# Supplemental information for: Single Molecule Localisation and Discrimination of DNA-Protein Complexes by Controlled Translocation Through Nanocapillaries

Roman D. Bulushev,[1,*] Sanjin Marion,[2,†] Ekaterina Petrova,[3,*]
Sebastian J. Davis,[1,*] Sebastian J. Maerkl,[3] and Aleksandra Radenovic[1]

[1] *Laboratory of Nanoscale Biology, Institute of Bioengineering, School of Engineering, EPFL, 1015 Lausanne, Switzerland*
[2] *Institute of Physics, Bijenička cesta 46, HR-10000 Zagreb, Croatia*
[3] *Laboratory of Biological Network Characterization, Institute of Bioengineering, School of Engineering, EPFL, 1015 Lausanne, Switzerland*

## S1. ANALYSIS OF FORCE AND CURRENT TRACES

Once the experimental force ($F_{ot}$) versus stage ($z$) traces were acquired we analysed them with a custom written script. The data extracted from experiments included: the DNA force level from which its linear charge density was calculated using $F_{DNA} \approx V\lambda_{DNA}$, the DNA exit point $y_x$, the DNA length, the distance from the protein jump to the exit point $\Delta y_p$, and the work $W_i$ involved in the protein exit from the capillary. To compute the work, each curve was shifted such that the force on DNA $F_{DNA}$ was 0. It was then numerically integrated to produce the curve $W_i(z) = \int_{z_A}^{z}(F_{ot} - F_{DNA})dz$ (See Fig. S1). The beginning of the jump $z_A$ was determined to be well before the protein event when the DNA level was flat so as to include any pre-jump changes in the force, or metastable jumping (see Fig. 1 in the main text). The final work attributed to the protein jumping and used in later analyses was $W_i \equiv W_i(z = z_B)$ where the end of the jump was determined based on the criteria that the curve needed to be locally flat, i.e that the force levels before and after the jump are at the same level.

In addition to force traces we obtained current traces. After their analysis we extracted a current change corresponding to DNA and protein and the location of the protein relative to the free end of the DNA.

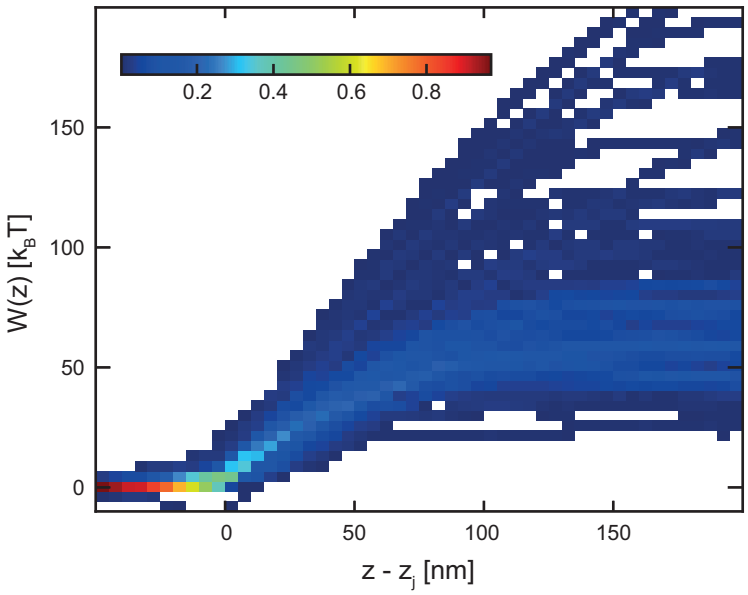


FIG. S1: **Density plot of calculated work versus stage W(Z) trajectories during a jump event.** The plot was normalized to a probability of 1 for each $z$ from 110 force versus stage curves for RNAP site at 750 nm. All force curves are shifted so that the jump location is at $z - z_j = 0$ before averaging.

*These authors contributed equally to this work
†These authors contributed equally to this work; All correspondence should be addressed to smarion@ifs.hr

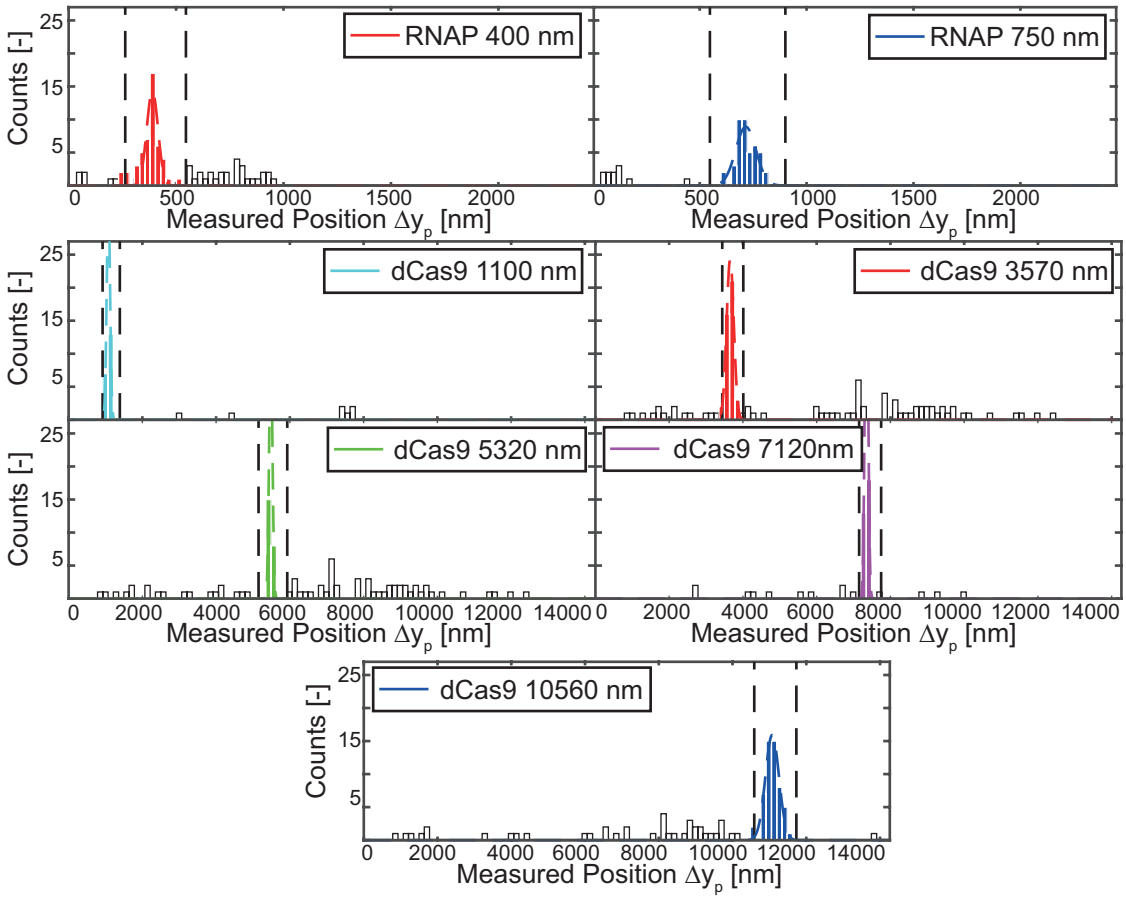


FIG. S2: **Histogram of protein localisation $\Delta y_p$ for each binding site.** Colored bins correspond to binding events identified as specific and white bins non-specifically bound proteins. Vertical, dashed, black lines represent the limit between specific and non-specific binding. See accompanying text for the criteria used to identify specific binding.

Once these values were obtained an elimination of traces based on the length of DNA was performed. Sometimes we observed DNA fragments shorter than expected. This fact can be a result of shearing due to experimental manipulation (pipetting, centrifugation, etc.)[1]. After this preliminary elimination a second round of analysis was used to discriminate specifically and non specifically bound proteins by implementing a modified $Z$ score[2],

$$Z_{score} = \left| \frac{0.6745 \cdot (x_i - \bar{x})}{MAD} \right|. \tag{S1}$$

If $Z_{score}$ was above 3.5 for a given trajectory then that protein event was classified as non specific. Here $x_i$ is the location of the protein for trace $i$, $\bar{x}$ is the median of the set, and $MAD$ is the median absolute deviation of the set given by the median of $|x_i - \bar{x}|$. The $Z$ score criteria was implemented to identify specific binding in all our cases (see Fig. S2)

## S2. LOCALISATION AND WORK HYSTERESIS

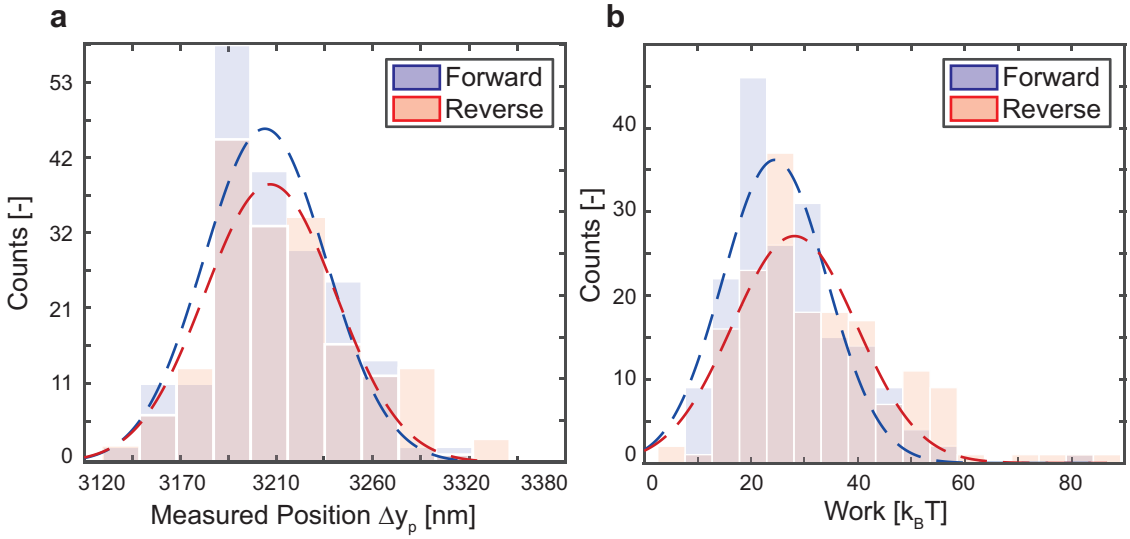


FIG. S3: **Localisation and work hysteresis measured for a DNA-protein complex a)** Forward (translocation from in- to outside the capillary) and reverse (from out- to inside) histograms of protein localisation $\Delta y_p$ performed to verify that localisation does not depend on the direction of the translocation protocol. Experiments were performed for dCas9 binding site at 3570 nm with 10 distinct protein-DNA complexes where each complex was probed with a forward and backward protocol multiple times producing from 20 to 40 individual protein jump events. For these experiments we used a sawtooth pattern for the nanopositioning stage, without allowing the DNA to exit. **b)** By performing back and forth experiments as described previously we compared the work performed in the forward and reverse events. A difference was observed between forward and reverse works which hints at the hysteresis of the cycle, but it was not significant enough for further analysis. Note that the definition of forward/backwards work will be different for proteins with a different charge sign.

## S3. ANALYTICAL SOLUTION FOR LOCALISATION SHIFT

In the experiment we do not measure the true position of the DNA-protein complex binding site, but it's apparent position. In order to determine the exact position of the binding site on the DNA, we needed to determine the relation between the stage position and the protein position when the jump is initiated. We connected the measured position of the protein at $y_p$ to the internal coordinate at $s_p$, also taking into account that we measure relative distances in respect to the DNA exit point at $y_x$. The measured protein position is the distance between the locations of the protein at $y_p$ and the DNA exit at $y_x$ locations:

$$\Delta y_p = y_x - y_p. \tag{S2}$$

The shift between the measured and theoretical position ($s_p$) is expressed as:

$$\Delta_{shift} = (L - s_p) - \Delta y_p, \tag{S3}$$

where $L - s_p$ comes from the fact that the experiments measure from the free end at $s = L$, while we measure contour distances from the bead at $s = 0$.

It remains to determine the relation between the DNA contour length $s$ and the distance between the bead and capillary opening $y$. This is, for any $s$, given by the extension $\mu$:

$$\mu(s) = \frac{y}{s}, \tag{S4}$$

and is always in the regime that $\mu(s) < 1$, i.e. the DNA is never "completely" extended. We can now calculate $\Delta y_p$ by transforming the bead to capillary distances into internal strand coordinates,

$$\Delta y_p = s_x \mu(s_x) - s_p \mu(s_p), \tag{S5}$$

where $s_x$ is the contour length between the bead and capillary at which the DNA exits. Without fluctuations this exit will happen when all the DNA is used for coiling at the tension and bead-opening distance such that $s_x = L$ and $y_x^0 = L\mu(L)$. In practice, as it is a single molecule experiment, the DNA will exit earlier due to fluctuations such that $y_x^0 = L\mu(L) > y_x = s_x\mu(L)$.

The length of a DNA strand that jumps out $\Delta s_x$ is not accessible for binding site localisation as the DNA jumps out even without the presence of the protein. This inaccessible length at the free end is the difference between the length at the moment of the ideal exit without fluctuations and the real exit:

$$\Delta s_x = s_x^0 - s_x = L - s_x. \tag{S6}$$

Here, the relationship between the length of strand where the jump took place and the experimental distance is:

$$y_x = s_x \mu(s_x). \tag{S7}$$

So that the experimental (or numerical) position is related to the inaccessible length $\Delta s_x$ through:

$$\Delta y_x = L\mu(L) - s_x\mu(s_x), \tag{S8}$$

or

$$\Delta y_x = L\mu(L) - (L - \Delta s_x)\mu(s_x). \tag{S9}$$

In practice the unprobeable region $\Delta s_x$ is not a problem since one can easily attach a piece of DNA to the end of the probed region thus rendering the DNA of interest wholly probeable. We also note here that the process of threading the free end of DNA is harder to predict and thus the entrance of DNA cannot be used instead of the exit as a reference point[3,4].

It remains to determine the DNA extension. In the strong stretching regime[5] the relation between the force extending the DNA, $F_{DNA} = \lambda_{DNA}\Delta V f(s)$, and the extension $\mu$ can be derived as:

$$\mu(s) = \frac{y}{s} = 1 - \frac{1}{\sqrt{A\, f(s)}}, \tag{S10}$$

with

$$f(s) = \begin{cases} 1 & \text{nanopore} \\ 1 - \frac{1}{1+\frac{L-s}{\xi}} & \text{nanocapillary,} \end{cases} \tag{S11}$$

and

$$A = \frac{4\lambda_{DNA} V L_p}{k_B T}. \tag{S12}$$

by taking into account the known force profile inside nanocapillaries[6]. Here $\lambda_{DNA}$ is the effective DNA linear charge density, $V$ the driving potential, $L_p$ the persistence length of DNA and $k_B T$ the thermal energy at room temperature. Equations (S3), (S5) and (S10) enable the conversion from the measured positions to the internal positions on the strand for bare DNA controlled translocation.

The previous equations neglect that the protein changes the force extending the DNA at the moment of the jump. In the regime when the force on the protein is small comparable to the force that is pulling the DNA into the nanocapillary, we can assume that, at the moment when the jump happens, the force extending the DNA is changed by the presence of the protein. We can approximate the force as $F_{ext} = F_{DNA} - F_p(s = s_p) = F_{DNA} - qV/\xi$ and obtain the corrected extension relation at $s_p$

$$\mu(s_p) = \frac{y}{s_p} = 1 - \frac{1}{\sqrt{4L_p F_{ext}/k_B T}}, \tag{S13}$$

with which we can predict the measured protein localisation shift for small forces on the protein. This formula is valid until the force on the protein becomes comparable to the force extending the DNA when we are no longer in the strong extension regime.

## S4. STOCHASTIC MODEL

### Coupled Langevin equations

We used a stochastic modelling scheme[7] previously implemented to explain controled translocation events in both nanopores[8] and nanocapillaries[9]. 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 DNA contour located between the bead and the capillary opening $s$. Thus we solved two Langevin equations with an external force determined by the total free energy $G(r,s)$. For the position of the bead $r = y - \rho$, with the bead radius $\rho = 1.5\ \mu m$, the temporal evolution is given by:

$$\eta_b \dot{r} = -\frac{\partial G(r,s)}{\partial r} + \sqrt{2k_B T \eta_b} g(t), \tag{S14}$$

and for the contour length of DNA $s$:

$$\eta_p \dot{s} = -\frac{\partial G(r,s)}{\partial s} + \sqrt{2k_B T \eta_p} g(t). \tag{S15}$$

Here $g(t)$ is a random, time-dependent, Gaussian, $\delta$-correlated noise of magnitude unity. The constants $\eta_b = 5 \cdot 10^{-5} pNs/nm$ and $\eta_p = 1 \cdot 10^{-5} pNs/nm$ are values for the Stokes friction for the bead and polymer, respectively, and $k_B T \approx 25.7$ meV is the thermal energy at room temperature. As in our previous work[9] the total free energy is:

$$G(r,s,z) = G_{DNA}(s) + G_{wlc}(r,s) + G_{ot}(r,z) + G_p(r,s), \tag{S16}$$

and consists of contributions from the optical trap $G_{ot}(r,z) = \frac{1}{2}\kappa(zr)^2$, the entropic free energy of a worm-like chain of DNA $G_{wlc}(r,s)$, the free energy of a charged DNA molecule in the nanocapillary $G_{DNA}(s)$, and the free energy contribution from a protein bound to the DNA $G_p(s)$ at the position $s_p$. 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 $\sim 100$ averages of pulling out protocols with random starting conditions.

### Free energy of the system

The elastic-entropic free energy $G_{wlc}$ for extending a DNA strand[5] is

$$G_{wlc}(r,s) = \frac{k_B T}{L_p}\left\{\frac{s}{4}\left[\frac{1}{1-(r-\rho)/s} - 1\right] - \frac{r-\rho}{4} + \frac{(r-\rho)^2}{2s}\right\}, \tag{S17}$$

with $L_p \approx 50$ nm the persistence length of DNA. In our regime, the force on a DNA persistence length segment is much larger than the thermal energy, resulting in a full extension of the DNA of contour length $s$ from the bead surface to nanocapillary opening. In the case of much longer capillary tips ($\xi \gg 100\ nm$) it might be necessary to include the details of the DNA extension.

Additionally, we have the free energy from the harmonic pulling force of the optical trap:

$$G_{ot}(r) = \frac{1}{2}\kappa(z-r)^2, \tag{S18}$$

where $\kappa$ is the optical trap spring constant $\kappa \approx 0.05$–$0.1\ pN/nm$. The bead equilibrium position $z$ was taken to be a function of time $z = vt$, where $v \approx 500\ nm/s$ was the experimental pulling speed.

The DNA electrostatic free energy is directly calculated from the electrostatic potential distribution $\Phi(z)$ inside the nanopore/capillary by approximating that the DNA is a charged extended rod with the effective linear charge density $\lambda$, with a constant contribution coming from the DNA contour length $s$ being on the energetically unfavourable side of the opening[8,9]

$$G_{DNA}(s) = \lambda_{DNA} \int_{-(L-s)}^{s} \Phi(x) dx. \tag{S19}$$

Here the position $x = 0$ is the location of the nanocapillary opening (exit) or the middle of the nanopore channel. The effective linear charge density of the DNA $\lambda_{DNA}$ includes both effects of counterion condensation and drag force from

any electroosmotic flow[8,10,11]. The effects of electroosmotic flow are larger in nanocapillaries than in nanopores[9,12,13]. Similarly to the DNA we can obtain the free energy for a point-like charged protein at a position $s_p$ along the DNA contour from the bead:

$$G_p(s) = \int \Phi(x) q^* \delta(x - (s - sp)) dx, \tag{S20}$$

here $q^*$ is the effective charge of the protein with both electrostatic and electroosmotic flow contributions[8]. The only difference in respect to our previous paper is in regards to the potential distributions used for nanocapillaries.

*Electrostatics of nanopores*

In the case of nanopores, the electric field acting on the DNA and protein is mostly localised around the pore opening[14]. We approximate the electrostatic potential in nanopores by using

$$\Phi(x) = V(\tanh(x/l_{eff}) + 1)/2, \tag{S21}$$

because when the effective electrostatic thickness of the nanopore $l_{eff}$ goes to 0 the potential transforms into a step function. Here $V$ is the driving electrostatic potential. From this the the free energy of a DNA strand is

$$G^{np}_{DNA} = -\frac{1}{2} l_{eff} \lambda_{DNA} V \left( \log\left( \cosh\left( \frac{L-s}{l_{eff}} \right) \right) - \log\left( \cosh\left( \frac{s}{l_{eff}} \right) \right) - L \right), \tag{S22}$$

and the resulting force $F_{DNA} = -\partial G_{DNA}/\partial s$:

$$F^{np}_{DNA} = -\frac{1}{2} \lambda_{DNA} V \left( \tanh\left( \frac{L-s}{l_{eff}} \right) + \tanh\left( \frac{s}{l_{eff}} \right) \right). \tag{S23}$$

For a protein represented as a point at location $s_p$ of charge $q$, the free energy is:

$$G^{np}_{p} = qV(\tanh((s - s_p)/l_{eff}) + 1)/2, \tag{S24}$$

leading to a force on the protein of:

$$F^{np}_{p} = \frac{qV}{2l_{eff}} (1 - \tanh^2((s - s_p)/l_{eff})). \tag{S25}$$

*Electrostatics of nanocapillaries*

The simplest model for the spatial dependence of the electrostatic potential valid inside a nanocapillary approximates the capillary as a cone[6,9]:

$$\Phi(x) = \frac{V}{1 - x/\xi} \theta(-x) + V\theta(x), \tag{S26}$$

with $\theta(x)$ the Heaviside step function. The first contribution comes into consideration when DNA is being pulled outside of the nanocapillary, giving a specific decay in the force profile[6]. The second contribution is the driving potential on the DNA due to its negative charge. Alternatively, a more accurate description of nanocapillaries can be made using a two cone geometry in the case of shrinking of nanocapillaries under SEM[9,15]. In our current case we found that one cone is sufficient for explaining the experimental results.

From the electrostatic potential of a nanocapillary follows the free energy of a DNA strand

$$G^{ncp}_{DNA} = V\lambda_{DNA} \left( \log\left( 1 + \frac{L-s}{\xi} \right) + s \right), \tag{S27}$$

and subsequently the force on the DNA:

$$F^{ncp}_{DNA} = -\lambda_{DNA} V \left( 1 - \frac{1}{1 + \frac{L-s}{\xi}} \right). \tag{S28}$$

The free energy of a point protein at location $s_p$ with charge $q^*$ is:

$$G_p^{ncp} = -q^* V \frac{1}{1+\frac{s-s_p}{\xi}}\theta(s_p - s) + q^* V \theta(s - s_p), \tag{S29}$$

and the force on the protein:

$$F_p^{ncp} = -\frac{q^* V}{\xi}\frac{1}{\left(1+\frac{s-s_p}{\xi}\right)^2}. \tag{S30}$$

**Controlled translocation of DNA-protein complexes**

Using the above relations for the free energy we can simulate the experimental conditions with a nanocapillary, and provide comparisons to nanopores. When translocating bare DNA (Fig. S4) the only striking difference between nanopores and nanocapillaries is the final part of the force profile. In the case of nanocapillaries a characteristic decay of the force acting on DNA is observed[6]. This is also seen in the increased coiling (inverse extension) during the exit. We also note that the simulations show that the fluctuations are largest at the beginning of the pulling, i.e for protein positions close to the bead.

Addition of a protein on the DNA produces a jump event at the location of the protein and relaxation of the coiled DNA strand (Fig. S5). The event is characterized by an abrupt change of the force during which the protein jumps trough the opening of the capillary to minimize its free energy. This is followed by a change of the contour length $s$ and correspondingly in the coiling (inverse extension). For the same charge, nanocapillaries provide a longer jump event due to the force on the protein being more spatially spread than in nanopores. Nanocapillaries are also more susceptible to induce metastable jumping of complexes, at the beginning of the jump, attributed to a energy barrier comparable to the thermal energy. We can also compare how the protein localisation shift depends on the geometry (Fig. S6a). For nanocapillaries the shift is larger comparing to nanopores as the DNA exit position exhibits a lower tension due to the specifics of the electrostatic potential decay. Fig. S6b shows a comparison of how in our case the analytical fit produces an underestimation of the true protein charge value due to the strong stretching approximation we used. Fig. S7 shows how the shift depends on a different driving voltage, which influences the DNA tension and thus extension, and on the shape of the capillary via the electrostatic decay length $\xi$.

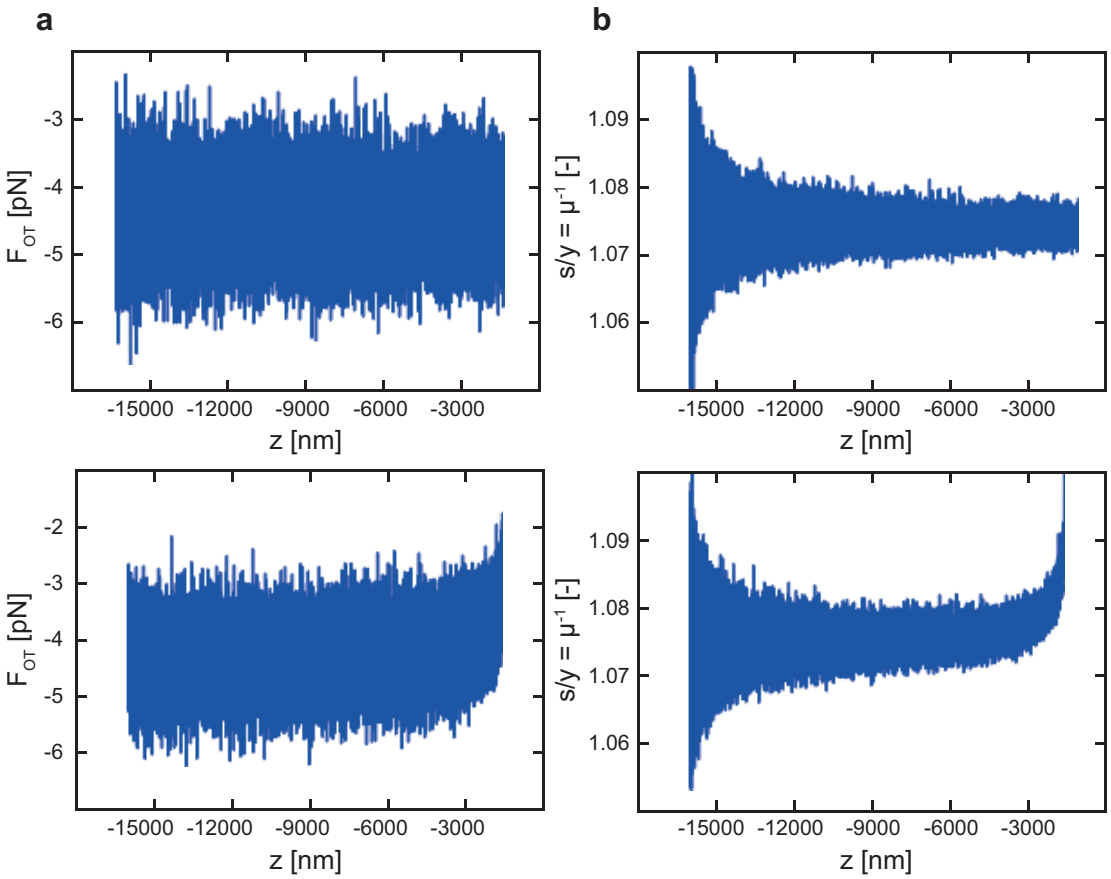


FIG. S4: **Comparison of DNA controlled translocation simulations for nanopore (top row) and nanocapillary (bottom row) geometries a)** A characteristic force versus stage position plot. **b)** Characteristic DNA coiling $y/s$ (inverse extension) versus stage position $z$. The parameters used were $\lambda_{DNA} = -0.04\ e/bp$ and $V = 200\ mV$ with the electrostatic decay length $\xi = 200\ nm$ for nanocapillaries and thickness $l_{eff} = 10\ nm$ for nanopores.

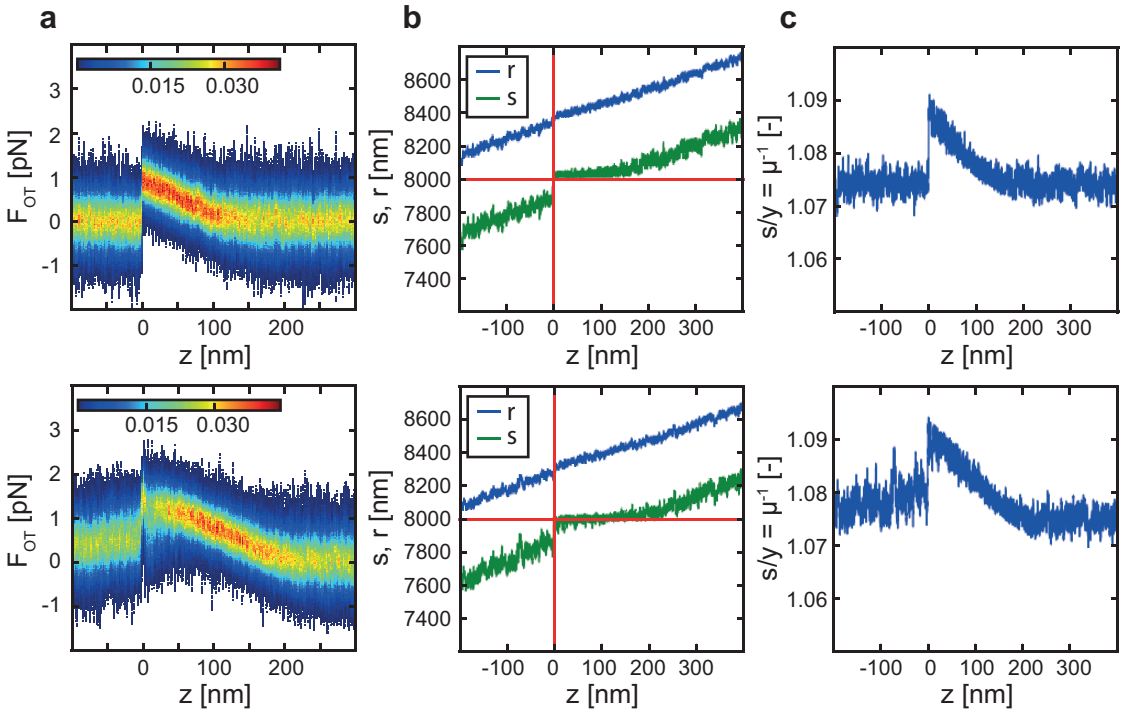


FIG. S5: **Comparison of characteristic DNA-protein jumps during controlled translocation in nanopores (top row) and nanocapillaris (bottom row) a)** Force versus stage position $z$ averaged for 100 different numerical starting conditions and normalized so that the total probability for each $z$ is equal to 1 (scale on the density plot shows this probability) **b)** Internal state variables $r, s$ versus stage position $z$. A single experiment is demonstrated with the horizontal red line marking the protein position $s_p$ and vertical red line the protein location which corresponds to $z = 0$, **c)** DNA coiling $s/y$ (inverse extension) versus stage position $z$. All stochastic model parameters are the same as in Fig. S4 except the protein was added with a charge of $q^* = 10\ e$

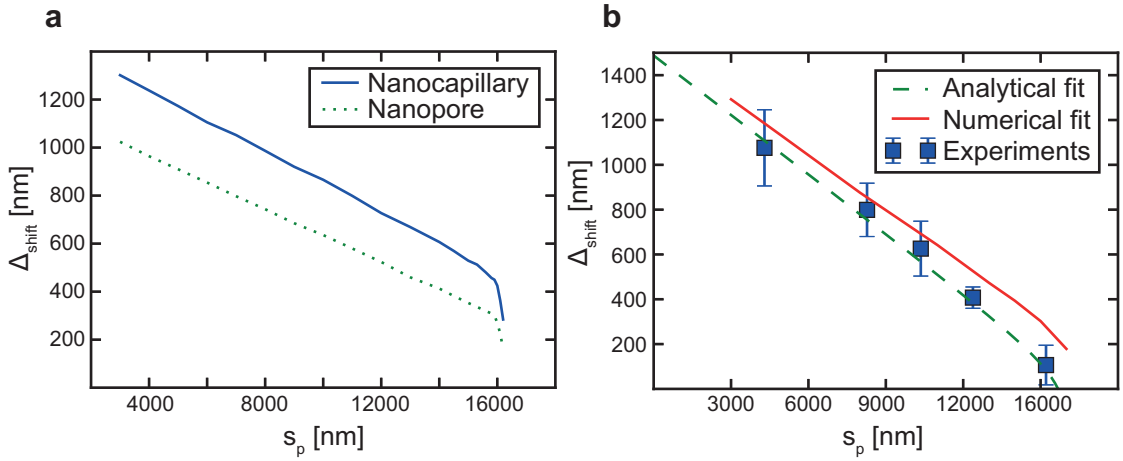


FIG. S6: **Variability of the experimental set-ups and fitting function a)** Comparison of nanopore and nanocapillary localisation shifts for parameters $V = 200\ mV$, $\lambda_{DNA} = -0.04\ e/bp$, $q^* = 5e$ and $\xi = 75\ nm$ for nanocapillary while $l_{eff} = 10\ nm$ for nanopore. **b)** Comparison of nanocapillary analytic fit ($V = 175\ mV$, $\lambda_{DNA} = -0.065\ e/bp$, $l_{eff} = 10\ nm$) and stochastic model numerical fit ($V = 175\ mV$, $\lambda_{DNA} = -0.04\ e/bp$, $\xi = 75\ nm$) for the same protein of charge $q^* = 5e$.

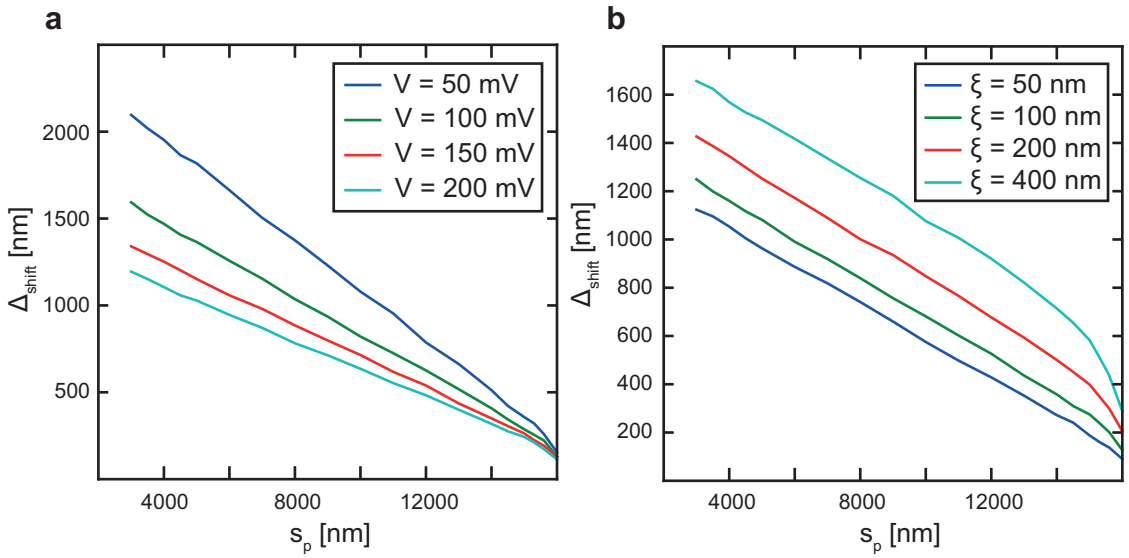


FIG. S7: **Numerically derived protein localisation shift while varying voltage and electrostatic decay length a)** Effect of different driving voltages on the localisation shift. **b)** Effect of different electrostatic decay lengths $\xi$ on the localisation shift. Parameters used are the same as in the preceding figure except when noted differently on the legend.

## S5. EQUILIBRIUM INFORMATION FROM PROTEIN JUMP EVENTS

The Jarzynski equality (JE)[16] enables us to determine the free energy difference $\Delta G_{AB} = G(r_B, s_B, z_B) - G(r_A, s_A, z_A)$ between two positions before and after the jump with $G(r, s, z)$ the total free energy of the system (S16). The system is considered to be in an equilibrium state before the jump happens at state variables $(r_A, s_A, z_A)$. After the non-equilibrium jump it relaxes into a new equilibrium state at $(s_B, r_B, z_B)$. The work $W_i$ measured for the individual protein jump was averaged over different experimentally realized trajectories $\langle W_i \rangle = \bar{W} \geq \Delta G_{AB}$, where the equality sign is valid when the number of samples goes to infinity. We obtained the difference between the two states with the JE:

$$e^{-\frac{\Delta G_{AB}}{k_B T}} = \langle e^{-\frac{W_i}{k_B T}} \rangle. \tag{S31}$$

In order to use the JE to obtain a meaningful physical quantity related to the protein we computed the difference between the free energy before and after the jump. We assumed that the measured force before and after the jump event is the same $F_{ot}(r_A, z_A) = F_{ot}(r_B, z_B)$, which implies that differences between stage and bead coordinates are the same $\Delta z = \Delta r$ and that the free energy of the optical trap does not change $\Delta G_{ot} = 0$. What remains is the free energy of the DNA worm-like chain (S17), DNA electrostatic free energy (S19) and protein free energy (S20).

The difference of free energies for the DNA before and after the jump for thin nanopores ($l_{eff} \ll 10nm$) is

$$\Delta G_{DNA} = V \lambda_{DNA} \Delta s, \tag{S32}$$

and for nanocapillaries

$$\Delta G_{DNA} = V \lambda_{DNA} \Delta s - V \lambda_{DNA} \xi \log \frac{\xi + L - s_B}{\xi + L - s_A}. \tag{S33}$$

Nanocapillaries have the additional term which is relevant when the DNA's free end is near the opening so that it is influenced by the characteristic potential (S26). If the protein location is far away from the DNA free end ($s_p \ll L$) this is negligible.

The protein free energy difference is $\Delta G_p = q^* V$ for nanopores, and

$$\Delta G_p = q^* V - \frac{q^* V}{1 + \frac{s_p - s_b}{\xi}}, \tag{S34}$$

for nanocapillaries. We can assume that the DNA strand position after the jump is the same as the protein position, $s_B \approx s_p$. This is confirmed by our stochastic model in Fig. S5 where $s_B \approx s_p$ even after the DNA has again extended and the force level has returned to its value before the jump. We can then further simplify by also writing $s_A \approx s_p - \Delta s$ leading to

$$\Delta G_p = q^* V \frac{\Delta s}{\xi + \Delta s}. \tag{S35}$$

If the protein is far away from the free end the change in DNA extension before and after the jump is negligible so that $\mu_A \approx \mu_B = \mu$. This is valid, unless we are dealing with proteins positioned near to the free end of the DNA for nanocapillaries. Finally the change in the free energy of the DNA worm like chain is

$$\Delta G_{wlc} \approx \frac{k_B T}{L_p} \frac{\Delta z}{4} \left[ \frac{\mu}{1 - \mu} - (1 - 2\mu) \right], \tag{S36}$$

which amounts to $\Delta G_{wlc} \ll 1\ k_B T$ and is negligible for discrimination of proteins

To summarize, the change of the free energy before and after the jump can be approximated by eq. (S35) which can be further simplified with $\Delta s \approx \Delta z$, valid for short jumps as the extension which links the two coordinates $\mu \approx 1$. With this equation we are able to use the free energy difference obtained from non-equilibrium work analysis (JE) and connect it to the effective charge of the bound protein.

In order to check that our use of the JE on our experimental data is warranted we observe the convergence of the resulting free energy $\Delta G_{AB}$. We computed the free energy of a total of $N$ measured works using an average over $n < N$ possible combinations of works. While the number of included works $n$ grows from 1 to $N$ we should expect a convergence of the JE with $n$ since the ensemble average should more precisely reflect the infinite average expressed in equation S31. On Fig. S8 we see the expected convergence. The fact that the curves, both for RNAP and dCas9 become flat with our number of experimentally measured works shows that we can use the JE to estimate the free energy and thus the effective charge of the proteins.

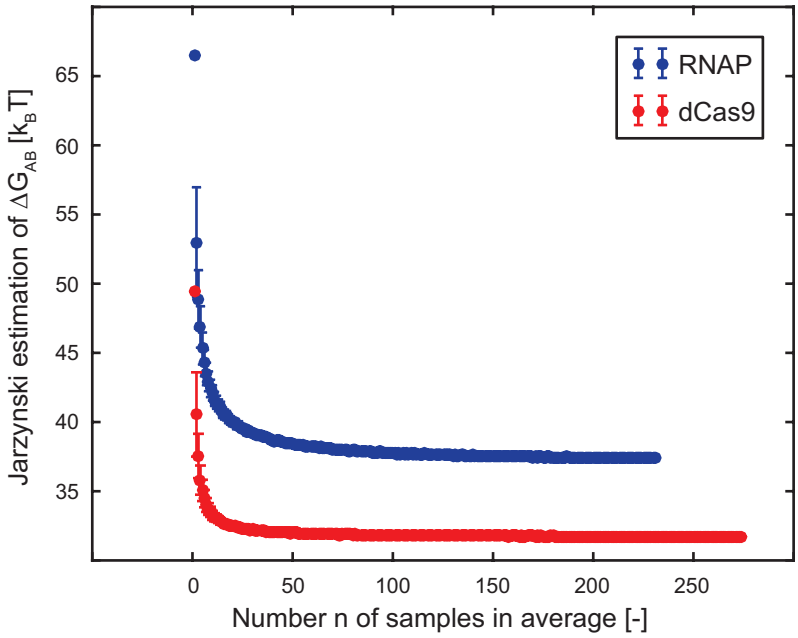


FIG. S8: **Convergence of Jarzinsky estimator for Fig. 4a in the main text.** Value of $\Delta G_{AB}$ from the JE for all possible $n$ combinations of works with $n = 1$ to $N$. A clear convergence is observed for a growing number of samples $n$. The curve being flat in the region of $N$ number of measurements we consider that, in our case, the JE is a good estimator of $\Delta G_{AB}$.

## S6. ADDITIONAL FIGURES

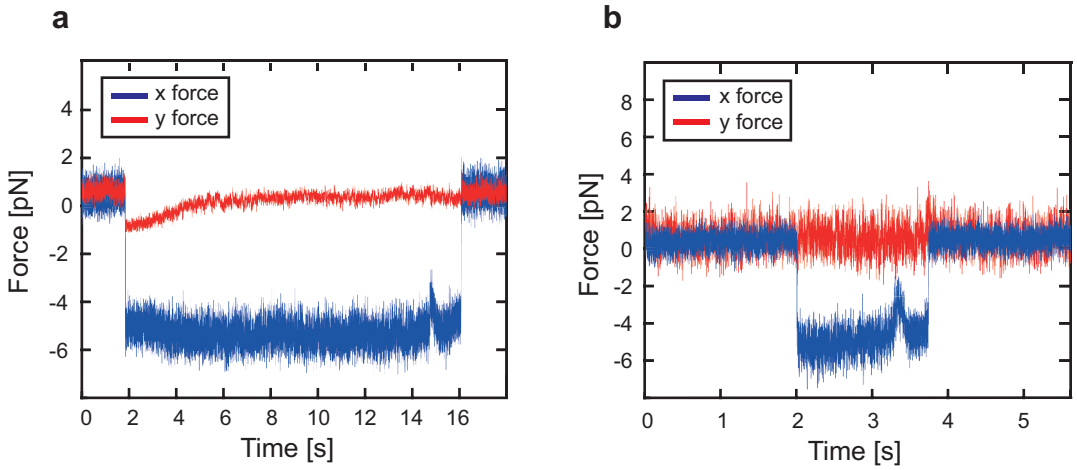


FIG. S9: **Examples of y force traces.** a) Typical example traces of x and y force as a function of time for dCas9. X force shows typical trace as described in FIG. 1 of the main text. Y force trace shows little response throughout the protocol apart from a small variation on DNA capture. b) As in panel a) but for typical RNAP trace and no response at all in the y direction.

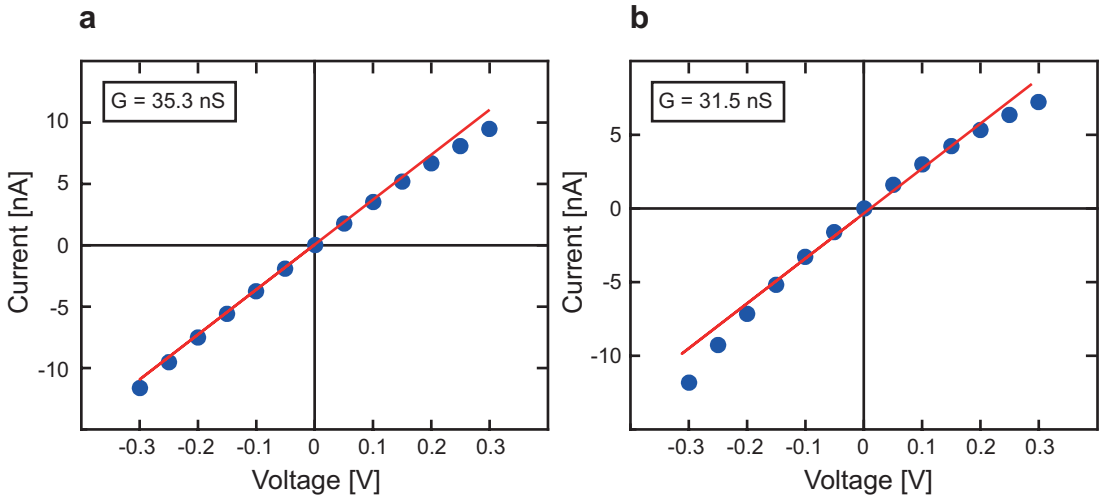


FIG. S10: **Current versus voltage characteristics for two typical nanocapillaries.** a) Typical current versus voltage curve of a capillary used in experiments with dCas9. Full line represents a linear fit around zero voltage. The current grows sublinearly at positive voltages, and superlinearly at negative voltages, as expected for rectification in nanopores of conical geometry[17]. b) As in panel a) but for typical capillary used with RNAP.

[1] H. Yoo, H. Lim, I. Yang, S. Kim, and S. Park, J. Biophys. Chem. **2**, 102 (2011).
[2] B. Iglewicz and D. C. Hoaglin, *How to Detect and Handle Outliers*, ASQC basic references in quality control (ASQC Quality Press, 1993).
[3] E. H. Trepagnier, A. Radenovic, D. Sivak, P. Geissler, and J. Liphardt, Nano Lett. **7**, 2824 (2007).
[4] M. Wanunu, W. Morrison, Y. Rabin, A. Y. Grosberg, and A. Meller, Nat. Nanotechnol. **5**, 160 (2010).
[5] J. F. Marko and E. D. Siggia, Macromolecules **28**, 8759 (1995).
[6] R. D. Bulushev, L. J. Steinbock, S. Khlybov, J. F. Steinbock, U. F. Keyser, and A. Radenovic, Nano Lett. **14**, 6606 (2014).
[7] G. Hummer and A. Szabo, Biophys. J. **85**, 5 (2003).
[8] A. Spiering, S. Getfert, A. Sischka, P. Reimann, and D. Anselmetti, Nano Lett. **11**, 2978 (2011).
[9] R. D. Bulushev, S. Marion, and A. Radenovic, Nano Lett. **15**, 7118 (2015).
[10] S. Ghosal, Phys. Rev. E **76**, 1 (2007).
[11] S. van Dorp, U. F. Keyser, N. H. Dekker, C. Dekker, and S. G. Lemay, Nat. Phys. **5**, 347 (2009).
[12] A. R. Hall, S. van Dorp, S. G. Lemay, and C. Dekker, Nano Lett. **9**, 4441 (2009).
[13] N. Laohakunakorn, B. Gollnick, F. Moreno-Herrero, D. G. A. L. Aarts, R. P. A. Dullens, S. Ghosal, and U. F. Keyser, Nano Lett. **13**, 5141 (2013).
[14] B. Lu, D. P. Hoogerheide, Q. Zhao, and D. Yu, Phys. Rev. E **86**, 011921 (2012).
[15] L. J. Steinbock, R. D. Bulushev, S. Krishnan, C. Raillon, and A. Radenovic, ACS Nano **7**, 11255 (2013).
[16] C. Jarzynski, Phys. Rev. Lett. **78**, 2690 (1997).
[17] N. Laohakunakorn and U. F. Keyser, Nanotechnology **26**, 275202 (2015).