NANO LETTERS

Letter

pubs.acs.org/NanoLett

Single Molecule Localization and Discrimination of DNA−Protein Complexes by Controlled Translocation Through Nanocapillaries

Roman D. Bulushev,[†] Sanjin Marion,[*,‡] Ekaterina Petrova,[§] Sebastian J. Davis,[†] Sebastian J. Maerkl,[§] and Aleksandra Radenovic[†]

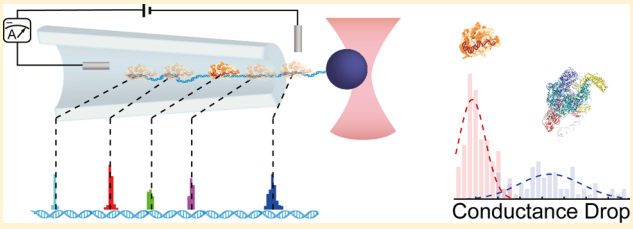
[†]Laboratory of Nanoscale Biology, Institute of Bioengineering, School of Engineering, EPFL, 1015 Lausanne, Switzerland

[‡]Institute of Physics, Bijenička cesta 46, HR-10000 Zagreb, Croatia

[§]Laboratory of Biological Network Characterization, Institute of Bioengineering, School of Engineering, EPFL, 1015 Lausanne, Switzerland

Supporting Information

ABSTRACT: Through the use of optical tweezers we performed controlled translocations of DNA−protein complexes through nanocapillaries. We used RNA polymerase (RNAP) with two binding sites on a 7.2 kbp DNA fragment and a dCas9 protein tailored to have five binding sites on λ-DNA (48.5 kbp). Measured localization of binding sites showed a shift from the expected positions on the DNA that we explained using both analytical fitting and a stochastic model. From the measured force versus stage curves we extracted the non-equilibrium work done during the translocation of a DNA−protein complex and used it to obtain an estimate of the effective charge of the complex. In combination with conductivity measurements, we provided a proof of concept for discrimination between different DNA−protein complexes simultaneous to the localization of their binding sites.



KEYWORDS: *Nanopore, nanocapillary, force measurements, DNA−protein complex, protein binding site, Jarzynski equation*

DNA−protein interactions ubiquitously regulate almost all aspects of cellular function, such as DNA chromosome maintenance, replication, transcriptional regulation, and DNA repair.[1] Most of these interactions occur after complex protein search and binding to a sequence specific DNA target. Because of the broad spectrum of possible interactions it is crucial to gain a better understanding of DNA−protein association, specifically the intricacies of binding.

Over the years, numerous methods have been developed to elucidate the role of DNA−protein interactions in cellular processes and to facilitate the translation of research into biotechnological applications.[2] Eletrophoretic mobility shift assay (EMSA),[3] nuclease footprinting,[4] SELEX-based approaches,[5] protein-binding microarray approaches,[6,7] chromatin immunoprecipitation-based microarray (ChIP), ChiP-seq[8−10] inhibition of enzymatic degradation are among the most frequent ensemble methods used to reveal sequence-specific protein binding to DNA. While quantitative analysis of thermodynamic and kinetic parameters can be assessed by surface plasmon resonance (SPR),[11] by EMSA, or, more recently, by mechanically induced trapping of molecular interactions (k-MITOMI).[12] These methods, even though they are successful in probing average interaction characteristics, disregard fine details of DNA−protein complex formation. In addition, most of them employ the use of chemical cross-linking agents, labeling tags and/or complex amplification protocols that can interfere with the proper protein conformation and as a result the proper DNA−protein interaction.

Single-molecule (SM) techniques have emerged to complement these methods and are suitable to characterize rare DNA−protein interactions with high sensitivity and reveal interesting phenomena deriving from the complex mechanisms and inhomogeneous dynamics of DNA−protein interactions. For example, atomic force microscopy (AFM)[13] and fluorescence resonance energy transfer (FRET)[14] have been used to describe the global structure of DNA−protein complexes and to probe assembly dynamics. Besides localization, SM techniques such as optical[15] or magnetic[16] tweezers can reveal binding dynamics, elongation rates, and pausing of DNA−protein complexes.

Nanopore-based sensing and force spectroscopy are the latest, label-free, additions to the growing plethora of SM methods used to detect and characterize DNA−protein interactions.[17−24] In free translocation experiments, local analyte characteristics can be directly related to the time-dependent ion conductance. Because of their tunable size, solid-state nanopores are well-suited to detect proteins attached to a long double strand of DNA.[25,26] Although free translocation experiments allow high throughput detection, the high speed of

Received: October 5, 2016
Revised: November 6, 2016
Published: November 7, 2016

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX

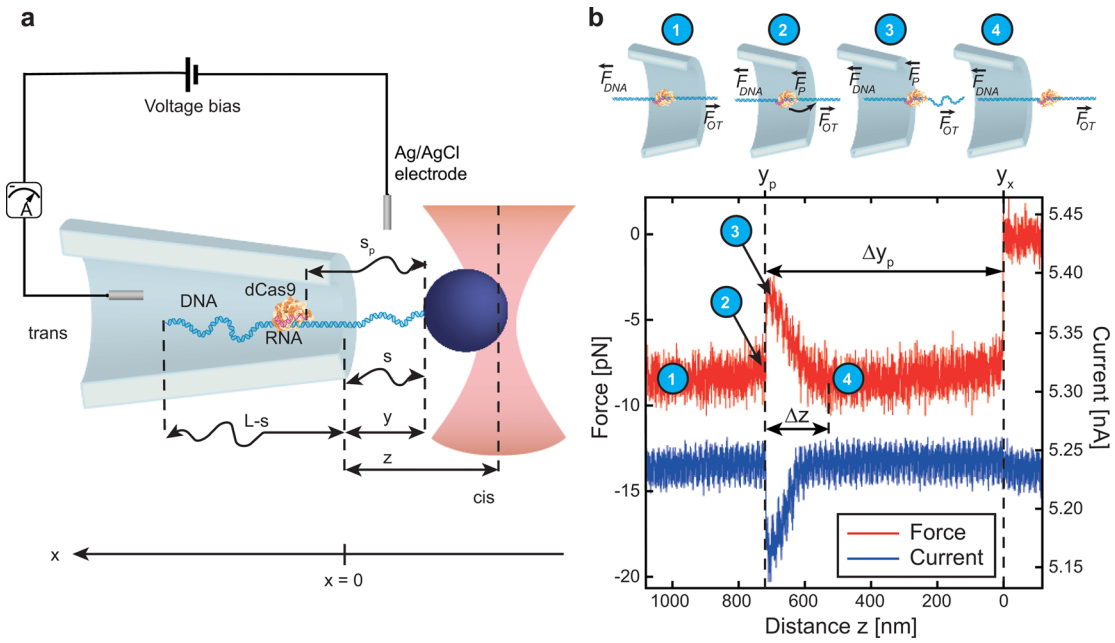


Figure 1. Setup and detected signals during controlled translocations of DNA−protein complexes through nanocapillaries. (a) Schematic of the experimental setup. The distance between the optical trap equilibrium position and the capillary opening (stage distance) is marked as z, the distance between the capillary opening and the surface of the bead is marked as y. The total length of the DNA is L, while s is the length of the DNA contour between the bead and the capillary opening. The bead-to-protein distance, as measured along the contour, is marked s_p. (b) Typical force and current curves for RNAP are shown as a function of the stage z with the jump location marked as y_p and the DNA exit out of the capillary marked as y_x. The protocol used to move the stage and thus pull out the DNA was a linear increase of z with a speed $v \approx 500$ nm/s. We measured the relative distance Δy_p between two points y_x and y_p, and the protein jump width marked as Δz. The force curve is marked at locations corresponding to the individual schematics placed above: (1) before the protein approaches the area of the jump, when the force on the protein becomes comparable to the force on the DNA; (2) location at the moment of the jump when the force on the protein is larger than the force exerted by the DNA on the protein; (3) the moment the protein has finished the jump, and the DNA is less extended (schematically represented as DNA coiling, for details see later discussion on DNA tension and the Supporting Information); (4) after the jump when the DNA extension has returned to the prejump value.

translocations and the large distribution of translocation times complicate localization of protein sites.[21,27] In addition, nanopore-based sensing is performed in nonphysiological conditions in a buffer with high salt content (>0.4 M) allowing for detection only of certain DNA−protein complexes that are capable of sustaining such conditions.

To overcome these hurdles we have recently used a combination of glass nanocapillaries and optical tweezers to detect and characterize different proteins bound to DNA.[17] We were able to elucidate their effective charge and role of the drag force, present due to electroosmotic flow in glass nanocapillaries.[28−31] Compared to free translocations, our system allows us to isolate and trap single DNA/DNA−protein complexes inside a nanocapillary and translocate them back and forth with a controlled speed in order to obtain biophysical information from force and current data. In addition, experiments in this system are performed in physiological conditions imitating the natural environment for DNA−protein complexes.

This study extends and complements our previous work by focusing on precise protein localization, charge, and size discrimination based on force and current traces, aided by robust analytical and numerical modeling. In this work we made use of two well-characterized DNA binding proteins *E. coli* RNA polymerase (RNAP) and dCas9. RNAP-DNA complex formation is critical for the initiation of transcription. On the other hand, dCas9, as a part of the CRISPR-Cas9 complex, has been utilized in various genome engineering studies alongside transcription activator-like effector nucleases (TALENs), zinc finger nucleases (ZFNs),[32] and so forth. One of the biggest advantages of dCas9 is its ease of programmability; dCas9-DNA complexes can be formed at different sites by means of different guide RNA molecules thus allowing us to investigate in details the fine resolution of our method.

Our method probes specific or high affinity sites while characterizing their size and charge and, as any SM force spectroscopy-based technique, could be adapted to monitor DNA−protein interaction dynamics.

Controlled Translocation of DNA−Protein Complexes Trough Nanocapillaries. The setup used in this work is the same as in our previous study on controlled translocations of DNA−protein complexes[17,33] and is similar to what has been used with solid state nanopores[18,34−36] and nanocapillaries.[37,38] A simple PDMS fluidic cell allows the separation of two chambers by a glass nanocapillary. A polystyrene bead, previously coated in DNA−protein complexes and trapped with optical tweezers, is brought outside the opening of the nanocapillary (see Figure 1a). By applying a voltage bias, the DNA is driven inside the nanocapillary that is confirmed by a corresponding drop in current and force (Figure 1b). Once the DNA is inside the capillary, a nanopositioning stage moves the capillary in the opposite direction until the DNA exits. Noteworthy, the conversion between stage position z and bead-to-capillary opening distance y is linear in all our data so we use relative distances in z and y interchangeably.

The controlled translocation of DNA continues smoothly with a change of the stage position z, marked by a flat force profile (Position 1 on Figure 1b) until a DNA−protein

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX

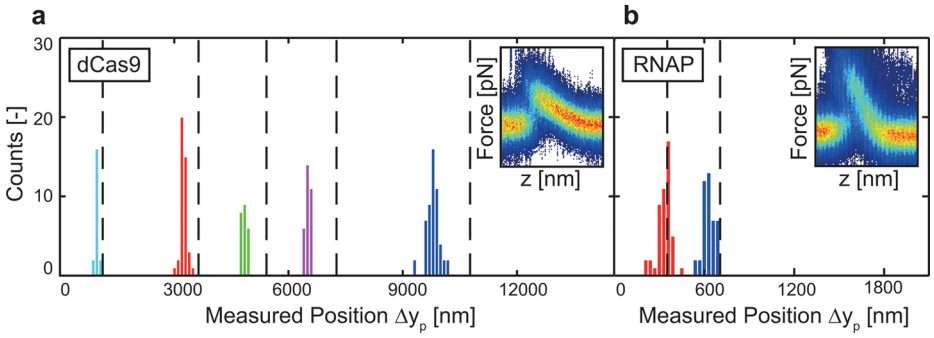


Figure 2. Localization of protein binding sites on DNA in physiological conditions. (a) Localization histograms obtained from force curves for five dCas9 sites along λ-DNA in 100 mM KCl, 10 mM HEPES, 5% glycerol, pH 7.5. Black, vertical, dashed lines represent expected (theoretical) binding sites at 1100, 3570, 5320, 7120, and 10560 nm from the free end. Different colored histograms represent different binding sites with only specific events shown. In total, 17 different capillaries were used across the five sites and the number of single specific complexes used was 20, 42, 24, 31, and 54 for each site, respectively. (b) Localization of two RNAP binding sites along a 2448 nm long DNA in 100 mM KCl, 10 mM HEPES, 5% glycerol, 0.01% TWEEN 20, pH 7.5. Black, vertical, dashed lines are theoretical binding sites at 400 and 750 nm from the free end. In total, 4 different capillaries were used across the 2 sites and the number of single specific complexes used was 49 and 43 for each site, respectively. Inset in panel a (b) represents the density plot of 60 (122) averaged force curves for dCas9 at site 1100 nm (RNAP at site 750 nm) that were shifted to the same DNA force level and normalized to a probability of 1 for each z. Force curves also include events without an equal DNA base level before and after the jump that were not used for later nonequilibrium work analysis.

complex approaches the opening of the capillary (Position 2). Near the opening there is the largest gradient of the electrostatic potential and the largest electroosmotic flow and thus the largest electrostatic force on the complex. In thin nanopores, for example, solid state nanopores that are close to two-dimensional, the potential gradient is well localized within the width of the pore. For capillaries, it depends on the manufacturing process[17,28] and causes an electrostatic potential more extended in space of the form $\sim 1/(1 + x/\xi)$ with ξ as the electrostatic decay length[17,33] and x growing inside the capillary with the origin at the opening. In our experiments, we found $\xi \approx 75$ nm by fitting to the tails of the force profiles for bare DNA.[33] When the force exerted by the external potential on the DNA strand is comparable to the force exerted on the DNA−protein complex, the protein "jumps" to the other side of the capillary opening (change from position 2 to 3 on Figure 1b).[17] This is identified by a characteristic peak at y_p in both force and current (Position 3 on Figure 1b). If the local force on the complex is directed outside of the nanocapillary, as in our case from an effectively positive local charge, the protein jumps outside the nanocapillary (positive work) followed by an increase in the length s (coiling) of DNA between the capillary opening and the bead. As the stage position z is increased, the complex stays at the same position until the extension of the DNA returns to the prejump level (Position 4 on Figure 1b). After the contour length s has become tense again, a change in the stage z also moves the DNA−protein complex away from the opening. If the local force on the protein, for example, from an effectively negatively charged DNA−protein complex, is directed into the nanocapillary, it first causes a reduction in the DNA tension, before jumping when the DNA tension has reduced sufficiently resulting in a mirrored jump shape (negative work).[18] The jump profile in current and force can be used to discriminate bound proteins as is shown later but both can also be used as a tool to localize the binding site.

Localizing Protein Binding Sites on DNA. To study localization of proteins, we used dCas9 and RNAP, which are known to bind to specific sites along DNA.[39,40] In contrast to free translocation experiments,[20,22] we carried out our experiments in physiological ionic and pH conditions. All DNA−protein complexes were probed in the same buffer (100 mM KCl, 10 mM HEPES, 5% glycerol, pH 7.5) barring the addition of 0.01% TWEEN 20 to RNAP to prevent sticking of the protein to the capillary. Applied voltage bias varied between 150 and 200 mV to optimize the capture rate during bead approach and signal-to-noise ratio of the current measurement. In the case of dCas9, we took advantage of the single guide RNA technique to engineer five distinct sites along the 16.5 μm (48.5 kbp) λ-DNA, whereas for RNAP we probed two distinct sites along a shorter, 2.45 μm (7.2 kbp) DNA (see Methods). In the case of dCas9, we performed experiments both with a single RNA guide present in the mixture as well as with several (two and three). The presence of multiple binding site possibilities did not affect the localization of each individual site and shows the potential of using our method to perform multiplexed localization. Indeed in some cases we were able to observe multiple binding events on the same DNA. Measured protein binding locations for all sites are shown in Figure 2a,b as histograms. The binding site position $\Delta y_p = y_x - y_p$ was determined as the distance between the DNA exit out of the capillary (y_x) and the position of the protein jump inside the capillary (y_p) (Figure 1b). For each binding site studied, we obtained a histogram that included both specific and unspecific events with clustering seen at positions close to the theoretically predicted ones. We eliminated all events that deviated strongly from these clusters as nonspecific binding (see Section S1 in the Supporting Information). With larger statistics, this method may be able to better assess nonspecific binding of proteins and determine if any differences are present in size or charge which would hint at variation in binding interactions between specific and nonspecific binding. Nonspecific binding has been seen both in vivo and in vitro[41−44] and results, for example in the case of dCas9, from mismatching between the RNA guide sequence and the DNA, bulge formations, or transient binding.[45] The obtained localization has an average standard deviation of 97 and 49 nm, respectively, for dCas9 and RNAP with a smaller error near the free end of DNA and thus a smaller error for shorter DNAs, as seen with RNAP.

We assume the error comes from a wide distribution of drag forces on the DNA strand, the stochastic nature of the process, unspecific interactions not taken into account by our statistics,

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX

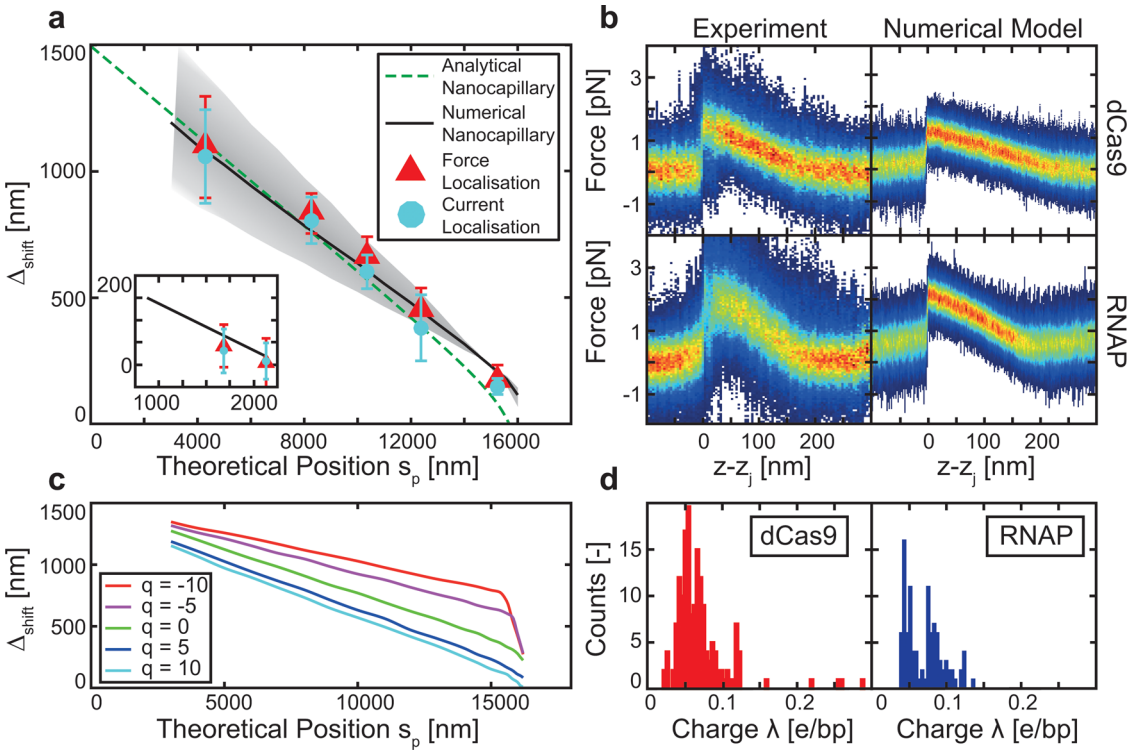


Figure 3. Protein binding site localization shift. (a) Difference between measured and expected location $\Delta_{shift} = (L - s_p) - \Delta y_p$ for dCas9 binding sites on λ-DNA. Red triangles correspond to localization shift obtained from force curves, and blue circles from current curves, all for five tailored binding sites on λ -DNA. Green dashed line represents the best fit obtained with the analytically derived nanocapillary shift formula. Full black line represents the numerically obtained fit with $\xi = 75$ nm, $V = 200$ mV, $\lambda_{DNA} = -0.04$ e/bp, and the effective protein charge $q^* = 10e$, while shaded area shows the variation of the numerical fit curve if λ_{DNA} varied from -0.02 e/bp (upper border) to -0.08 e/bp (lower border) as in panel d. Inset shows the same for two RNAP binding sites on a short 2448 nm DNA with an analytical fit with the same parameters as in the full panel except $\xi = 10$ nm (most likely due to boundary effects). (b) Comparison of 47 and 110 force versus stage position curves plotted as density plots for dCas9 and RNAP respectively (left column) with numerically obtained plots (right column). The sites shown correspond to the same as in the insets of Figure 2 (i.e., 1100 and 800 nm for dCas9 and RNAP, respectively). Before averaging, all curves were shifted so that the jump position is at $z = 0$ and all noisy curves without a well-defined DNA force level before and after the jump were not taken into consideration. (c) Plot of localization shift obtained from the numerical stochastic model using different values of the effective charge of the complex for the same parameters as the numerical fit in panel a. Charge $q = 0$ corresponds to no protein, that is, bare DNA localization shift. (d) Histograms of measured DNA linear charge density λ_{DNA} for dCas9 and RNAP attributed to changes in the electroosmotic flow induced drag force. The histograms correspond to a total of 10 (2) nanocapillaries used for dCas9 (RNAP) localization.

friction between the DNA and the capillary (if the pulling occurs slightly off axis, see Supporting Information Figure S9), and a minor contribution also comes from the different voltages used. To verify that there is no dependence on the translocation protocol, we performed a reverse protocol (threading the complex into the capillary) but found no difference in the localization value (see S2 in the Supporting Information), which we attribute to a small or negligible hysteresis.[17,18]

Localization Shift and Analytical Modeling. Although we localized protein binding sites, both thanks to force and current traces (Figure 3a), the positions were shifted in comparison to the expected locations (Figure 2a,b). For positions close to the free end of the DNA strand, the shift is small, or even negligible in the case of one RNAP binding site, and it grows as the expected position is closer to the tethered end of DNA. Figure 3a shows how the difference between the measured and expected positions for dCas9 and RNAP (inset) $\Delta_{shift} = (L - s_p) - \Delta y_p$ depends on the theoretical location of the complex. This shift can be explained by taking into account that the experimentally measured distance of the protein Δy_p is not the same as its position s_p along the DNA contour. At any distance y between the capillary opening and the bead, the DNA "coil" is extended to a contour length s such that the extension $\mu(s) = y/s < 1$. The shift Δ_{shift} is due to an early exit of the DNA strand out of the capillary (the DNA strand's exit length s_x is shorter than the full length L of the DNA strand), caused by fluctuations, as well as the DNA being extended by the pulling force. For each contour length s there is a corresponding y such that

$$\Delta y_p = s_x \mu(s_x) - s_p \mu(s_p) \tag{1}$$

with s_x the contour length corresponding to the DNA exit location y_x. The effect of the localization shift is to make the apparent protein position smaller than expected.

The extension of the DNA molecule, $\mu(s) = y/s$, which is required to explain the shift, can be obtained analytically in the strong stretching regime[46] which corresponds to the range of forces on DNA measured in typical experiments.[33,47] It is useful to compare how the results change between nanopores and nanocapillaries, as both can be used for localization. The force on the DNA can be written as $F_{DNA} = \lambda_{DNA} \Delta V f(s)$ where $f(s)$ is a general functional dependence of the potential on the length of DNA inside the capillary, such that $f(s) = 1$ for nanopores and $f(s) = 1 - \frac{1}{1 + \frac{L-s}{\xi}}$ for nanocapillaries.[17,33] λ_{DNA} is the

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX

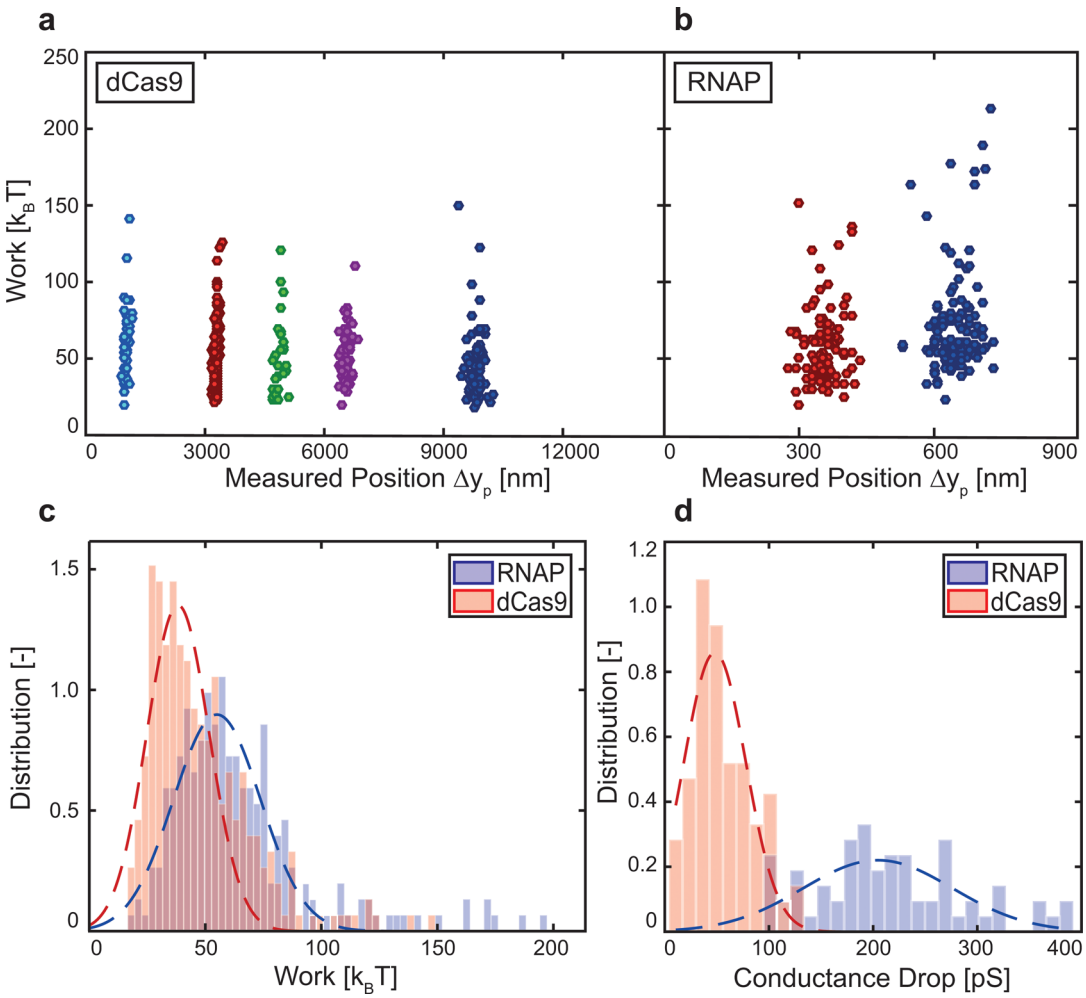


Figure 4. Analysis of works invested into protein jumps and conductance drops. (a) Scatter plots of measured works W_i for five different dCas9 binding sites (as in Figure 2a). (b) Scatter plots of measured works W_i for two different RNAP binding sites (as in Figure 2b). Here the scale of the x-axis is 10 times larger than in panel (a). (c) Comparison of jump event works obtained for dCas9 and RNAP. All binding site locations have been put together for all measurements done at $V = 200$ mV. Buffer conditions (100 mM KCl, 10 mM HEPES, 5% glycerol, pH 7.5) were the same for all dCas9 sites and with added 0.01% TWEEN 20 for RNAP. From these distributions, ΔG_{AB} values were computed using the Jarzysnki equality to be 31.7 and 37.9 k_BT for dCas9 and RNAP, respectively (for convergence plot see Supporting Information Figure S8). (d) Comparison of conductance changes for dCas9 and RNAP obtained from the maximal change of the conductance during a protein translocation relative to the DNA level.

effective linear charge density of DNA reduced by screening and drag. We assume that the extending force comes from both the force on the DNA and the force exerted by the complex, where, if it is small, we can continue in the strong stretching regime. Approximating that the protein jump is located at the point where the force is largest we can write for the extending force $F_{ext} = F_{DNA} - F_p(s = s_p) = F_{DNA} - q^*V/\xi$. The extension in the high stretching regime is then

$$\mu(s) = \frac{y}{s} = 1 - \frac{1}{\sqrt{\frac{4L_pF_{ext}}{k_BT}}} \tag{2}$$

with $L_p = 50$ nm the persistence length of DNA. Figure 3a shows the analytically obtained fit (from (eq 1) and (eq 2)) to the shift for dCas9 assuming a nanocapillary geometry with $\xi = 75$ nm and $V = 175$ mV (details in S3 in the Supporting Information). The fit results in $\lambda_{DNA} = -0.06$ e/bp, $L - s_x = 200$ nm and the effective DNA−protein complex charge of $q^* = 5e$. The obtained λ_{DNA} is larger than the experimental mean of −0.04 e/bp (Figure 3d) that we attribute to the strong stretching regime approximations used. Our analytical formulas can predict the shift in the range of small charges but in order to better fit the shift, and thus accurately predict the localization, a more precise modeling of how the complex charge affects translocation is necessary.

Numerical Modeling Scheme. To better explain the influence of the DNA−protein complex charge on the shift, we used a stochastic modeling scheme[48] previously implemented to explain controlled translocation events in both nanopores[18] and nanocapillaries.[17] The model is based on two coupled Langevin equations for two state variables of the system, the bead position away from equilibrium at the stage position z, denoted r, and the length of the contour between the bead and the capillary opening s. We solved two Langevin equations with an external force determined by the total free energy of the system $G(r, s)$, which has contributions from the charged DNA wormlike chain, optical trap, and a point protein with an effective charge q^* (coming from both electrostatics and electroosmotic flow induced Stokes drag[17]). Using our stochastic model we can obtain numerical fits to the experimentally determined shift. Figure 3a shows the fit obtained with parameters $\lambda_{DNA} = -0.04$ e/bp, $V = 200$ mV, ξ

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX

= 75 nm and an effective charge of the complex $q^* = 10e$ for dCas9 and the same except $\lambda_{DNA} = -0.02\ e/\text{bp}$ and $q^* = 12e$ for RNAP. λ_{DNA} values in the fit were taken as the means of the experimentally measured ones. We can also predict the general shape of the force curves using our simplified cone geometry (Figure 3b). Figure 3c shows that protein charge influences the shift for both positively and negatively charged bound proteins. The shaded area in Figure 3a shows the numerical fit variation if the linear charge density of DNA changes from its smallest to its largest experimentally measured value. This can easily account for the large standard deviations seen on the localization shift as explained by the wide distribution of effective linear charge densities of DNA $\lambda_{DNA} = \lambda^0_{DNA} - \lambda_{DRAG}$ seen in the DNA force levels (see Figure 3d). This is caused by a variability of electroosmotic flow induced drag on the DNA (λ_{DRAG}) in nanocapillaries.[17,28−31] Despite the drag force variability we account for the experimentally determined localization shift and localize proteins in a comparable range to other single molecule methods while in physiological conditions. In practice, the electroosmotic flow can be controlled by coating capillaries with lipids[49] or polymers[50,51] thus reducing the experimental error.

Discrimination of DNA-Bound Proteins Using Nonequilibrium Work and Conductivity. In order to simultaneously discriminate DNA−protein complexes while localizing, two methods, able to detect differences in charge or size, were used. The first is based on nonequilibrium work analysis of the force curves during the jump[17,18] while the second uses the characteristic of nanocapillaries to obtain current versus stage curves and uses them to differentiate proteins using conductivity drops.[52]

Because the stochastic protein jumps are nonequilibrium events we utilize work analysis in order to extract equilibrium quantities used for protein discrimination. Our experiment starts in an equilibrium state at time t_A and at stage position z_A and ends in an equilibrium state at time t_B and stage position z_B. The work done by a variation of the control parameter z is then

$$\tilde{W}_i = \int_{t_A}^{t_B} v\frac{\partial H}{\partial z}\,\mathrm{d}t = \int_{z_A}^{z_B} F_{ot}(z)\mathrm{d}z \tag{3}$$

with $v = \mathrm{d}z/\mathrm{d}t \approx 500$ nm/s the speed with which the stage is moving the capillary, and H the Hamiltonian of the system. This is connected to an integral of the measured optical tweezers force over the stage from the state at z_A to the state at z_B, where the protein event is wholly located in between these two stage positions. The work W_i done in an individual protein jump curve is computed as the area under the peak of the force−distance curve from the point it starts to change from the base DNA level until the point it returns to this level. This work is corrected by subtracting the average force on the bare DNA $W_i = \tilde{W}_i - F_{DNA}\Delta z$ (the value before and after the jump event should be the same, see S1 and S5 in the Supporting Information). The measured works W_i obtained from force curves can be inserted into the nonequilibrium work equation first shown by Jarzysnki,[53]

$$\left\langle \exp\left(-\frac{W_i}{k_BT}\right)\right\rangle = \exp\left(-\frac{\Delta G_{AB}}{k_BT}\right) \tag{4}$$

and connected to the free energy difference between the state before and after the jump ΔG_{AB}. With knowledge of our setup geometry and parameters, we can connect free energy differences to physical properties of the DNA−protein complexes. Assuming that the change of the wormlike chain free energy before and after the jump is negligible, that the force profile returns to the same value after the jump, and that changes in stage and contour length are approximately equal $y_B - y_A \approx s_B - s_A$, the difference in free energies before and after the jump can be simplified to

$$\Delta G_{AB} = q^*V\cdot\frac{\Delta z}{\xi + \Delta z} \tag{5}$$

where q^* is the effective charge of the complex, V is the voltage at which the experiment was carried out, and Δz is the width of the jump event (see S6 in the Supporting Information for details).

Using the Jarzynski equality we can determine the effective charges from our nonequilibrium measurements for discrimination of proteins. Work analysis for both dCas9 and RNAP results in a wide clustering (Figure 4a,b) due to the nonequilibrium and stochastic nature of the translocation events as well as a wide distribution of electroosmotic flow induced drag. Assuming that there is no difference between binding sites, we can group all the work values obtained at the same voltage for both dCas9 and RNAP (Figure 4c). Although similar, it is possible to discern two different DNA−protein complex distributions. In order to quantify the difference, we can apply the Jarzynski equality (eq 4) to the obtained distributions. Using the free energy relation (eq 5) with a determined average event width Δz = 220 and 170 nm for dCas9 and RNAP, respectively, and an electrostatic decay length ξ = 75 nm it is possible to obtain a value for the effective charge of the complex. For dCas9, we obtain $q^*_{dCas9} = 3.1\ e$ and for RNAP $q^*_{RNAP} = 3.4e$. We note that the charge value ∼3e obtained from the Jarzynski equality is much lower than the charge values obtained in the stochastic fit to the localization shift (∼10e). As we have a wide distribution of drag forces, most likely due to different capillaries used and thus different flow velocity distributions, we will have a distribution of effective charges q^* due to drag going into the work analysis. Because the Jarzynski equality is biased toward small works due to the factor e^{-W_i/k_BT}, we will preferentially obtain contributions from smaller charge values, that is smaller works, and the distribution will be skewed to show the lower bound for the complex effective charge. Both RNAP and dCas9 (with GFP attached, see Methods), having similar hydrodynamic radii (∼5 nm[54,55]), we expect them to experience a similar drag force, which would explain the similarity of the determined effective charge. The determined effective charges will also have a contribution from the bare electrostatic charge, but we expect a minor contribution from it as drag overcomes electrostatic charge.[17] Despite the similarities of dCas9 and RNAP, the distributions of work values and effective charge from the Jarzynski equality support discrimination.

To complement the small difference in effective charges of RNAP and dCas9, we can additionally discriminate them by comparing the current drops obtained simultaneously with force peaks for both proteins (Figure 4d). In a buffer of same ionic strength, RNAP exibits a larger current drop than dCas9 and enables a clearer discrimination between the proteins than the nonequilibrium work analysis. The origin of the conductance drop is known to depend strongly on salt conditions[56] and can be attributed to several different scenarios (current blockage,[57] molecular friction,[58] or a change in the advective current from current polarization[17]). A combination of conductance changes and nonequilibrium work analysis thus

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX

enables us to discriminate two proteins even if they have a similar charge or shape. In practice, one could use either one, the other, or both methods to discriminate proteins while tailoring the sensitivity to either charge or size.

Conclusions. We have demonstrated the viability of single molecule force spectroscopy using optical tweezers and glass nanocapillaries as a sensing tool for discrimination and localization of DNA−protein complexes on two proteins in physiological conditions: RNAP and dCas9. We detected the proteins shifted from their expected positions but were able to explain and correct for it by using both analytical and numerical modeling and localize them to within an accuracy of 50 nm.

In the context of the wider applicability of this technique to other types of geometries, we note the nanopore limit of (eq 5). For thin nanopores (<5 nm), like graphene[59−62] and MoS_2,[63] the free energy difference for a jump simplifies under the same assumptions to $\Delta F = qV$ with q as the bare electrostatic charge, making the analysis of experiments much simpler as the dependence on geometry is negligible. However, this comes at the cost of a worse signal-to-noise ratio for current measurements and harder experiments because combining optical tweezers with nanopores is significantly more complex than with nanocapillaries. In addition, nanocapillaries' sensitivity to the size of proteins could be tailored by reducing and controlling the electroosmotic flow induced drag with lipid or polymer surface coating.[49−51] Nanocapillaries can thus be tailored to be either sensitive to charge or size of proteins, enabling discrimination while simultaneously localizing them.

We believe that further implementation of optical tweezers combined with nanocapillaries is robust enough to precisely measure binding sites of previously uncharacterized proteins, which make up a large part of proteins known to bind to DNA, while simultaneously inferring their apparent hydrodynamic radius and charge. It could also study differences between different binding conformations, for example, binding of RNAP in the initiation stage versus the elongation stage, or Cas9 versus dCas9. Not only does our method recognize specific or high affinity sites while characterizing their size and charge, but with enough statistics could be used to infer low affinity or nonspecific binding sites. The use of the technique to detect multiple proteins bound to a single DNA (multiplexed detection) could also be put to use in detecting the presence of tandem repeats, such as the expansion of a CGG repeat occurring in Fragile X syndrome.[64]

Methods. *Combination of Nanocapillaries and Optical Tweezers.* Quartz capillaries with a 0.4 mm outer diameter and 0.3 mm inner diameter (Hilgenberg) were pulled using a P-2000 laser-assisted puller (Sutter Instrument) with a program containing 2 lines (Table 1). The pulled capillaries had a diameter of 200−300 nm. They were shrunk to diameters of 43−58 nm under SEM accordingly to[65] using a 500 pA beam current and 3 kV accelerating voltage. Only nanocapillaries with a circular, symmetric opening were used in this work. The shrunken nanocapillaries were embedded in a PDMS fluidic cell consisting of two chambers as in refs 17 and 33 and treated by oxygen plasma at ∼50 W for 2 min. Afterward they were filled with a buffer, containing 100 mM KCl, 10 mM HEPES, pH 7.5, 5% glycerol filtered through an Anotop 25 filter (Watman) beforehand. The buffer used for experiments with RNAP also contained 0.01% TWEEN 20 to reduce proteins sticking to the glass walls. The capillaries were used in the setup combining optical tweezers and nanocapillaries described in refs 17 and 33. The stiffness of the optical trap estimated using the power spectrum density method was in the range of 60−120 pN/μm.[66]

Table 1. Laser Puller Program[a]

line	heat	filament	velocity	delay	pull
1	600	4	10	145	0
2	600	4	10	145	140

[a]A program used to pull capillaries to diameters of 200−300 nm.

Formation of Bead-DNA−Protein Complexes. In order to obtain bead-DNA−protein complexes, first DNA−protein complexes were formed and then they were incubated for 15 min at 37 °C, 250 rpm with 10 μL (6.7×10^6) 3 μm streptavidin-coated polystyrene beads (Bangs Laboratories) suspended in the experimental buffer. Afterward the beads were centrifuged for 5 min, 4 °C, 1000g and the solution was exchanged to the experimental buffer. The formed complexes were stored on ice during experiments. For dCas9, fresh complexes were prepared every day, whereas for RNAP the same complexes could be used for up to 2 days due to their stability. The experiments were performed at room temperature. After 45−60 min, the buffer in the cis chamber was exchanged and new bead-DNA−protein complexes were introduced.

In the case of dCas9, first, 140 nM of guide RNA was mixed with ∼7 nM of dCas9 in the buffer containing 20 mM HEPES, pH 7.5, 100 mM KCl, 5 mM $MgCl_2$,1 mM DTT, 5% glycerol and in the presence of 0.5 Units of RNase inhibitors (Roche) for 15 min at 37 °C, 250 rpm.

To form a complex of RNAP on DNA, a T7A1 promoter was used. On this promoter, in the absence of UTP in the solution RNAP stalls on DNA after 20 bp.[67] A 7.2 kb ds DNA fragment containing a biotin tag on one end was obtained by PCR of the plasmid pRL574 kindly provided by Landick's lab. 50 nM RNAP (NEB) was mixed with ∼500 pM of DNA with a single T7A1 promoter, 500 μM ApU (TriLink), 100 μM of ATP, CTP, and GTP (Invitrogen) in a buffer, containing 25 mM Tris/HCl, pH 8.0, 100 mM KCl, 4 mM $MgCl_2$, 1 mM DTT, 3% glyceol, 0.15 mg/mL BSA for 30 min at 37 °C, 250 rpm.

Preparation of dCas9. DNA encoding the catalytically inactive dCas9 (Cas9(D10A/H840A)) version from *S. pyogenes* was PCR amplified from the pMJ841 plasmid (Addgene). The PCR product was fused at the 3′ end with yeGFP construct and cloned by Gibson assembly into a custom T7 RNAP expression vector whose sequence was verified by DNA sequencing (Microsynth). The generated vector was used for the in vitro synthesis of dCas9 yeGFP tagged protein by PURExpress In Vitro Protein Synthesis Kit (NEB) according to the manufacturer's conditions. The protein expression levels were monitored over time using a spectrophotometer (SynergyMx, Biotek) and the final dCas9 yeGFP concentration was quantified based on the GFP calibration curve.

Design and Preparation of Single Guide RNAs(sgRNAs). λ-DNA was screened for the presence of PAM motifs (5′X20NGG3′) separated by more than 5000 bp. Five sgRNAs (single guide RNAs) were designed bearing complementarity to the 20 bp 5′ adjacent PAM motif sites. sgRNAs secondary structure predictions were performed by Mfold online software. sgRNAs were prepared by in vitro transcription of dsDNA templates carrying a T7 promoter sequence. Transcription templates were generated by PCR amplification of ssDNA

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX

templates (IDT) containing the T7 binding site, 20 bp sequence complementary to the DNA target site, and sgRNA scaffold sequence. The sgRNAs were synthesized by in vitro transcription using the MEGAshortscript T7 Transcription kit (Ambion) according to the manufacturer's conditions. The sgRNAs were treated with Turbo DNase (Ambion), purified by lithium chloride precipitation method, and resuspended in nuclease-free water supplemented with RNase inhibitors (Promega). Prior to use, the quality of the sgRNA was assessed by agarose gel electrophoresis.

Stochastic Model. We used a stochastic modeling scheme[48] previously implemented to explain controlled translocation events in both nanopores[18] and nanocapillaries.[17] The model is based on two coupled Langevin equations for two state variables in the system, the bead position $r = y + \rho$ away from equilibrium at the stage position z with $\rho = 1.5$ μm as the bead radius, and the length of the contour located between the bead and the capillary opening as s. We solved two Langevin equations with an external force determined by the total free energy of the system $G(r, s)$. The free energy contains contributions from the optical trap $G_{ot}(r,z) = (1/2)\kappa(z - r)^2$, the free energy of a charged DNA molecule in the nanocapillary is $G_{DNA}(s)$, the free energy contribution from a protein bound to the DNA at position s_p is $G_p(s)$, and the wormlike chain (WLC) free energy of DNA is $G_{wlc}(r,s)$, which couples the equations (see Section S4 of the Supporting Information). We numerically solved the two coupled equations, while slowly varying the stage position z with a speed $v \approx 500$ nm/s from the DNA being almost completely inside the capillary until it exits. In order to determine any numerical parameter, we made ~100 averages of pulling protocols with random starting conditions.

■ ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/acs.nanolett.6b04165.

Analysis of force and current traces, localisation and work hysteresis, analytical solution for localisation shift, stochastic model details, equilibrium information from protein jump events (PDF)

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: smarion@ifs.hr.

ORCID

Sanjin Marion: 0000-0002-9892-7378

Aleksandra Radenovic: 0000-0001-8194-2785

Author Contributions

R.D.B., A.R., and S.J.M. designed the study. R.D.B. carried out experiments. S.J.D. wrote analysis scripts and performed data analysis. S.M. performed simulations and implemented analytical and numerical models that allowed data interpretation. E.P. synthesized dCas9 and guide RNAs. A.R. supervised the research. S.J.D. and S.M. wrote the manuscript. All authors provided important suggestions for the experiments, discussed the results, and contributed to the manuscript. R.D.B., S.M., E.P., and S.J.D. contributed equally to this manuscript.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

The authors thank Robert Landick and Rachel Anne Mooney for a plasmid containing promoter T7A1 and Michelle Wang and Chuang Tan for valuable discussion regarding formation of a stalled RNAP-DNA complex. This work was financially supported by a Swiss National Science Foundation (SNSF) Consolidator Grant (BIONIC BSCGI0 157802), by SNSF Grant (200021 153653, Glass nanopores as a versatile tool for single molecule analysis) and by HORIZON 2020 Grant (686271 AD-gut). S.M. was supported by the "Unity through Knowledge" Fund, Croatia under Grant 17/13 "Confined DNA".

■ REFERENCES

(1) McGhee, J. D.; von Hippel, P. H. *J. Mol. Biol.* **1974**, *86*, 469−89.
(2) Geertz, M.; Maerkl, S. J. *Briefings Funct. Genomics* **2010**, *9*, 362−373.
(3) Garner, M. M.; Revzin, A. *Nucleic Acids Res.* **1981**, *9*, 3047−3060.
(4) Galas, D. J.; Schmitz, A. *Nucleic Acids Res.* **1978**, *5*, 3157−3170.
(5) Roulet, E.; Busso, S.; Camargo, A. A.; Simpson, A. J.; Mermod, N.; Bucher, P. *Nat. Biotechnol.* **2002**, *20*, 831−835.
(6) Bulyk, M. L.; Huang, X.; Choo, Y.; Church, G. M. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 7158−7163.
(7) Maerkl, S. J.; Quake, S. R. *Science* **2007**, *315*, 233−237.
(8) Johnson, D. S.; Mortazavi, A.; Myers, R. M.; Wold, B. *Science* **2007**, *316*, 1497−1502.
(9) Barski, A.; Cuddapah, S.; Cui, K.; Roh, T.-Y.; Schones, D. E.; Wang, Z.; Wei, G.; Chepelev, I.; Zhao, K. *Cell* **2007**, *129*, 823−837.
(10) Nutiu, R.; Friedman, R. C.; Luo, S.; Khrebtukova, I.; Silva, D.; Li, R.; Zhang, L.; Schroth, G. P.; Burge, C. B. *Nat. Biotechnol.* **2011**, *29*, 659−664.
(11) Fägerstam, L. G.; Frostell-Karlsson, A.; Karlsson, R.; Persson, B.; Rönnberg, I. *Journal of chromatography* **1992**, *597*, 397−410.
(12) Geertz, M.; Shore, D.; Maerkl, S. J. *Proc. Natl. Acad. Sci. U. S. A.* **2012**, *109*, 16540−16545.
(13) Kasas, S.; Thomson, N. H.; Smith, B. L.; Hansma, H. G.; Zhu, X.; Guthold, M.; Bustamante, C.; Kool, E. T.; Kashlev, M.; Hansma, P. K. *Biochemistry* **1997**, *36*, 461−468.
(14) Hillisch, A.; Lorenz, M.; Diekmann, S. *Curr. Opin. Struct. Biol.* **2001**, *11*, 201−7.
(15) Woodside, M. T.; Anthony, P. C.; Behnke-Parks, W. M.; Larizadeh, K.; Herschlag, D.; Block, S. M. *Science* **2006**, *314*, 1001−1004.
(16) Fan, J.; Leroux-Coyau, M.; Savery, N. J.; Strick, T. R. *Nature* **2016**, *536*, 234−237.
(17) Bulushev, R. D.; Marion, S.; Radenovic, A. *Nano Lett.* **2015**, *15*, 7118−7125.
(18) Spiering, A.; Getfert, S.; Sischka, A.; Reimann, P.; Anselmetti, D. *Nano Lett.* **2011**, *11*, 2978−82.
(19) Sischka, A.; Spiering, A.; Khaksar, M.; Laxa, M.; König, J.; Dietz, K.-J.; Anselmetti, D. *J. Phys.: Condens. Matter* **2010**, *22*, 454121.
(20) Soni, G. V.; Dekker, C. *Nano Lett.* **2012**, *12*, 3180−3186.
(21) Plesa, C.; Ruitenberg, J. W.; Witteveen, M. J.; Dekker, C. *Nano Lett.* **2015**, *15*, 3153−3158.
(22) Raillon, C.; Cousin, P.; Traversi, F.; Garcia-Cordero, E.; Hernandez, N.; Radenovic, A. *Nano Lett.* **2012**, *12*, 1157−1164.
(23) Hornblower, B.; Coombs, A.; Whitaker, R. D.; Kolomeisky, A.; Picone, S. J.; Meller, A.; Akeson, M. *Nat. Methods* **2007**, *4*, 315−317.
(24) Sternberg, S. H.; Redding, S.; Jinek, M.; Greene, E. C.; Doudna, J. A. *Nature* **2014**, *507*, 62−67.
(25) Bell, N. A. W.; Keyser, U. F. *J. Am. Chem. Soc.* **2015**, *137*, 2035−2041.
(26) Squires, A.; Atas, E.; Meller, A. *Sci. Rep.* **2015**, *5*, 11643.
(27) Yu, J.-S.; Lim, M.-C.; Huynh, D. T. N.; Kim, H.-J.; Kim, H.-M.; Kim, Y.-R.; Kim, K.-B. *ACS Nano* **2015**, *9*, 5289−5298.

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX

(28) Laohakunakorn, N.; Gollnick, B.; Moreno-Herrero, F.; Aarts, D. G. A. L.; Dullens, R. P. A.; Ghosal, S.; Keyser, U. F. *Nano Lett.* **2013**, *13*, 5141−5146.
(29) Laohakunakorn, N.; Ghosal, S.; Otto, O.; Misiunas, K.; Keyser, U. F. *Nano Lett.* **2013**, *13*, 2798−2802.
(30) Laohakunakorn, N.; Thacker, V. V.; Muthukumar, M.; Keyser, U. F. *Nano Lett.* **2015**, *15*, 695−702.
(31) Laohakunakorn, N.; Keyser, U. F. *Nanotechnology* **2015**, *26*, 275202.
(32) Blackburn, M. C.; Petrova, E.; Correia, B. E.; Maerkl, S. J. *Nucleic Acids Res.* **2016**, *44*, e68−e68.
(33) Bulushev, R. D.; Steinbock, L. J.; Khlybov, S.; Steinbock, J. F.; Keyser, U. F.; Radenovic, A. *Nano Lett.* **2014**, *14*, 6606.
(34) Keyser, U. F.; Koeleman, B. N.; van Dorp, S.; Krapf, D.; Smeets, R. M. M.; Lemay, S. G.; Dekker, N. H.; Dekker, C. *Nat. Phys.* **2006**, *2*, 473−477.
(35) Keyser, U. F.; van der Does, J.; Dekker, C.; Dekker, N. H. *Rev. Sci. Instrum.* **2006**, *77*, 105105.
(36) Trepagnier, E. H.; Radenovic, A.; Sivak, D.; Geissler, P.; Liphardt, J. *Nano Lett.* **2007**, *7*, 2824−2830.
(37) Steinbock, L. J.; Otto, O.; Skarstam, D. R.; Jahn, S.; Chimerel, C.; Gornall, J. L.; Keyser, U. F. *J. Phys.: Condens. Matter* **2010**, *22*, 454113.
(38) Otto, O.; Steinbock, L. J.; Wong, D. W.; Gornall, J. L.; Keyser, U. F. *Rev. Sci. Instrum.* **2011**, *82*, 086102.
(39) Mulligan, M. E.; Hawley, D. K.; Entriken, R.; McClure, W. R. *Nucleic Acids Res.* **1984**, *12*, 789−800.
(40) Cho, S. W.; Kim, S.; Kim, J. M.; Kim, J.-S. *Nat. Biotechnol.* **2013**, *31*, 230−232.
(41) Wu, X.; Scott, D. A.; Kriz, A. J.; Chiu, A. C.; Hsu, P. D.; Dadon, D. B.; Cheng, A. W.; Trevino, A. E.; Konermann, S.; Chen, S.; Jaenisch, R.; Zhang, F.; Sharp, P. A. *Nat. Biotechnol.* **2014**, *32*, 670−676.
(42) Kuscu, C.; Arslan, S.; Singh, R.; Thorpe, J.; Adli, M. *Nat. Biotechnol.* **2014**, *32*, 677−683.
(43) Cencic, R.; Miura, H.; Malina, A.; Robert, F.; Ethier, S.; Schmeing, T. M.; Dostie, J.; Pelletier, J. *PLoS One* **2014**, *9*, e109213.
(44) O'Geen, H.; Henry, I. M.; Bhakta, M. S.; Meckler, J. F.; Segal, D. J. *Nucleic Acids Res.* **2015**, *43*, 3389−3404.
(45) Sternberg, S. H.; Redding, S.; Jinek, M.; Greene, E. C.; Doudna, J. A. *Nature* **2014**, *507*, 62−67.
(46) Marko, J. F.; Siggia, E. D. *Macromolecules* **1995**, *28*, 8759−8770.
(47) Wang, M.; Yin, H.; Landick, R.; Gelles, J.; Block, S. *Biophys. J.* **1997**, *72*, 1335−1346.
(48) Hummer, G.; Szabo, A. *Biophys. J.* **2003**, *85*, 5−15.
(49) Hernández-Ainsa, S.; Muus, C.; Bell, N. A. W.; Steinbock, L. J.; Thacker, V. V.; Keyser, U. F. *Analyst* **2013**, *138*, 104−106.
(50) Sola, L.; Chiari, M. *Journal of Chromatography A* **2012**, *1270*, 324−329.
(51) He, H.; Xu, X.; Jin, Y. *Anal. Chem.* **2014**, *86*, 4815−4821.
(52) Steinbock, L. J.; Krishnan, S.; Bulushev, R. D.; Borgeaud, S.; Blokesch, M.; Feletti, L.; Radenovic, A. *Nanoscale* **2014**, *6*, 14380−14387.
(53) Jarzynski, C. *Phys. Rev. Lett.* **1997**, *78*, 2690−2693.
(54) Rivetti, C.; Guthold, M.; Bustamante, C. *EMBO Journal* **1999**, *18*, 4464−4475.
(55) Jinek, M.; Jiang, F.; Taylor, D. W.; Sternberg, S. H.; Kaya, E.; Ma, E.; Anders, C.; Hauer, M.; Zhou, K.; Lin, S.; Kaplan, M.; Iavarone, A. T.; Charpentier, E.; Nogales, E.; Doudna, J. A. *Science* **2014**, *343*, 1247997−1247997.
(56) Smeets, R. M. M.; Keyser, U. F.; Krapf, D.; Wu, M.-Y.; Dekker, N. H.; Dekker, C. *Nano Lett.* **2006**, *6*, 89−95.
(57) Kim, S. C.; Kannam, S. K.; Harrer, S.; Downton, M. T.; Moore, S.; Wagner, J. M. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics* **2014**, *89*, 1−9.
(58) Kesselheim, S.; Müller, W.; Holm, C. *Phys. Rev. Lett.* **2014**, *112*, 018101.
(59) Aksimentiev, A. *Nanoscale* **2010**, *2*, 468−483.
(60) Dimitrov, V.; Mirsaidov, U.; Wang, D.; Sorsch, T.; Mansfield, W.; Miner, J.; Klemens, F.; Cirelli, R.; Yemenicioglu, S.; Timp, G. *Nanotechnology* **2010**, *21*, 065502.
(61) Wanunu, M.; Dadosh, T.; Ray, V.; Jin, J.; McReynolds, L.; Drndić, M. *Nat. Nanotechnol.* **2010**, *5*, 807−814.
(62) Venkatesan, B. M.; Estrada, D.; Banerjee, S.; Jin, X.; Dorgan, V. E.; Bae, M.-H.; Aluru, N. R.; Pop, E.; Bashir, R. *ACS Nano* **2012**, *6*, 441−450.
(63) Liu, K.; Feng, J.; Kis, A.; Radenovic, A. *ACS Nano* **2014**, *8*, 2504−2511.
(64) Fu, Y.-H.; Kuhl, D. P.; Pizzuti, A.; Pieretti, M.; Sutcliffe, J. S.; Richards, S.; Verkert, A. J.; Holden, J. J.; Fenwick, R. G.; Warren, S. T.; Oostra, B. A.; Nelson, D. L.; Caskey, C. *Cell* **1991**, *67*, 1047−1058.
(65) Steinbock, L. J.; Steinbock, J. F.; Radenovic, A. *Nano Lett.* **2013**, *13*, 1717−1723.
(66) Neuman, K. C.; Block, S. M. *Rev. Sci. Instrum.* **2004**, *75*, 2787−2809.
(67) Krummel, B.; Chamberlin, M. J. *Biochemistry* **1989**, *28*, 7829−42.

DOI: 10.1021/acs.nanolett.6b04165
Nano Lett. XXXX, XXX, XXX−XXX