

SCIENTIFIC REPORTS



OPEN

Nanopore sensing of individual transcription factors bound to DNA

Received: 16 April 2015

Accepted: 02 June 2015

Published: 25 June 2015

Allison Squires¹, Evrim Atas¹ & Amit Meller^{1,2}

Transcription factor (TF)-DNA interactions are the primary control point in regulation of gene expression. Characterization of these interactions is essential for understanding genetic regulation of biological systems and developing novel therapies to treat cellular malfunctions. Solid-state nanopores are a highly versatile class of single-molecule sensors that can provide rich information about local properties of long charged biopolymers using the current blockage patterns generated during analyte translocation, and provide a novel platform for characterization of TF-DNA interactions. The DNA-binding domain of the TF Early Growth Response Protein 1 (EGR1), a prototypical zinc finger protein known as zif268, is used as a model system for this study. zif268 adopts two distinct bound conformations corresponding to specific and nonspecific binding, according to the local DNA sequence. Here we implement a solid-state nanopore platform for direct, label- and tether-free single-molecule detection of zif268 bound to DNA. We demonstrate detection of single zif268 TFs bound to DNA according to current blockage sublevels and duration of translocation through the nanopore. We further show that the nanopore can detect and discriminate both specific and nonspecific binding conformations of zif268 on DNA via the distinct current blockage patterns corresponding to each of these two known binding modes.

Transcription Factors (TFs) are the primary gene expression regulators in all known biological systems. Canonically, TFs operate by binding to their sequence-specific targets along genomic DNA, employing a complex search and binding mechanism that remains an intensive subject for research¹. The TFs' binding affinities and sequence specificity must be maintained within bounds that permit highly dynamic yet robust activation and deactivation of specific sets of genes during the cell cycle. Therefore, probing and characterizing TF binding to their DNA targets continues to be a major area of interest in molecular biology, and the implementation of novel experimental and computational approaches has drawn substantial focus. A wide host of techniques, both direct and indirect, commonly have been used to characterize sequence-specific protein binding to DNA, for example electrophoretic mobility shift assays (EMSA)², inhibition of enzymatic degradation³, chromatin immunoprecipitation-based microarray⁴ or sequencing assays⁵, NMR^{6,7}, x-ray crystallography⁸, bioinformatics^{9,10}, and recently more advanced techniques such as optical tweezers¹¹, AFM¹², and even direct fluorescent visualization¹³, among others. However, most of these techniques require some combination of chemical cross linking of the TFs to the DNA, modification or tagging of the TF and DNA, and/or amplification assays, and may not have the capacity to resolve fine details of the TF/DNA complex, such as partial versus full binding of the TF domains to the DNA.

Solid-state nanopores are label- and tether-free, single-molecule biosensors for characterization of charged biopolymers, such as DNAs, RNAs, or proteins. An electrical field guides the biopolymers towards and into the nanopore, allowing them to probe individual molecules within an extremely dilute pool of analytes. Provided that the nanopore diameter is much smaller than the characteristic coil size, as each biopolymer passes through the pore, it must uncoil and thread through the narrow constriction in

¹Department of Biomedical Engineering Boston University Boston, Massachusetts 02215 U.S.A. ²Department of Biomedical Engineering The Technion – Israel Institute of Technology Haifa, Israel, 32000. Correspondence and requests for materials should be addressed to A.M. (email: ameller@bm.technion.ac.il)

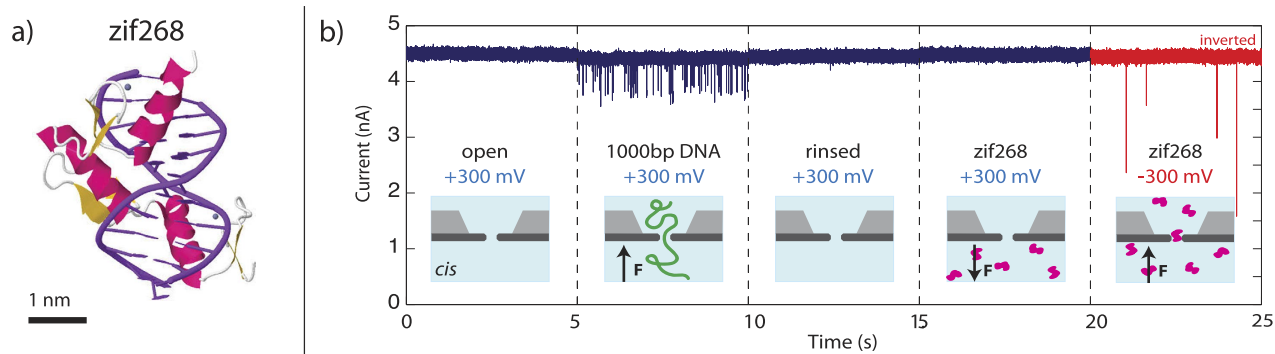


Figure 1. Continuous ion current traces of a 4 nm solid-state nanopore. **a)** Crystal structure of zif268 (PDB: 1AAY). **b)** Nanopore current at 300 mV bias, from left to right: Open pore current (4.5 nA); after the addition of 10 pM 1 kbp DNA to *cis* (transient blockages); clean after rinse with buffer (4.5 nA); after adding 10 nM zif268 (positively charged) to *cis* (no blockages); after reversing of the voltage to -300 mV (transient blockages).

a linear fashion. Thus, the time-dependent ion conductance of the partially blocked pore can be directly related to local analyte characteristics as the nanopore scans the length of the translocating biopolymer.

Previous studies using solid-state nanopores focused on the detection of protein coatings on DNA, such as RecA^{14–19}, or on individual large proteins such as streptavidin or antibodies bound to covalently-functionalized DNA^{20–22}. Nanopore Force Spectroscopy (NFS)²³ experiments have successfully employed voltage-driven removal of proteins from DNA to investigate the binding of restriction enzymes^{24,25}, restriction endonucleases²⁶, ligands for DNA aptamers²⁷, and histones^{28–30} along DNA. Discrimination of non-protein changes in local analyte structure, for example in DNA structure^{31–33} or by adding PNA tags³⁴, has also been demonstrated. In addition to many existing applications of nanopores, such as nucleic acid sequencing and barcoding^{34–38}, protein characterization^{22,39–46} and detection of epigenetic modifications^{47–49}, the ability to scan and identify features along the length of an analyte is particularly well-suited to determine the presence and binding interactions of protein-DNA complexes.

Due to their ability to rapidly scan hundreds of individual biopolymers, and resolve fine structural details along the analyte contour, solid-state nanopores have the potential to provide unique and rich information to facilitate our understanding of genetic regulation by TFs. However, to date the ability to detect small DNA-binding proteins reversibly bound to a specific recognition site along DNA, without the need for covalent functionalization or cross-linking, has proved elusive.

We demonstrate here that single-molecule measurements using solid-state nanopores can resolve single small proteins – the canonical zinc-finger DNA-binding domain of the Early Growth Response 1 (zif268) – bound natively to a recognition site. By performing single-molecule analysis of the ion current sublevels that emerge from the translocation of zif268/DNA complexes, we show that the nanopore can resolve and discriminate specific vs. nonspecific binding conformations of zif268, which produce distinct translocation current blockage patterns that correspond well to the known conformation and behavior of zif268 in both recognition (specific) and search (nonspecific) binding modes, respectively.

zif268

The DNA-binding domain of Early Growth Response 1 (EGR1), a zinc finger protein known as zif268, was selected as a prototypical transcription factor. Zif268 binds tightly to DNA ($0.2 \text{ nM} < k_D < 5 \text{ nM}$), has been used extensively as a model system for studying how TFIIIA-like zinc fingers recognize DNA, and has even served as a basis for engineering several types of artificial DNA-binding proteins⁵⁰. The crystal structure of zif268, solved via NMR by Pavletich⁶ and later improved upon by Elrod-Erickson⁵¹, is shown in Fig. 1a bound to DNA (PDB ID #1AAY)⁵². When bound, the three zinc fingers are situated in the major groove of the DNA and contact both bases and the phosphate backbone. The zinc fingers wrap $\sim 180^\circ$ around the DNA, forming a complex that is ~ 3 nm in diameter.

Figure 1b illustrates that the unbound zif268 does not interfere with normal operation of a small (~ 4 nm) solid state nanopore sensor. The pore is initially clean with a steady current of 4.5 nA at $+300$ mV bias (relative to *cis*). Upon addition of 1000 bp DNA to the *cis* chamber, transient current blockades indicate translocation of DNA through the nanopore. After a 10x volume rinse with buffer, no further translocations are observed, indicating that the DNA can be removed by dilution. As expected, addition of 10 nM of the positively charged zif268 ($pI \approx 10.5$) to *cis* produces no translocations under $+300$ mV bias, while at -300 mV translocations are observed.

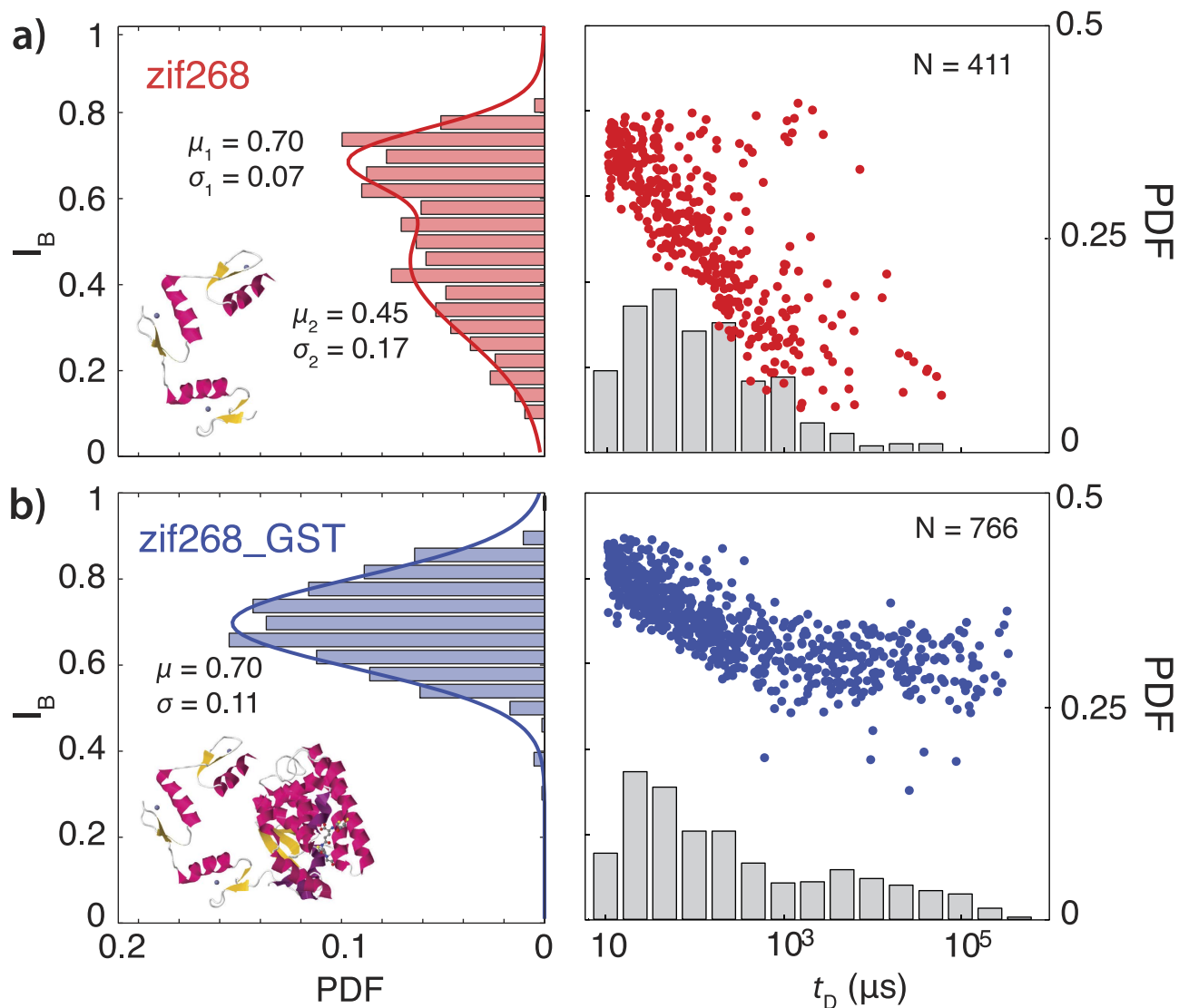


Figure 2. Translocations of unbound protein. **a)** zif268, showing two distinct blockage levels ($I_B=0.70$ and $I_B=0.45$) according to its probability density function (PDF), **b)** zif268_GST, showing only a single, broad blockage level ($I_B=0.70$). Insets (left): PDB of zif268 and composite PDB of zif268 with GST. To facilitate visualization of population density, a random white noise offset below the acquisition rate of this data ($-2 < \Delta t < +2 \mu\text{s}$, acquisition rate 250 kHz) has been added to the duration of each blockage level in the event diagram.

Unbound zif268 and zif268_GST characterization

We began our study by characterization of the unbound zif268 (10.6 kDa) and the bulkier zif268_GST (37 kDa, GST purification tag left uncleaved) at negative bias. In Fig. 2 we display event diagrams (fractional blockade current, $I_B = i_{\text{blocked}}/i_{\text{open}}$ versus translocation dwell-times t_D) as well as histograms of I_B and t_D , of the two proteins. These results reveal a surprising pattern: while translocations of zif268 ($N=411$) are relatively deep and exhibit at least two distinct blockage levels (Fig. 2a: $I_B=0.45$ and 0.70), translocations of zif268_GST ($N=766$) exhibit only a single blockage level ($I_B=0.70$), as shown in Fig. 2b. This suggests that free zif268 can adopt at least geometrically distinct two conformations, while zif268_GST adopts one. One possible source of this difference is the charge on each domain: although both zif268 and zif268_GST both have net positive charge, each individual zinc finger is predicted to have a positive charge ($pI=8.8, 9.9,$ and 11.1 for ZF_1, ZF_2, and ZF_3, respectively, per EMBOSS calculation)⁵³ while the GST is slightly negative at neutral pH ($pI=6.2$). Thus the smaller but highly charged zif268 may adopt a more expanded conformation due to mutual repulsion of the three zinc fingers, while the oppositely charged domains of zif268_GST may tend to stick together in a more compact conformation with lower net charge that can explain the shift to longer times in t_D ⁵⁴.

Characterization of zif268 bound to DNA

Translocation of bound zif268 may be observed in the fine structure of translocation events using smaller, thinned nanopores ($d = 3.5$ nm, effective thickness ~ 7 nm) for improved signal-to-noise ratio. Unlike the zif268_GST + DNA complexes (Supplementary Information, Figure S8), the small zif268 + DNA events exhibit a diverse behavior that may be classified into at least five distinct categories according to the observed current blockage pattern. Figure 3a shows schematically the design of two DNA samples used in this study, (-)DNA (left), which does not contain any predicted binding sites for zif268, and (+)DNA (right), for which one centered binding site for zif268 is predicted. These DNA fragments (960 bp and 1009 bp, respectively) were prepared by PCR amplification from the plasmid M13KO7 (NEB N0315S), with one centered binding site for zif268 (5' GCGTGGGCG 3'), as detailed in the Supplementary Information. The location of the binding site is marked in Fig. 3a with a red square. Binding sites were predicted by scoring both sequences using a position weight binding matrix for zif268 (Transfac database 7.0)^{55,56}. Prior to detection with a nanopore, DNA samples were pre-incubated at a ratio of 1:200 with purified recombinant zif268 (see Supplementary Information for details).

Figure 3b shows sample translocation events for (+)DNA incubated with zif268 (data using the (-)DNA and zif268 is shown in Figure S9). An analysis of our data suggests that three main blockage sublevels, A, B, and C, exist in the majority of events, and that these three levels appear in five distinct patterns rather than in random combinations. The five most frequently observed patterns are: A only (light blue, 60.5%), ABA (red, 24.1%), AC (green, 7%), ABAC (purple, 2.7%) and C only (gold, 1.7%). The remaining 4% of events in this data set represent a wide variety of blockage patterns, none of which account for more than 1% of the total translocation events observed. Blockage level patterns were determined using a maximum-likelihood level-finding algorithm to define an initial set of many possible sublevels. These sublevels were grouped into three major populations, representing the A, B, and C sublevels, which could be separated by selecting thresholds between them. The initial sublevels were then used to reconstruct the final event pattern by combining any neighboring sublevels that had not crossed one of these thresholds (see Supplementary Information for details).

Figure 3c shows event diagrams for the dwell time t_D and blockage level I_B associated with the sublevels for events exhibiting one of the five translocation patterns (A: blue, ABA: red, AC: green, ABAC: purple, C: orange). Each colored point represents a sublevel blockage level, I_B^i , and its corresponding sublevel dwell time, t_D^i , where the superscript i denotes one of the three sublevels (A, B, or C). The background gray scatter plot shown in all frames and at top left represents the total dwell time t_D and mean blockage level I_B for all observed translocations. The large number of events with unusually deep blockage sub-levels suggests additional observed patterns in the DNA-TFs complexes as compared to bare DNA, which typically displays a single blockage level in small nanopores which exclude folded DNA translocation⁵⁷.

In Fig. 4 we further analyze our classification of the translocation event types by examining sublevel properties across the entire data set. Figure 4a is a superposition of all event diagrams from Fig. 3b (colors consistent with Fig. 3b), and demonstrates that the A, B, and C sublevels from different translocation patterns also overlap in their durations, forming three distinct populations, highlighted here with ovals. For example, although both ABA:(t_D^{A1} , I_B^{A1}) sublevels (pink) and ABA:(t_D^{A2} , I_B^{A2}) sublevels (maroon) mostly coincide with the event-averaged current level and dwell time, ABA:(t_D^B , I_B^B) sublevels (red) form a new cluster at short dwell times $t_D^B = 50 \mu\text{s}$ and deeper blockage $I_B^B = 0.71$. Similarly, the AC (green) sub levels (t_D^C , I_B^C) forms a separate cluster at moderate t_D^C (256 μs) and deep $I_B^C = 0.46$, which are not otherwise observable in the event-averaged analysis. When superimposed as shown in Fig. 4a, it is clear that the clusters formed for the B and C sublevels exhibit consistent timing and blockage level across all five event patterns, suggesting that these sublevels represent three different states and/or geometries of the analyte molecule that arise during translocation. This conclusion is further illustrated by plotting the histograms of each group of sublevels (Fig. 4b), across all observed patterns, which clearly show three distinct current blockage sublevels.

In order to interpret our results shown in Fig. 3 and Fig. 4, we further investigated the sublevel dwell-times. It is evident that the A blockage level may be attributed to the translocation of double-stranded DNA without bound protein (Supplementary Information, Figure S8 and Figure S9). But what do the B and C sublevels represent? The nanopore should be able to accommodate translocation of the bound zif268 + DNA, and the binding site is located at the center of the DNA. It is therefore reasonable to hypothesize that passage of the complex through the nanopore would create a transient extra blockage such as B sublevel during the course of an event. As shown in Fig. 5a, the duration of the B sublevel is significantly correlated to the total dwell time of the DNA only sublevels ($t_D^A = t_D^{A1} + t_D^{A2}$) with a Pearson correlation coefficient $\rho = 0.45$ (high significance, two-tailed $p = 2.3 \times 10^{-17}$ relative to null). Moreover, the timing of the B sublevel within the overall event is consistent with a single binding site at the center of the dsDNA. Figure 5b shows the relative position of the B sublevel within the event, where $x_{\text{rel}} = t_{B,\text{center}} / (t_D^{A1} + t_D^{A2})$ so that $x_{\text{rel}} = 0$ corresponds to a B sublevel at the start of the event, and $x_{\text{rel}} = 1$ corresponds to a B sublevel at the end of the translocation event. The binding site on the sample described here is located at $x_{\text{rel}} = 0.55$ within the 1009 bp double-stranded (+)DNA. We find an expected value of $\langle x_{\text{rel}} \rangle = 0.49 \pm 0.14$, which places the actual binding site well within our estimation, with an effective position error of ± 140 bp for this sample. In the case of (-)DNA + zif268, the B sublevels, and most notably the ABA pattern (Figure S9), are much less frequent as compared with the (+)DNA + zif268

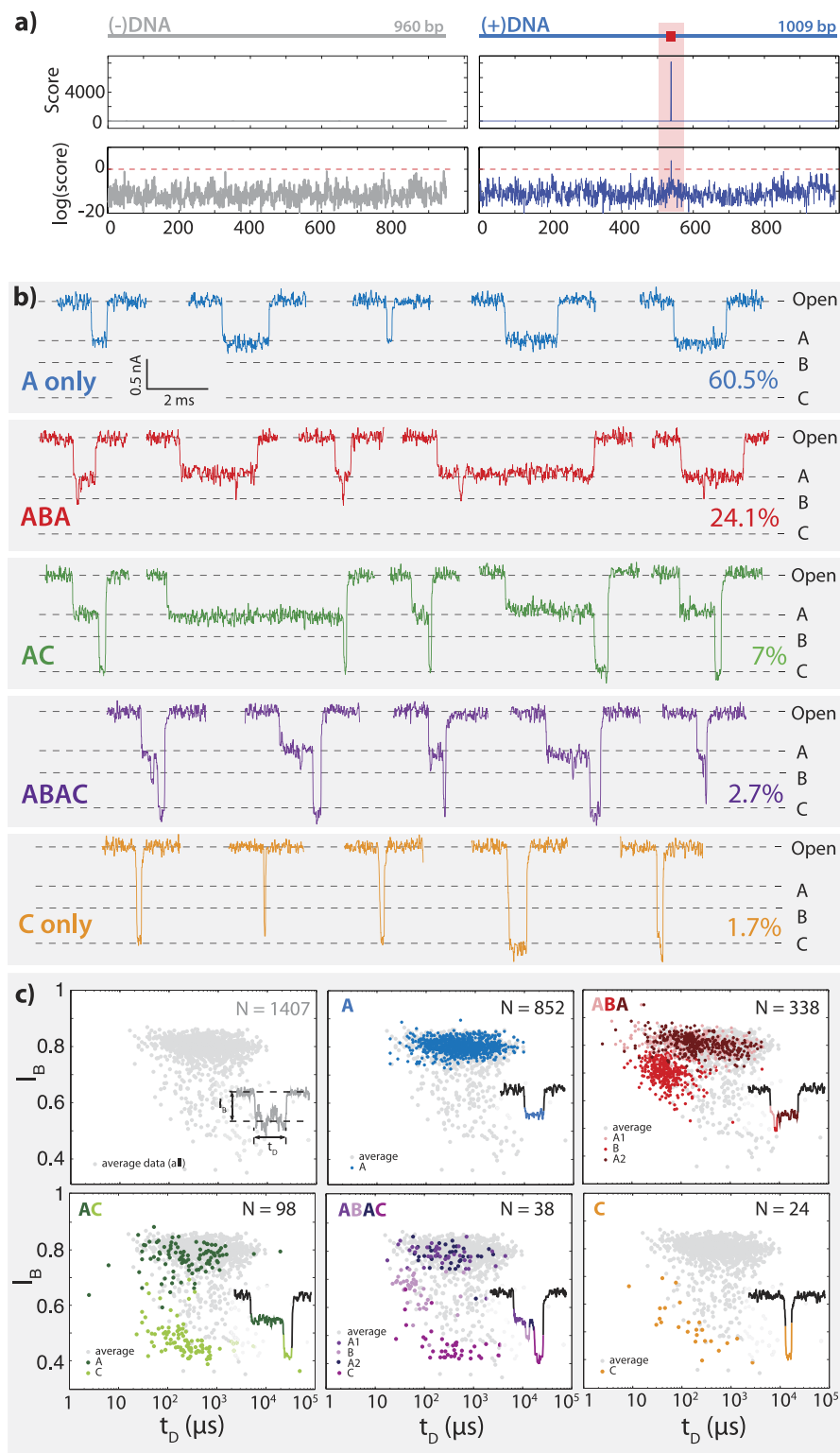


Figure 3. Blockage level patterns for zif268 + DNA in a thinned nanopore. a) DNA samples for binding: (-)DNA contains no binding sites and (+)DNA contains one centered binding site (red square), as identified by scoring the sequences according to a zif268 position weight matrix (see Supplementary Information for details, Table S1). **b)** DNA sample events showing all observed blockage sublevel patterns: A, ABA, AC, ABAC, and C. **c)** Scatter plots showing event average depth I_B and duration t_D (gray) as compared to sublevel depth I_B^i and duration t_D^i (by color, as indicated on inset sample events). To facilitate visualization of population density, a random white noise offset below the acquisition rate of this data ($-2 < \Delta t < +2 \mu\text{s}$, acquisition rate 250 kHz) has been added to the duration of each blockage level in the event diagram.

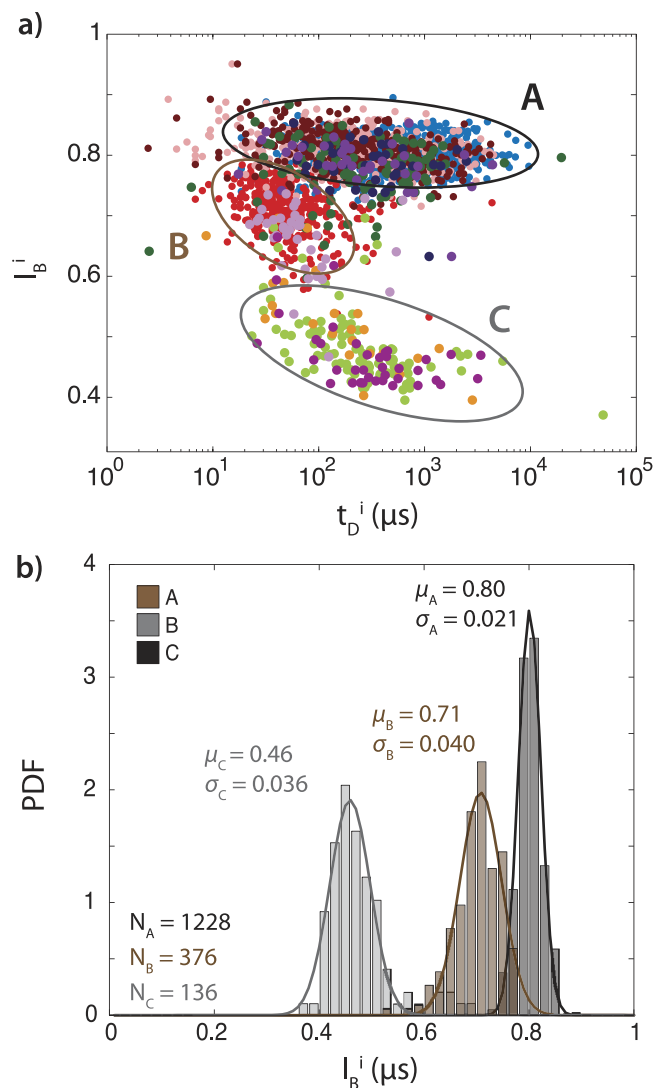


Figure 4. Cumulative A, B, and C sublevels from all categorized blockage patterns. a) Superimposed scatter plot of A, B, and C sublevels from all blockage patterns. Ellipses indicate A, B, and C populations. b) Blockage level distributions for the three sublevels across all blockage patterns.

(4.3% and 11.5%, respectively). Additionally as shown in Figure S11, in the $(-)$ DNA + zif268 case the B sublevels are not found in the center; and rather fall into two populations at $x_{\text{rel}} = 0.49 \pm 0.016$ and $x_{\text{rel}} = 0.79 \pm 0.014$, perhaps suggesting the existence of one or more preferred binding sites, not predicted by the binding models, for zif268 located off-center along the $(-)$ DNA.

In contrast to the B sublevel, the duration of the C sublevel is *not* significantly correlated to the DNA dwell time t_D^A (Fig. 5c), with a Pearson correlation coefficient $\rho = 0.17$, which is not statistically significant at 95% (two-tailed $p = 0.09$ relative to null). Based on Fig. 5a, the dwell time of the section of the analyte that causes the B sublevel is related to the overall velocity of the DNA through the nanopore, and is preceded and followed by the passage of bare DNA (A sublevels). These evidences are consistent with passage of a bound zif268 through the nanopore, located roughly at the predicted binding site (Fig. 5b). However, because the duration of the C sublevel is weakly related to the velocity of the DNA, and occurs after the DNA has passed mostly through the pore, we can conclude that the state of the analyte resulting in the C sublevel is substantially different from that involved in the ABA events.

Figure 5d shows log scale histograms of the duration of sublevels A, B, and C across all types of recorded events. The B sublevels have short dwell times ($49 \mu\text{s}$) compared to both the A and C sublevels (519 and $256 \mu\text{s}$, respectively), which is consistent with a bound complex with a short footprint (<10 bp) passing through the nanopore on a long (1 kbp) piece of DNA. However, if the bound complex were to travel roughly at the same velocity as the rest of the DNA, then the expected dwell time for the B sublevel would be $5 \mu\text{s}$, a timescale that is shorter than our current sensing capability. This raises the possibility that a fraction of actual ABA events are misclassified as A events (bare DNA) because the B sublevel is too short to be detected. This observation may explain the relatively small fraction of the ABA

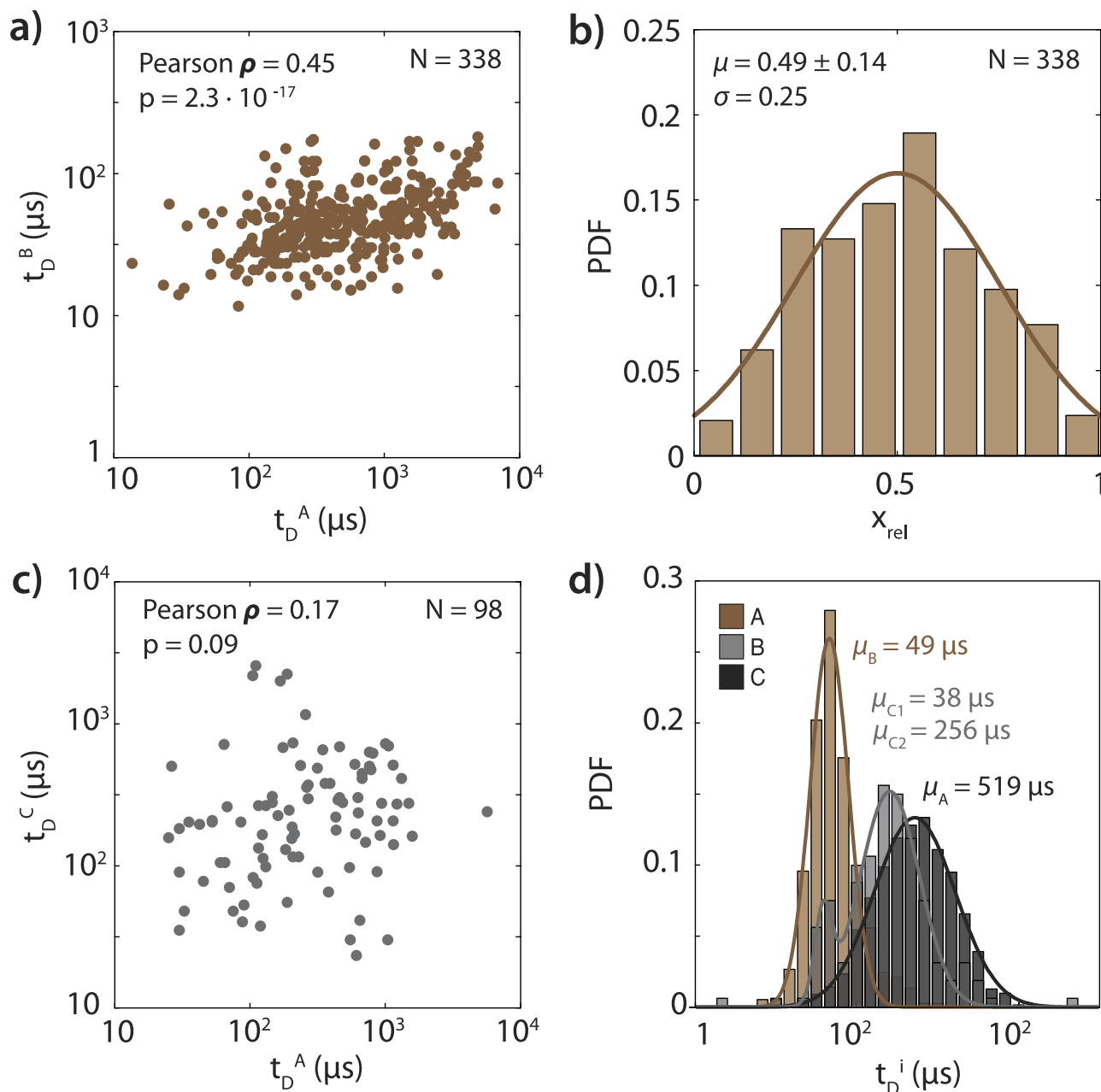


Figure 5. Timing of A, B, and C sublevels. **a)** Correlation of B level duration to total A level duration in ABA events (Pearson $\rho=0.45$ correlation, with significant $p=2.3\times 10^{-17}$). **b)** Distribution of observed relative position x_{rel} of the B level within ABA events with Gaussian fit, $N=338$. **c)** Correlation of C level duration to A level duration in AC events (Pearson $\rho=0.17$ correlation, with non-significant $p=0.09$). **d)** Distribution of observed durations of A, B, and C levels from all classifiable blockage level patterns observed.

events (roughly 24%) and relatively large fraction of A events (61%), whereas according to EMSA (200:1 protein:DNA incubation ratio) nearly 100% of the DNA used in this study should have one specifically bound zif268 at center (Figure S6 and Figure S7).

But what is the source of the C sublevel observed at the end of many events, and is it related to zif268 binding? A direct comparison of bare DNA with (−)DNA + zif268 and (+)DNA + zif268 shown in Figure S9 also reveals that: 1) patterns including B or C sublevels are very rarely detected for bare DNA, and 2) C sublevels, and most notably the AC pattern, are detected with similar frequency for both (−)DNA + zif268 and (+)DNA + zif268 samples (15.8% and 13.9%, respectively). Recently published findings on a zif268 conformational switch from its specific to non-specific bound forms on DNA suggest that we should indeed expect to see two very different conformations – and therefore different blockage

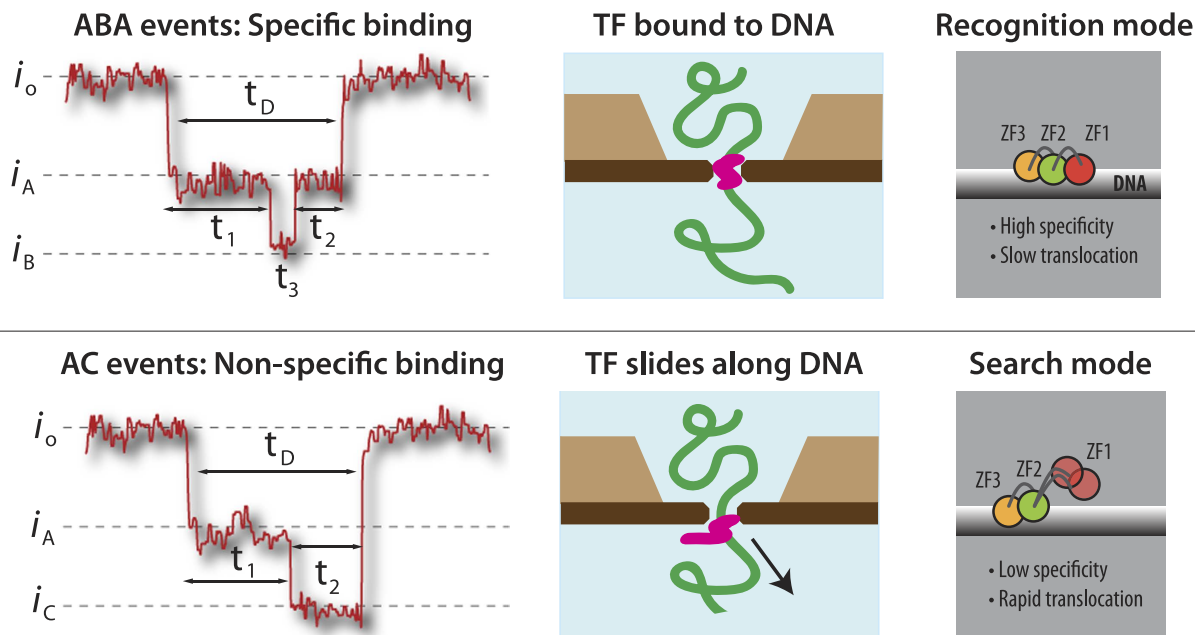


Figure 6. Suggested model and interpretation of results. Left: Diagram of blockage levels and timing for ABA and AC events. Center: Cartoon representation of suggested mechanism. Right: Recognition mode (ABA) and search mode (AC) conformations, as determined by Zandarashvili *et al.*

levels – for zif268 +DNA translocations. Zandarashvili and co-workers present NMR data that reveals fast, highly dynamic motions of the zinc finger 1 (ZF1) domain of zif268 away from the duplex DNA when the transcription factor is non-specifically bound to DNA⁵⁸, and that in this state zif268 slides easily and quickly along DNA. In its bound recognition state, zif268 takes on the structure previously determined by Elrod-Erickson, Pavletich, and Pabo^{6,51}, where all three zinc fingers are bound tightly in the major groove of the DNA. From studies of the kinetics of zif268's movement between nonspecific DNA duplexes, Zandarashvili *et al.* also infer that the constant motion of ZF1 and its position lifted away from the DNA duplex facilitates intersegment transfer of the transcription factor between distant regions of DNA, speeding up the binding site search process⁵⁸. The timescales associated with the larger, nonspecific binding conformation also provide key clues to the expected blockage pattern: In “search” mode, the zif268 zinc finger 1 explores its expanded conformation on a nanosecond timescale, and the zif268 slides along DNA with a microsecond timescale (in the absence of force), both of which occur much faster than typical translocation speeds⁵⁸.

These known conformations allow us to more clearly identify the sublevels and observed blockage patterns for zif268 +DNA translocations, as depicted in Fig. 6: We propose that the B sublevel represents translocation of the bound zif268 +DNA complex, while the C sublevel represents removal of nonspecifically bound zif268 from the DNA. The short duration and extra blockade of the B sublevel compared to bare DNA blockage level are both consistent with the temporary presence in the sensing volume of a locally bulky section of DNA, which is precisely the description of zif268 bound to a 9bp region of DNA. Similarly, the deep blockage of the C level is consistent with the larger “search” conformation, while its presence at the end of events is consistent with its ability to slide easily along DNA.

The translocation of unbound zif268 described earlier in this paper further supports this possibility; we note that the untagged zif268 protein exhibited deep blockages and multiple conformations (resulting in the broad I_B distribution, Fig. 2). This also implies that if zif268 is responsible for the C sublevel, its conformation when creating this blockage must be more similar to its conformation during free translocation than to its specifically bound conformation in the major groove of DNA, which agrees with the “search” conformation established by Zandarashvili *et al.* This expanded conformation would also enhance any potential interaction with the nanopore walls, potentially aiding shear-induced rupture of the zif268 from the DNA. The electrophoretic force applied on double-stranded DNA threaded in a pore can reach $\sim 10^2$ pN at 300 mV (depending on various variables, such as the effective charge of the DNA and interactions with the pore walls) which is comparable to the force required to rupture zif268 off of DNA as measured by AFM^{12,23}. We note, however, that the rupture force depends strongly on the loading rate and directionality, which require further investigation for the nanopore system.

Summary and Discussion

The ability to detect the presence of bound complex, determine the protein conformation, and optionally remove transcription factors from single molecules of DNA opens a wide range of possibilities for

nanopore characterization of transcription factors bound to DNA. In our study, the combination of high salt and the high protein:DNA ratio employed in this experiment enabled concurrent observation of both specific and nonspecific binding modes. Solid-state nanopores are chemically robust, so this technique could be extended to a wide range of binding and detection conditions.

Our results indicate that nanopore sensors can not only identify the presence of single small transcription factors bound to DNA, but that they can detect and discriminate between specific and nonspecific binding modes. The observed blockage patterns associated with translocation of bound zif268 +DNA correspond well to the established recognition and search binding modes. These results highlight the rich data available upon examining the fine structure of blocked current sublevels within a single translocation event. This unique technique provides access to entirely new types of single-molecule information about DNA binding to transcription factors.

Methods

zif268 (courtesy of Scot Wolfe, University of Massachusetts Medical School) was cloned into a pGex2t plasmid (GE Healthcare Life Sciences) along with a cleavable glutathione S-transferase (GST) tag. Fusion protein was expressed in BL21 cells grown in LB media at 37 °C, and purified using a glutathione sepharose column (GE Healthcare Life Sciences). GST tags were cleaved from zif268 via room-temperature overnight incubation with thrombin (GE Healthcare Life Sciences) and then eluted. Fusion GST_zif268 was eluted from the purification column using excess glutathione. Purified protein (stock concentration 50 μM) was aliquoted and stored at −80 °C until use.

Specific binding of both zif268_GST and zif268 to DNA was first verified using 50 bp oligomers for EMSA. Both zif268_GST and zif268 show similar binding affinity for the wild-type binding site, GCGTGGGCG (Supplementary Information, Figure S4). The DNA used in nanopores experiments consists of a 1000 bp sequence containing a single, centered wild-type binding site, PCR-amplified from a commercially available plasmid (M13KO7). A 1000 bp negative control without this binding site, as defined by a maximum variation of two substitutions, deletions, or insertions as compared to the consensus binding site, was computationally identified and PCR-amplified (from λ-DNA, see Supplementary Information for details, Figure S3). All PCR primer sequences and amplified sequence details are available in the Supplementary Information. Binding affinity for the specific and nonspecific sites on the 1000 bp DNA samples in the nanopore buffer conditions (1M KCl) was verified by EMSA (Supplementary Information, Figure S6). All binding was performed immediately prior to nanopore experiments, with a 30 min incubation of 200:1 protein:DNA.

Nanopores were fabricated in freestanding low-stress amorphous silicon nitride (SiN_x) membranes, deposited using Low Pressure Chemical Vapor Deposition (LPCVD) to a thickness of either 25 nm (DNA + zif268_GST experiments) or 60 nm (DNA + zif268 experiments) on <100> 350 μm thick silicon wafers. The 60 nm thick membranes were subsequently locally thinned to improve signal-to-noise ratio⁴⁶ by controlled Reactive Ion Etching (RIE) in ~1.5 μm diameter circular regions patterned by full-wafer optical lithography, leaving ~15 nm thick wells in which pores were later fabricated (see Supplementary Information). Nanopores were drilled using a JEOL 2010F transmission electron microscope as previously described⁵⁹. Nanopores were cleaned using piranha and assembled in a custom Teflon cell, and low-noise electrical measurements were performed in a custom-built electrically shielded setup using Ag/AgCl electrodes and an Axopatch 200B patch-clamp amplifier, all as previously described^{34,37,57,60}. All measurements were performed in 1M KCl, 10 mM Tris, pH 7.5.

The current flowing through the nanopore under the applied bias was measured using two homemade Ag/AgCl electrodes connected to an Axon 200 amplifier and sampled at 250 kHz/16 bits using a DAQ card (PCI-6534, National Instruments). Custom LabVIEW software was used to detect/save events and control the voltage applied across the pore. Event classification was performed off-line using custom Matlab code to identify and separate current blockage sublevels.

References

1. Spitz, F. & Furlong, E. E. M. Transcription factors: from enhancer binding to developmental control. *Nat Rev Genet* **13**, 613–626, (2012).
2. Garner, M. M. & Revzin, A. A gel electrophoresis method for quantifying the binding of proteins to specific DNA regions: application to components of the Escherichia coli lactose operon regulatory system. *Nucleic Acids Research* **9**, 3047–3060, (1981).
3. Galas, D. J. & Schmitz, A. DNase footprinting: a simple method for the detection of protein-DNA binding specificity. *Nucleic Acids Research* **5**, 3157–3170, (1978).
4. Ren, B. *et al.* Genome-wide location and function of DNA binding proteins. *Science* **290**, 2306–2309, (2000).
5. Johnson, D. S., Mortazavi, A., Myers, R. M. & Wold, B. Genome-wide mapping of *in vivo* protein-DNA interactions. *Science* **316**, 1497–1502, (2007).
6. Pavletich, N. P. & Pabo, C. O. Zinc finger-DNA recognition: crystal structure of a Zif268-DNA complex at 2.1 Å. *Science* **252**, 809–817, (1991).
7. Takayama, Y., Sahu, D. & Iwahara, J. NMR studies of translocation of the Zif268 protein between its target DNA Sites. *Biochemistry* **49**, 7998–8005, (2010).
8. Frederick, C. A. *et al.* Kinked DNA in crystalline complex with EcoRI endonuclease. *Nature* **309**, 327–331, (1984).
9. Messina, D. N., Glasscock, J., Gish, W. & Lovett, M. An ORFeome-based analysis of human transcription factor genes and the construction of a microarray to interrogate their expression. *Genome Res* **14**, 2041–2047, (2004).
10. Bulyk, M. L. Computational prediction of transcription-factor binding site locations. *Genome Biol* **5**, 201, (2003).
11. Heller, I., Hoekstra, T. P., King, G. A., Peterman, E. J. & Wuite, G. J. Optical tweezers analysis of DNA-protein complexes. *Chem Rev* **114**, 3087–3119, (2014).

12. Wang, Y. *et al.* Analysis of DNA and Zinc finger interactions using mechanical force spectroscopy. *Nanobiotechnol* **2**, 87–93, (2006).
13. Wang, Y. M., Austin, R. H. & Cox, E. C. Single Molecule Measurements of Repressor Protein 1D Diffusion on DNA. *Physical Review Letters* **97**, 048302, (2006).
14. Smeets, R. M. M., Kowalczyk, S. W., Hall, A. R., Dekker, N. H. & Dekker, C. Translocation of RecA-Coated Double-Stranded DNA through Solid-State Nanopores. *Nano Letters* **9**, 3089–3095, (2009).
15. Kowalczyk, S. W., Hall, A. R. & Dekker, C. Detection of Local Protein Structures along DNA Using Solid-State Nanopores. *Nano Letters* **10**, 324–328, (2010).
16. Hall, A. R., van Dorp, S., Lemay, S. G. & Dekker, C. Electrophoretic Force on a Protein-Coated DNA Molecule in a Solid-State Nanopore. *Nano Letters* **9**, 4441–4445, (2009).
17. Sischka, A. *et al.* Dynamic translocation of ligand-complexed DNA through solid-state nanopores with optical tweezers. *J Phys Condens Matter* **22**, 454121, (2010).
18. Spiering, A., Getfert, S., Sischka, A., Reimann, P. & Anselmetti, D. Nanopore Translocation Dynamics of a Single DNA-Bound Protein. *Nano Letters* **11**, 2978–2982, (2011).
19. Venkatesan, B. M. *et al.* Stacked Graphene-Al₂O₃ Nanopore Sensors for Sensitive Detection of DNA and DNA-Protein Complexes. *ACS Nano* **6**, 441–450, (2012).
20. Tabard-Cossa, V. *et al.* Single-Molecule Bonds Characterized by Solid-State Nanopore Force Spectroscopy. *ACS Nano* **3**, 3009–3014, (2009).
21. Carlsen, A. T., Zahid, O. K., Ruzicka, J. A., Taylor, E. W. & Hall, A. R. Selective Detection and Quantification of Modified DNA with Solid-State Nanopores. *Nano Letters* **14**, 5488–5492, (2014).
22. Bell, N. A. & Keyser, U. F. Specific protein detection using designed DNA carriers and nanopores. *J Am Chem Soc* **137**, 2035–2041, (2015).
23. Dudko, O. K., Mathé, J., Szabo, A., Meller, A. & Hummer, G. Extracting Kinetics from Single-Molecule Force Spectroscopy: Nanopore Unzipping of DNA Hairpins. *Biophysical Journal* **92**, 4188–4195, (2007).
24. Hornblower, B. *et al.* Single-molecule analysis of DNA-protein complexes using nanopores. *Nat Meth* **4**, 315–317, (2007).
25. Zhao, Q. *et al.* Detecting SNPs using a synthetic nanopore. *Nano Letters* **7**, 1680–1685, (2007).
26. Dorvel, B. *et al.* Analyzing the forces binding a restriction endonuclease to DNA using a synthetic nanopore. *Nucleic Acids Res* **37**, 4170–4179, (2009).
27. Arnaut, V., Langecker, M. & Simmel, F. C. Nanopore force spectroscopy of aptamer-ligand complexes. *Biophys J* **105**, 1199–1207, (2013).
28. Ivankin, A., Carson, S., Kinney, S. R. M. & Wanunu, M. Fast, Label-Free Force Spectroscopy of Histone-DNA Interactions in Individual Nucleosomes Using Nanopores. *Journal of the American Chemical Society* **135**, 15350–15352, (2013).
29. Langecker, M. *et al.* Nanopores Suggest a Negligible Influence of CpG Methylation on Nucleosome Packaging and Stability. *Nano Letters* **15**, 783–790, (2015).
30. Soni, G. V. & Dekker, C. Detection of Nucleosomal Substructures using Solid-State Nanopores. *Nano Letters* **12**, 3180–3186, (2012).
31. Plesa, C., van Loo, N., Ketterer, P., Dietz, H. & Dekker, C. Velocity of DNA during translocation through a solid state nanopore. *Nano Letters*, (2015).
32. McNally, B., Wanunu, M. & Meller, A. Electromechanical Unzipping of Individual DNA Molecules Using Synthetic Sub-2 nm Pores. *Nano Letters* **8**, 3418–3422, (2008).
33. Mihovilovic, M., Hagerty, N. & Stein, D. Statistics of DNA Capture by a Solid-State Nanopore. *Physical Review Letters* **110**, 028102, (2013).
34. Singer, A., Rapireddy, S., Ly, D. H. & Meller, A. Electronic Barcoding of a Viral Gene at the Single-Molecule Level. *Nano Letters* **12**, 1722–1728, (2012).
35. Cherf, G. M. *et al.* Automated forward and reverse ratcheting of DNA in a nanopore at 5-A precision. *Nat Biotechnol* **30**, 344–348, (2012).
36. Manrao, E. A. *et al.* Reading DNA at single-nucleotide resolution with a mutant MspA nanopore and phi29 DNA polymerase. *Nat Biotech* **30**, 349–353, (2012).
37. McNally, B. *et al.* Optical recognition of converted DNA nucleotides for single-molecule DNA sequencing using nanopore arrays. *Nano Letters* **10**, 2237–2244, (2010).
38. Atas, E., Singer, A. & Meller, A. DNA sequencing and bar-coding using solid-state nanopores. *Electrophoresis* **33**, 3437–3447, (2012).
39. Firnkes, M., Pedone, D., Knezevic, J., Döblinger, M. & Rant, U. Electrically Facilitated Translocations of Proteins through Silicon Nitride Nanopores: Conjoint and Competitive Action of Diffusion, Electrophoresis, and Electroosmosis. *Nano Letters* **10**, 2162–2167, (2010).
40. Fologea, D., Ledden, B., McNabb, D. S. & Li, J. Electrical characterization of protein molecules by a solid-state nanopore. *Applied Physics Letters* **91**, 053901–053901–053903, (2007).
41. Han, A. *et al.* Sensing protein molecules using nanofabricated pores. *Applied Physics Letters* **88**, 093901, (2006).
42. Larkin, J., Henley, R. Y. & Muthukumar, M., Rosenstein, Jacob K. & Wanunu, M. High-Bandwidth Protein Analysis Using Solid-State Nanopores. *Biophysical Journal* **106**, 696–704, (2014).
43. Plesa, C. *et al.* Fast Translocation of Proteins through Solid State Nanopores. *Nano Letters* **13**, 658–663, (2013).
44. Nivala, J., Marks, D. B. & Akeson, M. Unfoldase-mediated protein translocation through an [alpha]-hemolysin nanopore. *Nat Biotech* **31**, 247–250, (2013).
45. Rodriguez-Larrea, D. & Bayley, H. Multistep protein unfolding during nanopore translocation. *Nat Nano* **8**, 288–295, (2013).
46. Nir, I., Huttner, D. & Meller, A. Direct sensing and discrimination among Ubiquitin and Ubiquitin chains using solid-state nanopores. *Biophys J*, in press, (2015).
47. Laszlo, A. H. *et al.* Detection and mapping of 5-methylcytosine and 5-hydroxymethylcytosine with nanopore MspA. *Proceedings of the National Academy of Sciences* **110**, 18904–18909, (2013).
48. Shim, J. *et al.* Detection and Quantification of Methylation in DNA using Solid-State Nanopores. *Scientific Reports* **3**, 1389, (2013).
49. Wallace, E. V. B. *et al.* Identification of epigenetic DNA modifications with a protein nanopore. *Chemical Communications* **46**, 8195–8197, (2010).
50. Elrod-Erickson, M. & Pabo, C. O. Binding Studies with Mutants of Zif268: Contribution of Individual Side Chains to Binding Affinity and Specificity in the Zif268 Zinc Finger-DNA Complex. *Journal of Biological Chemistry* **274**, 19281–19285, (1999).
51. Elrod-Erickson, M., Rould, M. A., Nekudova, L. & Pabo, C. O. Zif268 protein-DNA complex refined at 1.6Å: a model system for understanding zinc finger-DNA interactions. *Structure* **4**, 1171–1180, (1996).
52. Schrodinger, LLC. *The PyMOL Molecular Graphics System, Version 1.3r1* (2010).
53. Rice, P., Longden, I. & Bleasby, A. EMBOSS: the European Molecular Biology Open Software Suite. *Trends Genet* **16**, 276–277, (2000).

54. Record, M. T., Jr., Anderson, C. F. & Lohman, T. M. Thermodynamic analysis of ion effects on the binding and conformational equilibria of proteins and nucleic acids: the roles of ion association or release, screening, and ion effects on water activity. *Q Rev Biophys* **11**, 103–178, (1978).
55. Swirnoff, A. H. & Milbrandt, J. DNA-binding specificity of NGFI-A and related zinc finger transcription factors. *Molecular and Cellular Biology* **15**, 2275–2287, (1995).
56. Pfenning, A. R., Schwartz, R. & Barth, A. L. A comparative genomics approach to identifying the plasticity transcriptome. *BMC Neuroscience* **8**, 20–20, (2007).
57. Wanunu, M., Sutin, J., McNally, B., Chow, A. & Meller, A. DNA Translocation Governed by Interactions with Solid-State Nanopores. *Biophysical Journal* **95**, 4716–4725, (2008).
58. Zandarashvili, L. *et al.* Asymmetrical roles of zinc fingers in dynamic DNA-scanning process by the inducible transcription factor Egr-1. *Proc Natl Acad Sci U S A* **109**, E1724–1732, (2012).
59. Kim, M. J., Wanunu, M., Bell, D. C. & Meller, A. Rapid fabrication of uniformly sized nanopores and nanopore arrays for parallel DNA analysis. *Advanced Materials* **18**, 3149–3153, (2006).
60. Wanunu, M., Morrison, W., Rabin, Y., Grosberg, A. Y. & Meller, A. Electrostatic focusing of unlabelled DNA into nanoscale pores using a salt gradient. *Nat Nano* **5**, 160–165, (2010).

Acknowledgements

The authors thank Prof. Zhiping Weng (University of Massachusetts Medical School) for helpful discussions and input regarding transcription factor-DNA binding, and Prof. Scot Wolfe (University of Massachusetts Medical School) for providing the zif268_GST plasmid. We are grateful for financial support from the National Institutes of Health (NHGRI grant no. R01 HG-005871), from the Marie Curie People award (GA-2010-277060, ERC) and from the Israeli Centers of Research Excellence (I-CORE) program (Center #1902/12). We also thank the staff at the Harvard University Center for Nanoscale Sciences (CNS) and the Boston University Photonics Center Optoelectronic Processing Facility for their support.

Author Contributions

A.H.S. and A.M. conceived and designed the experiments. E.A. expressed and purified protein samples. A.H.S. and E.A. designed and amplified DNA samples and performed binding experiments. A.M. fabricated the setup and the custom LabView software. A.H.S. fabricated nanopores and performed nanopore experiments. A.H.S., E.A. and A.M. performed data analysis and prepared the manuscript and supplementary information.

Additional Information

Supplementary information accompanies this paper at <http://www.nature.com/srep>

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Squires, A. *et al.* Nanopore sensing of individual transcription factors bound to DNA. *Sci. Rep.* **5**, 11643; doi: 10.1038/srep11643 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>