Advanced Studies in Medical Sciences, Vol. 5, 2017, no. 1, 13 - 53
HIKARI Ltd, www.m-hikari.com
https://doi.org/10.12988/asms.2017.722

Biological System Analysis Using a Nanopore Transduction Detector: from miRNA Validation, to Viral Monitoring, to Gene Circuit Feedback Studies

Stephen Winters-Hilt

Biology and Computer Science Department
Connecticut College
270 Mohegan Ave.
New London, CT 06320, USA
&
Meta Logos Inc.
124 White Birch Dr.
Guilford, CT 06437, USA

Abstract

The system biologist is currently lacking a general-use method for a gene circuit 'voltmeter' or gene system algorithm 'print statement'. What is needed is a non-destructive, carrier non-modifying, means of testing 'live' biological systems at the single-molecule level. A method using the nanopore transduction detector (NTD) is demonstrated for single-molecule characterization in some situations, so may provide what is lacking. An important aspect of this approach is that use can be made of inexpensive antibody, protein, aptamer, duplex nucleic acid, or nucleic acid annealing molecules (for miRNA and viral monitoring) that have *specific* binding to the system component of interest. The NTD transducer's specific binding can also be designed to have low affinity binding as needed, such that there can be a 'catch and release' on low copy-number molecular components, such that there is not a disruption to the molecular system under study. NTD transducers are typically constructed by linking a binding moiety of interest to a

nanopore current modulator, where the modulator is designed to be electrophoretically drawn to the channel and partly captured, with its captured end distinctively modulating the flow of ions through the channel. Using inexpensive (commoditized) biomolecular components, such as DNA hairpins, this allows for an easily constructed, versatile, platform for biosensing. High specificity high affinity binding also allows a very versatile platform for assaying at the single molecule level, even down to the single isoform level, including molecular substructure profiling, such as glycosylation profiling in antibodies. An inexpensive commoditized pathway for constructing nanopore transducers is demonstrated. Nanopore transduction detector based reporter/event-transducer molecules may serve as a means to perform multicomponent mRNA-miRNA-protein and protein-protein systems analysis in general settings.

Keywords: Biological System; Gene Circuit; Nanopore Detection; miRNA

1 Introduction

A growing number of questions facing molecular and medicinal biology experts are systems biology questions, where the complex interaction of genes, mRNAs, proteins, miRNAs, and various metabolites is described at the 'system level'. System level problems are often described in terms of 'gene circuits' or 'metabolic algorithms'. These comparisons to system descriptions in electrical engineering and computer science offer some insights due to actual parallels, and some misleading comparisons due to oversimplification in comparison to actual biological systems.

A reductionist analysis of a biological system, not surprisingly, reveals that the sum is greater than its parts. But this is actually found to be the case in electrical circuits as well, where emergent properties, especially emergent noise and communications properties, are often found in circuits with feedback. Even simple physical systems involving just three bodies in classical orbital dynamics gives rise to chaotic behavior, which was not expected in early physics, where the sum was originally NOT thought to be greater than its parts. Iterative dynamical systems in general are found to exhibit chaotic behavior and emergent constructs such as strange attractors and limit cycles. Systems with feedback, thus, can do surprising things, and biological systems definitely have done some surprising things ranging from living systems in their amazing variety to complex phenomena such as intelligence, language, and consciousness.

The nanopore transduction detector (NTD) method is typically based on a single protein-channel biosensor implemented on a lipid bilayer (synthetic cell membrane), but it could also be implemented as a live cell assay by using the original patch clamp protocol for measuring current through a channel on a live cell [1] (the invention of the patch clamp amplifier resulted in the Nobel Prize for

Medicine or Physiology in 1991 [2]). In order for the NTD 'voltmeter' operating on the biological system to work, cell-based or not, the normal operational buffer of the NTD must accommodate a change to the physiological or cellular buffer environment of the biological system of interest, and, if cell-based, the 'carrier signal' that is the basis of the analysis can no longer be channel-current based, but channel-noise based with use of laser modulations for noise state excitation. Work with robust NTD operation with a variety of buffer pH and in the presence of high concentrations of interference agents reveals that operational stability with a wide range of buffers has been achieved [3]. Laser modulations have also been introduced to improve the NTD mechanism to have more general applicability [4], and for purposes of establishing an improved 'stochastic carrier wave' molecular state tracking capability [5], so many of the complications with returning to the single-cell application are mostly solved. What remains to be resolved for general applicability of the NTD system analyzer method, for both *in vitro* and an possible *in vivo* studies (see Disc.), is a standardized method for NTD transducer construction and operation (see Suppl. and [6]), and progress along these lines will be shown in the Results. An inexpensive method for a NTD-based biological system 'voltmeter' is thus possible for both *in vivo* and *in vitro* applications.

2 Background

2.1 Nanopore Transduction Detection (NTD)

The nanopore transduction detection (NTD) platform [5, 7] includes a single nanometer scale channel and an engineered, or selected, channel blockading molecule. The channel blockading molecule is engineered to provide a current modulating blockade in the detector channel when drawn into the channel, and held, by electrophoretic means. The channel has inner diameter at the scale of that molecule. For most biomolecular analysis implementations this leads to a choice of channel that has inner diameter in the range 0.1-10 *nanometers* to encompass small and large biomolecules, where the inner diameter is 1.5 nm in the alpha-hemolysin protein based channel used in the results that follow (see Fig. 1 and Suppl. Fig.s 1 & 2). Given the channel's size it is referred to as a nanopore in what follows. In efforts by others 'nanopore' is sometimes used to describe 100-1000 nm range channels, which are here referred to here as micropores.

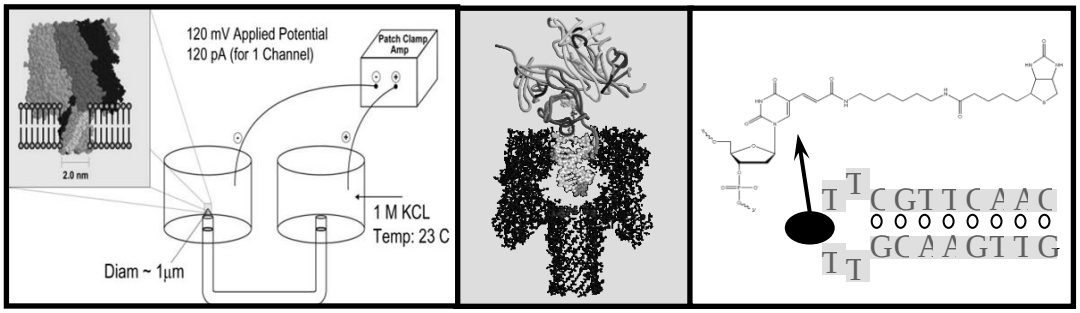


Figure 1. Schematic diagram of the Nanopore Transduction Detector.

Reprinted with permission of [7]. **Left**: shows the nanopore detector consists of a single pore in a lipid bilayer which is created by the oligomerization of the staphylococcal alpha-hemolysin toxin in the left chamber, and a patch clamp amplifier capable of measuring pico Ampere channel currents located in the upper right-hand corner. **Center**: shows a biotinylated DNA hairpin molecule captured in the channel's cis-vestibule, with streptavidin bound to the biotin linkage that is attached to the loop of the DNA hairpin. **Right**: shows the biotinylated DNA hairpin molecule (Bt-8gc).

In order to have a *capture* state in the channel with a *single* molecule, a true nanopore is needed, not a micropore, and to establish a coherent capture-signal exhibiting non-trivial stationary signal statistics, which is the modulating-blockade desired, the nanopore's limiting inner diameter typically needs to be sized at approximately 1.5nm for duplex DNA channel modulators (precisely what is found for the alpha-hemolysin channel). The modulating-blockader is captured at the channel for the time-interval of interest by electrophoretic means, which is established by the applied potential that also establishes the observed current flow through the nanopore.

The NTD molecule providing the channel blockade has a second functionality, typically to specifically bind to some target of interest, with blockade modulation discernibly different according to binding state (DNA annealing examples are shown in Suppl. Fig. 3 [8] and Fig. 2 [9]). NTD modulators are engineered to be bifunctional: one end is meant to be captured and modulate the channel current, while the other, extra-channel-exposed end, is engineered to have different states according to the event detection. Examples include extra-channel ends linked to binding moieties such as antibodies, antibody fragments, or aptamers. Examples also include 'reporter transducer' molecules with cleaved/uncleaved extra-channel-exposed ends, with cleavage by, for example, UV or enzymatic means [5]. By using signal processing with pattern recognition to manage the streaming channel current blockade modulations, and thereby track the molecular states engineered into the transducer molecules, a biosensor or assayer is enabled.

Fundamentally, the weaknesses of the standard ensemble-based binding analysis methods are directly addressed with this single-molecule approach. The role of conformational change during binding, in particular, could potentially be directly explored in this setting. This approach also offers advantages over other translation-based nanopore detection approaches in that the transduction-based apparatus introduces two strong mechanisms for boosting sensitivity on single-molecule observation: (i) engineered sensitivity in the transduction molecule itself; and (ii) machine learning based signal stabilization and highly sensitive state resolution. NTD used in conjunction with recently developed pattern recognition informed sampling capabilities greatly extends the usage of the single-channel apparatus [10] (including learning the avoidance of blockades associated with channel failure, when contaminants necessitate; and nanomanipu-

lation, where we have a single-molecule under active control in a nanofluidics-controlled environment). For medicine and biology, NTD methods may aid in understanding multi-component interactions (with co-factors or adjuvants), and aid in designing co-factors according to their ability to result in desired binding or modified state.

Nanopore transduction detection (NTD) works at a scale where physics, chemistry, and biomedicine methodologies intersect. In some applications the NTD platform functions like a biosensor, or an artificial nose, at the single-molecule scale, e.g., a transducer molecule rattles around in a single protein channel, making transient bonds to its surroundings, and the binding kinetics of those transient bonds is directly imprinted on a surrounding, electrophoretically driven, flow of ions. The observed channel current blockade patterns are engineered or selected to have distinctive stationary statistics, and changes in the channel blockade stationary statistics are found to occur for a transducer molecule's interaction moiety upon introduction of its interaction target. In other applications the NTD functions like a 'nanoscope', e.g., a device that can observe the states of a single molecule or molecular complex. With the NTD apparatus the observation is not in the optical realm, like with the microscope, but in the molecular-state classification realm. NTD, thus, provides an unprecedented new technology for characterization of transient complexes. The nanopore detection method uses the stochastic carrier wave signal processing methods developed and described in prior work [5,7], and comprises machine learning methods for pattern recognition that can be implemented on a distributed network of computers for real-time experimental feedback and sampling control [10]. Details on engineering NTD transducers is given in the Suppl. and [6].

2.2 Electrical/Biological circuits: the biosystem extra element theorem (BEET)

A reductionist analysis of electrical circuits involves a reduction to circuit elements that have linear responses. In this regard biology only compares weakly, as the components of a biological circuit are generally non-linear over much of their operational range. Even so, for some biological system settings sufficiently small perturbations in the biological components can often be made such that they provide a linear system response. Given the complexity of the biological feedback systems, however, this might seem to be small progress. It is very significant, however, given the existence of a sophisticated method from advanced circuit design and analysis that is applicable for linear response systems known as the 'extra element theorem' [11]. It interesting to note that this important circuit method from electrical system theory has not been imported into biological system discussions given its likely significant role in molecular evolutionary theory. The extra element theorem from electrical circuit theory allows simpler circuits, that are more easily understood, to have new components added (the 'extra' element), and if the new component happens to create a feedback loop, then the complexity of the feedback loop analysis can be much more easily evaluated and understood directly by way of the extra element theorem.

In practice, very complex electrical amplifier circuits can be built-up and analyzed in this way, by repeated use of the extra element theorem. This offers the means to have a reductionist analysis while capturing the growing complexity of holistic irreducible systems. For a biological variant of the extra element theorem a patchwork of linear response regimes could be used in understanding a particular biological system.

The 'messengers' in biological and electrical systems differ greatly in many respects, which can make some gene circuit intuition entirely misguided. The carriers in an electrical circuit, for example, are remarkably simple by comparison with biological system signal carriers. Electrical charge moves through wires like a fluid. Granted, the electrical charge moves at a sizable fraction of the speed of light, but it is so like a fluid flow that some current flow discussions are basically plumbing discussions, where the description of the current flow is often compared to flow of water through pipes where pipe narrowness is akin to resistance, etc. The flow/interaction topology of electrical current is also self-evident in the connectivity that can be seen in the wiring of the circuit diagram. If the biological system is too interconnected in this comparison this is often where the analogy is shifted to discussions of a gene system *algorithm*. The electrical messengers, or charge carriers, are also vastly simpler than the biological system messengers. Electrical current carriers are of only one type (electrons), and don't have attractive self-interaction molecular carriers (as with dimerization … unless you are talking superconductivity), and don't have internal state (in the sense of the circuit model) like with biological secondary messengers. Biological system messengers, on the other hand, come in a huge variety, operate at the single molecule level, and depending on perspective, everything in the biological system might be considered a system messenger in a massive, living, autocatalytic cascade. The biological system carriers or messengers are also much fewer in number compared to their electrical counterparts. This actually makes things more complicated. In electronics having small currents is modelled as a noise source, where once the discreteness of the charge carriers begins to be discernible this puts one in the realm of stochastic 'shot' noise. In the biological comparison this stochastic underpinning, if significant, again favors a shift to the 'algorithm' analogy instead of the circuit analogy. To further complicate matters, the biological carriers of the system interactions interact with each other, and typically have internal states (e.g., proteins and riboswitches often have conformational states), so the picture of the carriers for biology introduces vastly greater complexity and interaction interconnectivity.

In electrical circuit analysis a good voltmeter is something that will not significantly 'load' or alter the circuit while measuring a particular component's voltage drop. Likewise, in analyzing a computer program, or resolving a runtime error (the closest analogy to analyzing a 'live' biological algorithm), one of the best tools available is to simply introduce a 'print statement' to track any internal state behavior of interest in the program. This is where the weakness of the circuit

or algorithm analogy in biological systems is most profound. The system biologist doesn't have a gene circuit voltmeter or gene system algorithm print statement. Some of the closest biochemistry methods to offer such capabilities are fluorescence based, and in certain specialized applications remarkable results have been obtained along these lines, but they typically involve the introduction of constructs with a great deal of effort that won't scale well to the vast number of biological systems that need to be studied in the post-genomic era. What is needed is a non-destructive, carrier non-modifying, means of testing 'live' biological systems, possibly in their native cellular environment.

2.3 Validation of miRNA's and miRNA binding sites using a nanopore transduction nanoscope

The discovery of the RNA interference (RNAi) immune response and translational regulation mechanism has led to an explosion in the number of identified microRNAs (miRNAs) and their mRNA binding sites. An understanding of miRNAs and their binding sites, typically in the 3' untranslated region (3' UTRs) of mRNAs, is helping to explain a wide range of complex phenomena, ranging from latency control by viruses during infection (such as with HIV) [12], to complex regulation in system syndromes such as in diabetes and in the effects of aging [13], to the general trans-regulation of mRNAs at the translational level (complementing transcription factor and promoter cis-regulation at the transcriptional level) [14]. The examination of miRNAs, and especially miRNA binding sites, is confounded by the small size of the miRNAs, however: 21-25 nucleotides in length for typical mature miRNAs, and only 7-8 base ssRNA seed regions in the guide-strand RNA incorporated into the RNAi's RISC complex for actual binding/repression to complementary 7-8 base sequence in the 3'UTR region of the target mRNA [15]. For the latter case of verification for miRNA/RISC derived sequence binding with a 7 base sequence in a mRNA's 3'UTR there is further complication given possible posttranscriptional modifications, such as via inosine substitution for adenosine due to adenosine deaminases with inosine recognition as guanine in terms of base-pairing that can alter the actual target sequence of the miRNA/RISC binding [16]. This is in addition to the obvious complication of identifying the presence of RNA annealing when the annealing only involves 7 bases of RNA.

Preliminary work with NTD-based detection on short DNA annealing suggests a possible means to examine the miRNA/RISC binding to target 3'UTR region with or without the RISC complexes argonaute proteins intact, where results are expected to improve even more upon refinement using locked nucleic acid transducer/reporter probes (see Discussion). NTD based detection of DNA annealing has been demonstrated on DNA sequences as short as 5 bases [8], and in the presence of a variety of interference agents and chaotropes [3]. NTD based detection has also been demonstrated in a variety of buffer conditions so could be established in a buffer conducive to the RISC complex remaining intact and where the annealing to 3'UTR complement sequence occurs with the binding strength

found *in vivo*. NTD detection can also operate on small volumes since it makes use of a *single* protein channel interaction, thereby inherently operating at the single-molecule interaction level. NTD detection can, thus, identify single-molecule binding events in a non-destructive manner that may be conducive to the 'live' characterization of many critical, transient, interactions.
For biosensing or bioassays applications in general, not all miRNA or miRNA binding site analyses need be in cellular or physiological buffer either. In a 'destructive setting' more forceful miRNA validation assays, and analysis of annealing-based events, can be pursued by use of chaotropes such as urea. Clearer identification of collective binding events, such as for highly complementary annealing interactions, is found to occur upon introduction of chaotropes that eliminate non-specific DNA interactions, and many 'simple' binding interactions, not involving collective interactions of many components as with annealing [9].

2.4 Possible us of a nanopore transduction detector for rapid viral testing

The rapid development, deployment, and evaluation of a virus diagnostic would afford a patient the critical time needed to undergo a basic anti-viral therapy during the critical early infectious stages of a viral illness. A method and system appears to be possible for targeted DNA annealing tests using a nanopore transduction detector (NTD) where use is made of a DNA-annealing reporter molecule that is engineered, via a simple design process akin to probe designs for use in nucleic acid microarrays, to transduce strongly matched, and annealed, segments of the a virus genome to associated channel-current blockade events. The rapid viral test procedure is equally applicable to fungal and bacterial infection and can be designed to distinguish the strain of the infection if the genome for that strain is known. Further details on targeted nucleic acid assaying with a NTD are in the Suppl.

2.5 Use of chaotropes to improve signal resolution

In the nucleic acid annealing studies on the NTD platform described in [7] (see Fig. 2), the introduction of chaotropes allows for improved nucleic acid annealing identification.

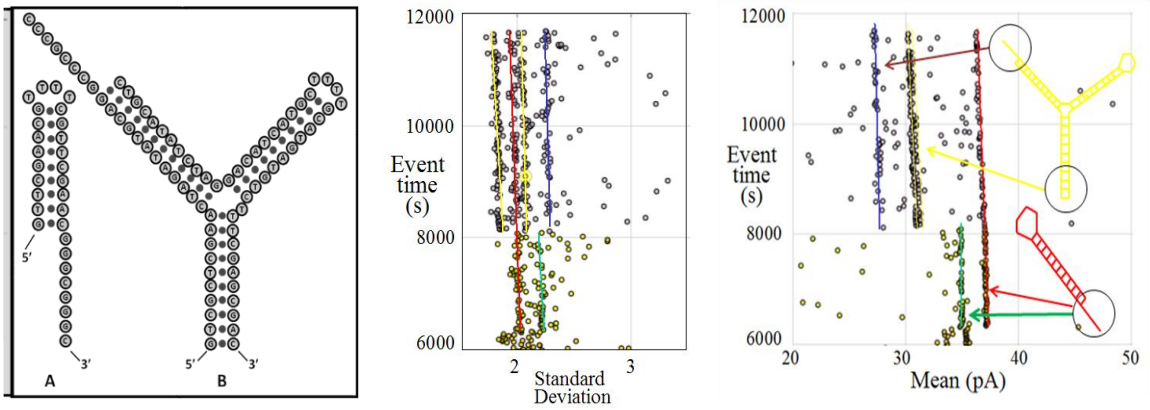


Figure 2. Eight-base annealing using a NTD Y-transducer.

Reprinted with permission of [7]. **Left:** The DNA hairpin and DNA Y-nexus transducer secondary structures with sequence information shown. **Center and Right: Y-shaped DNA transducer with overhang binding to DNA hairpin with complementary overhang.**

The ability of the of the NTD apparatus to tolerate high chaotrope concentration, up to 5M urea, was demonstrated in [3]. DNA hairpin control molecules have demonstrated a manageable amount of isoform variation even at 5M urea (Fig. 3).

In Fig. 2, only a portion of a repetitive validation experiment is shown, thus time indexing starts at the 6000th second. From time 6000 to 6300 seconds (the first 5 minutes of data shown) only the DNA hairpin (sequence details in [7,9]) is introduced into the analyte chamber, where each point in the plots corresponds to an individual molecular blockade measurement. At time 6300 seconds urea is introduced into the analyte chamber at a concentration of 2.0 M. The DNA hairpin with overhang is found to have two capture states (clearly identified at 2 M urea). The two hairpin channel-capture states are marked with the green and red lines, in both the plot of signal means and signal standard deviations. After 30 minutes of sampling on the hairpin+urea mixture (from 6300 to 8100 seconds), the Y-shaped DNA molecule is introduced at time 8100. Observations are shown for an hour (8100 to 11700 seconds). A number of changes and new signals now are observed: (i) the DNA hairpin signal class identified with the green line is no longer observed – this class is hypothesized to be no longer free, but annealed to its Y-shaped DNA partner; (ii) the Y-shaped DNA molecule is found to have a bifurcation in its class identified with the yellow lines, a bifurcation clearly discernible in the plots of the signal standard deviations. (iii) the hairpin class with the red line appears to be unable to bind to its Y-shaped DNA partner, an inhibition currently thought to be due to G-quadruplex formation in its G-rich overhang. (iv) The Y-shaped DNA molecule also exhibits a signal class (blue line) associated with capture of the arm of the 'Y' that is meant for annealing, rather than the base of the 'Y' that is designed for channel capture.

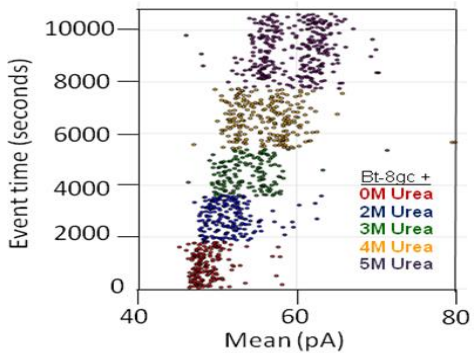


Figure 3. Bt-8gc transducer blockade signals in the presence of high urea concentrations. Reprinted with permission [3]. Sufficiently strong Urea concentration (5M) results in racemization of the two loop capture-variants, while weaker urea (<2M) does not. The results show Bt-8gc measurements at 30 minute

intervals (1800 s on vertical axis) with urea concentration 0, 2, and 3M, 45 minutes at 4M, and 60 minutes at 5 M, with signal blockade mean on the x-axis, with results consistent with the two-state loop hypothesis, and consistent with the observation of such in Fig. 1 (see [7]) not due to zero or weak urea content but due to high strain due to mass and charge effects upon binding to the large streptavidin molecule.

2.6 Managing common interference agents, and aptamers and antibodies as easily identifiable interference or transducer

Preliminary results involving interference tests are provided in the Results section, and a brief background on interference issues is given there. For further details on aptamers in the context of nanopore transduction detection see [17-19], and in general see [20-24]. Further details on antibodies pertinent to nanopore detection are placed in the Suppl., including diagrammatic figures and structural details from various references [25-32].

2.7 Transduction enables Channel Current Cheminformatics (CCC)

In the NTD platform, sensitivity increases with observation time in contrast to translocation technologies where the observation window is fixed to the time it takes for a molecule to move through the channel [5]. Part of the sensitivity and versatility of the NTD platform derives from the ability to couple real-time adaptive signal processing algorithms to the complex blockade current signals generated by the captured transducer molecule. The NTD system, deployed as a biosensor platform, possesses highly beneficial characteristics from multiple technologies: the specificity of antibody or aptamer binding, the sensitivity of an engineered channel modulator to specific environmental change, and the robustness of the electrophoresis platform in handling biological samples.

3 Methods

3.1 Nanopore Detector Experiments

Each experiment is conducted using one alpha-hemolysin channel inserted into a diphytanoyl-phosphatidylcholine/hexadecane bilayer across a, typically, 20-micron-diameter horizontal Teflon aperture. The alpha-hemolysin pore has a 2.0 nm width allowing a dsDNA molecule to be captured (while a ssDNA molecule translocates). The effective diameter of the bilayer ranges mainly between 5-25 μm (1 μm is the smallest examined). This value has some fluctuation depending on the condition of the aperture, which station is used (each nanopore station, there are four, has its own multiple aperture selections), and the bilayer applied on a day to day basis. Seventy microliter chambers on either side of the bilayer contain 1.0 M KCl buffered at pH 8.0 (10 mM HEPES/KOH) except in the case of buffer experiments where the salt concentration, pH, or identity may be

varied. Voltage is applied across the bilayer between Ag-AgCl electrodes. DNA control probes are typically added to the *cis* chamber at 10-20 nM final concentration. All experiments are maintained at room temperature (23 ± 0.1 °C), using a Peltier device.

3.2 NTD control probes

The five DNA hairpins studied in [33, 34] have been carefully characterized, so are used as highly sensitive controls (obtained from IDT DNA with PAGE purification). The nine base-pair hairpin molecules share an eight base-pair hairpin core sequence, with addition of one of the four permutations of Watson-Crick base-pairs that may exist at the blunt end terminus, i.e., 5'-G|C-3', 5'-C|G-3', 5'-T|A-3', and 5'-A|T-3'. Denoted 9GC, 9CG, 9TA, and 9AT, respectively. The full sequence for the 9GC hairpin is 5'-GTTCGAACGTT TTCGTTCGAAC-3'. The eight base-pair DNA hairpin (8GC) is identical to the core eight base-pair part of the 9GC sequence, except the terminal base-pair is changed to be 5'-G|C-3' (e.g., 5'-GTCGAACGTT TTCGTTCGAC-3'). Each hairpin was designed to adopt one base-paired structure.

NTD Y-transducer/Reporter probe

The Y-shaped NTD-transducer molecule design used in the SNP experiments [5,9], and described in the Discussion, has a three-way DNA nexus geometry: 5'-CTCCGTCGAC GAGTTTATAGAC TTTT GTCTATAAACTC GCAGTCATGC TTTT GCATGACTGC GTCGACGGAG-3'. Two of the junctions' arms terminate in a 4T-loop and the remaining arm, of length 10 base-pairs, is usually designed to be blunt ended. The blunt ended arm, or 'stem', has been designed such that when it is captured by the nanopore it produces a toggling blockade. Variants of the Y-transducer sequence are indicated in the figures for the Y-transducer annealing experiments.

Biotinylated DNA probes (from IDT DNA, purification by PAGE)

8GC-BiodT: 5'- GTCGAACGTT/iBiodT/TTCGTTCGAC -3'
9GC-BiodT: 5'- GTTCGAACGTT/iBiodT/TTCGTTCGAAC -3'

Biotinylated LNA/DNA Chimeric probes (from Exiqon, purification by HPLC)

8GC-BiodT: 5'- +G+TCGAA+C+GTT/iBiodT/TT+CGT+T+CG+AC -3'. The LNA version of 8GC-Bt has 8 LNA bases shown preceded by '+', 12 DNA bases, and 1 biotin dT base.
9GC-BiodT: 5'- +G+CTTGAA+C+GT/iBiodT/TT+CGTT+CAA+GC -3'. The LNA version of 9GC-Bt stem does not have the same exactly the sequence as the DNA-based 9GC, and has only a 3dT loop aside from the modified dT with biotin attachment, and has 7 LNA bases shown preceded by '+', 14 DNA bases, and 1 biotin dT base.

Laser Trapping probes (from IDT DNA, purification by HPLC)

The 20bp hairpin with 4dT loop:

9GC-ext:
5’-GTTCGAACGGGTGAGGGCGCTTTTGCGCCCTCACCCGTTCGAAC -3’
The 20bp hairpin with 5dT loop , where the central loop dT was modified to have a linker to biotin: 9GC-BiodT-ext:
5’-GTTCGAACGGGTGAGGGCGCTT/iBiodT/TTGCGCCCTCACCCGTTCGAAC -3’

3.3 Conjugation to Magnetic Beads

The streptavidin-coated magnetic bead diameters were approximately 1 micron and the mass about 1 pg. Some of the bead preparations involved use of BSA buffer, which required tolerance of BSA at the nanopore detector. This was separately confirmed for the concentrations of interest, up to the level of 8mg/mL BSA, in preliminary tests for bead usage.

3.4 Laser Setup

Laser illumination provided by a Coherent Radius 635-25. Output power before fiber optic was 25mW at a wavelength of 635 nm. The beam was chopped at 4Hz. During laser excitation studies the Faraday cage was removed. Significant 60 Hz wall-power noise was not seen with case removed when there was no laser illumination, but with cage removed and under laser illumination 60Hz line noise could clearly be seen. The 60 Hz line noise was, thus, picked up at the laser’s power supply and transmitted via the laser excitation process into the detector environment as a separate modulatory source. After fiber optic, approximately 5-10mW illumination is focused into in an approximate 1mm illumination diameter was produced at the nanopore detector’s aperture.

3.5 Data acquisition and FSA-based Signal acquisition

Data is acquired and processed in two ways depending on the experimental objectives: (i) using commercial software from Axon Instruments (Redwood City, CA) to acquire data, where current was typically filtered at 50 kHz bandwidth using an analog low pass Bessel filter and recorded at 20 μs intervals using an Axopatch 200B amplifier (Axon Instruments, Foster City, CA) coupled to an Axon Digidata 1200 digitizer. Applied potential was 120 mV (*trans* side positive) unless otherwise noted. In some experiments, semi-automated analysis of transition level blockades, current, and duration were performed using Clampex (Axon Instruments, Foster City, CA). (ii) using LabView based experimental automation. In this case, ionic current was also acquired using an Axopatch 200B patch clamp amplifier (Axon Instruments, Foster City, CA), but it was then recorded using a NI-MIO-16E-4 National Instruments data acquisition card (National Instruments, Austin TX). In the LabView format, data was low-pass filtered by the amplifier unit at 50 kHz, and recorded at 20 μs intervals. Signal acquisition from the 20 μs sample stream was done using a Finite State Automaton (FSA) [5, 33].

3.6 Machine Learning based Signal Processing

A brief description of the ML processing and control methods, including HMM-based Signal Feature extraction, SVM-based classification, Pattern Recognition Informed (PRI) Sampling, is provided in the Suppl, and further implementation details can be found in [5].

4 Results

Results are shown in Sec. 4.1 for a LNA/DNA chimeric transducer with magnetic bead attachment that modulates channel current while both bound and unbound when exposed to laser-tweezer pulsing. This demonstrates a method for generic NTD engineering. In Sec. 4.2, results are given for NTD operation in the presence of standard cellular and blood serum interference agents, but at very high test concentrations, with applications to rapid viral testing (via pin-prick blood testing). In Sec. 4.3 the largely electrophoretic-based advantages to selecting molecules of interest, and avoiding interference molecules, is complemented by results demonstrating a filter method that is related to size exclusion chromatography.

4.1 Biotinylated 8 base-pair LNA hairpin binding experiments with streptavidin (a streptavidin biosensor)

The results of the LNA/DNA chimera based NTD transducer/reporter redesign are shown in a series of screen captures of representative blockade events. Automated signal analysis has been demonstrated in prior work [10] with similar DNA variants [33, 34] (see Suppl. Fig. 8), so this won't be repeated here.

The biotinylated 8 base-pair DNA hairpin (DNA 8GC-Bt, shown in Fig. 1) has lifetime (until melting and channel translocation event) about 6s on average, with a wide range of observations from a fraction of a second to 15s that is dependent on buffer, and temperature, etc. (consistent with early work on DNA hairpin gauges in the nanopore [34]). The biotinylated 8 base-pair LNA/DNA chimeric hairpin (LNA 8GC-Bt), on the other hand, has lifetime 12 minutes on average, ranging from about 3minutes to over 30 minutes for individual melting times. Similarly, 9 base-pair DNA hairpins have lifetimes going from about one minute with individual lifetimes from 2s to 120s. Compare this with LNA 9GC-Bt lifetimes that are typically greater than 60 minutes, even in 2M urea.

In Fig. 4 the nanopore detector software is set to only capture the first 5s of a blockade trace, then perform a polarity reversal to eject the captured analyte and proceed with a new capture.

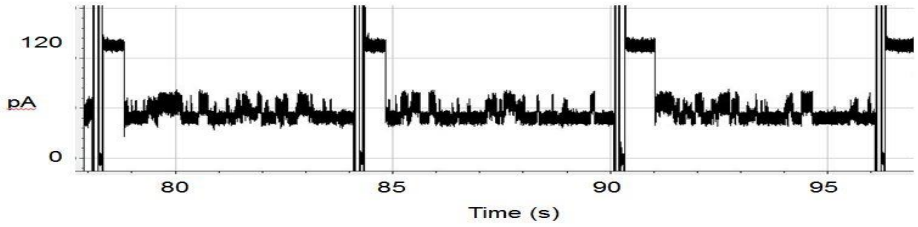


Figure 4. LNA 8GC-Bt blockade signals, 5s blockade before auto-eject (shown as the vertical current reset pulses that occur during the polarity reversal). Very little open channel (less than 1s at 120pA) occurs before the next capture event. The concentration of LNA in the detector well is 2uM. Concentration of 12nM in the detector well produces similar blockades, but with significantly greater (~200s) open channel time between blockade events.

In Fig. 5 streptavidin is added in a 1:1 ratio to the LNA 8GC-Bt already present (e.g., the streptavidin concentration in the detector well is 2uM). The timescale is longer, but the hold time for the hairpin blockades is still held at 5s when comparing to Fig. 4. The result shown is typical for the first 10 minutes after introduction of streptavidin. The blockade signal structure is unaltered from that shown in Fig. 4, it is merely compressed by the larger timescale shown. Note the much longer intervals of open channel even though the LNA concentration hasn't changed. This is due to the streptavidin binding some of the LNA and sequestering it in solution, leaving effectively lower concentration of LNA free to report to the channel detector. The signals produced will continue to change as more LNA is sequestered, and eventually bound streptavidin is pulled to the nanopore detector (to 'report'). Unbound streptavidin is almost never seen to interact with the channel. Streptavidin has pI 7-8, so this was initially thought to be due to it having a possibly positive charge in the pH 8 of the standard experimental buffer setting, but in studies at pH 9 there is still no streptavidin blockade signal even in mM concentrations. Basically, non-glycosylated proteins, even if very negatively charged at pH 8, such as albumin with pI 4.7, will not interact with the channel. Certain proteins are found to strongly interact, however, such as some classes of antibodies (even with pI 8.5 in pH8 buffer), but this is not described further here.

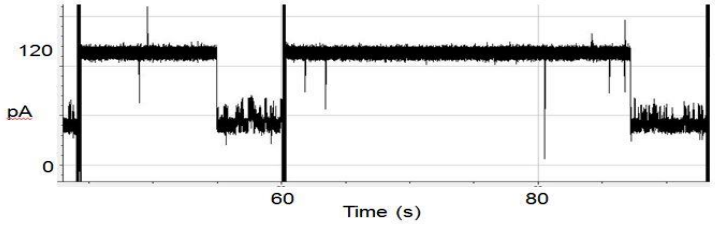


Figure 5. LNA 8GC-Bt blockade signals in the presence of streptavidin during the first 10 minutes after introduction of streptavidin. LNA and streptavidin are in a 1:1 ratio, with both at 2uM concentration in the detector well.

After another 10 minutes has passed since the introduction of streptavidin a new class of blockade begins to be seen (Fig. 6). The new class does not 'toggle' and is never seen (in runs with over 2000 LNA 8GC-Bt blockades) if streptavidin has not been added. After another 10 minutes has passed (30 minutes since the introduction of streptavidin) the free LNA sequestration is nearly complete (even though 1:1 streptavidin can bind up to 4 biotins). Fig. 7 shows one free LNA blockade (in middle), and two bound LNA blockades (one on either side).

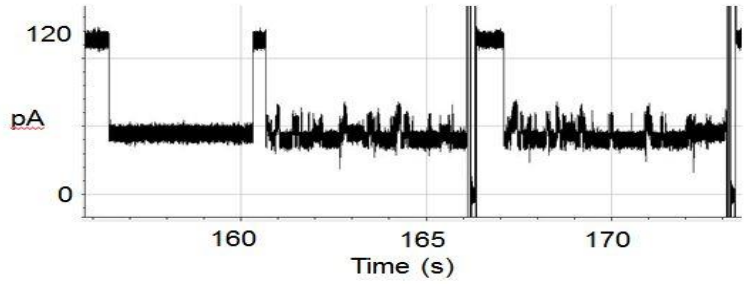


Figure 6. LNA 8GC-Bt blockade signals in the presence of streptavidin during the second 10 minutes after introduction of streptavidin. A bound reporter signal is shown as the leftmost blockade event.

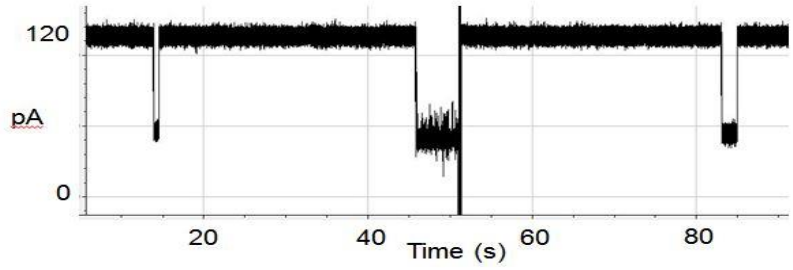


Figure 7. LNA 8GC-Bt blockade signals in the presence of 1:1 streptavidin after about 30 minutes of reaction time. The central blockade is an unbound reporter signal, the much shorter left and right blockades are bound reporter blockades.

After another 10 minutes has passed (roughly 40 minutes since the introduction of streptavidin) the free LNA sequestration is complete, free LNA will now be seen only rarely, with bound signal dominating (Fig. 8). Bound signal will now often be captured for sufficiently long that it reaches the 5s auto-eject time. This is likely because the captures will be dominated by streptavidin that is multiply-bound with biotinylated LNAs (providing an even greater pI shift than the singly bound streptavidin, thereby dominating the blockade events seen, and more strongly electrophoretically held at the channel). At later times and at the larger timescales (2.5 minutes shown in Fig. 8) 'melted' ssLNA translocation events are seen as short blockade 'spikes'.

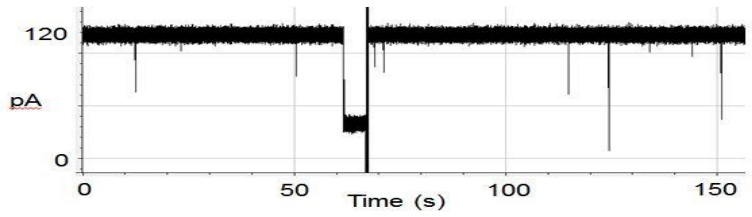


Figure 8. Streptavidin bound LNA 8GC-Bt blockade signals after about 40 minutes of reaction time.

4.2 Interference Testing

The electrophoretic part of the NTD detector provides a huge advantage when dealing with possible contaminants. Electrophoresis drives strong negative charges to the nanopore detector during normal operation. Nucleic acids in particular will be separated and driven to the detector, along with certain proteins and other molecules that have a low pI. Most proteins with low pI are found to have very little interaction with the nanopore channel, as already mentioned, the main exception being antibodies. Consider the common level of interference agents used to demonstrate robust medical testing applications (see Table 1). Actual levels of interference agents seen in (healthy) human blood samples are far lower (see Table 2). Consider working with a 1uL sample (such as with a pinprick sample) that contains high levels of common interference agents from blood, or other biological sources, Table 3 shows the very high contaminant levels that have been tested on the NTD with very low concentrations of reporter molecule. Most interference agents pose little channel interaction and the occasional channel blockade that does occur is short and non-modulatory. As a group antibodies are the exception, where a single monoclonal antibody (mAb) is found to produce a variety of distinct channel modulation signals types (see Suppl. Sec. S.3 for details).

Bilirubin:	10mg/dL = 0.10mg/mL
Cholesterol:	800 mg/dL = 8.00 mg/mL
Hemoglobin:	250mg/dL = 2.50 mg/mL
Triglyceride:	500mg/dL = 5 mg/mL

Table 1. Common level of interference agents used to demonstrate robust medical testing applications.

Bilirubin	5mg/L (10uM)
Cholesterol (healthy)	< 2mg/mL (5mM)
Hemoglobin in plasma	2mg/dL = 0.02mg/mL (300nM)
Hemoglobin in whole blood	150mg/mL(2.5mM)
Triglyceride	1g/L (1mM)
Serum DNA (no cell ruptures)	1-200ng/ml

Albumin	35-50 g/L (600uM)
Immunoglobulin G (IgG)	15mg/mL (at 160kDa → 93.75nmol/mL)
Urea	15 mg/dL (3mM)
Glucose (fasting)	100 mg/dL (5mM)

Table 2. Actual levels of interference agents seen in (healthy) human blood samples.

Cholesterol (healthy)	8mg/mL > 2mg/mL
Hemoglobin	4mg/mL > 2.5mg/mL
Immunoglobulin G (IgG)	30mg/mL > 15mg/mL
Urea	> 5M >> 3mM
Glucose	>> 50mM > 5mM

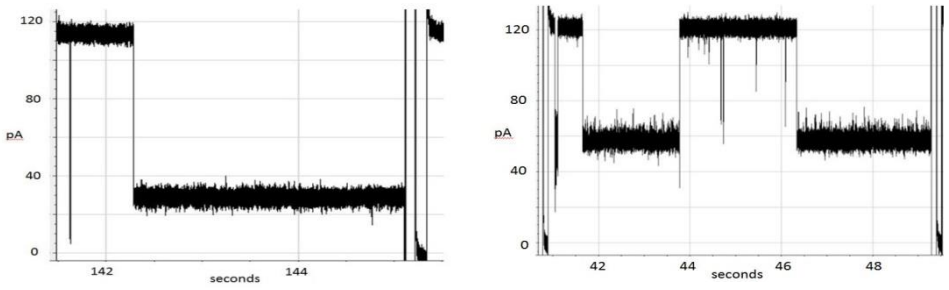
Table 3. Contaminant levels that have been tested where reporter molecules are easily discerned.

In studies with interference on the control 9GC molecule it is found that 1uL of 0.7nM 9GC can easily be seen in the detector (that has 70uL wells) in presence of 1uL of 1uM 7GC (approximately a 1:1000 ratio of 9GC to 7GC but easily discerned due to the distinctive channel modulation of the 9GC molecule). If analyzing the trace amounts of DNA present in blood serum (such as for early fungal pathogen identification), suppose 10ng/mL of total DNA is present of which 1/1000 is due to fungal pathogen. If the fungal pathogen is 'reported' by a modified form of the 9GC molecule (or a Y-transducer) then it is necessary to 'see' 1/1000 of 10ng/ml 9GC at the detector. Since 10ng/mL concentration of 9GC is 1.5nM, and we can see even less, 0.7nM, when the rest of the serum DNA is interference (from accidental cell ruptures, etc.), then it is clear that we can detect on trace DNA targets. Interference from other biomolecules that have higher pI is handled much more easily: 1 uL of 0.7nM 9GC in the presence of 4mg/mL hemoglobin (Hb) is easily resolved. Hb has a pI = 6.87 (normal, sickle cell pI=7.09), so in the standard pH=8 buffer it is expected that some Hb should be delivered to the channel, but even when this occasionally occurs, it has no apparent interaction. This is in agreement with albumin interference results, where concentration = 8mg/mL, and with a pI of 4.7, it is expected that many of the albumin molecules should be delivered to the channel, but no significant channel blockade events or even brief 'noise-spike' blockades are seen (this is thought to be because albumin is not glycosylated). In practice, an albumin capture matrix could be used to prevent the normally high levels of blood albumin (the main protein in blood plasma) from entering the nanopore detector. This would not be to prevent interference with the channel detection per se, but to prevent bilayer interactions. Having entered the nanopore detector albumin can still potentially be blocked from bilayer interference by having a surface scaffolding on the bilayer from PEG linked albumin [17, 18].

Cholesterol acts similarly to albumin, where high concentrations are not found to have observable channel blockade effect. This is not to say that albumin and cholesterol have no effect whatsoever, they appear to have a beneficial effect at physiological or lower concentration via stabilizing the bi-layer against rupture and to overall reduced current leakage (membrane permeability), and result in a lower RMS noise to the overall single channel current (no cholesterol, typical channel current RMS noise is 1.32 pA; with cholesterol it drops to 1.02 pA). The suspected role of albumin in channel nucleation is also revealed in these studies as late channel additions (bad news for single channel experiments) are observed to occur with introduction of albumin. Bilirubin has similar isoelectric point to albumin and similar non-reactivity with the channel.

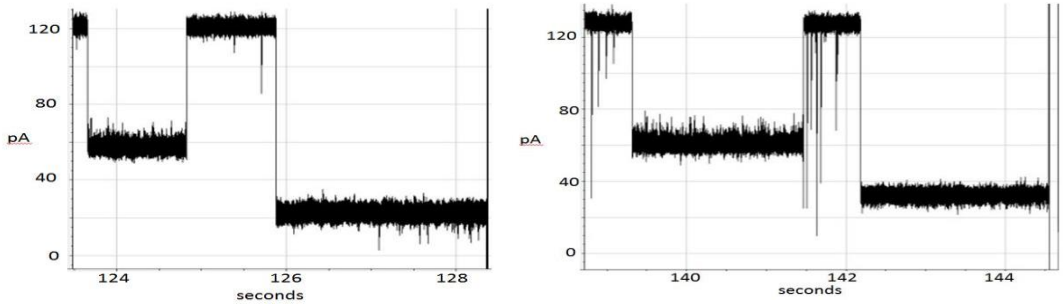
4.3 Polyethylene glycol (PEG) for size exclusion chromatography and filtering

Introducing PEG into the buffer reveals strong size-exclusion chromatography fractionation effects, allowing species to be computationally grouped according to their PEG shift measurements then presented as an ordered 'computational gel-separated' list of species (affording gel-separation and blot-identification entirely on the NTD apparatus). In the results shown in Fig.s 9 & 10 we see representative channel blockades for two types of DNA hairpins (Fig. 9), each with 4dT loops capping one end, one with seven base-pair stem (7CG molecule in Methods), and one with a twelve base-pair stem (12CG molecule in the Methods). Fig. 10 shows observations on mixtures of 7CG and 12CG before and after addition of PEG. The PEG-shift in this instance should see a shift in channel events to favoring more channel events with the larger nucleic acid, 12CG over 7CG in these experiments. Before addition of PEG hundreds of 7CG and 12CG events were observed with the ratio of 12CG to 7CG events: 0.82. After addition of PEG the ratio favors 12CG: 1.33. There are also more counts overall. So have the overall appearance of greater concentration of 12CG (roughly twice), when it should be halved by the removal of volume to accommodate the dilute PEG solution addition. In other words, an effective ionic concentration increase due to the volume excluding effect of PEG on charged analytes, with increased volume exclusion effect on larger charged molecules like 12CG vs. 7CG.



A B

Figure 9. DNA hairpin blockade signals before addition of PEG. (A) 12CG blockade; (B) 7CG blockades.



A B

Figure 10. 7CG and 12CG DNA hairpin mixture blockade signals before and after addition of PEG. (A) Before. **(B)** After.

5 Discussion

5.1 Antibodies as easily identifiable interference or transducer

Some mAb blockades produce a very clean toggling between two levels (see Suppl. Fig.'s 4-7 for antibody description, and some typical blockade signals). The mAb interference modulatory signals are easily discerned from a modulatory signal of interest, however, especially with increased observation time as needed. Aside from being an interference agent, antibodies offer a direct means for having a NTD transducer since their modulatory blockade signals are observed to change upon introduction of antigen. The problem with using an antibody directly as a transducer in a biosensor arrangement is that the antibody produces multiple blockade signal types (a dozen or more) just by itself (without binding). This weakness for use directly as a biosensor (they can still be linked indirectly as in [9]) is because the antibody is a glycoprotein that has numerous heterogeneous glycosylations and glycations, with many molecular side-groups that might be captured by the nanopore detector to produce modulatory blockades. If the purpose is to study the post-translational modifications (PTMs) themselves, a glyco-profile of the antibody in other words, then the numerous signal types seen are precisely the information desired. A more complete analysis of antibody blockades on the nanopore detector is beyond the scope of this paper, and will be in a separate paper. Some further details on the Antibody structure and its direct glyco-profiling is still given next, however, since similar PTMs can be analyzed on other proteins of critical biomedical interest.

Further details on antibodies pertinent to nanopore detection, including diagrammatic figures (Suppl. Fig.s 5-7), and structural details from various references [25-32] are placed in the Suppl.

5.2 NTD Transducer Design

The *bound* state of the transducer/reporter molecule is sometimes found to not transduce to a different toggling ionic current flow blockade, but to a fixed-level blockade (i.e., the transducer provides distinctive channel modulation when unbound, but not so distinctive fixed-level channel blockades when bound [5]).

It is important for *both* the bound and unbound transducers to have distinctive channel modulations in order to have automated high-precision state identification and tracking (and allow for multiplex assaying). In this instance, the switch to a fixed-level blockade was thought to be an effect of the large electrophoretically held complex forcing the channel-captured end to reside in one blockade state. This was previously explored in experiments where a streptavidin-coated magnetic bead was attached to biotinylated DNA hairpins known to be good modulators or poor channel modulators [4]. Once a streptavidin coated magnetic bead was attached to the biotinylated hairpins, it was found that gently pulsing the nanopore channel environment with a chopped laser beam (a laser-tweezer tugging) allowed a distinctive channel modulation to result (see Fig. 11). It was found more recently that the induced blockade modulations occur in two types (described in detail in [3] for chaotrope induced, see Fig. 3; and in [4], for early laser-tweezer induced results. Further laser tweezer results showing the different, overlapping, modes will be given in the Results, where the experiments are performed with a DNA-hairpin transducer as in previous studies. In terms of the convenient Y-transducer, however, the same could be done by simply making use of the unused arm, as shown in Fig. 12 (further discussion in Suppl. Sec. S.5 and [6]).

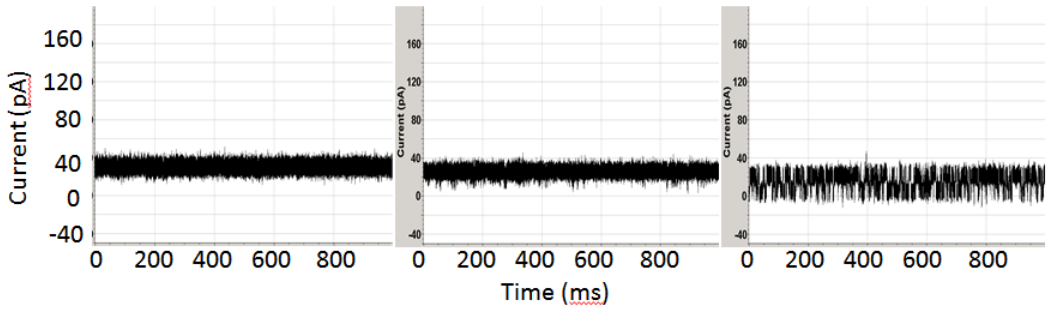


Figure 11. A (Left) Channel current blockade signal where the blockade is produced by 9GC DNA hairpin with 20 bp stem. Reprinted with permission [4]. (Center) Channel current blockade signal where the blockade is produced by 9GC 20 bp stem with magnetic bead attached. (Right) Channel current blockade signal where the blockade is produced by c9GC 20 bp stem with magnetic bead attached and driven by a laser beam chopped at 4 Hz. Each graph shows the level of current in picoamps over time in milliseconds.

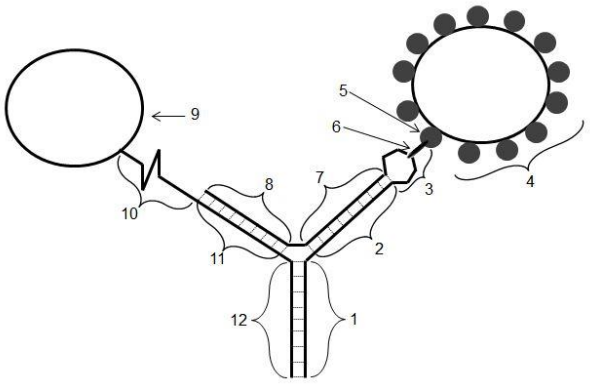


Figure 12. Y-laser transducer for high-specificity binding detection or individual protein binding & conformational change study. Reprinted with permission of [6]. The Y-transducer is meant to have a study molecule, region 9, attached by a single stranded nucleic acid linker, region 10, that is possibly abasic (non-base-pairing), that is linked to a single stranded nucleic acid region, region 11 & 12, that is meant to anneal to a second nucleic acid to create the Y-shaped nucleic acid construct shown.

In Fig. 12 the annealed Y-transducer is comprised of two, possibly LNA/RNA/DNA chimeric, nucleic acids, where the first single stranded nucleic acid is indicated by regions 1-3 and 7-8 and the second nucleic acid is indicated by regions 10-12. The paired regions {1,12}, {2,7}, and {8,11} are meant to be complements of one another (with standard Watson-Crick base-pairing), and designed such that the annealed Y-transducer molecule is meant to be dominated by one folding conformation (as shown). Region 3 is a biotin-modified thymidine loop, typically 4-5 dT in size (here 5dT shown with 2 dT, a biotinylated dT, then another 2 dTs), that is designed to be too large for entry and capture in the alpha-hemolysin channel, such that the annealed Y-transducer only has one orientation of capture in the nanopore detector (without bead, region 4, attached). Region 4 is a streptavidin coated magnetic bead (that is susceptible to laser-tweezer impulses). The base region, comprising regions {1,9}, is designed to form a duplex nucleic acid that produces a toggling blockade when captured in a nanopore detector. The typical length of the base-paired regions is usually 8, 9 or 10 base-pairs. The study molecule (region 9), an antibody for example, has linkage to single stranded nucleic acid via a commoditized process due to the immuno-PCR industry so is an inexpensive well-established manufacturing approach for the molecular construction. The Y-transducer on the left will not form if the ‘immuno-PCR tagged’ antibody is not present, which provides an additional level of event detection validation. If region 9 is a DNA enzyme that is processively acting on a DNA substrate this may provide a new means for nucleic acid sequencing.

NTD transducers are typically constructed by covalently linking a binding moiety of interest to a nanopore current modulator, where the modulator is designed to be electrophoretically drawn to the channel and partly captured, with its captured end distinctively modulating the flow of ions through the channel. Using inexpensive (commoditized) biomolecular components, such as DNA hairpins, this allows for a very versatile platform for biosensing, and given the high specificity high affinity binding possible, this also allows a very versatile platform for assaying at the single molecule level, even down to the single isoform level, e.g., molecular substructure profiling, such as glycosylation profiling. (Glycosylation profiling can also be done directly for some molecules that directly produce toggling blockades, antibodies in particular [5]. Glycosylation profiling is of critical importance in the development of the most effective antibody treatments [26-31].) Two complications with the transducer design, however, are (1) the convenient DNA-based modulators are often short-lived; and (2) the overall transducer's bound state often doesn't modulate. The first is shown to be solved using locked nucleic acid (LNA) nucleosides, the second is solved by introducing a third functionality for receiving laser-tweezer impulses by means of a covalently attached magnetic bead (another commoditized component). A description of the detector's robust performance in the presence of numerous interference agents with very low analyte concentration was also needed, and this is now much more clearly affirmed. LNA Y-transducers with magnetic bead attachment and laser pulsing gives rise to a generic modulator arrangement (see Fig. 12), that modulates even when bound, to allow NTD probing over long timescales on biological system components. An inexpensive commoditized pathway for constructing nanopore transducers is thereby obtained.

5.3 Biosystem Extra-Element Theorem (BEET)

In the electrical engineering setting the extra element theorem (EET) allows circuits without feedback to be understood in the presence of feedback by choosing the extra element to be the feedback element. In electrical engineering this gives rise to an updated, quantitative, solution. In the stochastic Biosystem Extra-Element Theorem (BEET) setting, feedback complexity can be handled similarly. The BEET method allows a balance to be struck between reductionist and holistic approaches. In this setting it is possible to work with the 'black box' giving rise to the emergent behavior and consider perturbations to that system. BEET also shows how to evolve to gene circuits with more components via a series of small (evolutionary) changes.

Using the NTD-method to perform analysis of "gene circuits" it is, thus, possible to have a 'voltmeter for the circuit' in a circuit analogy. The NTD-quantified gene-circuit analysis can then be enhanced with use of (BEET) method for analysis. In the NTD BEET setting, a collection of NTD reporter molecules with specific binding to different molecules can be used to perform multiplex analysis of the system molecular profile by differentiating the reporter molecules according to their different channel modulation signals. The NTD BEET system

could also employ multiple component modulation, and molecular knock-outs (by having strong binding) to effect double null injection to the equivalent gene circuit for a variety of extra element theorem testing procedures akin to their electrical engineering counterparts.

In the nucleic acid annealing studies on the NTD platform described in [7] (see Fig. 2), the critical role of chaotropes for robust nucleic acid annealing studies on the NTD platform was revealed. The ability of the NTD apparatus to tolerate high chaotrope concentration, up to 5M urea, was demonstrated more recently in [3], where the DNA hairpin control molecules demonstrated a manageable amount of isoform variation even at 5M urea (see Fig. 3). This allows a variety of annealing-based experiments to be robustly performed with nucleic acids, including miRNA binding site profiling in the presence of both known and unknown miRNA molecules, with or without complexation with argonaute proteins that occur in the RISC complex (but this would part of a destructive assaying application of the NTD, not the non-destructive system analysis mainly being discussed).

Preliminary work examining TBP binding to TATA binding site sequences placed in one arm of the Y-transducer construct [18] suggest a similar construct could be employed for purposes of miRNA binding site validation. The Y-transducer for miRNA binding site profiling on mRNAs would take the hypothesized sequence of the miRNA binding region, typically from the mRNA's 3'UTR region, and incorporate it either into one arm of a Y-transducer, or incorporate it such that it crosses the Y-nexus, the latter case potentially offering the greatest sensitivity to binding events, as was seen in the Y-SNP construct described previously. The latter case may not allow sufficient steric freedom for miRNA binding, however, when complexed with argonaute protein, so the arm variant may still be necessary for analysis of some miRNAs. This approach to miRNA target validation also benefits from validation at the actual annealing step of the interaction, thereby accounting for possible modification to the miRNA such as may occur with adenosine deaminases, where adenosine deaminases that act on RNA catalyze the conversion of adenosine to inosine residues in some double-stranded RNA substrates. A subset of miRNAs have been found to have modulated processing efficiency when deaminated at particular residues [16], and this is now thought to impact a significant fraction of miRNAs.

The RNAi probe examination could also be reversed, where the miRNA is sought that is associated with a suspected miRNA binding site (such as when the 3'UTR motif has an anomalous rate of occurrence and is shared across homologous genes in multiple organisms). Software to perform the aforementioned motif analysis has been developed for when only (pre-genomic) EST data is available, but this discussion is outside the scope of this paper so won't be discussed further.

6 Conclusion

LNA/DNA chimeras are shown to allow a much more robust long-lived NTD reporter molecule. The engineered NTD transducer/reporter molecule, minimally, has two functions, specific-binding and channel-modulation, and in the general setting, a third function to receive excitations such that channel modulation can be induced for all states of the transducer whether bound or not. Results shown here introduce excitations using a magnetic bead attachment in the presence of a laser-tweezer pulsing. A simple NTD transducer design via LNA/DNA chimeras or via mAb selection is also described. Operability of the NTD platform over a wide range of chaotrope concentration and inexpensive LNA/DNA transducer design allows simple nucleic acid testing via the NTD platform for purposes of miRNA profiling or viral RNA monitoring. For protein or nondestructive system type monitoring, the LNA/DNA transducer chimeras offer the necessary long-lived reporting capability needed to have a biological gene circuit 'voltmeter'.

Acknowledgements. The author, SWH, would also like to thank META LOGOS Inc., for research support and a research license. (META LOGOS was co-founded by SWH in 2009.) The author would like to thank the Meta Logos nanopore technicians Eric Morales, Joshua Morrison, Evenie Horton, and early work by the nanopore technicians based out of the SWH Lab at Children's Hospital New Orleans: Amanda Alba and Andrew Duda.

References

[1] B. Sakmann, E. Neher, Patch clamp techniques for studying ionic channels in excitable membranes, *Annu. Rev. Physiol.*, **46** (1984), 455-472. https://doi.org/10.1146/annurev.ph.46.030184.002323

[2] The Nobel Prize in Physiology or Medicine 1991, Nobel Media AB. nobelprize.org

[3] S. Winters-Hilt, A. Stoyanov, Nanopore Event-Transduction Signal Stabilization for Wide pH Range under Extreme Chaotrope Conditions, *Molecules*, **21** (2016), no. 3, 346. https://doi.org/10.3390/molecules21030346

[4] S. Winters-Hilt, Nanopore Detector based analysis of single-molecule conformational kinetics and binding interactions, *BMC Bioinformatics*, **7** (2006), Suppl 2, S21. https://doi.org/10.1186/1471-2105-7-s2-s21

[5] S. Winters-Hilt, *Machine-Learning Based Sequence Analysis, Bioinformatics & Nanopore Transduction Detection*, 2011.

[6] S. Winters-Hilt, Nanopore Transducer Engineering and Design, Submitted March 2017.

[7] S. Winters-Hilt, E. Horton-Chao, E. Morales, The NTD Nanoscope: potential applications and implementations, *BMC Bioinformatics*, **12** (2011), Suppl 10, S21. https://doi.org/10.1186/1471-2105-12-s10-s21

[8] K. Thomson, I. Amin, E. Morales, S. Winters-Hilt, Preliminary nanopore cheminformatics analysis of aptamer-target binding strength. *BMC Bioinformatics*, **8** (2007), Suppl 7, S11. https://doi.org/10.1186/1471-2105-8-s7-s11

[9] S. Winters-Hilt, The α-Hemolysin nanopore transduction detector – single-molecule binding studies and immunological screening of antibodies and aptamers, *BMC Bioinformatics*, **8** (2007), Suppl 7, S9. https://doi.org/10.1186/1471-2105-8-s7-s9

[10] A.M. Eren, I. Amin, A. Alba, E. Morales, A. Stoyanov, and S. Winters-Hilt, Pattern Recognition Informed Feedback for Nanopore Detector Cheminformatics, *Adv. Exp. Med. Biol.*, **680** (2010), 99-108. https://doi.org/10.1007/978-1-4419-5913-3_12

[11] R.D. Middlebrook, Null Double Injection and the Extra Element Theorem, *IEEE Trans. Educ.*, **32** (1989), no. 3, 167-180. https://doi.org/10.1109/13.34149

[12] H. Zhang, Reversal of HIV-1 Latency with Anti-microRNA Inhibitors, *The International Journal of Biochemistry & Cell Biology*, **41** (2009), no. 3, 451–454. https://doi.org/10.1016/j.biocel.2008.07.016

[13] Ruqiang Liang, David J. Bates and Eugenia Wang, Epigenetic Control of MicroRNA Expression and Aging, *Current Genomics*, **10** (2009), no. 3, 184–193. https://doi.org/10.2174/138920209788185225

[14] I. Bantounas, L.A. Phylactou1 and J. Uney, RNA interference and the use of small interfering RNA to study gene function in mammalian systems, *J. Mol. Endocrinology*, **33** (2004), no. 3, 545-557. https://doi.org/10.1677/jme.1.01582

[15] Vikram Agarwal, George W. Bell, Jin-Wu Nam, David P. Bartel, Predicting effective microRNA target sites in mammalian mRNAs, *eLife*, **4** (2015), e05005. https://doi.org/10.7554/elife.05005

[16] W. Yang, T.P. Chendrimada, Q. Wang, M. Higuchi, P.H. Seeburg, R. Shiekhattar, K. Nishikura, Modulation of microRNA processing and expression through RNA editing by ADAR deaminases, *Nat. Struct. Mol. Biol.*, **13** (2005), 13-21. https://doi.org/10.1038/nsmb1041

[17] S. Winters-Hilt, Isomer-specific trace-level biosensing using a nanopore transduction detector and a simple process for engineering inexpensive biosensing, diagnostic, and therapeutic transducer molecules, Submitted July 2016.

[18] S. Winters-Hilt, Amanda Davis, Iftekhar Amin and Eric Morales, Nanopore current transduction analysis of protein binding to non-terminal and terminal DNA regions: analysis of transcription factor binding, retroviral DNA terminus dynamics, and retroviral integrase-DNA binding, *BMC Bioinformatics*, **8** (2007), Suppl 7, S10. https://doi.org/10.1186/1471-2105-8-s7-s10

[19] S. Winters-Hilt, Exploring protein conformation-binding relationships and antibody glyco-profiles using a nanopore transduction detector, Submitted June 2016.

[20] D. Xiang, S. Shigdar, G. Qiao, T. Wang, A.Z. Kouzani, S-F. Zhou et al., Nucleic Acid Aptamer-Guided Cancer Therapeutics and Diagnostics: the Next Generation of Cancer Medicine, *Theranostics*, **5** (2015), no. 1, 23-42. https://doi.org/10.7150/thno.10202

[21] K. Germer, M. Leonard, X. Zhang, RNA aptamers and their therapeutic and diagnostic applications, *Int. J. Biochem. Mol. Biol.*, **4** (2013), no. 1, 27-40.

[22] X. Ming, B. Laing, Bioconjugates for targeted delivery of therapeutic oligonucleotides, *Adv. Drug Deliv. Rev.*, **87** (2015), 81-89. https://doi.org/10.1016/j.addr.2015.02.002

[23] C. Meyer, U. Hahn and A. Rentmeister, Cell-Specific Aptamers as Emerging Therapeutics, *J. Nucl. Acids*, **2011** (2011), 1-18, 904750. https://doi.org/10.4061/2011/904750

[24] J.W. Lee, H.J. Kim & K. Heo, Therapeutic aptamers: developmental potential as anticancer drugs, *BMB Reports*, **48** (2015), no. 4, 234-237. https://doi.org/10.5483/bmbrep.2015.48.4.277

[25] B. Wagner, D.C. Miller, T.L. Lear and D.F. Antczak, The Complete Map of the Ig Heavy Chain Constant Gene Region Reveals Evidence for Seven IgG Isotypes and for IgD in the Horse, *J. Immunol.*, **173** (2004), 3230-3242. https://doi.org/10.4049/jimmunol.173.5.3230

[26] K-T. Shade and R.M. Anthony, Antibody Glycosylation and Inflammation, *Antibodies*, **2** (2013), 392-414. https://doi.org/10.3390/antib2030392

[27] S. Radaev and P.D. Sun, Recognition of IgG by Fcγ Receptor: The role of Fc glycosylation and the binding of peptide inhibitors, *J. of Biological Chemistry*, **276** (2001), no. 19, 16478–16483. https://doi.org/10.1074/jbc.m100351200

[28] S. Ha, Y. Ou, J. Vlasak, Y. Li, S. Wang, K. Vo et al., Isolation and Characterization of IgG1 with Asymmetrical Fc Glycosylation, *Glycobiology*, **21** (2011), no. 8, 1087-1096. https://doi.org/10.1093/glycob/cwr047

[29] G. Zauner, M.H.J. Selman, A. Bondt, Y. Rombouts, D. Blank, A.M. Deelder, et al., Glycoproteomic Analysis of Antibodies, *Mol. Cell. Proteomics*, **12** (2013), no. 4, 856-865. https://doi.org/10.1074/mcp.r112.026005

[30] Jerrard M. Hayes, Eoin F. J. Cosgrave, Weston B. Struwe, Mark Wormald, Gavin P. Davey, Roy Jefferis and Pauline M. Rudd, Glycosylation and Fc Receptors, Chapter in *Fc Receptors*, Vol. 382, 2014, 165-199. https://doi.org/10.1007/978-3-319-07911-0_8

[31] D. Fernandes, Demonstrating Comparability of Antibody Glycosylation during Biomanufacturing, *Eur. Biopharma Rev.*, (2005), 106 -110.

[32] S. Winters-Hilt, E. Morales, I. Amin, A. Stoyanov, Nanopore-based kinetics analysis of individual antibody-channel and antibody antigen interactions, *BMC Bioinformatics*, **8** (2007), Suppl 7, S20. https://doi.org/10.1186/1471-2105-8-s7-s20

[33] S. Winters-Hilt, W. Vercoutere, V. S. DeGuzman, D. Deamer, M. Akeson and D. Haussler, Highly Accurate Classification of Watson-Crick Base-Pairs on Termini of Single DNA Molecules, *Biophys J.*, **84** (2003), 967-976. https://doi.org/10.1016/s0006-3495(03)74913-3

[34] W. Vercoutere, S. Winters-Hilt, H. Olsen, D. Deamer, D. Haussler, and M. Akeson, Rapid Discrimination Among Individual DNA Molecules at Single Nucleotide Resolution Using an Ion Channel, *Nat Biotechnol*, **19** (2001), no. 3, 248-252. https://doi.org/10.1038/85696

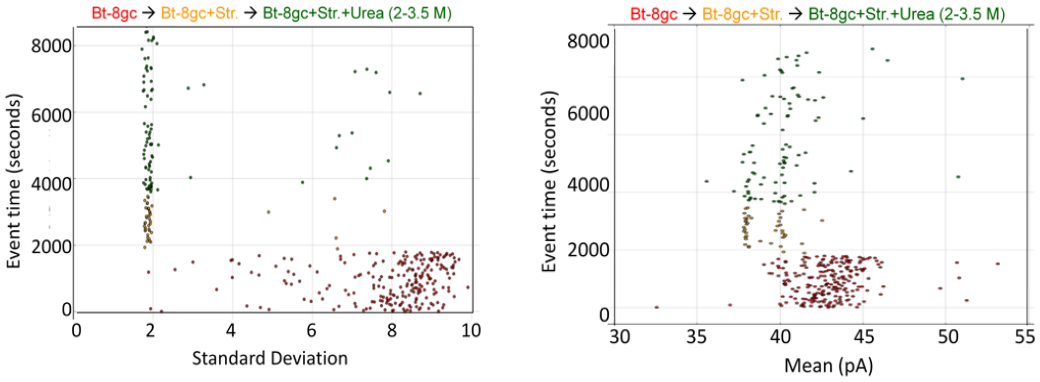
[35] Sonja A. Rasmussen, Denise J. Jamieson, Margaret A. Honein and Lyle R. Petersen, Zika Virus and Birth Defects — Reviewing the Evidence for Causality, *New England Journal of Medicine Special Report*, **374** (2016), 1981-1987. https://doi.org/10.1056/nejmsr1604338

[36] Van-Mai Cao-Lormeau, Alexandre Blake, Sandrine Mons, Stéphane Lastère, Claudine Roche, Jessica Vanhomwegen et al., Guillain-Barré Syndrome outbreak associated with Zika virus infection in French Polynesia: a case-control study, *Lancet*, **387** (2016), 1531–1539. https://doi.org/10.1016/s0140-6736(16)00562-6

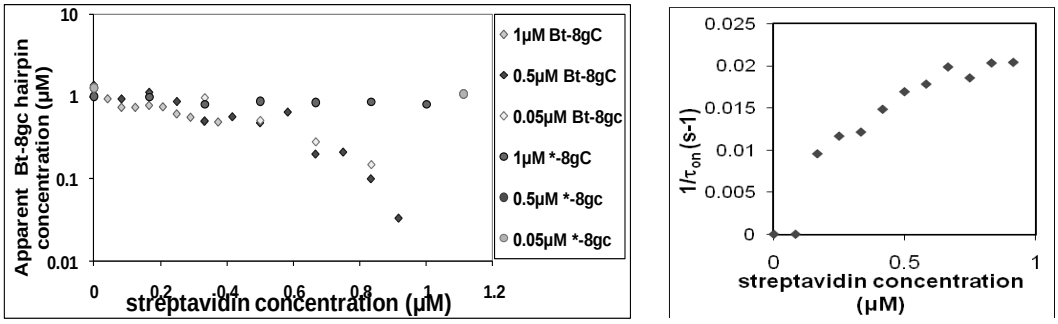
[37] Chi-Yu Xu, Hong-Mei Zhu, Jian-Hua Wu, Hai Wen and Cui-Jie Liu, Increased permeability of blood–brain barrier is mediated by serine protease during Cryptococcus meningitis, *J. of International Medical Research*, **42** (2014), no. 1, 85–92. https://doi.org/10.1177/0300060513504365

Supplement

S.1 Biotin-Streptavidin binding studies via nanopore event transduction



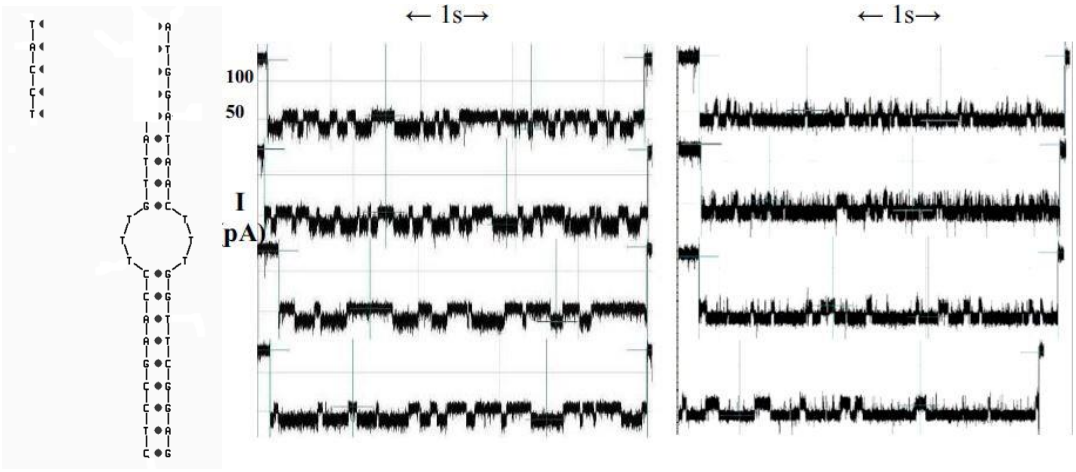
Suppl. Figure 1. Left. Observations of individual blockade events are shown in terms of their blockade standard deviation (x-axis) and labeled by their observation time (y-axis) Reprinted with permission of [7]. The standard deviation provides a good discriminatory parameter in this instance since the transducer molecules are engineered to have a notably higher standard deviation than typical noise or contaminant signals. At T=0 seconds, 1.0 μM Bt-8gc is introduced and event tracking is shown on the horizontal axis via the individual blockade standard deviation values about their means. At T=2000 seconds, 1.0 μM Streptavidin is introduced. Immediately thereafter, there is a shift in blockade signal classes observed to a quiescent blockade signal, as can be visually discerned. The new signal class is hypothesized to be due to (Streptavidin)-(Bt-8gc) bound-complex captures. **Right**. As with the Left Panel on the same data, a marked change in the Bt-8gc blockade observations is shown immediately upon introducing streptavidin at T=2000 seconds, but with the mean feature we clearly see two distinctive and equally frequented (racemic) event categories. Introduction of chaotropic agents degrades first one, then both, of the event categories, as 2.0 M urea is introduced at T=4000 seconds and steadily increased to 3.5 M urea at T=8100 seconds.



Suppl. Figure 2. Left. The apparent Bt-8gc concentration upon exposure to Streptavidin.

Reprinted with permission [7]. The vertical axis describes the counts on unbound Bt-8gc blockade events and the above-defined mapping to "apparent" concentration is used. In the dilution cases, a direct rescaling on the counts is done, to bring their "apparent" concentration to 1.0 μM concentration (i.e., the 0.5 μM concentration counts were multiplied by 2). For the control experiments with no biotin (denoted '*-8gc'), the *-8gc concentration shows no responsiveness to the streptavidin concentration. **Right. The increasing frequency of the blockades of a type associated with the streptavidin-Bt-8gc bound complex.** The background Bt-8gc concentration is 0.5 μM, and the lowest clearly discernible detection concentration is at 0.17 μM streptavidin.

S.2 DNA Annealing studies via nanopore event transduction

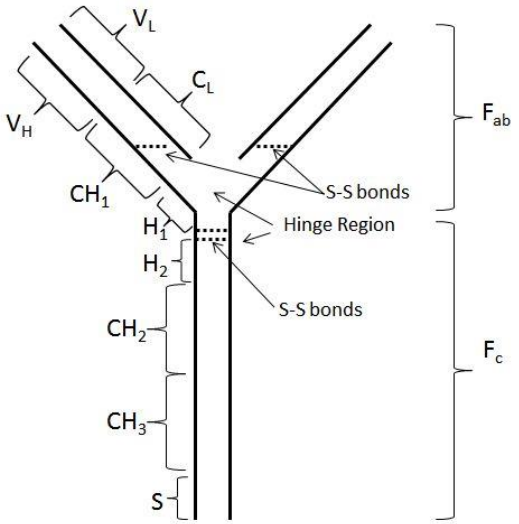


Suppl. Figure 3. Five-base annealing using a pseudo-aptamer NTD transducer. Reprinted with permission [8]. **Left:** The preliminary aptamer experiments are based on the DNA molecule obtained from annealing ssDNA1: 5'-CAAGCTTGGTTTCGATAGGTA-3' with ssDNA2: 5'-ATCGTTTCCAAGCTTG-3'.

For the pseudo-aptamer binding experiment a solution of annealed ssDNA1 and SSDNA2 molecules was exposed to ssDNA3: 5'-TACCT-3' (which anneals to the remaining AGGTA complement on ssDNA1). The target 5-base ssDNA is introduced subsequent to obtaining a toggler-type capture of the aptamer molecule (properly annealed). (The transducer is referred to as a pseudo-aptamer experiment due to its simplification to a DNA annealing based detection.) **Center:** A collection of toggle signals from the captured pseudo-aptamer. **Right:** A collection of toggle signals from the pseudo-aptamer soltion upon exposure to the ssDNA3 five-base target sequence. A distinctive blockade feature only observed in the blockade signals after ssDNA3 is introduced, aside from the level dwell-time changes, are the much higher frequency of upward "spike" transitions, from the lower level to the upper level.

S.3 Managing common interference agents and antibodies as easily identifiable interference or transducer

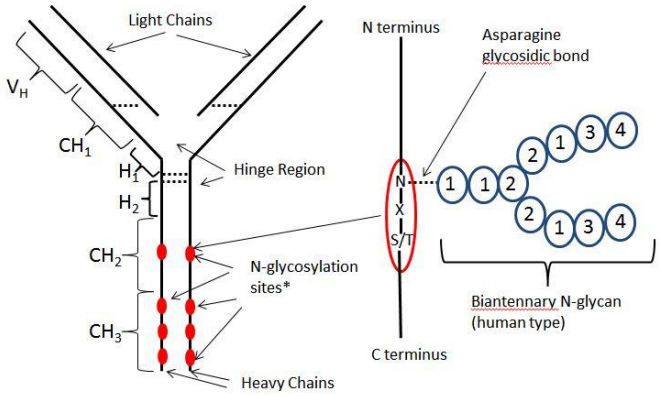
Antibodies are the secreted form of a B-cell receptor, where the difference between forms is in the C-terminus of the heavy chain region. Suppl Fig. 4 shows the standard antibody schematic. Standard notation is shown for the constant heavy chain sequence (‘CH’, ‘H’, and ‘S’ parts), variable heavy chain region (‘VH’ part), the variable light chain region (‘VL’ part), and constant light chain region (‘CL’ part). The equine IGHD gene for the constant portion of the heavy chain has exons corresponding with each of the sections CH1,H1,H2,CH2,CH3,CH4(S), and for the membrane-bound form of IGHD, there are two additional exons, M1 and M2 for the transmembrane part, thus, CH1, H1, H2, CH2, CH3, CH4(S), M1, M2 [25]. In Suppl. Fig. 4, the long and short chains are symmetric from left to right, their glycosylations, however, are generally not symmetric. Critical di-sulfide bonds are shown connecting between chains, each of the VH and CH regions typically have an internal disulfide bond as well. The lower portion of the antibody is water soluble and can be crystallized (denoted Fc). The upper portion of the antibody is the antigen binding part (denoted Fab).



Suppl. Figure 4. The standard antibody schematic. Reprinted with permission of [18]. Standard notation is shown for the constant heavy chain sequence ('CH', 'H', and 'S' parts), variable heavy chain region ('VH' part), the variable light chain region ('VL' part), and constant light chain region ('CL' part). The full heavy chain sequence is derived from recombination of the VH part and {CH,H,S} parts (where the secretory region S is also called CH4). The long and short chains are symmetric from left to right, their glycosylations, however, are generally not symmetric. Critical di-sulfide bonds are shown connecting between chains, each of the VH and CH regions typically have an internal disulfide bond as well. The lower portion of the antibody is water soluble and can be crystallized (denoted Fc). The upper portion of the antibody is the antigen binding part (denoted Fab).

Suppl. Fig. 5 shows a typical antibody N-glycosylation (exact example for equine IGHD [25]). One possible N-glycosylation site is indicated in region CH2, and three possible N-glycosylation sites are indicated in region CH3. N-glycosylation consists of a covalent bond (glycosidic) between a biantennary N-glycan (in humans) and asparagine (amino acid 'N', thus N-glycan). The covalent glycosidic bond is enzymatically established in one of the most complex post translational modifications on protein in the cell's ER and Golgi organelles, and usually only occurs in regions with sequence "NX(S/T) – C-terminus" where X is 'anything but proline' and the sequence is oriented with the C-terminus as shown. Licensed therapeutic antibodies typically display 32 types of biantennary N-glycans [26-31], consisting of N-acetyl-glucosamine residues (GlcNAc, regions '1'); mannose residues (Man, regions '2'); galactose residues (Gal, regions '3'), and Sialic Acid Residues (NeuAc, regions '4'), as shown in Suppl. Fig. 5. The N-glycans are classified according to their degree of sialylation and number of galactose residues: if disialylated (shown) have A2 class. If asymmetric and monosialylated have A1 class. If not sialylated then neutral (N class). If two galactose residues (shown) then G2 class, if one, then G1 class, if zero, then G0

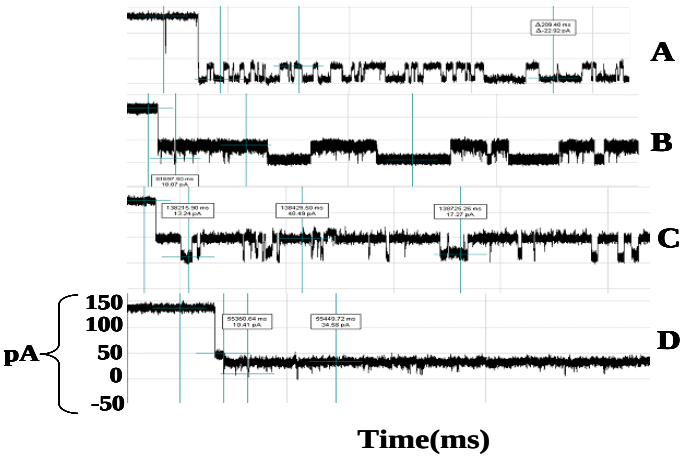
class. If there is an extra GlcNAc residue bisecting between the two antennae +Bi class (–Bi shown). If a core fucose is present (location near GlcNAc at base), then +F (–F shown). So the class shown is G2-A2. The breakdown on the 32 types is as follows: 4 G2-A2; 8 G2-A1; 4 G1-A1; 4 G2-A0; 7 G1-A0; 4 G0-A0 [31]. The N-glycans with significant acidity (A2 and A1) are 16 of the 32, so roughly half of the N-glycans enhance acidity. The other main glycosylation, involving O-glycans, occurs at serine or threonine (S/T). The main non-enzymatic glycations occur spontaneously at lysines ('K') in proteins in the blood stream upon exposure to glucose via the reversible Maillard reaction to form a Schiff Base (cross-linking and further reactions, however, are irreversible and associated with the aging process).



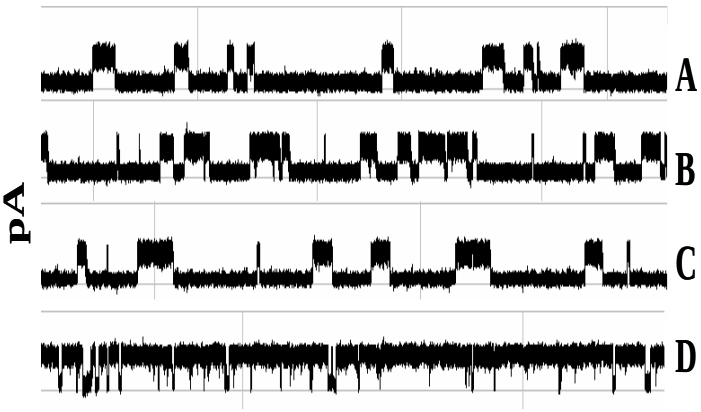
Suppl. Figure 5. Typical antibody N-glycosylation. Reprinted with permission of [18]. A schematic for typical antibody N-glycosylation is shown (drawn from results on the equine IGHD gene [22]), where one possible N-glycosylation site is indicated in region CH2, and three possible N-glycosylation sites are indicated in region CH3. N-glycosylation consists of a covalent bond (glycosidic) between a biantennary N-glycan (in humans) and asparagine (amino acid 'N', thus N-glycan). The covalent glycosidic bond is enzymatically established in one of the most complex post translational modifications on protein in the cell's ER and Golgi organelles, and usually only occurs in regions with sequence "NX(S/T) – C-terminus" where X is anything but proline and the sequence is oriented with the C-terminus as shown. Licensed therapeutic antibodies typically display 32 types of biantennary N-glycans, consisting of N-acetyl-glucosamine residues (GlcNAc, regions '1'); mannose residues (Man, regions '2'); galactose residues (Gal, regions '3'), and Sialic Acid Residues (NeuAc, regions '4'). The N-glycans are classified according to their degree of sialylation and number of galactose residues: if disialylated (shown) have A2 class. If asymmetric and monosialylated have A1 class. If not sialylated then neutral (N class). If two galactose residues (shown) then G2 class, if one, then G1 class, if zero, then G0

class. If there is an extra GlcNAc residue bisecting between the two antennae +Bi class (–Bi shown). If a core fucose is present (location near GlcNAc at base), then +F (–F shown). So the class shown is G2-A2. The breakdown on the 32 types is as follows: 4 G2-A2; 8 G2-A1; 4 G1-A1; 4 G2-A0; 7 G1-A0; 4 G0-A0. The N-glycans with significant acidity (A2 and A1) are 16 of the 32, so roughly half of the N-glycans enhance acidity. The other main glycosylation, involving O-glycans, occurs at serine or threonine (S/T). The main non-enzymatic glycations occur spontaneously at lysines ('K') in proteins in the blood stream upon exposure to glucose via the reversible Maillard reaction to form a Schiff Base (cross-linking and further reactions can be irreversible).

The base of the antibody plays the key role in modulating immune cell activity. The base is called the Fc region for 'fragment, crystallizable', which is the case, and to differentiate it from the Fab region for 'fragment, antigen-binding' that is found in each of the arms of the Y-shaped antibody molecule (see Suppl. Fig. 4). The Fc region triggers an appropriate immune response for a given antigen (bound by the Fab region). The Fab region gives the antibody its antigen specificity; the Fc region gives the antibody its class effect. IgG and IgA Fc regions can bind to receptors on neutrophils and macrophages to connect antigen with phagocyte, known as opsonization (opsonins attach antigens to phagocytes). This key detail may explain the modulatory antibody interaction with the nanopore channel. IgG, IgA, and IgM can also activate complement pathways whereby C3b and C4b can act as the desired opsonins. The C-termini and Fc glycosylations of an antibody's heavy chain, especially for IgG, is thus a highly selected construct that appears to be what is recognized by immune receptors, and is evidently what is recognized as distinct channel modulator signals in the case of the NTD (mAb channel blockade signals are shown in Suppl. Fig.s 6 & 7). Using NTD we can co-opt the opsonization receptor-binding role of the Fc glycosylations (and mAB glycations and glycosylations in general), and C-terminus region, to be a channel modulating role. This may also permit a new manner of study of the critical opsonization role of certain classes of antibodies (and possibly differentiate the classes in more refined ways) by use of the nanopore detector platform. The channel may provide a means to directly measure and characterize antibody Fc glycosylations, a critical quality control needed in antibody therapeutics to have correct human-type glycosylation profiles in order to not (prematurely) evoke an immunogenic response.



Suppl. Figure 6. Multiple Antibody Blockade Signal Classes (1s traces). Reprinted with permission [4]. Examples of the various IgG region captures and their associated toggle signals: the four most common blockade signals produced upon introduction of a mAb to the nanopore detector's analyte chamber (the *cis*-channel side, typically with negative electrode). Other signal blockades are observed as well, but less frequently or rarely.



Suppl. Figure 7. Antibody-Antigen binding – clear example from specific capture orientation. Reprinted with permission [4]. Each trace shows the first 750 ms of a three minute recording, beginning with the blockade signal by an antibody molecule that has inserted (some portion) into the Alpha-hemolysin channel to produce a toggle signal (A). Antigen is introduced at the beginning of frame A (100 μg/ml of 200 kD multivalent synthetic polypeptide (Y,E)-A—K).

A preliminary description of antibody blockade studies on the nanopore detector, for a well defined synthetic polypeptide antigen, are given in [9, 32]. Similar results are found IgG subclass 1 monoclonal antibodies for biotin, HIV, and anti-GFP, and all have produced similar signals. The critical role of Fc glycosylation has already been mentioned, but there is also the critical role in understanding antigen-antibody binding in the Fab region. As elaborated on in

[9,32], hydrophobic bonds are very difficult to characterize by existing crystallographic and other means, and often contribute half of the overall binding strength of the antigen-antibody bond. Hydrophobic groups of the biomolecules exclude water while forming lock and key complementary shapes. The importance of the hydrophobic bonds in protein-protein interactions, and of critically placed waters of hydration, and the complex conformational negotiation whereby they are established, may be accessible to direct study using nanopore detection methods. Further work on antibody studies are beyond the scope of this paper, however, and will be presented elsewhere.

S.4 Viral monitoring via targeted nucleic acid assaying

The explosive geographic expansion of the Zika virus provides another reminder that rapid diagnostic tools for new viral infections is an ever increasing need. The rapid deployment of a fast diagnostic tool in the example of the Zika virus is all the more pertinent given that the virus has been shown to be the cause of microcephaly in the fetuses of exposed pregnant women [35], along with results indicating possible brain damage (Guillain-Barre reaction) to a significant fraction of those exposed [36]. A rapid development, deployment, and evaluation of a Zika virus diagnostic would afford the patient the critical time needed to undergo aggressive prophylactic measures. Similarly, certain fungal infections need to be diagnosed as early as possible (cryptococcus neoformans, for example, can disrupt and cross the blood-brain barrier [37]). The treatments for many fungal infections are highly toxic, however, such that they will only be undertaken if infection is highly likely.

Pathogens that are suspected can potentially be probed in a matter of hours using an NTD platform with the methods described here using probes designed according to the pathogen's genomic profile. Unknown pathogens would first need to either have their genomes sequenced (less than a day) if sufficient DNA already available, or a sample directly measured via a test assay template (same procedure as for biomarker discovery) for assay-level fingerprint determination, then testing for that pathogen fingerprint in the patient.
The NTD platform can be enhanced to be a rapid annealing-based detection platform due a recently established ability [3] to operate under high chaotropic conditions (up to 5M urea), which allows measurement of collective binding interactions such as nucleic acid annealing with other simpler binding and related complexes thereby eliminated and effectively filtered from the analysis task. What remains to be done is to establish a general production method for creating a NTD transducer for the sequence of interest, and this is described in [6].

S.5 Engineering or Selecting NTD Transducer/Reporter molecules and use of laser excitation

A related complication with using DNA-based channel modulators has been their short lifetimes until melting. This problem has been eliminated by use of locked nucleic acid nucleosides (LNAs), as shown in the Resultse. LNAs serve to reduce

twist modes by locking the nucleic acid and thereby restricting its internal degrees of freedom in term of twist/stretch. This can be a good thing in that it will simplify the SCW signal training mentioned above. A simpler SCW analysis is not critical, however, so the main optimization to be accomplished by 'locking up' the modulator with increased LNA is effectively a tuning over molecular variants with greater or lesser twist mode event transmission. For annealing-based detection this is a big deal since the properly annealed nucleic acid duplex will transmit twist mode excitations notably differently than improperly annealed DNA (if even present). For this reason some modulator arrangements with laser-tweezer pulsing may have their bead attachment on the same arm as the annealing binding site (further details to follow), and have a low number of LNA bases in the LNA/DNA chimeras in the binding template (keeping blunt terminus and Y-nexus regions strongly LNA based to prevent melting as much as possible, but permitting twisting). Further discussion and transducer design in [6].

S.6 Machine Learning based Signal Processing

S.6.1 HMM-based Signal Feature extraction

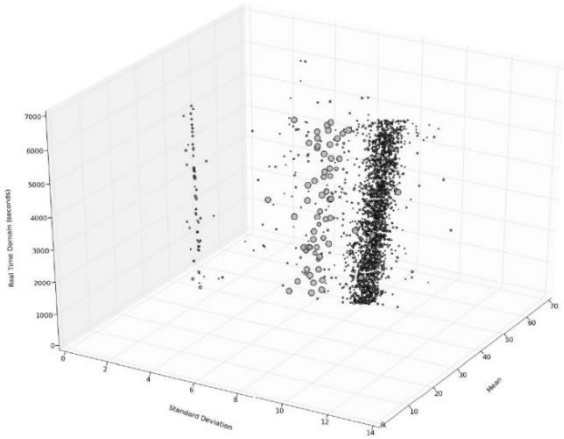
With completion of FSA preprocessing, an HMM is used to remove noise from the acquired signals, and to extract features from them. The HMM in one configuration (for control probe validation) is implemented with fifty states, corresponding to current blockades in 1% increments ranging from 20% residual current to 69% residual current [5,33]. The HMM states, numbered 0 to 49, corresponded to the 50 different current blockade levels in the sequences that are processed. The state emission parameters of the HMM are initially set so that the state j, $0 <= j <= 49$ corresponding to level $L = j+20$, can emit all possible levels, with the probability distribution over emitted levels set to a discretized Gaussian with mean L and unit variance. All transitions between states are possible, and initially are equally likely. Each blockade signature is de-noised by 5 rounds of Expectation- Maximization (EM) training on the parameters of the HMM. After the EM iterations, 150 parameters are extracted from the HMM. The 150 feature vectors obtained from the 50- state HMM-EM/Viterbi implementation are: the 50 dwell percentage in the different blockade levels (from the Viterbi trace-back states), the 50 variances of the emission probability distributions associated with the different states, and the 50 merged transition probabilities from the primary and secondary blockade occupation levels (fits to two-state dominant modulatory blockade signals). Variations on the HMM 50 state implementation are made as necessary to encompass the signal classes under study.

S.6.2 SVM-based classification

The 150-component feature vector extracted for each blockade signal is then classified using a trained Support Vector Machine (SVM). The SVM training is done off-line using data acquired with only one type of molecule present for the training data (bag learning). Further details on the SVM and overall channel current cheminformatics signal processing are detailed in [5, 33].

S.6.3 Pattern Recognition Informed (PRI) Sampling

For experiments with PRI sampling, a channel blockade capture signal is filtered and amplified before it is sent through the DAQ. The first 200 ms detected after drop from baseline are sent via TCP-IP protocol to the HMM software, which generates a profile for each signal sent. The HMM-generated profile is processed with the SVM classifier to determine whether the signal is acceptable. If the signal is acceptable, the message to continue recording is sent to the LabWindows software to continue recording, and the molecule is not ejected from the channel by the amplifier. If not, the amplifier briefly reverses the polarity to eject the molecule from the channel. The nanopore experiments with PRI sampling described in [15] are done with a 1:70 mixture of 9GC:9TA. In Suppl. Figure 8 the PRI sampling acquisition results are shown, with the rarer 9GC molecules properly identified, and sampled for a full 5 second duration, while others molecules are rejected, typically in 100 ms (with the prototype network setup used here).



Suppl. Figure 8. Standard deviation vs. Mean vs event-observation time vs PRI-informed sample observation time (4th dimension represented as the radius of the data point). Reprinted with permission [10]. This figure shows a successful real-time operation on the PRI-sampling method on the ND platform. 9GC signal is selected for observation and it is at a 1:70 lower concentration than the decoy 9TA DNA hairpins. As can be seen, only 9GC signals are held for the lengthier observation time, all other molecules being rejected promptly upon identification (the smaller diameter events points correspond to short lived events), where the brief duration of the event is dictated by the active, PRI-control, of the device voltage.

In the interference Results (Sec. 4.2) we see further evidence of the robustness of the channel when exposed to interference agents. Some agents if present with sufficiently high concentration, however, can damage the bilayer. Albumin is an example of such and it is the main protein found in blood samples. Albumin can intercalate into the bilayer (cholesterol also) and initially this strengthens the bilayer and lowers the system RMS current noise, but eventually there is too much

of a good thing and the albumin is probably agglomerating and causing bilayer disruption, which, in turn, can compromise the entire experiment. There are a variety of buffer modifications that can be introduced that are protective of the bilayer, including blocking the albumin intercalation. In doing so, however, new interference molecules are introduced that can damage the channel. It is observed, however, that the new interference problem is only a problem if the protocol is non-responsive, e.g. if the blockade is not recognized as a 'bad' blockade and ejected promptly (if not ejected promptly the molecule gets 'stuck'). What is needed is an auto-eject cycle for whatever minimal observation time is needed per blockade in the experiment of interest, to minimize channel blockade time regardless (we will see this feature used in the Results that follow). What is also needed is good and bad signal recognition. Generally any signal that is modulating is good, so if all signals are rejected if non-modulatory in their first 0.5 seconds is a pretty good operational setting. The PRI sampling can thus be employed, indirectly, to provide channel protection and maintain operational status for prolonged periods.

S.7 Possible NTD Nanoscope applications

The nanopore transduction detector (NTD) offers a means to examine the binding and conformational changes of individual biomolecules in a non-destructive manner that is well-suited to non-destructive analysis of biomolecular systems. The critical choice of transducer in system biology NTD applications is for one with very high specificity but that is only weakly binding so as not to be disruptive to the biological system or gene circuit. It may also be possible to use the NTD method in live cell assays as well, via use of laser modulations, not for fluorophore excitation, however, but for noise state excitation for use by the NTD where the need to generate a steady channel *current* is avoided in detector operation (which would be destructive to the cell). The NTD method is typically based on a single protein-channel biosensor used with a patch clamp amplifier on a (synthetic cell membrane) lipid bilayer. In the live cell assay the patch clamp application would return to its origins, where it was developed for patch clamp measurements of currents and current gating through channels on live cells. In order for the NTD 'voltmeter' to operate on the biological system to work, however, the normal operational buffer of the NTD must also accommodate a change to the physiological or cellular buffer environment of the biological system of interest. Recent work with robust NTD operation with a variety of buffer pH and in the presence of high concentrations of interference agents [3] reveals that this capability has been achieved.

In addition to the study of DNA, DNA-DNA interactions, and DNA-Protein interactions, the nanopore experimental setup has significant potential *vis-à-vis* the study of protein-protein interactions on the single molecule level. DNA-protein and protein-protein interactions are an integral component of gene-regulation and the cellular signaling apparatus. Cell signaling networks, gene regulation, and pathogen-induced genomic or transcriptome modifications,

are areas of intense current study since they are the basis for many disease states (ranging from metabolic disease, to cancer to autoimmunity). Fundamentally, the scientific benefits to molecular biology and a number of other fields (nanobiotechnology) are significantly impacted if nanopore detection methods can be utilized successfully in the system biology setting.

While cell biological, genetic, and structural biological approaches have contributed significantly to our understanding of signaling networks, we still do not have a clear understanding of the how these networks are regulated because of their inherent complexity. System wide approaches (yeast two-hybrid screens, bioinformatics approaches, for example) have emerged as powerful tools to map topologies of these signaling networks, but, unfortunately, are unable to tell us much about the nature of the links between individual nodes (activities). A complete understanding, therefore, requires that attention be paid to the single-molecule biochemistry and biophysics of the individual interacting species.

S.7.1 NTD application in live cell assays

The NTD methods proposed are compatible with using the NTD method in live cell assay settings as well, with use of laser modulations for noise state excitation for use by the NTD. The NTD method is typically based on a single protein-channel biosensor implemented on a (synthetic cell membrane) lipid bilayer, but in the live cell assay it would be based on patch clamp measurements of current through a channel on a live cell. Measurement of single channels at the cellular level has been done for 30 years [1], since the development of the patch clamp amplifier (that was originally designed for use in channel studies on single cells, for which the Nobel was obtained in 1991 [2]). The biosensor conformation used in the typical nanopore detector, however, is based on channel current blockades at discernibly different *levels*, which implies that there is at least one current that isn't zero, which is incompatible with using the standard cell patch clamp for channel biosensor applications (the cell would rupture). In the nanopore transducer setting, however, a minimal charge current could be used that could be non-destructive to the cell if periodically reversed, where most of the critical signal information would now reside in the noise profile (where the noise state would be driven by a laser-tweezer tugging at a covalently attached magnetic bead). The key signal analysis method to use in reading the changing noise states has already been developed [5], and involves a collection of machine learning based signal processing methods comprising the stochastic carrier wave (SCW) platform.

S.7.2 NTD application in programmable nanoblot

The NTD platform could be described as a programmable microarray. In essence, a programmable Southern Blot, Northern Blot, Western Blot, etc., is provided by the NTD given its direct computational coupling. Previous work with introducing PEG into the buffer also reveals strong size-exclusion chromatography fractionation effects, allowing species to be computationally grouped according to

their PEG shift measurements [5] then presented as an ordered 'computational gel-separated' list of species (affording gel-separation and blot-identification entirely on the NTD apparatus, when the destructive aspect of adding a bunch of PEG is permissible). A method and system for using the nanopore transduction detector (NTD) is, thus, described for examining the binding and conformation changes of individual biomolecules in a non-destructive manner, and by (destructive) assay methods, involving urea and PEG for example, that provides a general tool for analysis of biomolecular systems.

S.7.3 NTD in relation to other quantitative approaches, such as qPCR

The strengths of the device in terms of single molecule detection are also the weakness in the sense of event detection throughput. The previously mentioned PRI informed sampling can eliminate blocking conditions at the (single) channel detector akin to having a Maxwell Demon for purposes of single-molecule classification and rejection; such that a nearly optimal use of the single-channel's sequential sampling operation can be accomplished, but this only goes so far. An array of nanopore detectors would significantly resolve this problem, and such has been done by other researchers and companies to some extent. It is unclear if the nanopore detectors in an array configuration have the necessary bandwidth for observing channel transduction enhancements, however, so this is still a largely unexplored area.

The strengths of the NTD apparatus with stochastic carrier wave (SCW) event encoding are most evident when trying to have a discussion of noise problems. Transduced events have carrier-waves representations that are easily discerned under high noise conditions just as with any carrier-wave based communication scheme. It's as if an error-correcting encoding scheme is already built-in (that is realized using machine learning methods via an automated HMM feature extraction process and a SVM classification apparatus). Sensitivity and specificity for resolving highly similar control molecules is greater than 99.99% [5].

When the NTD transduction method makes use of transducer molecules that are DNA based and have annealing-based specific binding to a DNA target of interest things begin to sound like quantitative real-time PCR (qPCR). In qPCR the presence of a DNA molecule is revealed via a highly specific DNA probe annealing event where the DNA probe has a fluorophore attached that can be revealed under laser illumination at the appropriate frequency. Fluorophore excitation is a quantum statistical event at the single-fluorophore level, so this analysis is still typically a 'bulk', or aggregate, molecular analysis to some extent (although experiments to have truly single-molecule fluorescence have been performed, such as with FRET). The NTD probe comparison to qPCR probe would be much the same insofar as the annealing section of the probe nucleic acid. Instead of a fluorophore attachment for 'read-out', however, the NTD probe would have a portion that would be favored for channel-capture and for channel modulation. Just as different fluorophores offer multiplex capability, different

channel modulators offer resolving capabilities for multiple applications. The NTD event transduction is inherently a single molecule detection event and has ‘quantitative’ at the single event level, so may offer more detailed evaluation of relative gene expression. The PCR part of qPCR can be co-opted on the NTD platform for enzymatically amplified detection events, for more discussion along these lines see the ELISA-like nanopore detector methods TERISA and TARISA in [5].

Received: March 5, 2017; Published: March 23, 2017