (#NBPX5004Z-HR) and mounted in custom 3D-printed flow cells, sandwiching the membrane between two fluidic reservoirs filled with buffered electrolyte solution (1 M KCl, 10 mM HEPES, pH 8). A high transmembrane potential (~10 V) is then applied via Ag/AgCl electrodes connected to a custom fabrication circuit that also monitors the current across the membrane. As soon as a large jump in current is detected, signalling the perforation of the membrane (breakdown) and the onset of ionic conduction, the voltage is quickly switched off, resulting in a single nanoscale hole across the membrane.[49,50]

The size of each nanopore is estimated by first taking a series of current measurements at lower applied transmembrane voltages (≤ 1 V) and fitting the Ohmic I-V response to a straight line in order extract a pore conductance (*G*). The pore conductance is then converted into an estimate of pore diameter (*d*) using the following conductance model[3]:

$$d = \frac{G}{2\sigma}\left(1 + \sqrt{1 + \frac{16\sigma L}{\pi G}}\right)$$

where σ is the measured conductivity of the electrolyte solution and *L* is the nanopore length, estimated in this work as the nominal, manufacturer-supplied value of the membrane thickness.

Prior to a sensing experiment, both flow cell reservoirs were flushed with fresh sensing electrolyte (3.6 M LiCl, 10 mM HEPES, pH 8). A negatively-charged DNA sample (~1 nM per target and equilibrated to the same ionic strength/species as the sensing electrolyte) was then added to one reservoir and electrophoretically driven to the nanopore by application of a small transmembrane potential (0 – 200 mV). The resulting current signals from translocating molecules are recorded by a high-bandwidth amplifier (VC100, Chimera Instruments) at 1 MHz instrument bandwidth (4.17 MHz sampling rate).

5.5.3 – Data Analysis

As detailed in the main text, raw current data from an experiment is first passed to the Nanolyzer software package (Northern Nanopore Instruments) for event detection and sublevel fitting to the internal event structure. Supporting Information Section S1 contains an example of the Nanolyzer settings used to configure this analysis. Event sublevels are then grouped by the values of metadata generated in the prior step ("sublevel blockage" and "sublevel duration", see Supporting Information Section S2) using Nanolyzer's Clustering tool. This associates a new numerical value ("sublevel label") to each sublevel in the dataset. Finally, event selection filters based on these sublevel labels (in addition to other metadata, see S2) are entered as SQL queries into Nanolyzer's Data Manager tool and used to create separately addressable subsets of events. A full list of

selection filters used is presented in Supporting Information Section S3. In addition to the metadata generated by Nanolyzer, two others ("truncated ECD" and "local max deviation", see main text for definitions) were calculated with custom scripts written in Python using the 'pandas' library. These scripts were applied to .csv files exported from Nanolyzer containing lists of all events/sublevels in each subset as well as the values of their associated metadata.

Acknolwedgements

The authors would like to acknowledge the support of the Natural Sciences and Engineering Research Council of Canada (NSERC), through funding from grant #CRDPJ 530554-18. The authors would like to thank Drs. Carlos Catalano and Qin Yang of the University of Colorado for generously providing the λ terminase enzyme used in this work.

5.6 – References

Note that an earlier draft version of this article is available as a preprint[51].

(1) Xue, L.; Yamazaki, H.; Ren, R.; Wanunu, M.; Ivanov, A. P.; Edel, J. B. Solid-State Nanopore Sensors. *Nat Rev Mater* **2020**, *5* (12), 931–951. https://doi.org/10.1038/s41578-020-0229-6.

(2) Tabard-Cossa, V. Instrumentation for Low-Noise High-Bandwidth Nanopore Recording. In *Engineered Nanopores for Bioanalytical Applications*; Edel, J. B., Albrecht, T., Eds.; Norwich, N.Y. : William Andrew ; Oxford : Elsevier Science distributor, 2013; pp 59–93. https://doi.org/10.1016/B978-1-4377-3473-7.00003-0.

(3) Kowalczyk, S. W.; Grosberg, A. Y.; Rabin, Y.; Dekker, C. Modeling the Conductance and DNA Blockade of Solid-State Nanopores. *Nanotechnology* **2011**, *22* (31), 315101. https://doi.org/10.1088/0957-4484/22/31/315101.

(4) Ying, Y. L.; Hu, Z. L.; Zhang, S.; Qing, Y.; Fragasso, A.; Maglia, G.; Meller, A.; Bayley, H.; Dekker, C.; Long, Y. T. Nanopore-Based Technologies beyond DNA Sequencing. *Nat Nanotechnol* **2022**, *17* (11), 1136–1146. https://doi.org/10.1038/s41565-022-01193-2.

(5) Forstater, J. H.; Briggs, K.; Robertson, J. W. F.; Ettedgui, J.; Marie-Rose, O.; Vaz, C.; Kasianowicz, J. J.; Tabard-Cossa, V.; Balijepalli, A. MOSAIC: A Modular Single-Molecule Analysis Interface for Decoding Multistate Nanopore Data. *Anal Chem* **2016**, *88* (23), 11900–11907. https://doi.org/10.1021/acs.analchem.6b03725.

(6) Raillon, C.; Granjon, P.; Graf, M.; Steinbock, L. J.; Radenovic, A. Fast and Automatic Processing of Multi-Level Events in Nanopore Translocation Experiments. *Nanoscale* **2012**, *4* (16), 4916. https://doi.org/10.1039/c2nr30951c.

(7) Plesa, C.; Dekker, C. Data Analysis Methods for Solid-State Nanopores. *Nanotechnology* **2015**, *26* (8), 084003. https://doi.org/10.1088/0957-4484/26/8/084003.

(8) Sun, Z.; Liu, X.; Liu, W.; Li, J.; Yang, J.; Qiao, F.; Ma, J.; Sha, J.; Li, J.; Xu, L.-Q. AutoNanopore: An Automated Adaptive and Robust Method to Locate Translocation Events in Solid-State Nanopore Current Traces. *ACS Omega* **2022**, *7* (42), 37103–37111. https://doi.org/10.1021/acsomega.2c02927.

(9) Carson, S.; Wilson, J.; Aksimentiev, A.; Wanunu, M. Smooth DNA Transport through a Narrowed Pore Geometry. *Biophys J* **2014**, *107* (10), 2381–2393. https://doi.org/10.1016/j.bpj.2014.10.017.

(10) Briggs, K.; Madejski, G.; Magill, M.; Kastritis, K.; de Haan, H. W.; McGrath, J. L.; Tabard-Cossa, V. DNA Translocations through Nanopores under Nanoscale Preconfinement. *Nano Lett* **2018**, *18* (2), 660–668. https://doi.org/10.1021/acs.nanolett.7b03987.

(11) Fologea, D.; Brandin, E.; Uplinger, J.; Branton, D.; Li, J. DNA Conformation and Base Number Simultaneously Determined in a Nanopore. *Electrophoresis* **2007**, *28* (18), 3186–3192. https://doi.org/10.1002/elps.200700047.

(12) Storm, A. J.; Chen, J. H.; Zandbergen, H. W.; Dekker, C. Translocation of Double-Strand DNA through a Silicon Oxide Nanopore. *Phys Rev E* **2005**, *71* (5), 051903. https://doi.org/10.1103/PhysRevE.71.051903.

(13) Bell, N. A. W.; Muthukumar, M.; Keyser, U. F. Translocation Frequency of Double-Stranded DNA through a Solid-State Nanopore. *Phys Rev E* **2016**, *93* (2), 022401. https://doi.org/10.1103/PhysRevE.93.022401.

(14) He, L.; Karau, P.; Tabard-Cossa, V. Fast Capture and Multiplexed Detection of Short Multi-Arm DNA Stars in Solid-State Nanopores. *Nanoscale* **2019**, *11* (35), 16342–16350. https://doi.org/10.1039/C9NR04566J.

(15) Smeets, R. M. M.; Keyser, U. F.; Krapf, D.; Wu, M.-Y.; Dekker, N. H.; Dekker, C. Salt Dependence of Ion Transport and DNA Translocation through Solid-State Nanopores. *Nano Lett* **2006**, *6* (1), 89–95. https://doi.org/10.1021/nl052107w.

(16) Verschueren, D. V; Jonsson, M. P.; Dekker, C. Temperature Dependence of DNA Translocations through Solid-State Nanopores. *Nanotechnology* **2015**, *26* (23), 234004. https://doi.org/10.1088/0957-4484/26/23/234004.

(17) Cao, C.; Krapp, L. F.; Al Ouahabi, A.; König, N. F.; Cirauqui, N.; Radenovic, A.; Lutz, J.-F.; Peraro, M. D. Aerolysin Nanopores Decode Digital Information Stored in Tailored Macromolecular Analytes. *Sci Adv* **2020**, *6* (50), 1–9. https://doi.org/10.1126/sciadv.abc2661.

(18) Boukhet, M.; König, N. F.; Ouahabi, A. Al; Baaken, G.; Lutz, J.-F.; Behrends, J. C. Translocation of Precision Polymers through Biological Nanopores. *Macromol Rapid Commun* **2017**, *38* (24), 1700680. https://doi.org/10.1002/marc.201700680.

(19) Chen, K.; Zhu, J.; Bošković, F.; Keyser, U. F. Nanopore-Based DNA Hard Drives for Rewritable and Secure Data Storage. *Nano Lett* **2020**, *20* (5), 3754–3760. https://doi.org/10.1021/acs.nanolett.0c00755.

(20) Zhu, J.; Ermann, N.; Chen, K.; Keyser, U. F. Image Encoding Using Multi-Level DNA Barcodes with Nanopore Readout. *Small* **2021**, *17* (28), 2100711. https://doi.org/10.1002/smll.202100711.

(21) Chen, K.; Kong, J.; Zhu, J.; Ermann, N.; Predki, P.; Keyser, U. F. Digital Data Storage Using DNA Nanostructures and Solid-State Nanopores. *Nano Lett* **2019**, *19* (2), 1210–1215. https://doi.org/10.1021/acs.nanolett.8b04715.

(22) Kong, J.; Bell, N. A. W.; Keyser, U. F. Quantifying Nanomolar Protein Concentrations Using Designed DNA Carriers and Solid-State Nanopores. *Nano Lett* **2016**, *16* (6), 3557–3562. https://doi.org/10.1021/acs.nanolett.6b00627.

(23) Bell, N. A. W.; Keyser, U. F. Digitally Encoded DNA Nanostructures for Multiplexed, Single-Molecule Protein Sensing with Nanopores. *Nat Nanotechnol* **2016**, *11* (7), 645–651. https://doi.org/10.1038/nnano.2016.50.

(24) Kong, J.; Zhu, J.; Chen, K.; Keyser, U. F. Specific Biosensing Using DNA Aptamers and Nanopores. *Adv Funct Mater* **2019**, *29* (3), 1807555. https://doi.org/10.1002/adfm.201807555.

(25) Sze, J. Y. Y.; Ivanov, A. P.; Cass, A. E. G.; Edel, J. B. Single Molecule Multiplexed Nanopore Protein Screening in Human Serum Using Aptamer Modified DNA Carriers. *Nat Commun* **2017**, *8* (1), 1552. https://doi.org/10.1038/s41467-017-01584-3.

(26) Cai, S.; Sze, J. Y. Y.; Ivanov, A. P.; Edel, J. B. Small Molecule Electro-Optical Binding Assay Using Nanopores. *Nat Commun* **2019**, *10* (1), 1797. https://doi.org/10.1038/s41467-019-09476-4.

(27) He, L.; Tessier, D. R.; Briggs, K.; Tsangaris, M.; Charron, M.; McConnell, E. M.; Lomovtsev, D.; Tabard-Cossa, V. Digital Immunoassay for Biomarker Concentration Quantification Using Solid-State Nanopores. *Nat Commun* **2021**, *12* (1), 5348. https://doi.org/10.1038/s41467-021-25566-8.

(28) Lam, M. H.; Briggs, K.; Kastritis, K.; Magill, M.; Madejski, G. R.; McGrath, J. L.; de Haan, H. W.; Tabard-Cossa, V. Entropic Trapping of DNA with a Nanofiltered Nanopore. *ACS Appl Nano Mater* **2019**, *2* (8), 4773–4781. https://doi.org/10.1021/acsanm.9b00606.

(29) Briggs, K. Solid-State Nanopores: Fabrication, Application, and Analysis. PhD Thesis, University of Ottawa, 2018. https://doi.org/10.20381/ruor-22794.

(30) Chien, C.-C.; Shekar, S.; Niedzwiecki, D. J.; Shepard, K. L.; Drndić, M. Single-Stranded DNA Translocation Recordings through Solid-State Nanopores on Glass Chips at 10 MHz Measurement Bandwidth. *ACS Nano* **2019**, *13* (9), 10545–10554. https://doi.org/10.1021/acsnano.9b04626.

(31) Misiunas, K.; Ermann, N.; Keyser, U. F. QuipuNet: Convolutional Neural Network for Single-Molecule Nanopore Sensing. *Nano Lett* **2018**, *18* (6), 4040–4045. https://doi.org/10.1021/acs.nanolett.8b01709.

(32) Arima, A.; Tsutsui, M.; Washio, T.; Baba, Y.; Kawai, T. Solid-State Nanopore Platform Integrated with Machine Learning for Digital Diagnosis of Virus Infection. *Anal Chem* **2021**, *93* (1), 215–227. https://doi.org/10.1021/acs.analchem.0c04353.

(33) Taniguchi, M.; Minami, S.; Ono, C.; Hamajima, R.; Morimura, A.; Hamaguchi, S.; Akeda, Y.; Kanai, Y.; Kobayashi, T.; Kamitani, W.; Terada, Y.; Suzuki, K.; Hatori, N.; Yamagishi, Y.; Washizu, N.; Takei, H.; Sakamoto, O.; Naono, N.; Tatematsu, K.; Washio, T.; Matsuura, Y.; Tomono, K. Combining Machine Learning and Nanopore Construction Creates an Artificial Intelligence Nanopore for Coronavirus Detection. *Nat Commun* **2021**, *12* (1), 3726. https://doi.org/10.1038/s41467-021-24001-2.

(34) Xia, K.; Hagan, J. T.; Fu, L.; Sheetz, B. S.; Bhattacharya, S.; Zhang, F.; Dwyer, J. R.; Linhardt, R. J. Synthetic Heparan Sulfate Standards and Machine Learning Facilitate the Development of Solid-State Nanopore Analysis. *Proceedings of the National Academy of Sciences* **2021**, *118* (11), 1–7. https://doi.org/10.1073/pnas.2022806118.

(35) Tsutsui, M.; Takaai, T.; Yokota, K.; Kawai, T.; Washio, T. Deep Learning-Enhanced Nanopore Sensing of Single-Nanoparticle Translocation Dynamics. *Small Methods* **2021**, *5* (7), 2100191. https://doi.org/10.1002/smtd.202100191.

(36) Zhang, D. Y.; Winfree, E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. *J Am Chem Soc* **2009**, *131* (47), 17303–17314. https://doi.org/10.1021/ja906987s.

(37) Zhu, J.; Zhang, L.; Zhou, Z.; Dong, S.; Wang, E. Molecular Aptamer Beacon Tuned DNA Strand Displacement to Transform Small Molecules into DNA Logic Outputs. *Chemical Communications* **2014**, *50* (25), 3321. https://doi.org/10.1039/c3cc49833f.

(38) Kong, J.; Zhu, J.; Keyser, U. F. Single Molecule Based SNP Detection Using Designed DNA Carriers and Solid-State Nanopores. *Chemical Communications* **2017**, *53* (2), 436–439. https://doi.org/10.1039/C6CC08621G.

(39) Beamish, E.; Tabard-Cossa, V.; Godin, M. Identifying Structure in Short DNA Scaffolds Using Solid-State Nanopores. *ACS Sens* **2017**, *2* (12), 1814–1820. https://doi.org/10.1021/acssensors.7b00628.

(40) King, S.; Briggs, K.; Slinger, R.; Tabard-Cossa, V. Screening for Group A Streptococcal Disease via Solid-State Nanopore Detection of PCR Amplicons. *ACS Sens* **2022**, *7* (1), 207–214. https://doi.org/10.1021/acssensors.1c01972.

(41) Chen, P.; Gu, J.; Brandin, E.; Kim, Y.-R.; Wang, Q.; Branton, D. Probing Single DNA Molecule Transport Using Fabricated Nanopores. *Nano Lett* **2004**, *4* (11), 2293–2298. https://doi.org/10.1021/nl048654j.

(42) Mihovilovic, M.; Hagerty, N.; Stein, D. Statistics of DNA Capture by a Solid-State Nanopore. *Phys Rev Lett* **2013**, *110* (2), 028102. https://doi.org/10.1103/PhysRevLett.110.028102.

(43) Kowalczyk, S. W.; Wells, D. B.; Aksimentiev, A.; Dekker, C. Slowing down DNA Translocation through a Nanopore in Lithium Chloride. *Nano Lett* **2012**, *12* (2), 1038–1044. https://doi.org/10.1021/nl204273h.

(44) Fologea, D.; Gershow, M.; Ledden, B.; McNabb, D. S.; Golovchenko, J. A.; Li, J. Detecting Single Stranded DNA with a Solid State Nanopore. *Nano Lett* **2005**, *5* (10), 1905–1909. https://doi.org/10.1021/nl051199m.

(45) Karau, P.; Tabard-Cossa, V. Capture and Translocation Characteristics of Short Branched DNA Labels in Solid-State Nanopores. *ACS Sens* **2018**, *3* (7), 1308–1315. https://doi.org/10.1021/acssensors.8b00165.

(46) Carlsen, A.; Tabard-Cossa, V. Mapping Shifts in Nanopore Signal to Changes in Protein and Protein-DNA Conformation. *Proteomics* **2022**, *22* (5–6), 2100068. https://doi.org/10.1002/pmic.202100068.

(47) Bell, N. A. W.; Keyser, U. F. Specific Protein Detection Using Designed DNA Carriers and Nanopores. *J Am Chem Soc* **2015**, *137* (5), 2035–2041. https://doi.org/10.1021/ja512521w.

(48) Catalano, C. E. The Terminase Enzyme from Bacteriophage Lambda: A DNA-Packaging Machine. *Cell Mol Life Sci* **2000**, *57* (1), 128–148. https://doi.org/10.1007/s000180050503.

(49) Kwok, H.; Briggs, K.; Tabard-Cossa, V. Nanopore Fabrication by Controlled Dielectric Breakdown. *PLoS One* **2014**, *9* (3), e92880. https://doi.org/10.1371/journal.pone.0092880.

(50) Briggs, K.; Kwok, H.; Tabard-Cossa, V. Automated Fabrication of 2-Nm Solid-State Nanopores for Nucleic Acid Analysis. *Small* **2014**, *10* (10), 2077–2086. https://doi.org/10.1002/smll.201303602.

(51) Roelen, Z.; Briggs, K.; Tabard-Cossa, V. Analysis of Nanopore Data: Classification Strategies for an Unbiased Curation of Single-Molecule Events from DNA Nanostructures. *ChemRxiv* **2023**. https://doi.org/10.26434/chemrxiv-2023-09wgb.

5.7 – SUPPORTING INFORMATION

Section S1: Configuration Details of Nanolyzer Analysis

Table 5S.1*: Example of Nanolyzer configuration details used to locate events and fit sublevels within the data of Table 5.1 / Figure 5.4 in the main text, starting from the raw current timecourse as input. Important highlights include a threshold of 6 standard deviations away from the baseline current to trigger the search for an event ("threshold") and the use of a low-pass 250 kHz Bessel filter during both the event detection and fitting stages ("data_cutoff" and "event_cutoff").*

[Event Segmentation Settings]	[Event Fitting Settings]	[Event Rejection Criteria]
event_direction=0	cusum_delta=6.0	event_minpoints=20.83
baseline_min=-15600.0	cusum_elasticity=0.0	event_maxpoints=20833.33
baseline_max=-14500.0	cusum_minstep=3.0	
manual_baseline_override=0	subevent_minpoints=20.83	
manual_baseline=0.0	intra_threshold=16.0	
manual_baseline_std=0.0	intra_hysteresis=9.0	
threshold=6.0	padding_wait=20.83	
hysteresis=7.0	stepfit_samples=20.83	
use_data_filter=1	max_sublevels=10	
data_cutoff=250000.0	attempt_recovery=1	
data_poles=8	use_event_filter=1	
fixed_event_length=0.0	event_cutoff=250000.0	
	event_poles=8	

Section S2: List of Metadata Used in Nanolyzer Data Manager ("Data Dictionary")

Table 5S.2*: List of the subset of metadata available to the "Data Manager" tool in Nanolyzer that were used in this work during the process of clustering sublevels, building selection filters, or creating plots. "Event statistics" refer to parameters that characterize entire translocation events while "sublevel statistics" refer to parameters relating to the individual fitted sublevels within each event (see Fig. 5.3a in the main text for the breakdown of a sample event).*

Name	Type	Units	Description
'duration_us'	Event statistic	µs	Event dwell time, time between first deviation from baseline and return to baseline
'area_pC'	Event statistic	pC	Event equivalent charge deficit, numerical time-integral of difference between raw current values and averaged baseline
'max_deviation_pA'	Event statistic	pA	Event max deviation, absolute value of largest difference in current from averaged baseline
'n_levels'	Event statistic		Number of fitted sublevels within event, including baseline levels
'level_duration_us'	Sublevel statistic	µs	Fitted sublevel duration
'blockages_pA'	Sublevel statistic	pA	Fitted sublevel blockage
'level_max_deviation_pA'	Sublevel statistic	pA	Sublevel max deviation
'level_num'	Sublevel statistic		Running count of sublevels within event, beginning at 0 and ending at ['n_levels' – 1]
'levels_left'	Sublevel statistic		Number of remaining sublevels in event, starting at ['n_levels' – 1] and ending at 0
'sublevel_labels'	Sublevel statistic		Numerical label assigned to sublevel during clustering

Section S3: Selection Filters Used – List of SQL Queries

Table 5S.3: *Example list of SQL queries used to implement selection filters in Nanolyzer "Data Manager" (see Figure 5.2 and Table 5.2 in the main text for filter definitions). In this example, sublevels have been assigned the following numerical labels during the clustering stage (see Fig. 5.3): 'unfolded carrier' → '1', 'folded carrier' → '2', 'cargo' → '3'. The threshold used for local max deviation ("level_max_deviation_pA") in Filters A & B is experiment-specific (see discussion surrounding Figure 5.7 in main text for guidelines used) and takes a value of 1500 pA in this example. All selection filters are implicitly prefaced within Nanolyzer by* "SELECT * FROM events WHERE".

Filter	SQL Query
Type 1U	((id IN (SELECT id FROM sublevels WHERE level_num=1 AND sublevel_labels = 3) AND id IN (SELECT id FROM sublevels WHERE levels_left=1 AND sublevel_labels = 1)) AND n_levels = 4)
Type 1P	(((id IN (SELECT id FROM sublevels WHERE level_num=1 AND sublevel_labels = 2) AND id IN (SELECT id FROM sublevels WHERE levels_left=1 AND sublevel_labels = 1)) AND id IN (SELECT id FROM sublevels WHERE level_num=2 AND sublevel_labels = 3)) AND n_levels = 5)
Type 2U	((id IN (SELECT id FROM sublevels WHERE level_num=1 AND sublevel_labels = 1) AND id IN (SELECT id FROM sublevels WHERE levels_left=1 AND sublevel_labels = 3)) AND n_levels = 4)
Type 2P	(((id IN (SELECT id FROM sublevels WHERE level_num=1 AND sublevel_labels = 2) AND id IN (SELECT id FROM sublevels WHERE levels_left=1 AND sublevel_labels = 3)) AND id IN (SELECT id FROM sublevels WHERE level_num=2 AND sublevel_labels = 1)) AND n_levels = 5)
Type 1/2F	((id IN (SELECT id FROM sublevels WHERE level_num=1 AND sublevel_labels = 2) AND id IN (SELECT id FROM sublevels WHERE levels_left=1 AND sublevel_labels = 3)) AND n_levels = 4)
Filter A	(id IN (SELECT id FROM sublevels WHERE levels_left=1 AND sublevel_labels = 1) AND id IN (SELECT id FROM sublevels WHERE levels_left=0 AND level_max_deviation_pA > 1500))
Filter B	((NOT (id IN (SELECT id FROM sublevels WHERE levels_left=1 AND sublevel_labels = 1)) AND (id IN (SELECT id FROM sublevels WHERE levels_left=0 AND level_max_deviation_pA > 1500) OR id IN (SELECT id FROM sublevels WHERE levels_left=1 AND level_max_deviation_pA > 1500))) AND id IN (SELECT id FROM sublevels WHERE sublevel_labels = 1))

Section S4: Removing Type 0 Events before Sublevel Clustering

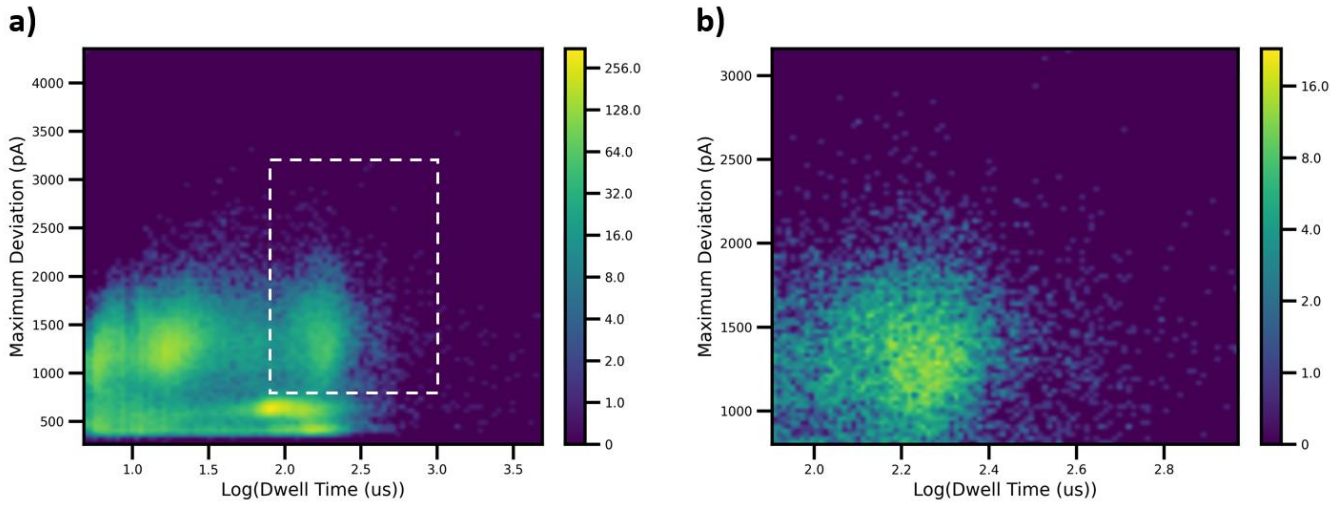
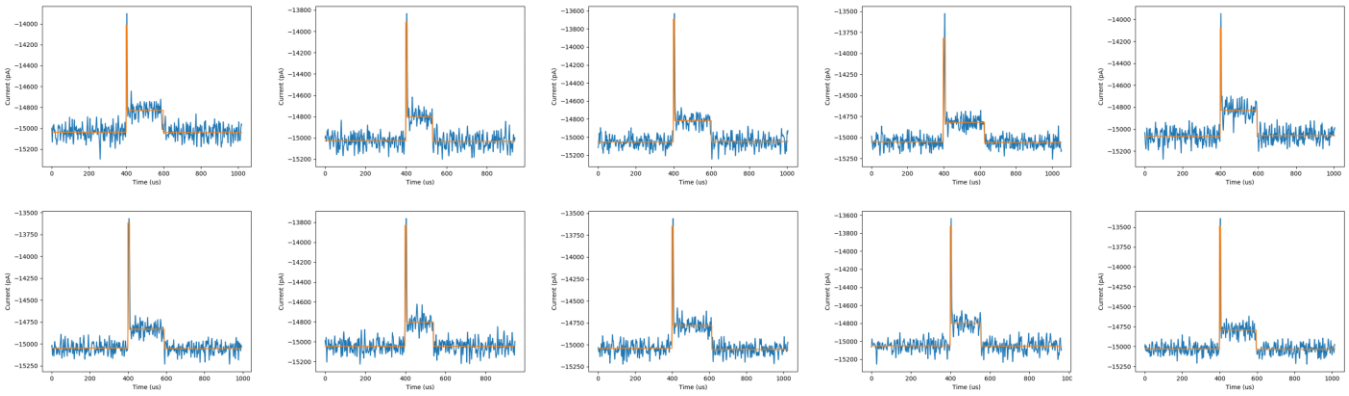


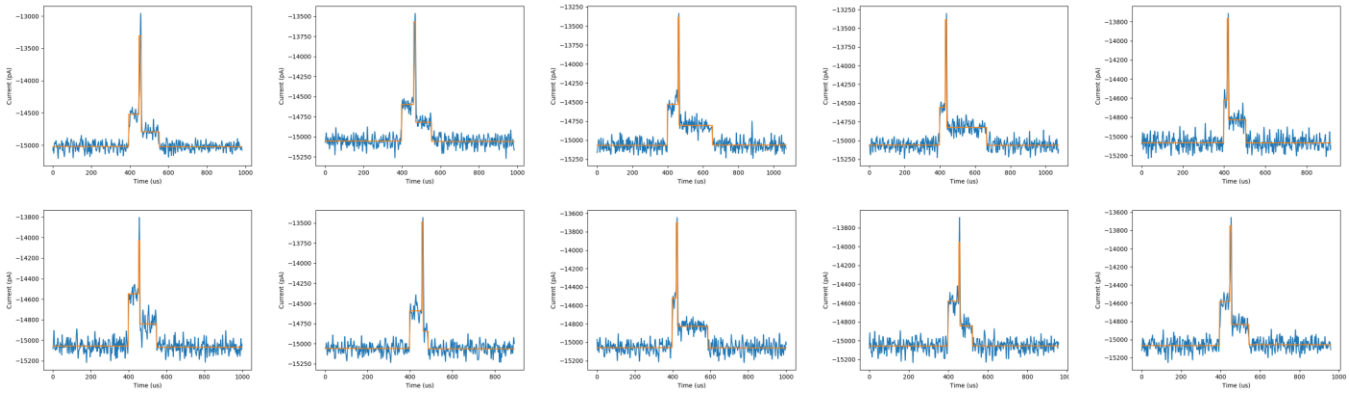
Figure 5S.1: *Illustration of a process to isolate translocation events resulting from hybridized carrier-cargo pairs prior to sublevel clustering analysis in Nanolyzer.* ***a)*** *2D histogram of max deviation vs. log[dwell time] for a full dataset of events (2-kbp carriers + 12-arm star junctions, ~13-nm pore diameter, 3.6 M LiCl, 75 mV transmembrane voltage). The dashed box represents a pair of selection filters used to remove short (in duration) or shallow (in current blockage) events that are not characteristic of those due to carriers successfully hybridized to cargo molecules. In this case, the filters select for events where dwell time > 80 µs and max deviation > 800 pA.* ***b)*** *2D histogram of max deviation vs. log[dwell time] for the result of the selection filters placed on the full dataset of (a). This subset of relatively deep and long-lived events was used to create Table 5.1 / Figure 5.4 in the main text.*

Section S5: First Ten Events from Each Category of Table 5.1 / Figure 5.4

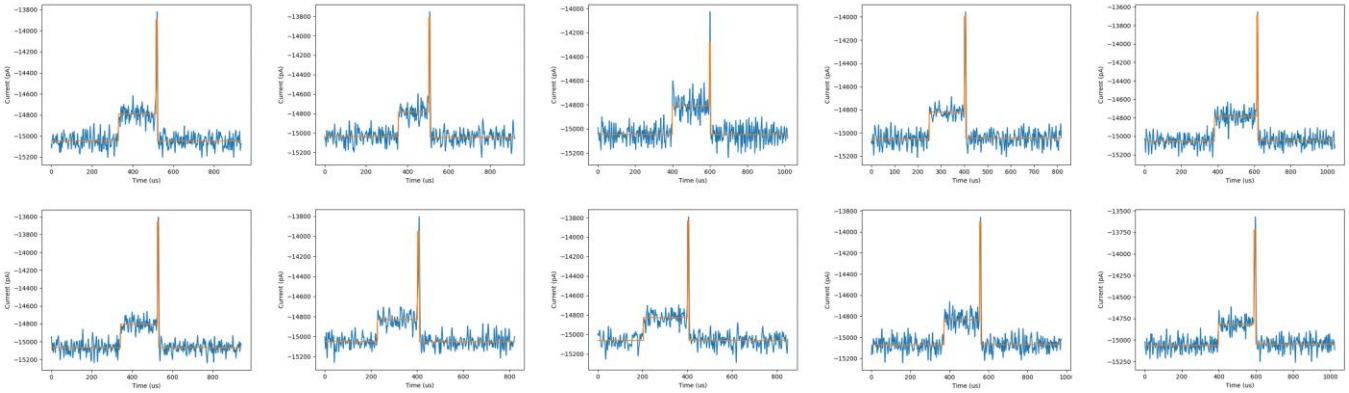
Type 1U



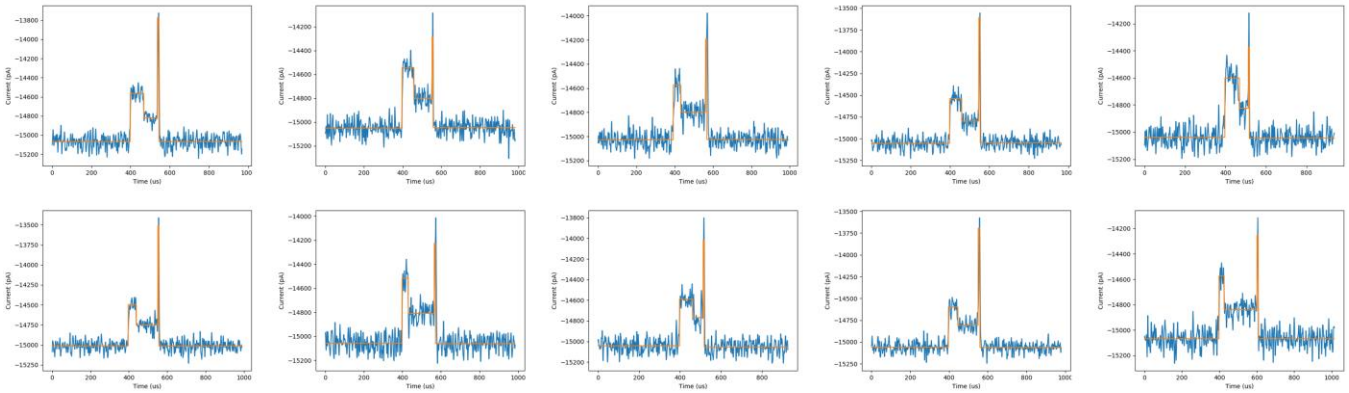
Type 1P



Type 2U



Type 2P



Type 1/2F

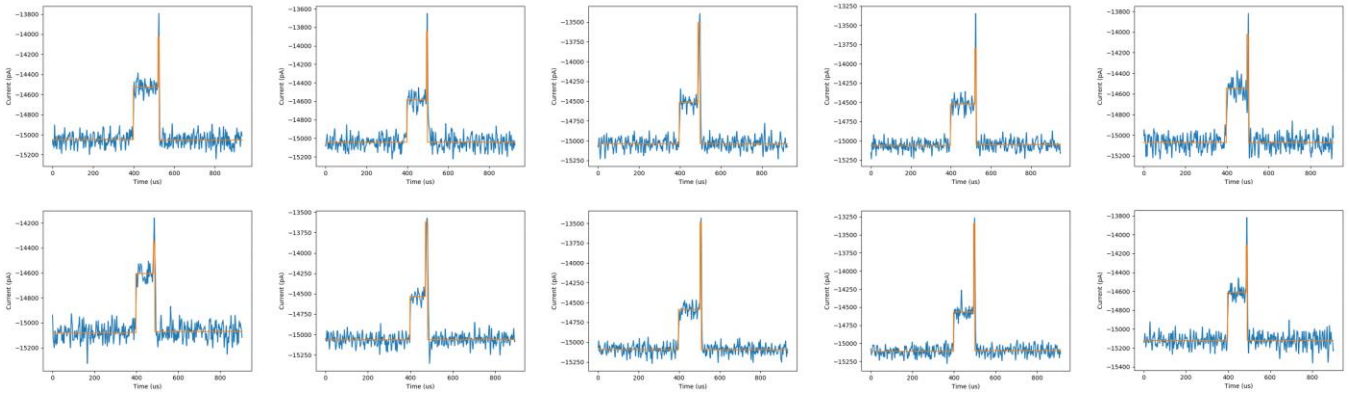
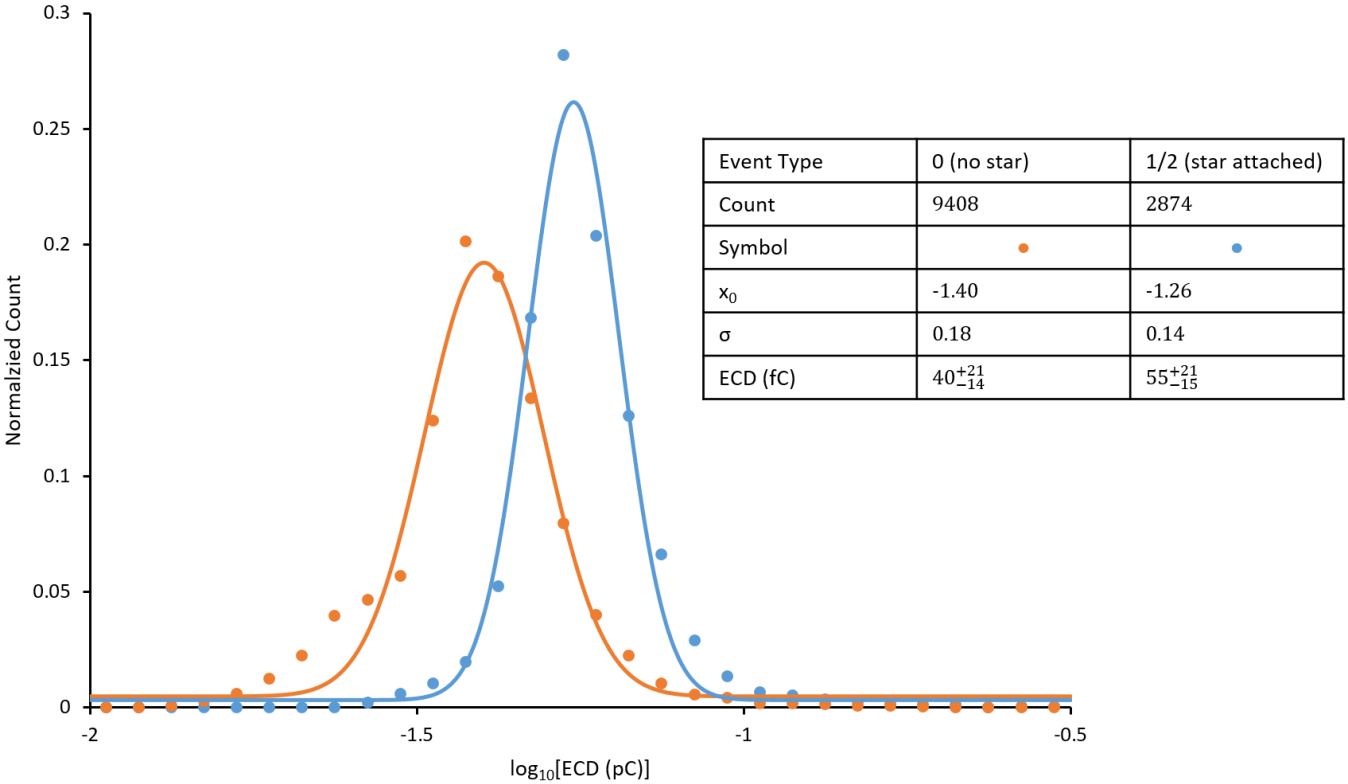


Figure 5S.2*: Gallery of current traces for the first ten events (earliest start times) sorted into each category of Table 5.1 / Figure 5.4 in the main text (~13-nm pore, 2-kbp carrier + 12-arm star junction, 75 mV transmembrane voltage, 3.6 M LiCl) using the selection filters represented in Figure 5.2 and the sublevel labelling approach outlined in Figure 5.3. The captured event shapes correspond well to those that were expected (when designing the filters) due to the generally successful degree of sublevel fitting within this dataset as well as the exact sequence of sublevel labels demanded by the selection filter of each category.*

Section S6: Comparing ECD Distributions of Type 0 vs. Type 1/2 Events



Event Type	0 (no star)	1/2 (star attached)
Count	9408	2874
Symbol	●	●
x_0	-1.40	-1.26
σ	0.18	0.14
ECD (fC)	40^{+21}_{-14}	55^{+21}_{-15}

Figure 5S.3*: Comparison of the ECD distributions of events with and without cargo blockages (Types 1 /2 vs. Type 0 in the scheme of Fig 5.2, main text) using the folded events of the same dataset as Table 5.1 / Figs 5.4, 5S.1. The distribution from events with cargo attachments is noticeably shifted toward larger ECD values, related to the extra integrated area under the cargo subevents (see Fig 5.6a for an illustrated ECD calculation). Folded events (Types P & F) were chosen for this comparison as unfolded Type 0 events contained a noticeable subpopulation from shortened carrier fragments (due to incomplete synthesis reactions or damage during handling, still faintly visible as the left "shoulder" of the folded Type 0 distribution in the figure), and this influence toward smaller ECD values is separate from that due to the absent cargo subevents.*

Section S7: 1D Histograms of ECD/TrECD for Events of Figure 5.6

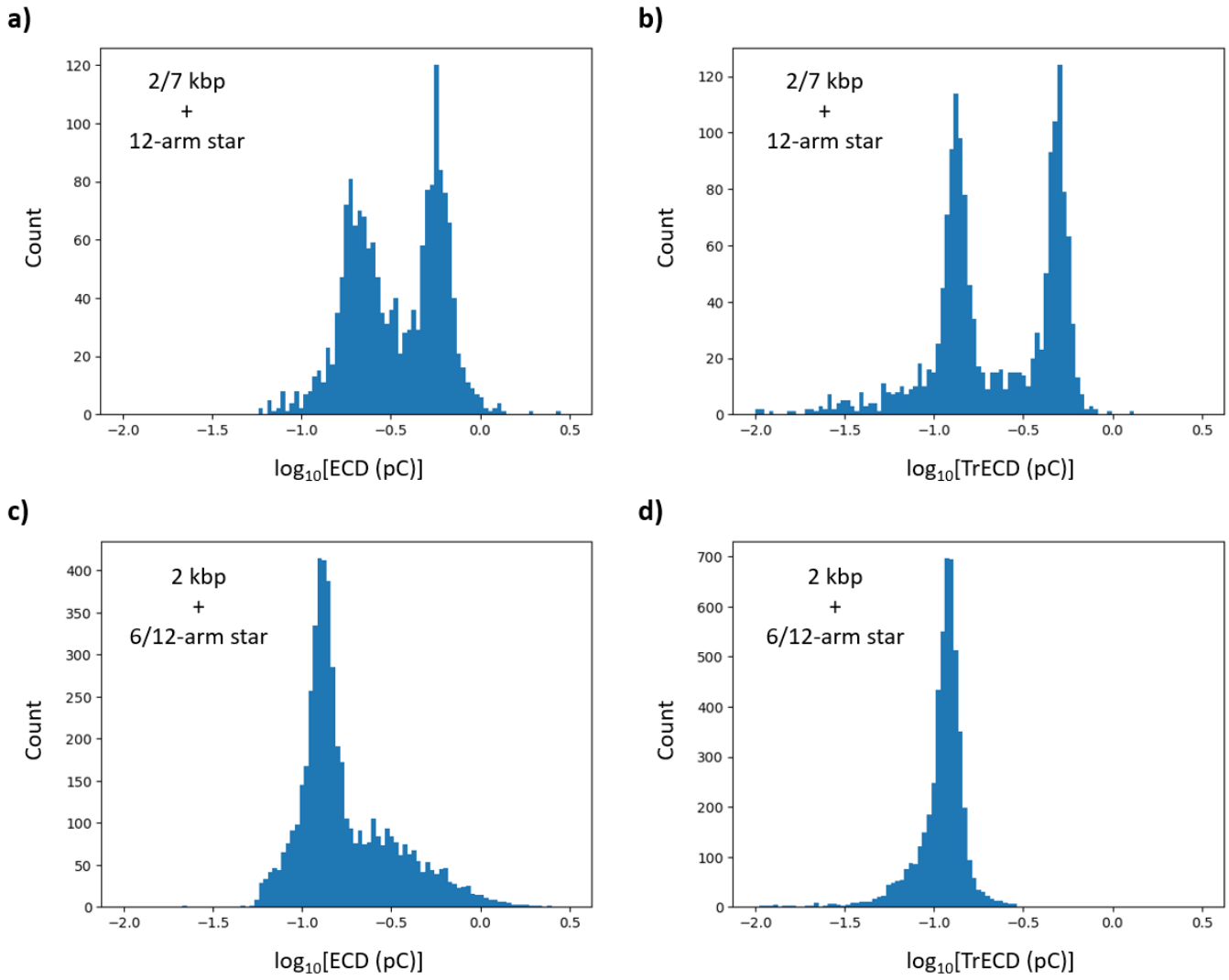


Figure 5S.4*: Distributions of equivalent charge deficit using the datasets of Figure 5.6 in the main text.* ***a)*** *1D histogram of raw, whole-event ECD ("ECD", see Fig. 5.6a), log-scaled, for the events of Figs 5.6c & 5.6d (mix of 2-kbp carrier attached to 12-arm star junction & 7-kbp carrier attached to 12-arm star junction, ~11-nm pore, 3.6 M LiCl, 200 mV).* ***b)*** *1D histogram of truncated ECD ("TrECD", see Fig. 5.6b), log-scaled, for the same events as in (a).* ***c)*** *1D histogram of raw, whole-event ECD, log-scaled, for the events of Figs 5.6e & 5.6f (mix of 2-kbp carrier attached to 12-arm star junction & 2-kbp carrier attached to 6-arm star junction, ~8-nm pore, 3.6 M LiCl, 200 mV).* ***d)*** *1D histogram of TrECD, log-scaled, for the same events as in (c).*

Section S8: Control – Two Carrier + Star Pairs Run Separately and Together on a Single Pore

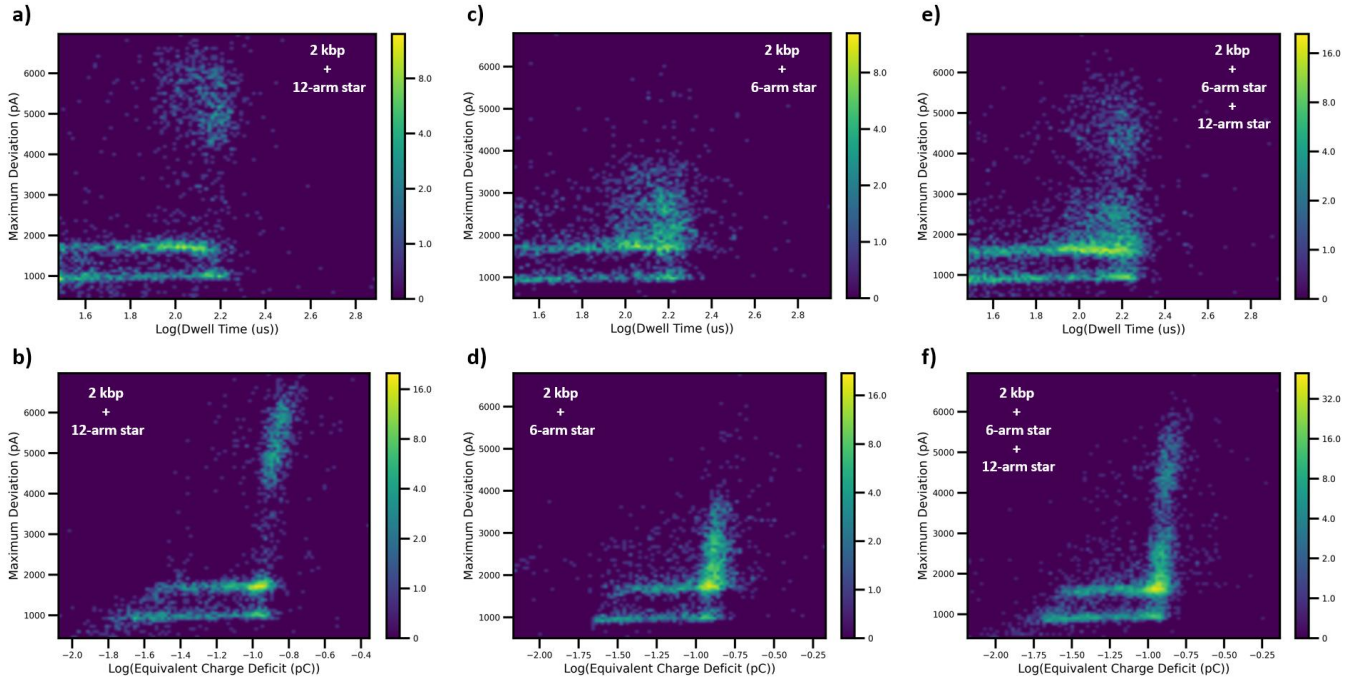


Figure 5S.5: *Comparison of current blockage signatures from two cargo-carrier pairs run separately and together on a single nanopore (~10.5 nm pore diameter, 3.6 M LiCl, 200 mV).* ***a)*** *2D histogram of max deviation vs. log[dwell time] for a sample of 2-kbp carriers + 12-arm star junctions.* ***b)*** *2D histogram of max deviation vs. log[ECD] of the same events as in (a).* ***c)*** *2D histogram of max deviation vs. log[dwell time] for a sample of 2-kbp carriers + 6-arm star junctions.* ***d)*** *2D histogram of max deviation vs. log[ECD] of the same events as in (c).* ***e)*** *2D histogram of max deviation vs. log[dwell time] for a sample of 2-kbp carriers + 12-arm star junctions & 2-kbp carriers + 6-arm star junctions, added together.* ***f)*** *2D histogram of max deviation vs. log[ECD] of the same events as in (f). Over the course of sequentially sensing each the of the three samples (including injecting each sample, recording 30 – 60 minutes of translocation events per sample, and flushing the fluidic cell with clean buffer between samples), the pore grew in diameter from ~10 nm to ~11 nm, resulting in slightly shallower blockage depths by the time the final sample was recorded (6/12-arm star mix, panels e & f).*

Section S9: Sample Events of 2 kbp + 6/12-Arm Stars through Smaller (8-nm) Pore

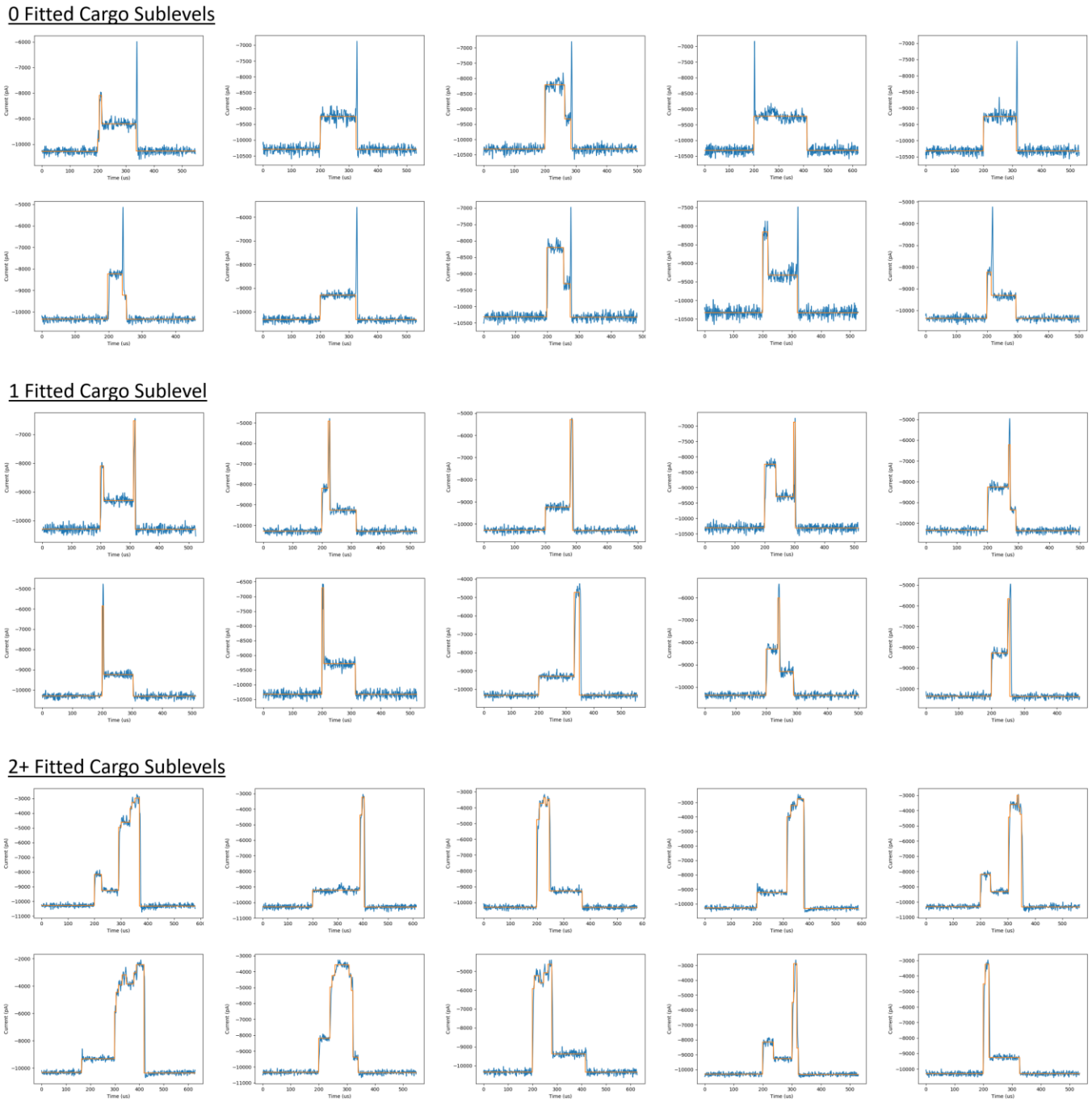


Figure 5S.6*: Gallery of current traces and fitted sublevels for events from a mixed sample of carrier-cargo pairs (2-kbp carriers + 12-arm stars & 2-kbp carriers + 6-arm stars) run through a smaller (~8-nm diameter) pore (200 mV, 3.6 M LiCl). The images are organized by the number of sublevels fitted to blockages of the cargo within each event (0, 1, or 2+) and ten representative examples are presented in each case.*

Section S10: Filters A & B Applied to 2 kbp + 6/12-Arm Stars through Larger (12-nm) Pore

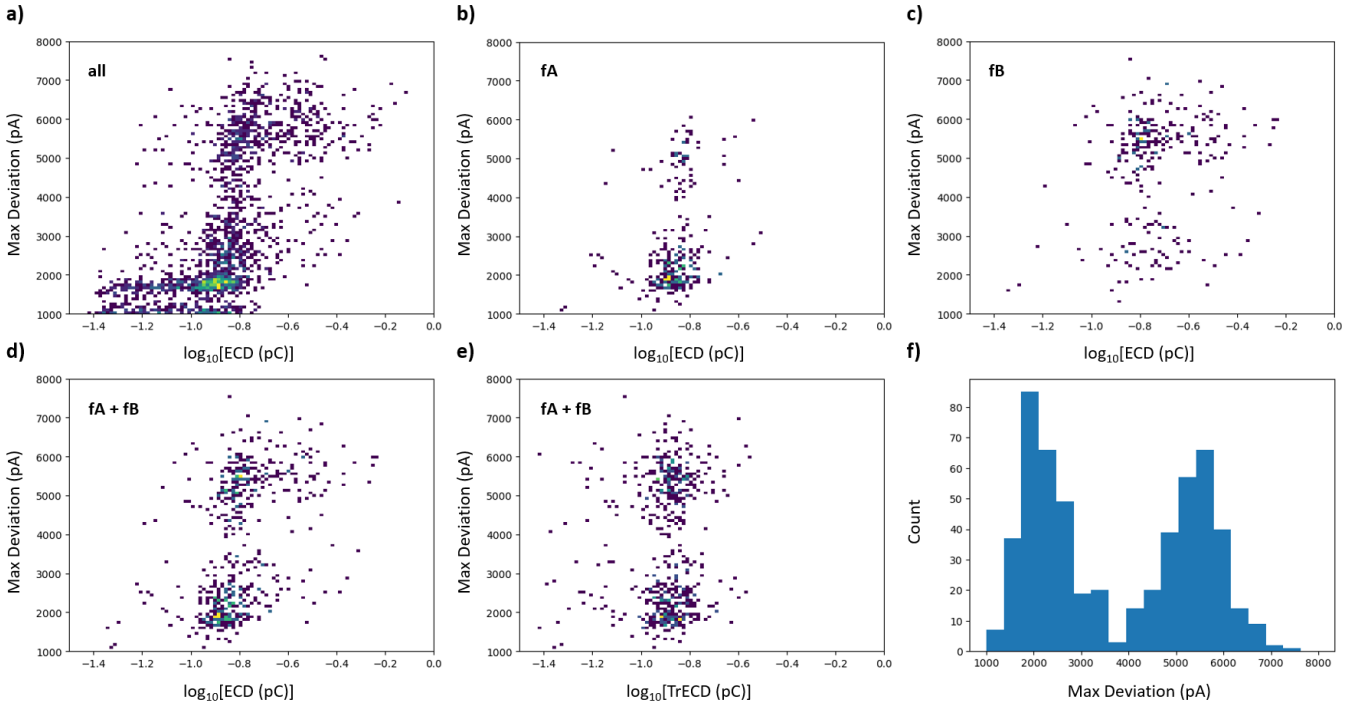


Figure 5S.7*: Application of Filters "A" and "B" (defined in main text, see Fig. 5.7 and Table 5.2) to extract Type 2 events from a dataset featuring a mix of carrier-cargo pairs (2-kbp carriers attached to 12-arm or 6-arm star junctions) run on a larger ~12-nm pore (180 mV, 3.6 M LiCl).* ***a)*** *2D histogram of total events plotted on max deviation vs. log[ECD].* ***b)*** *2D histogram (max deviation vs. log[ECD]) of the output of Filter A on the same dataset as (a).* ***c)*** *2D histogram (max deviation vs. log[ECD]) of the output of Filter B.* ***d)*** *2D histogram (max deviation vs. log[ECD]) of the union of Filters A and B.* ***e)*** *2D histogram (max deviation vs. log[truncated ECD]) of the union of Filters A and B. Two populations are visible, centred on the same TrECD value but different max deviation values, reflecting the single shared carrier length and two different cargo attachments (6- and 12-arm stars) present in the sample.* ***f)*** *1D histogram of max deviation for the events of (d) and (e) showing two separable peaks.*

Section S11: Rigid Sorting Approach on Events from 2 kbp + 6/12-Arm Stars in Figure 5.7

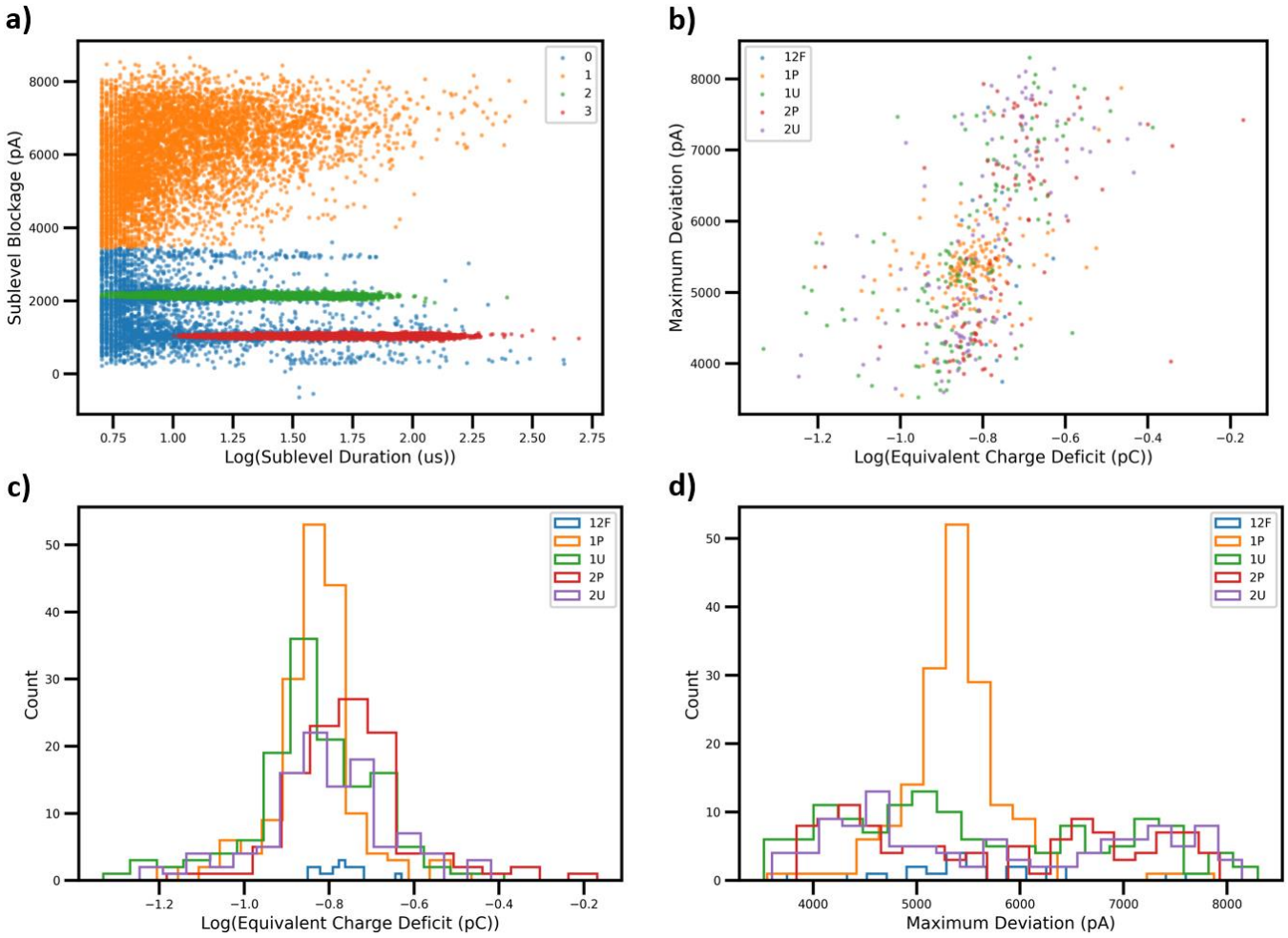


Figure 5S.8: *Application of rigid carrier-cargo filters (see Figure 5.2 in main text for definitions) to a dataset resulting from a mix of carrier-cargo pairs (2-kbp carriers attached to 12-arm or 6-arm star junctions) run on a smaller ~8-nm pore (200 mV, 3.6 M LiCl). An alternate analysis approach on this same dataset that is based on more "relaxed" selection filters (see Table 5.2 in main text) is presented in Figure 5.7.* ***a)*** *Scatter plot of sublevel blockage vs. log[sublevel duration] for all sublevels in dataset, showing labels assigned to subpopulations in clustering stage.* ***b)*** *Scatter plot of (event) max deviation vs. log[ECD] for events in the dataset extracted by rigid selection filters (see legend for colour coding)* ***c)*** *1D histogram of log[ECD] for events in the dataset extracted by rigid selection filters.* ***d)*** *1D histogram of max deviation for events in the dataset extracted by rigid selection filters. Note the apparent bias in these filters towards selecting specifically for Type 1P events from carriers attached to 6-arm star junctions (orange peak, deepest events of lower max deviation population) that is not present in the analysis approach of Figure 5.7 (compare Fig 5S.8b with Fig 5.7e).*

CHAPTER 6 – CONCLUSION

6.1 – Thesis Summary

In summary, solid-state nanopores have shown great potential as a versatile sensing platform due to their ability to characterize a wide range of unlabelled analytes at the single-molecule level. This has paved the way for important applications in medicine and science involving nanopores such as in detecting rare biomarkers of diagnostic value and in elucidating details of fundamental polymer physics. The work presented in this thesis aimed to help advance the development of this powerful platform by addressing two of its current points of weakness. Firstly, nanopores are a surface sensor and, in their basic implementation, rely on diffusion to deliver target analytes from a bulk surrounding volume – this slows down sample analysis by limiting the rate at which targets are captured and detection signals are generated. Secondly, while the signals that nanopores generate can be mapped to individual translocating molecules (and thus provide richer information about a sample when compared to ensemble measurements), these signals can lack detail about small important changes (e.g. DNA sequence, protein identity) among similar molecules especially in the large pore regime ($d \gtrsim 10$ nm).

In *Chapter 2*, we added a parallel optical detection channel to our nanopore system and minimized its noise contributions to the ionic current signal, thus laying some groundwork to better distinguish different targets (and different configurations of single targets) than with current blockage alone. *Chapter 3* made efforts toward combining solid-state nanopores with active target transport to the pore (to increase its capture rate) using dielectrophoresis, and uncovered critical factors to successfully implementing this functionality in future systems. In *Chapter 4*, we developed methods for producing sticky-ended DNA carriers, so as to facilitate the translation of similar targets into specific nanopore signals, and an emphasis was placed on ease of carrier synthesis and the ability to vary carrier length. Finally, *Chapter 5* investigated the analysis of nanopore signals produced by DNA carriers from the previous chapter, in order to gain insight into the challenges associated with detecting and characterizing structured, target-specific molecules, and developing refinements of existing analysis techniques in the process. In the remainder of this final chapter, we present several additional

preliminary results that document efforts to build upon the work in the previous chapters and move the ideas presented there closer to practical applications.

6.2 – Ongoing Work

Improved SE-DNA Synthesis Method

Figure 6.1 presents an alternative technique to producing sticky-ended DNA carriers that combines advantages from each of the two methods presented in *Chapter 4*, namely, using a reduced set of reagents (as in the terminase digest method), and providing the ability to modify the exact sequence or even the existence of the ssDNA overhangs at each end (as in the oligo assembly method) [1]. The basic scheme is outlined in Figure 6.1a and starts with producing a double-stranded amplicon bounded by two recognition sites for restriction enzymes. After a digestion step with these enzymes, a pair of "linker" oligos are ligated onto the short sticky ends created during cleavage. Because the sequences of the linker oligos internally contain (and thus determine) the sequences of the final overhangs and because short, chemically synthesized oligos can be purchased commercially (e.g. from IDT), each overhang can separately be designed with a desired length, polarity (5' or 3'), or series of chemical modifications. A sample gel of a 6.2 kbp SE-DNA carrier produced by this method (and designed to have 12-nt, complementary ssDNA overhangs for comparison with the previous methods), is displayed in Figure 6.1b and shows the molecules reversibly circularizing/linearizing with shifts in incubation temperature, as before. Figure 6.1c presents the output of a nanopore experiment featuring two carrier lengths (1.8 and 6.2 kbp) produced by this method, together with a 12-arm DNA star cargo molecule. As can be seen in the scatter plot on the left, the resulting translocation events are easily separable by carrier length and cargo attachment state through the metrics of truncated ECD and max deviation, respectively (see *Chapter 5*). Of the events that do not feature an attached DNA star (populations *i* and *ii* in Figure 6.1c) a substantial portion of these are the result of circularized molecules, as evidenced their single current blockage level at ~2× the magnitude of an unfolded fragment (e.g. compare sample events *i*, *ii* with *iii*, *iv* on the right of Figure 6.1c), which is a consequence of the (arbitrary) choice here to have each carrier feature complementary ends.

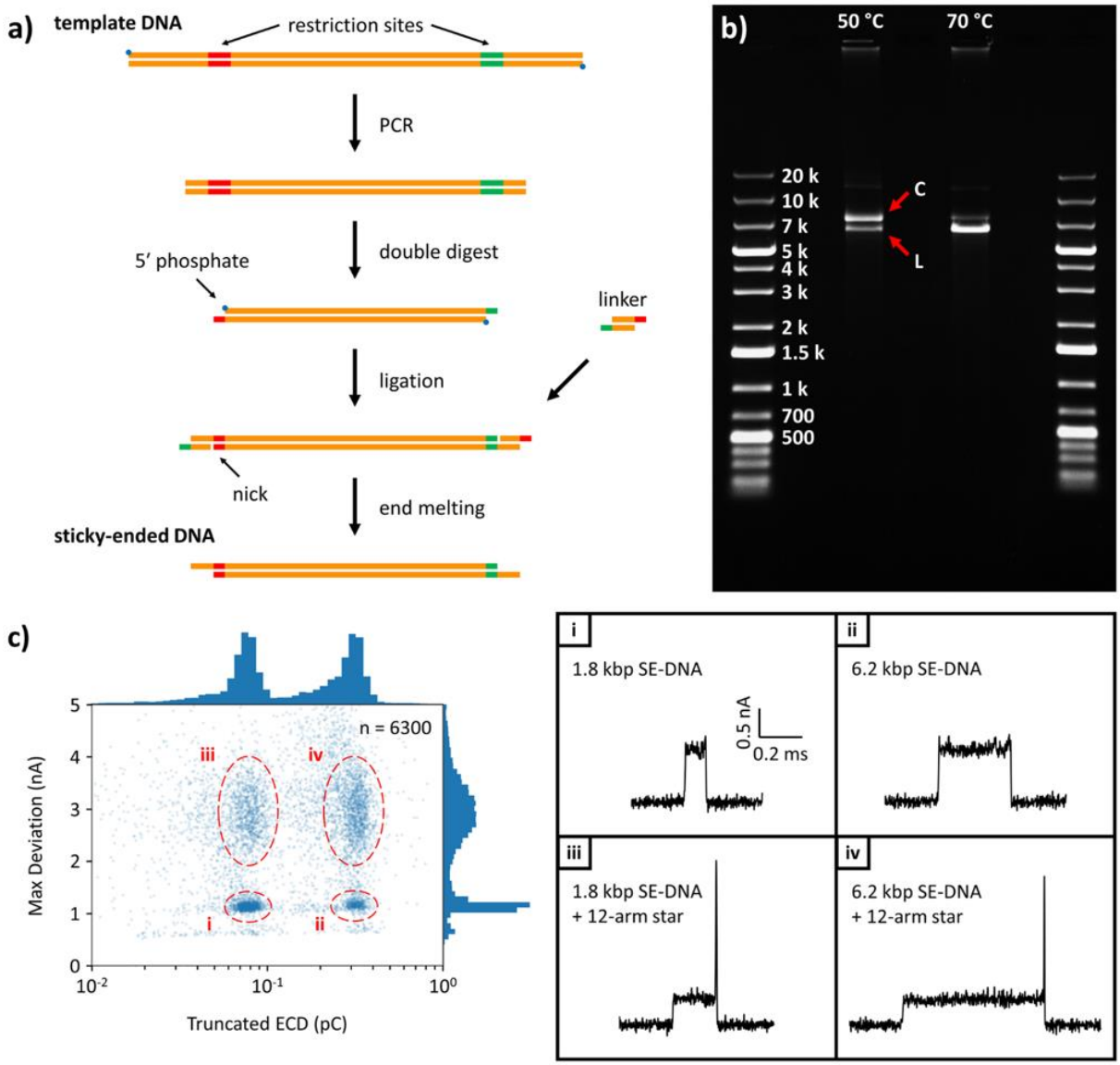


Figure 6.1: *Overview of improved SE-DNA synthesis method.* ***a)*** *Schematic of the general protocol. The length of the final product is determined primarily by the distance between the two restriction sites (which can be adjusted as desired during the PCR step) while the number of sticky ends and their sequences are flexibly determined by the choice of linker molecule(s).* ***b)*** *Agarose gel (0.5%, 0.5× TAE, pre-stained with Gel Red, 70 V) of 6.2 kbp SE-DNA produced by this method, here using λ-DNA template, KpnI and SacI restriction enzymes, and a single linker. The two main band positions correspond to: the linear product ("L") migrating between the 5 kbp and 7 kbp DNA ladder bands, and a slower-migrating circular product ("C") resulting from the annealing of the two complementary ends of a linear molecule.* ***c)*** *Scatter plot of max deviation vs. log[TrECD] for events from the nanopore detection of a DNA nanostructure attached to multiple SE-DNA lengths (1.8/6.2 kbp SE-DNA + 12-arm DNA stars, ~13 nm pore diameter, 3.6 M LiCl pH 8 buffer, 150 mV transmembrane potential). Two main peaks are visible in the histograms (along the margins of the scatter plot axes) of both statistics, resulting in four total event populations corresponding to: i) 1.8 kbp SE-DNA on its own, ii) 6.2 kbp SE-DNA on its own, iii) 1.8 kbp SE-DNA attached to a 12-arm star, iv) 6.2 kbp SE-DNA attached to a 12-arm star. Representative current traces of events from each of the four populations are presented in the panels on the right. Images adapted from* [1]*, with permission.*

Translating Target Analytes into Carrier-Based Nanopore Signals

While *Chapter 5* explored the analysis of nanopore translocation events from DNA carriers attached to cargo molecules, the cargos there (DNA star junctions) were chosen as a proof of concept and not for their intrinsic value to a real diagnostic application. If, instead, a specific biomolecule (e.g. a protein) was targeted for detection, then that target molecule could potentially be employed as the cargo in a sensing experiment by: 1) directly binding it to the carrier, for instance, through a functionalized DNA oligo as in the examples of *Section 1.5* in *Chapter 1* (involving DNA aptamers, antigens for antibodies, etc.), and 2) subsequently detecting the additional current it blocks during passage through the nanopore. An alternative approach, however, would be to retain a precisely structured molecule such as a DNA star as the cargo (e.g. for its enhanced abilities to generate recognizable blockage signals and to not clog pores), but instead separately couple the presence/ absence of this cargo on the carrier to the presence/absence of the specific target molecule in a given sample.

Figure 6.2 outlines one implementation of such a scheme that is based on toehold-mediated DNA strand displacement [2,3]. The idea is to split the assay for the target into a two-step process: the first step uses a bioassay design to output a DNA oligo strand in the presence of the target, and the second (separately optimized) step uses this output strand to modify the signals of carrier translocations during a nanopore sensing experiment. Figure 6.2a illustrates a “sandwich assay” approach to generating the output strand that has previously been utilized in our lab [4]. An antibody-coated magnetic bead specifically captures a target protein which then also binds a secondary antibody that is conjugated to the output strand through a photocleavable linker. Once the excess reagents are washed off, a UV lamp is used to break apart the oligo strands from the sandwich complexes, allowing for their recovery from the supernatant.

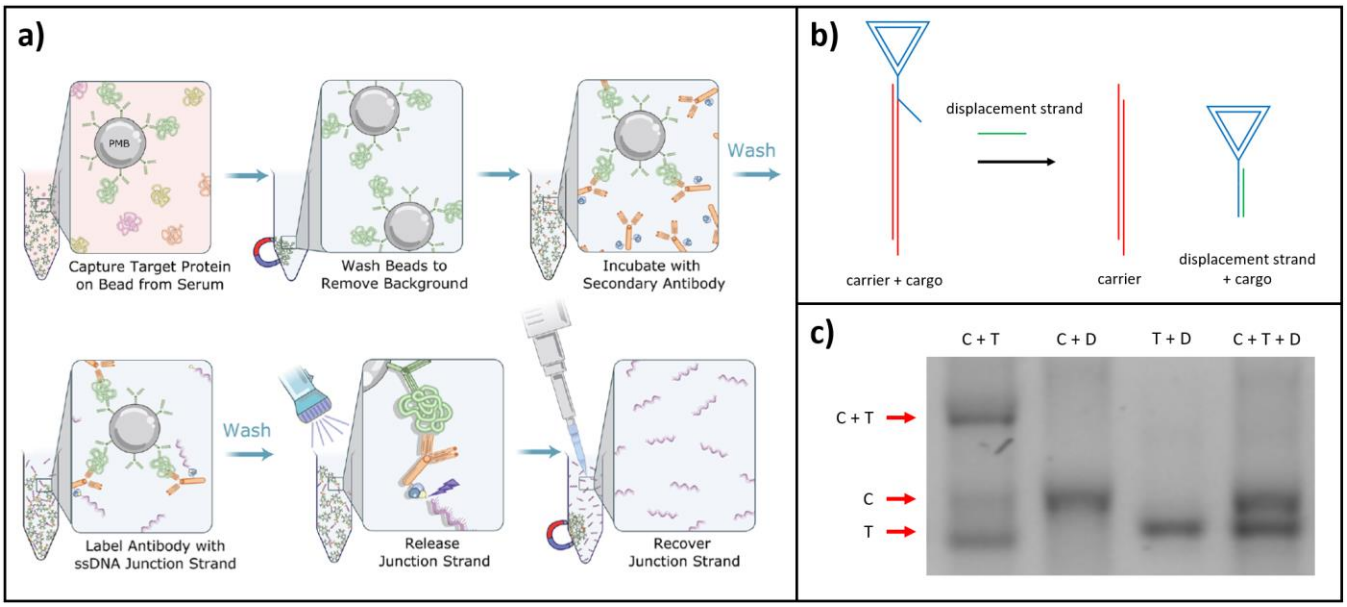


Figure 6.2: *Example scheme to translate a target molecule into a nanopore signal using DNA carriers.* ***a)*** *Method to generate an output oligonucleotide (here labelled as the "junction strand") from the presence of a target protein. The protein is "sandwiched" by the binding of both a magnetic bead-conjugated antibody and an output strand-conjugated antibody. The output strand is recovered after washing the magnetically captured sandwich complexes of excess reagents and photocleavage of the oligo-antibody linkers. Image adapted from* [4]*, with permission.* ***b)*** *Using toehold-mediated strand displacement to convert the output oligo of (a) into a carrier-based nanopore signal. The single-stranded overhang of a DNA nanostructure ("cargo") that hybridizes with a complementary overhang on the carrier is designed to feature a "toehold" extension. The cargo overhang (carrier complement + toehold) is fully complementary to the output strand (now acting as a "displacement strand") and so preferentially binds it, separating carrier and cargo.* ***c)*** *Sample agarose gel (2% in 0.5× TBE buffer) showing the displacement of a DNA tetrahedron nanostructure ("T") that features a 5-nt toehold extension from a 200 bp carrier ("C") in the presence of a 17-nt displacement strand ("D"). Gel data collected by undergraduate student Karolina Kropop.*

Once recovered, the output strands from the first step would then be transferred to a second solution volume to trigger a strand displacement reaction on a set of carrier-cargo molecules, as illustrated in Figure 6.2b. The system starts with a DNA nanostructure ("cargo") already bound to the end of a long dsDNA fragment ("carrier") through complementary single-stranded overhangs on each piece; notably, the overhang on the cargo is also extended by additional bases (comprising the "toehold") that are not complementary to the carrier. Meanwhile, the sequence of the output strand is designed to be complementary to the *entire* cargo overhang (carrier complement + toehold) so that its binding to the cargo is energetically favoured over that of the carrier. Because of this, output strands present in solution are able to initiate hybridization with the toeholds of pre-assembled carrier-cargo molecules and subsequently, as their bases anneal with the rest of the cargo overhang sequence, kick off the more weakly bound carriers from the complex (thereby acting as a "displacement strand").

When such a carrier-cargo mixture is loaded onto a nanopore sensor, the amount of strand displacement that occurred within the volume will translate into a proportional reduction in the number of carrier translocation events that feature an internal cargo sub-event (e.g. as in events *iii*, *iv* of Figure 6.1c) being recorded**.** Furthermore, since the amount of strand displacement scales in turn with the amount of target protein present during the sandwich assay step (via the amount of output strand released), this proportion of cargo-less carrier events is therefore also (indirectly) a measure of the original target concentration of the sample being tested. Figure 6.2c presents a sample gel (run by undergraduate student Karolina Kropop) of a strand displacement reaction on a DNA tetrahedron cargo with a 5-nt toehold that was attached to a 200-bp carrier, and shows the effective removal of the cargo from the carrier in the presence of the displacement strand.

We note that being able to make quantitative measurements of target concentration with the type of assay depicted in Figure 6.2 will depend on starting with relatively pure stocks of cargo-attached carriers, that is, before any displacement strands are added. If, for instance, cargo-free carriers exist in solution initially, this could lead to false positives in the form of nanopore signals lacking cargo sub-events that were not caused by a displacement strand. Similarly, if an excess of free cargo molecules exists in solution, these could sequester added displacement strands before they have a chance to interact with the carriers, leading to false negatives. The gel images in Figure 6.3 present a proof-of-concept purification step toward this goal of that involves molecular weight cut-off centrifugal filters (Pall Nanosep^{TM}), in this case aiming to remove an excess of free 12-arm star cargos. Both the 100k and 300k cut-off filters are observed to retain a significant portion of the 2-kbp carriers (especially the 100k filter) while almost entirely removing the (more compact) DNA stars in the flow-through (see the caption of Figure 6.3 for details).

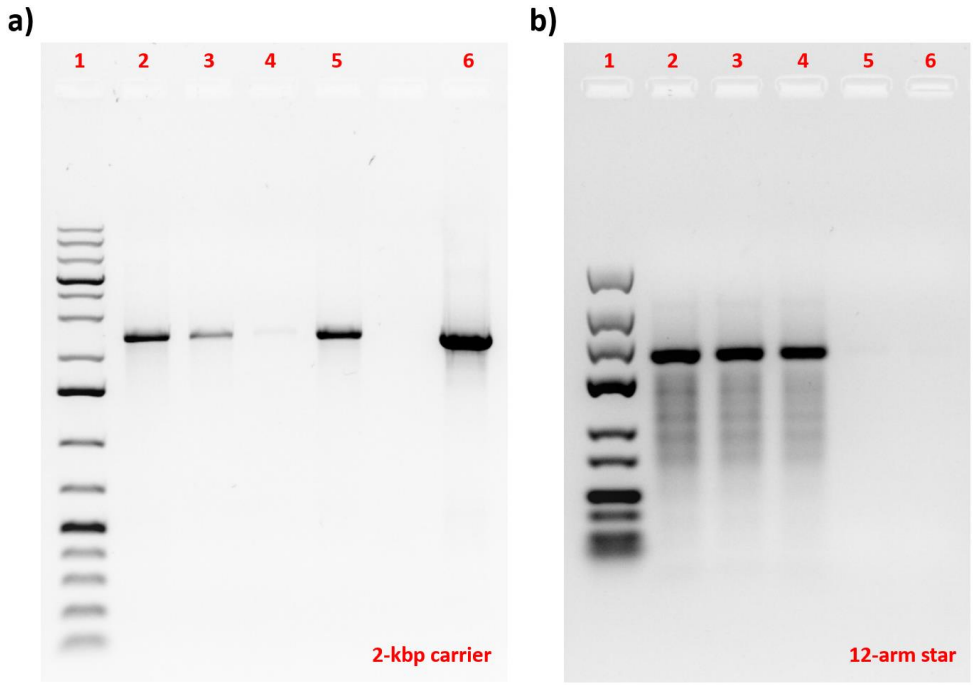


Figure 6.3: *Spin column separation of DNA carriers and cargo.* ***a)*** *Agarose gel (0.7%, 0.5× TBE, pre-stained with GelRed, 70 V) of 2-kbp DNA carriers passed through or retained by centrifugal filters (Pall Nanosep™). Lane 1: DNA ladder (GeneRuler 1 kb Plus). Lane 2: DNA carrier (control). Lane 3: Flow-through of DNA carriers in 300k filter. Lane 4: Flow-through of DNA carriers in 100k filter. Lane 5: Retention of DNA carriers in 300k filter. Lane 6: Retention of DNA carrier in 100k filter.* ***b)*** *Agarose gel (2%, 0.5× TBE, pre-stained with GelRed, 70 V) of 12-arm stars passed through or retained by centrifugal filters (Pall Nanosep™). Lane 1: DNA ladder (GeneRuler Low Range). Lane 2: Unpurified 12-arm star (control). Lane 3: Flow-through of DNA stars in 300k filter. Lane 4: Flow-through of DNA stars in 100k filter. Lane 5: Retention of DNA stars in 300k filter. Lane 6: Retention of DNA stars in 100k filter. From (a) and (b) the 100k filters appear to retain almost all of the 2-kbp carriers while passing almost all of the 12-arm stars and so could potentially be used to wash off excess stars prior to a strand-displacement assay involving these molecules (as depicted in Figure 6.2).*

Probing the Temperature Response of SE-DNA with Nanopores

As another application for the SE-DNA molecules produced earlier (outside of acting as molecular carriers), we explored using a temperature-controlled nanopore setup to probe the melting of their annealed single-stranded ends. Adding temperature control to a nanopore system widens its scope of possible studies to beyond room temperature phenomena and could be used, for instance, to monitor protein or DNA origami unfolding, to probe the interaction energies between two components, or even to carry out certain biochemical reactions (e.g. PCR) *in situ* near the pore. Previous gel studies characterizing our SE-DNA molecules in particular had displayed a sharp transition in molecular structure between moderate (e.g. 50 °C) and elevated (e.g. 70 °C) temperatures, from being able to form higher-order structures (such as circles and dimers) at the bottom of the temperature range to the molecules being completely linearized as the ssDNA end interactions stabilizing these

structures were disrupted by the increasing thermal energy in the system (e.g. see Figure 6.1b). If the different molecular structures found on either side of this temperature transition could be reliably distinguished by their translocation signals, then the progress of end melting could be monitored in real time with a nanopore. One possible application for this could be to include DNA carriers with different sticky end lengths (corresponding to different melting temperatures) in the same reservoir as a main nanopore analyte being tested – these could then act as a series of inert temperature reporters in solution by recording which of the sticky ends lengths had dissociated at various stages of the experiment.

Figure 6.4 presents an overview of a temperature sweep experiment performed on 7-kbp SE-DNA with 12-nt, complementary overhangs. For these tests, we used a temperature-controlled nanopore system built in our lab by Dmytro Lomovtsev [5]. 1D histograms of minimum blockage depth and average blockage depth are shown in Figures 6.4a and 6.4b, respectively, for SE-DNA translocation events recorded at room temperature. In both panels, clear evidence exists, in the form of sharp peaks at ~2× the level of unfolded dsDNA blockages (~900 pA vs. ~450 pA), of molecules that pass through the pore in a completely folded state. Presumably, many of these events can be attributed to circularized molecules, which can only translocate as pairs of dsDNA fragments due to their connected ends (e.g. see events *i* and *ii* in Figure 6.1c). These extra populations were also completely suppressed in a control nanopore experiment at room temperature using dsDNA carriers that lacked sticky ends (e.g. the average blockage distribution steadily decayed from the 1× level to the 2× level, without the appearance of a peak at 2×, data not shown). Figure 6.4c shows how both the baseline current through the pore and the blockage depth of unfolded dsDNA increased as the temperature was raised from 20 °C to 65 °C. Such increases are consistent with an expected rise in electrolyte conductivity with temperature, although the overall ratio of blockage depth to baseline current $\Delta I/I$ actually decreased slightly with temperature, an effect that has been observed previously [6]. Meanwhile, peak passage times through the pore were drastically lowered at higher temperatures (~3× from 20 °C to 65 °C), likely a consequence of the reduced buffer viscosity and thus drag opposing translocation in these conditions (see Figure 6.4d).

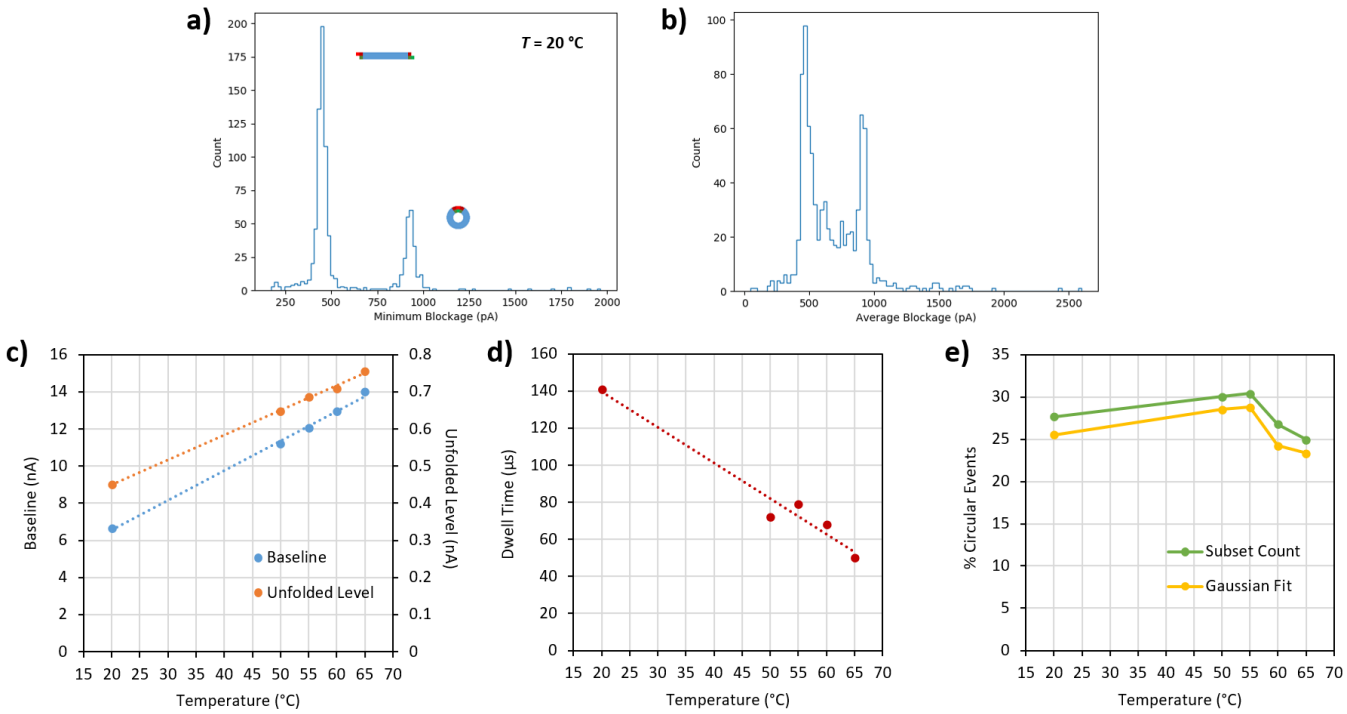


Figure 6.4: *Nanopore sensing of circularized 7-kbp SE-DNA under a temperature sweep (~9 nm pore diameter, 1.2 M LiCl, 200 mV). Histograms of* ***a)*** *minimum blockage and* ***b)*** *average blockage of analyzed SE-DNA translocations at room temperature (20 °C). Since circularized DNA is expected to only pass through the pore in (multiples of) two strands at a time, prominent populations centred on twice the blockage level of unfolded DNA should correlate with the proportion of circular vs. linear DNA strands present in the sample. Trends in* ***c)*** *open-pore current/blockage depth and* ***d)*** *most probable translocation time as the temperature is increased. Both the baseline and blockage currents increase (as expected from increased buffer conductivity) but not in a fixed ratio to one another. A drastic ~3× reduction in passage time is observed at 65 °C vs. room temperature.* ***e)*** *Percent "circular events" (defined via the ratio of fully folded events) vs. temperature. The results of two approaches to quantify this ratio are shown.*

The ratio of "circular events" to total events over the temperature range tested is presented in Figure 6.4e, as measured using the aforementioned population of events with minimum blockages at ~2× the unfolded dsDNA level (see Figure 6.4a). This value was calculated either by a raw count of events above a minimum blockage threshold (e.g. > 1.5× unfolded dsDNA) or by fitting the minimum blockage histograms to double Gaussian curves and extracting the area of 2× dsDNA peak, and both methods tracked similarly with temperature (although slightly offset from each other). Interestingly, the initial trend, from 20 °C to 55 °C, was for the percentage of fully folded events to rise slightly. This is unlikely the result of more circles being formed however, as prior to the nanopore experiment the samples had been equilibrated at 50 °C for > 1 hr and translocation data was acquired for only ~10 min at each temperature value. Instead, slight rises in the percentage of fully folded events are potentially indicative of some competing effects that arise during sensing and analysis, such as

faster translocations (Figure 6.4d) leading to poorly resolved signals of partially folded molecules that cannot be fitted with an unfolded sublevel, or greater levels of folding overall from the DNA generally becoming more flexible at higher temperatures (e.g. see Equation 1.6 in *Chapter 1*). On the other hand, as temperatures continued to be increased past 55 °C, these competing effects were seemingly overcome by the linearization of DNA circles and the fraction of fully folded events indeed began to drop before the pore became permanently clogged at 65 °C (a melting temperature ~67 °C is expected for the 12-nt end sequence under the tested experimental conditions [7]). Being able to extract accurate ratios of circular and linear carriers from a nanopore experiment will likely depend on adopting a more refined analysis approach (perhaps relying on all-points event histograms [8–10] rather than sublevel fitting) or at least quantifying and compensating for the competing effects by running a similar temperature sweep on blunt-ended molecules of the same length as a control.

It should be noted that running nanopore experiments at some of the higher temperatures featured in Figure 6.4 faced many challenges, including lower signal-to-noise ratios (Figure 6.4c), faster translocations (Figure 6.4d), bubbles appearing in solution that could de-wet the membrane, and increased rates of both target capture (good) and pore clogging (bad). There could therefore be a benefit to lowering the transition point of interesting, nanopore-resolvable phenomena to more accessible temperatures to the tool through the simultaneous use of chemical denaturants within the system. For instance, the additive urea is known to decrease the melting temperature of dsDNA by ~2.25 °C / (M urea) [11]. To this end, we also investigated the performance impact of urea on solid-state nanopore sensing.

Figure 6.5 shows the effect of urea concentration on the overall conductance through the pore. In the two pores tested, adding increasing amounts of urea at a fixed salt concentration (1.2 M LiCl) led to linear decreases in conductance (Figure 6.5a). While the rates of conductance decrease differed between the two pores, these trends collapsed to a single curve when normalized by the conductance at 0 M urea (Figure 6.5b), which is consistent with the effect being independent of pore geometry and isolated to the overall conductivity of the solution (see Eq. 1.3 in *Chapter 1*). This is likely related to the noticeably increased viscosities of urea solutions at concentrations above 1 M [12,13]. Interestingly, the reduction in pore conductance with urea was more drastic for solutions with

higher background salt concentrations as can be seen in Figure 6.5c, which plots the ratio of $G_{8M\ urea}$ / $G_{0M\ urea}$ through a pore at three different values of LiCl concentration. If the linear relationship extracted from Figure 6.5c is combined with that from Figure 6.5b, it results in an empirical interpolation formula for the conductivity of *any* LiCl/urea solution within the tested concentration ranges, which is useful for predicting the magnitude of ionic currents expected in an experiment. For instance, Figure 6.5d plots the interpolated conductivities of LiCl solutions containing 8 M urea, and predicts a maximum value at ~2.8 M LiCl, rather than the conductivity monotonically increasing over this range of LiCl concentration, as it does in the absence of urea [14].

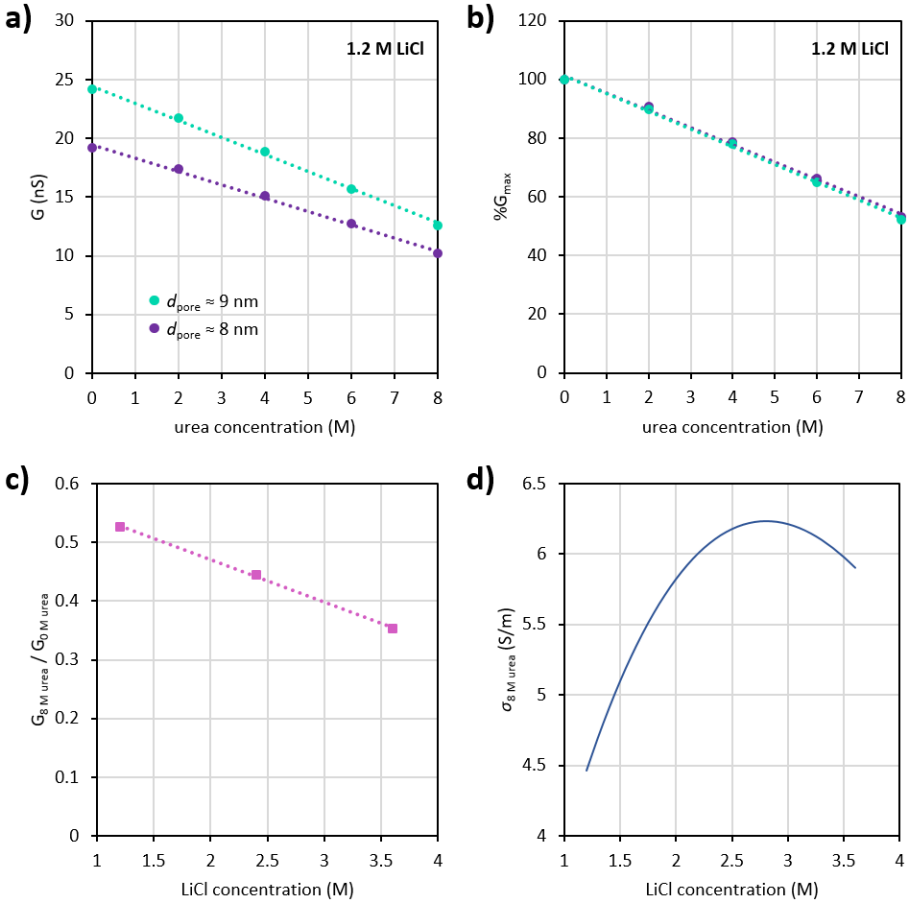


Figure 6.5: *Effect of urea on pore conductance.* ***a)*** *Pore conductance vs. urea concentration for two pores in a background of 1.2 M LiCl. The conductances decrease roughly linearly with increasing urea in both cases, but with different slopes.* ***b)*** *With the conductance plotted as a fraction of its value at 0 M urea however, both pores behave similarly, decreasing by ~6% of this max value for each 1 M addition of urea. This is consistent with the urea leading to an overall reduction in buffer conductivity, possibly due to increasing viscosity. The colour scheme for the two pores is maintained from (a).* ***c)*** *Decrease in pore conductance with 8 M urea relative to 0 M urea at different LiCl concentrations. Another (roughly) linear trend is observed, allowing for easy interpolation of pore conductance for any combination of urea and LiCl within the concentration limits tested (0 – 8 M urea, 1.2 – 3.6 M LiCl). An example of such an interpolation formula is plotted in* ***d)****, showing the predicted solution conductivity as a function of LiCl concentration in the presence of 8 M urea.*

Figures 6.6a and 6.6b present results from an experiment translocating (blunt-ended) 6-kbp DNA through a pore under two urea concentrations (0 M or 8 M urea in 2.4 M LiCl) at room temperature. As expected from the lowered solution conductivity (Figure 6.5), the blockage depths of both folded and unfolded translocations are noticeably reduced (~2×) in the presence of urea (Figure 6.6a). Event duration, on the other hand, increased dramatically (~4×) with added urea (Figure 6.6b), likely due to the higher viscosity of the solution [15]. These increased dwell times have interesting implications for a temperature-controlled nanopore experiment such as the melting of the DNA circles presented in Figure 6.4, since this should partially counteract the increases in translocation velocity that occurred at higher temperatures, which led to a loss in time resolution of folded/unfolded sublevels there (Figure 6.4d). Adding urea could also replace the need to lower the LiCl concentration of a such an experiment to prevent the baseline current from reaching the 20 nA saturation limit of our high-bandwidth amplifiers at elevated temperatures; while both approaches reduce the pore conductance, lowering [LiCl] decreases DNA translocation times [16] (compounding resolution issues) while raising [urea] increases them. Lastly, Figure 6.6c shows a gel image of SE-DNA circularizing at moderate temperatures (50 °C, slowly ramped down to 20 °C) both in the absence and presence of 7 M urea, lending further plausibility to the idea of a future temperature-controlled nanopore experiment to observe the full melting transition of SE-DNA molecules over a narrower temperature range (predicted reduction in T_M of ~2.25 °C/M × 7 M ≈ 16 °C).

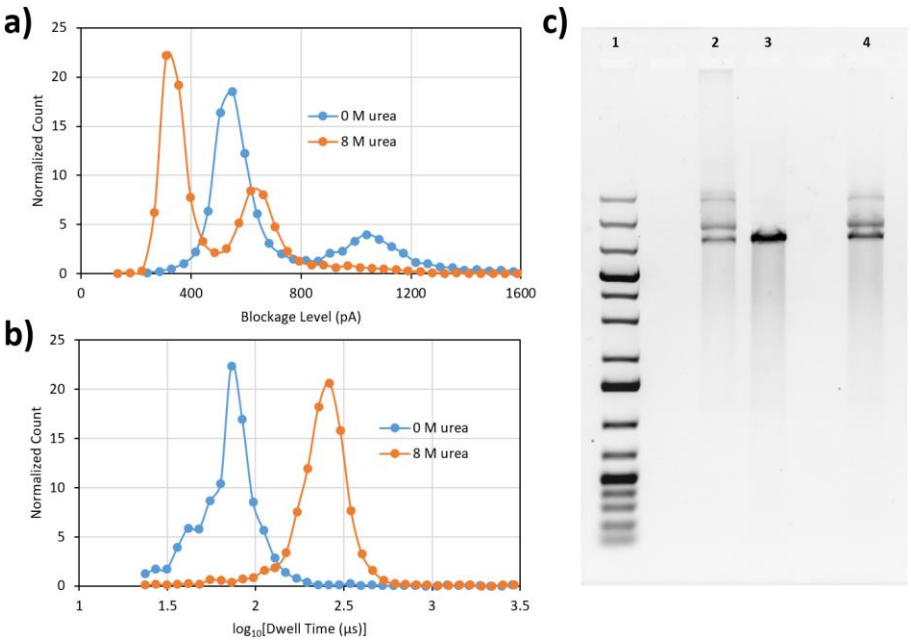


Figure 6.6: *DNA experiments with urea. Comparison of* ***a)*** *blockage depths and* ***b)*** *dwell times of translocation events corresponding to 6 kbp dsDNA driven through the same 9-nm pore (200 mV in 2.4 M LiCl) in the absence (blue) or presence (orange) of 8 M urea. As might be predicted from the lower overall solution conductivity observed above, the blockage depths in 8 M urea are reduced relative to the urea-free condition. Interestingly, this is coupled with an increase in dwell times for the samples in urea (3 – 4× shift in distribution peak).* ***c)*** *Agarose gel (0.5%, 0.5× TAE, pre-stained with GelRed, 70 V) showing effect of urea on circularization/dimerization of 7-kbp SE-DNA. Lane 1: DNA ladder (GeneRuler 1 kb Plus). Lane 2: SE-DNA incubated at 50° C in 10 mM $MgCl_2$. Lane 3: aliquot of the same 50° C incubation as Lane 2 followed by 10 min at 70° C to melt interactions of the 12-nt sticky ends. Lane 4: SE-DNA incubated at 50° C in 10 mM $MgCl_2$ + 7 M urea. The presence of elevated levels of urea does not completely prevent interactions of the sticky ends at moderate temperatures. This may allow for a future temperature-controlled nanopore experiment that starts at room temperature and finishes well below 70 °C to fully probe the melting of the end interactions.*

As a final note, nanopores can also be used to monitor the reverse process to end melting in real time, namely, end annealing. Figure 6.7 outlines an experiment where samples of 2-kbp DNA carriers and complementary 12-arm stars were combined in the reservoir of a nanopore fluidic cell and sensed by the pore (current trace presented in Figure 6.7a). As discussed previously (see *Chapter 5*), distinct populations of events from this experiment were visible in a scatter plot of blockage depth vs. dwell time, consisting of: deep and short-lived events (free stars), shallow and long-lived events (free carriers), and finally deep and long-lived events (star-attached carriers, circled red in Figure 6.7b). While the overall event rate of captured molecules remained relatively consistent over the course of the ~20 min experiment (blue points in Figure 6.7c), the specific rate for deep and long-lived events increased steadily with time (red points), signalling a rising proportion of hybridized carrier-cargo molecules in the reservoir. In this way, nanopores could be used to extract reaction kinetics *in situ*

from a molecular interaction of interest, so long as the products of this interaction lead to distinguishable current blockage signals from the reactants.

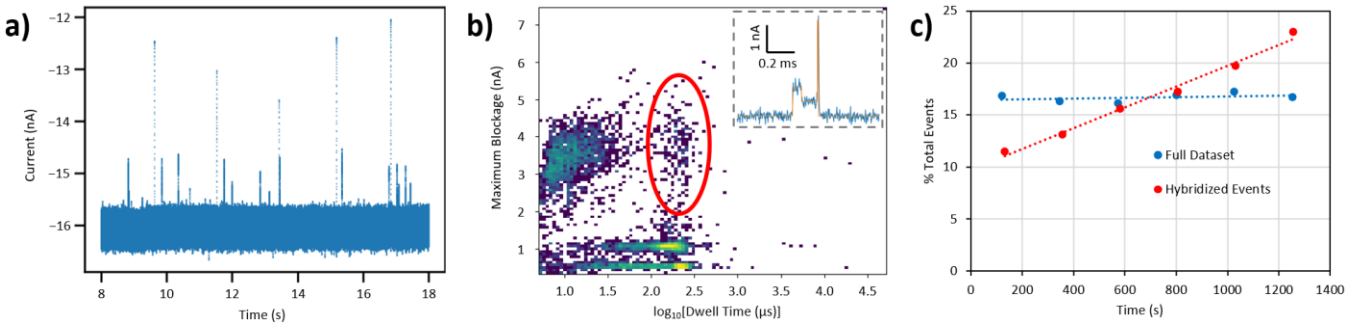


Figure 6.7: *Real time nanopore monitoring of sticky-end hybridization.* ***a)*** *Current trace (10 s) of 2-kbp DNA carriers and 12-arm DNA stars translocating a nanopore (~15 nm pore, 3.6 M LiCl, 125 mV). Blockage events corresponding to each species are distinguishable by eye in the trace as being either relatively deep (ΔI ≈ 3 – 4 nA, stars) or shallow (ΔI ≲ 1 nA, carrier).* ***b)*** *2D histogram of max blockage vs. log[dwell time] for all translocation events (~20 min) from the experiment in (a). The population of deep and long-lived events associated with hybridized carrier-star translocations (sample event trace in inset) is circled in red.* ***c)*** *Events binned by start time for both the full dataset (in blue) and the circled "hybridized" population (in red) from (b). The slope for the full dataset is relatively flat (0.0%/min), reflecting a constant overall nanopore capture rate over this ~20 min time interval, while there is a noticeable increase in the rate of "hybridized" events (0.6%/min) over time, which we attribute to increasing reaction progress of the in situ combination of stars and carriers outside the nanopore.*

6.3 – References

1. Roelen Z, Tabard-Cossa V. Synthesis of length-tunable DNA carriers for nanopore sensing. Sauli E, editor. PLoS One. 2023;18: e0290559. doi:10.1371/journal.pone.0290559

2. Zhang DY, Winfree E. Control of DNA Strand Displacement Kinetics Using Toehold Exchange. J Am Chem Soc. 2009;131: 17303–17314. doi:10.1021/ja906987s

3. Zhu J, Zhang L, Zhou Z, Dong S, Wang E. Molecular aptamer beacon tuned DNA strand displacement to transform small molecules into DNA logic outputs. Chemical Communications. 2014;50: 3321. doi:10.1039/c3cc49833f

4. He L, Tessier DR, Briggs K, Tsangaris M, Charron M, McConnell EM, et al. Digital immunoassay for biomarker concentration quantification using solid-state nanopores. Nat Commun. 2021;12: 5348. doi:10.1038/s41467-021-25566-8

5. Lomovtsev D. A platform for high-bandwidth , low-noise electrical nanopore sensing with thermal control. University of Ottawa. 2022. doi:10.20381/ruor-27924

6. Nicoli F, Verschueren D, Klein M, Dekker C, Jonsson MP. DNA Translocations through Solid-State Plasmonic Nanopores. Nano Lett. 2014;14: 6917–6925. doi:10.1021/nl503034j

7. Kibbe WA. OligoCalc: an online oligonucleotide properties calculator. Nucleic Acids Res. 2007;35: W43–W46. doi:10.1093/nar/gkm234

8. Li J, Gershow M, Stein D, Brandin E, Golovchenko JA. DNA molecules and configurations in a solid-state nanopore microscope. Nat Mater. 2003;2: 611–615. doi:10.1038/nmat965

9. Storm AJ, Chen JH, Zandbergen HW, Dekker C. Translocation of double-strand DNA through a silicon oxide nanopore. Phys Rev E. 2005;71: 051903. doi:10.1103/PhysRevE.71.051903

10. Steinbock LJ, Otto O, Chimerel C, Gornall J, Keyser UF. Detecting DNA Folding with Nanocapillaries. Nano Lett. 2010;10: 2493–2497. doi:10.1021/nl100997s

11. Hutton JR. Renaturation Kinetics and thermal stability of DNA in aqueous solutions of formamide and urea. Nucleic Acids Res. 1977;4: 3537–3555. doi:10.1093/nar/4.10.3537

12. Tonomura T, Okamoto K. The Viscosity and Conductivity of Aqueous Solutions of Electrolytes Mixed with Glycine and Urea. Bull Chem Soc Jpn. 1966;39: 1621–1627. doi:10.1246/bcsj.39.1621

13. Singer A, Kuhn H, Frank-Kamenetskii M, Meller A. Detection of urea-induced internal denaturation of dsDNA using solid-state nanopores. Journal of Physics: Condensed Matter. 2010;22: 454111. doi:10.1088/0953-8984/22/45/454111

14. Wu X, Gong Y, Xu S, Yan Z, Zhang X, Yang S. Electrical Conductivity of Lithium Chloride, Lithium Bromide, and Lithium Iodide Electrolytes in Methanol, Water, and Their Binary Mixtures. J Chem Eng Data. 2019;64: 4319–4329. doi:10.1021/acs.jced.9b00405

15. Fologea D, Uplinger J, Thomas B, McNabb DS, Li J. Slowing DNA Translocation in a Solid-State Nanopore. Nano Lett. 2005;5: 1734–1737. doi:10.1021/nl051063o

16. Kowalczyk SW, Wells DB, Aksimentiev A, Dekker C. Slowing down DNA Translocation through a Nanopore in Lithium Chloride. Nano Lett. 2012;12: 1038–1044. doi:10.1021/nl204273h