Detection of individual proteins bound along DNA using solid state nanopores

Calin Plesa, Justus W. Ruitenberg, Menno J. Witteveen, Cees Dekker*

Department of Bionanoscience, Kavli Institute of Nanoscience, Delft University of Technology, Lorentzweg 1, 2628 CJ Delft, The Netherlands.

*Corresponding author. E-mail: c.dekker@tudelft.nl
ABSTRACT:

DNA in cells is heavily covered with all types of proteins that regulate its genetic activity. Detection of DNA-bound proteins is a challenge that is well suited to solid-state nanopores as they provide a linear readout of the DNA and DNA-protein volume in the pore constriction along the entire length of a molecule. Here we demonstrate that we can realize the detection of even individual DNA-bound proteins at the single-DNA-molecule level using solid state nanopores. We introduce and use a new model system of anti-DNA antibodies bound to lambda phage DNA. This system provides several advantages since the antibodies bind individually, tolerate high salt concentrations, and will, due to their positive charge, not translocate through the pore unless bound to the DNA. Translocation of DNA-antibody samples reveals the presence of short 12 µs current spikes within the DNA traces, with amplitudes that are about 4.5 times larger than that of dsDNA, which are associated with individual antibodies. We conclude that transient interactions between the pore and the antibodies are the primary mechanism by which bound antibodies are observed. This work provides a proof-of-concept for how nanopores could be used for future sensing applications.

KEYWORDS: protein, nanopore, DNA, binding, resolution, antibody
Solid-state nanopores have facilitated the development of a number of novel techniques to study biological samples. Although the majority of the work has focused on DNA and proteins, there has been some research into DNA-protein complexes as well. Some of these studies have investigated the detection of DNA-bound nucleosomes\(^1\), biotinylated DNA-oligomer-bound monovalent streptavidin\(^2\), SSB bound to ssDNA\(^3\), or RNAP-DNA transcription complexes\(^5\) with short (<1kbp) strands where no positional information could be determined. Other studies have used small diameter solid-state pores\(^6, 7\) or biological pores\(^8, 9\), where the DNA-protein complex is too large to translocate though the pore, to determine the presence of the complex through its effects on the characteristics of the translocation events. Previous studies from our lab with longer DNA strands focused on RecA\(^10, 11\). This DNA repair protein binds cooperatively to DNA, forming long filaments of varying size which enabled, under optimal conditions, us to determine a best resolution of small protein patches of 5 adjacent proteins in a row. Here we present our experimental efforts to map out individual bound-protein positions on long DNA strands using solid-state nanopores in standard measurement conditions. We show that transient protein-pore interactions allow us to detect the presence of individual DNA-bound proteins.

In our solid-state nanopore measurements, a 20 nm thick silicon nitride membrane containing a ~20 nm diameter nanopore is placed in-between two reservoirs containing a buffered 1M KCl salt solution and Ag/AgCl electrodes. Upon the application of an electric field, negatively charged biomolecules are electrophoretically driven from the one chamber (Cis) towards the positive electrode in the other chamber (Trans) as shown in Figure 1a, while positively charged biomolecules experience the opposite effect. The passage of a biomolecule through the nanopore results in a temporary reduction in the ionic current (Figure 1d). The magnitude of this reduction is, at these high-salt concentrations, determined by the volume taken up by the part of the
molecule that resides in the pore. Consequently, in the case of a protein bound along a long piece
of DNA, the passage of DNA produces a distinct blockade level (I_1) while any bound protein is
expected to appear as an additional temporary increase in the blockade level, on top of the DNA
blockade level.

We introduce a new model system to study this concept in detail. It is based on anti-DNA
antibodies to serve as a proof-of-principle for this technique. Anti-DNA antibodies were first
discovered in patients with systemic lupus erythematosus in the late 1950s\textsuperscript{12,13}. These antibodies
bind randomly and independently to DNA with a measured K_D of 90 nM for 160 kDa mouse
monoclonal IgG2a antibodies against dsDNA in physiological conditions\textsuperscript{14}. We carried out
nanopore measurements (Supplementary Section S1) using only free antibodies and determined
that they are positively charged at pH 8 in 1M KCl. This property ensures that any free
antibodies present in the solution on the Cis side will translocate away from the nanopore unless
they are bound to a DNA molecule, eliminating the possibility that any of the observed events
are due to a free unbound antibody co-translocating with a DNA molecule. We verified the
binding of the antibodies to DNA molecules with AFM (Figure 1b) and gel-shift assays (Figure
1c). Our nanopore-based measurements also showed that the antibody-DNA complexes still bind
in high salt conditions, albeit with a much higher dissociation constant, as discussed later. This is
an important point, since most proteins do not have observable binding to DNA in the 1M salt
concentration that is standardly used for nanopore experiments, since the electrostatic screening
is high. These commercially available antibodies provide a simple system which can be used to
explore the capabilities and limits of mapping local structures along DNA molecules using solid-
state nanopores.
Nanopore measurements on DNA-antibody mixtures reveal the presence of current spikes on top of the DNA blockade signals, which can be directly correlated to the addition of the antibodies, see Fig 1e. Solutions containing both anti-DNA antibodies and lambda phage 48.5 kbp DNA were incubated at 37°C in low salt conditions, before being put to 1M KCl and added to the Cis chamber of the nanopore flowcell. Analysis of the resulting translocation events revealed events with very short-duration large-amplitude current spikes present within the DNA blockades as shown in Figure 1e. Analysis was carried out by first selecting unfolded events. Spikes were subsequently detected if they crossed a minimum amplitude threshold. Basically the analysis comes down to detecting events within events. This approach differentiates the high amplitude spikes produced by the antibodies from smaller-amplitude folds and knots that are always observed in high-bandwidth measurements on long DNA in large pores. Statistics on the percentage of events with spikes reveal a clear correlation between the addition of antibodies and the appearance of the spikes.

Figure 2a shows the fraction of DNA events containing spikes as the minimum amplitude threshold is increased. Two independent antibody-DNA experiments (red & blue) show the presence of many large amplitude spikes, compared with the DNA-only control (green). We chose to use a spike detection threshold of 3.5$I_1$ as it differentiates quite well between spikes caused by antibodies and spikes observed in DNA-only experiments, which are due to DNA knots and folds. The typical spike amplitudes are sufficiently large that the vast majority of spikes are captured at the 3.5$I_1$ threshold used (Fig S4). Figure 2b shows the fraction of events with spikes of amplitude larger than 3.5$I_1$ at voltages ranging from 100 mV to 400 mV for experiments with both DNA and antibodies as well as control experiments containing only DNA. At 100 mV approximately 82% of events have at least one spike present, while only 6% of the
events in the DNA-only controls have spikes, a clear indication that the observed current spikes can be attributed to bound antibodies. This fraction is observed to decrease as a function of applied voltage. This effect can be attributed to the increased force reducing the lifetime of transient antibody-pore interactions, but more importantly to the decreased spatial resolution since the DNA-protein construct translocates faster but the 52 kHz amplifier bandwidth cannot be increased further. These experimental results demonstrate that it is indeed possible to detect single DNA-bound proteins with solid-state nanopores.

Analysis of the spike position is consistent with a random binding process, while the spike duration and amplitude show that the antibody passage occurs very quickly and is at the limit of what is resolvable. Figure 3b shows the position of the current spikes, normalized using the translocation time of the full DNA event. The spikes are found to be distributed randomly over the entire duration of the events, as is expected when the antibodies bind at random positions along the molecule. Their most probable amplitude is 4.5 times higher than the current drop produced by a single dsDNA molecule (Figure 3d) at 100 mV applied voltage. This amplitude is around 2 times lower than the blockade expected from the excluded volume of the antibody (Supplementary Section S6), indicating that the spikes are being distorted by the filtering due to their very short duration. Indeed, the FWHM dwell time of the current spikes $\Delta t_{\text{Peak}}$ was found to be approximately 12 $\mu$s (Figure 3c), which is below the amplitude distortion threshold of the Gaussian low-pass-filter used and right on the edge of the resolution of our system (17 $\mu$s). Factoring in the effect of the filtering of the duration of the spikes for our conditions (Fig. S5) we estimate the unfiltered translocation time of the spikes to be around 9 $\mu$s. Indeed, we also observe that the amplitude is higher for longer events: analyzing a DNA-antibody data set measured at our 40 kHz bandwidth, we observe an average amplitude of $0.45 \pm 0.01$ nA for spikes that have a
FWHM time below 17 µs (the filter’s distortion point), but 0.61 ± 0.02 nA for those with FWHM times longer than 17 µs. These observations indicate that we may be missing antibodies due to the finite measurement resolution, which we address further below.

What current signatures should we expect to see for bound antibodies in these conditions? The evidence suggests that the current spikes are due to transient interactions between the DNA-antibody complex with the nanopore, effectively holding the complex within the nanopore long enough for it to be resolved. As a simple straightforward estimate, we can use the most-probable-translocation-time of the DNA events at 100 mV (~1.5 ms) and the known DNA length (16.5 µm) to estimate a mean translocation velocity of around 11 nm/µs. This implies that, on average, an unfiltered duration of 9 µs corresponds to roughly 100 nm, which is about nine times larger than the size of the antibody. In other words, the duration of the observed current spikes is significantly longer than what we would expect from a freely translocating antibody-DNA complex. These observations suggest that the antibodies are interacting with the SiN pore surface and holding the complex bound inside the pore for a sufficiently long time to detect it. Indeed, this idea is further supported by the occasional presence of events with very long blockades with larger amplitudes that we would expect from the antibody-DNA complex (Supplementary Section S8) sticking inside the pore for a longer period of time.

Increasing the spatial resolution by a factor of four does not significantly increase the number of antibodies observed. In our previous study with RecA, the resolution was maximized by measuring at very low applied voltages (10 mV). Reducing the voltage so much increases the translocation time of the DNA but at the cost of an extremely low signal-to-noise ratio and event rate. We carried out measurements with DNA-antibody solutions at 25 mV (Supplementary Section S4) as shown in Fig. S8, where lambda DNA has a most probable translocation time of
around 8.8 ms. Event analysis was carried out by splitting traces into a low-frequency (2kHz) DNA component and a high-frequency (20kHz) part for the spikes. Using this approach we are able to resolve 25 nm features. We found a mean of 2.5 spikes per event, compared to about 1.5 spikes per event at 100 mV. The small size of this increase relative to the resolution improvement suggests that we are observing most of the bound molecules at both voltages and that the $K_D$ of the antibodies is much higher in high salt compared to the measured $K_D$ (90 nM) in low salt. We used the event rate of the unbound antibodies to put a lower limit on the $K_D$ in high salt of 1 µM (Supplementary Section S10). Furthermore, we modeled the antibody binding as a Poisson process and showed that the probability of having more than one antibody in a 25 nm segment is only 3%, much lower than the observed rates. This further indicates that the spikes are due to individual antibodies, not multiple closely-bound antibodies. These results indicate that our approach is able to detect individual DNA bound protein.

We observed an increase in the total duration of the DNA translocation time as a function of the number of spikes that are observed within the event. Figure 4 shows the average dwell time of DNA events without spikes to be 1.48 ± 0.08 ms at 100 mV. This increases to an average of 1.60 ± 0.07, 1.78 ± 0.07, 2.02 ± 0.10, 2.2 ± 0.14 ms for events with one to four spikes, respectively. The linear fit to this data, shown in Figure 4, has a slope of 189 µs/spike. This surprisingly high value prompts the question: If the average FWHM dwell time of a spike is 12 µs, then why is the average time of the DNA translocation increased by a much higher value of 189 µs per spike? The observation suggests that the spikes have a higher probability of being observed within DNA events with inherently longer translocation times. In other words, since we know that DNA molecules have a wide distribution of translocation times, the molecules that happen to take longer to translocate and thus move at a slower average speed,
have a higher spatial resolution and increase the probability of a transient protein-pore interaction occurring. An additional contributing factor might be the presence of smaller DNA fragments that have smaller translocation times and fewer antibodies bound. As the majority of shorter fragments are removed from the dataset by setting a minimum integrated area for each event, it is unlikely that this causes an increase in the total duration of the DNA translocation time as a function of the number of spikes per the event. Overall, the observed increase in the total dwell time is consistent with a temporary interaction between the DNA-antibody complex and the pore.

Smaller F_{ab} fragments could not be observed in DNA translocation experiments. We created F_{ab} fragments from the anti-DNA antibodies using a standard Papain/Protein A technique. These smaller (50 kDa versus the 160 kDa of IgG2a) F_{ab} fragments were incubated with lambda DNA and the DNA-F_{ab} constructs were translocated through smaller 10 nm pores. Interestingly, no convincing differences were seen between these samples and DNA-only controls, i.e. we did not observe any significant spikes due to the presence of the F_{ab} fragments, as shown in Fig. S11. This could be due to low affinity of the F_{ab} fragments in high salt or may suggest that the transient protein-pore interactions observed in the IgG2a antibodies are mediated by the F_{c} region, which is no longer present in the F_{ab} fragments.

Could the observed current spikes be the result of mechanisms other than transient protein-pore interactions? We considered and ruled out a number of other possible alternatives to explain the observed current spikes. 1) Could the spikes be caused by transient DNA-pore interactions? Such interactions are a known issue in solid-state nanopore experiments. In this mechanism the DNA and the pore temporarily interact. If such an interaction occurs in the close vicinity of a bound protein, that protein may become visible if the duration of the interaction is greater than
about 10 µs. Although this is also consistent with the observed increase in dwell time as a function of the number of observed spikes, the interaction coincidentally would have to occur right at the site of a bound antibody. We deem this unlikely. Indeed, it would also have to occur multiple times for the events with multiple spikes and it does not explain why no current spikes are visible in the case of the Fab fragments. Furthermore, we would expect the observed amplitude to have a much wider distribution since the protein could sit at varying distances away from the pore constriction. 2) The local velocity during the course of a DNA translocation event is known to fluctuate. Could the DNA temporarily slow down to the extent that the antibodies are visible? We estimate that the local velocity would have to slow down by a factor of around 12 times relative to the mean translocation velocity, something that is unlikely to occur so consistently in all events. Alternatively, since the antibodies are positively charged which leads to an additional electrophoretic force in the direction opposite to the DNA translocation and they increase the drag force, we also considered and ruled out the possibility that the local translocation velocity could be slowed down significantly due to the presence of a bound antibody (Supplementary Section S10). 3) Each antibody has two binding sites, bringing up the possibility that they could be forming loops in the DNA molecule reminiscent of those formed by Lac repressor. This, however, also seems unlikely since as soon as the first site has bound, the probability of the second site binding to a nearby piece of DNA is far higher than for any DNA segment far away. Furthermore, such DNA loops would have amplitudes of 2I₁, well below the observed spike amplitudes. 4) Could two antibodies bind very closely together to form a complex larger than the diameter of the pore? We used the Poisson distribution to calculate 3% probability that two or more antibodies are found on the same 25 nm segment, far lower the amount of spikes observed at 25 mV. Alternatively, such a large complex could get stuck at the
pore constriction long enough to produce a spike signal. To check this hypothesis, we translocated the same antibody-DNA mixture through a pore with a 35 nm diameter, a large-diameter pore that would be sufficient to accommodate two or more antibodies and the DNA molecule (Supplementary Section S9). We were, however, still able to observe the current spikes in this situation, suggesting that the formation of a bulky complex is not the mechanism. All these considerations and observations together render the proposed mechanism of transient protein-pore interactions the most probable.

**Conclusion.**

In a series of proof-of-principle experiments, we have demonstrated that individual DNA-bound protein can be detected with solid-state nanopores. Although this work focused on anti-DNA antibodies that bind randomly, there is a potential to extend this technique to bound proteins of any arbitrary type, by using cross-linked site-specific bound proteins and primary antibodies, to visualize specific complexes. Several issues must be addressed for this technique to be generally applicable to any protein-DNA system. Since measurements at physiological salt conditions do not provide a high enough signal-to-noise ratio, the technique is currently limited to proteins which are resistant to high-salt (1M) conditions with a slow time scale of dissociation, or proteins which have been covalently cross-linked to the DNA. A further complication is the fast translocation velocity relative to the maximum measurement bandwidth achievable, which makes the detection of a single protein very challenging. The recent development of high-bandwidth amplifiers\textsuperscript{24, 25} and low capacitance glass membranes\textsuperscript{26} should help improve this limitation. Finally, in order to determine the positions of bound protein, the current signals which are recorded in the temporal domain must be converted to spatial information using a currently unknown mapping function which is dependent on the local velocity profile along the length of
the DNA. As we have recently shown, the stochasticity of the process introduces significant uncertainties into the position determined from single measurements, while ensemble measurements provide far more accurate information. None of these obstacles, however, present a fundamental roadblock to the continued development of this technique. With further improvements, it should be possible to develop a nanopore technique for identifying DNA-binding proteins which is complementary to approaches such as chromatin immuno-precipitation (ChIP) and DNA adenine methyltransferase identification (DamID).

**Methods - Nanopores.** SiN membranes were fabricated and 20 nm diameter nanopores were drilled with a TEM as described previously. After TEM drilling, membranes were manually painted with a layer of PDMS in order to reduce the capacitance and improve the signal-to-noise ratio. Chips were mounted in a PMMA flowcell, after which the two reservoirs were filled with 1M KCl 10mM Tris 1mM EDTA solution at pH8. The current was recorded with a standard electrophysiology setup consisting of an Axopatch 200B amplifier, digitized with a Digidata 1322A DAQ, and subsequently analyzed in the Transalyzer Matlab package. The nanopore measurements presented have a bandwidth of at least 30 kHz at 100 mV and are limited by the Axopatch’s bandwidth of 52 kHz at higher applied voltages.

**Methods - DNA with bound antibodies.** Anti-DNA antibodies (HYB331-01) purchased from Abcam (Cambridge, UK) at a concentration of 625 nM were incubated with 0.25 nM lambda DNA (Promega) (at a 932:1 ratio) in 18.75 mM NaCl, 2 mM Tris, pH8 for 10 min at 37°C. Right before starting a nanopore measurement, 20 uL of this DNA-antibody mixture was added to 10 uL of 3M KCl, 30 mM Tris, 3 mM EDTA at pH8 for final concentrations of 166 pM DNA and 156 nM antibodies in a 1 M KCl solution.
**Figure 1** – a) Schematic representation of a dsDNA molecule with a bound antibody translocating through a 20 nm pore. b) AFM image of a 2.2 kbp DNA molecule with three antibodies attached. c) Gel-shift-assay showing that antibodies binding to a 40 bp DNA fragment do not traverse into the gel in lane 1, free DNA in lane 2, and a 10 bp ladder in lane 3. d) Typical unfolded lambda DNA translocation event. e) Examples of translocation events containing spikes associated to anti-DNA antibodies. The applied voltage is 100 mV and the nanopore diameter is 20 nm.
Figure 2 – a) The fraction of events with spikes as a function of the spike detection threshold for two experiments with antibodies and lambda DNA (red and blue) as well as a control experiment with only lambda DNA (green). The dotted green line represents the “ghost” spikes that would be detected due to Gaussian noise. The dashed red line is the $3.5I_1$ spike detection threshold used in this study to differentiate antibody-induced spikes from spikes also present in the DNA-only control. b) The fraction of events with spikes as a function of applied voltage for two experiments with antibodies and lambda DNA (red and blue) as well as a control experiment with only lambda DNA (green).
Figure 3 – Statistics for spikes with an amplitude above 3.5 I/I₁ observed on mixtures of lambda DNA and antibody, translocating through a 20 nm pore at 100 mV. 

a) Schematic of the data extracted for each spike, which includes the amplitude (Iₚₑᵃᵏ) as measured from the first DNA level (I₁), the FWHM duration (Δtₚₑᵃᵏ), the absolute position (tₚₑᵃᵏ), and the normalized position (tₚₑᵃᵏ/Δtₐₙ₉). 

b) Normalized position (tₚₑᵃᵏ/Δtₐₙ₉) distribution shows uniform binding all over the length of the molecule. Part of the higher population at the start is attributed to brief large amplitude folds which sometimes occur at the start of the translocation process.

c) FWHM dwell time (Δtₚₑᵃᵏ) histogram for the current spikes, showing a peak around 12 µs.

d) Amplitude (Iₚₑᵃᵏ) of the current spikes showing a peak around 0.5 nA (4.5x larger than I₁ (0.11 nA)). Insert: distribution of the number of peaks per event.
Figure 4 – The average dwell time and standard error of DNA translocation events as a function of the number of spikes present within the event. The solid line is a linear fit to the average of the two datasets and has a slope of 0.189 ms/spike.
ACKNOWLEDGMENT

The authors would like to thank Meng-Yue Wu for TEM drilling of nanopores and Christophe Danelon for helpful discussions. This work was supported by the Netherlands Organisation for Scientific Research (NWO/OCW), as part of the Frontiers of Nanoscience program, and a European Research Council Advanced grant NanoforBio (no. 247072), and by the Koninklijke Nederlandse Akademie van Wetenschappen (KNAW) Academy Assistants Program.

Supporting Information. Nanopore characterization of anti-DNA antibodies, AFM data and methods, filtering effects, 25 mV data, antibody binding in high salt, excluded volume estimates, F_{ab} fragment data, long-blockade events, 35 nm data, additional discussion. This material is available free of charge via the Internet at http://pubs.acs.org.
REFERENCES


for TOC only