Supporting Information for

Single-Molecule Reaction Chemistry in Patterned Nanowells

Delphine Bouilly, Jason Hon, Nathan S. Daly, Scott Trocchia, Sefi Vernick, Jaeeun Yu, Steven Warren, Ying Wu, Ruben L. Gonzalez, Jr.*, Kenneth L. Shepard*, and Colin Nuckolls*

Correspondence to: cn37@columbia.edu, rlg2118@columbia.edu, and shepard@ee.columbia.edu

Table of content:

1. Materials and methods

Device fabrication. Long carbon nanotubes were grown by chemical vapor deposition across a 1 x 1 cm oxidized silicon substrate (525 μm thick degenerately-doped silicon covered with 285 nm of thermal oxide). Horse-spleen ferritin was deposited on the substrate edge and annealed to form iron catalyst particles for nanotube nucleation. Directional growth of nanotubes was achieved by flowing a mixture of argon and hydrogen gas through ethanol and over the substrate at 890 °C[1,2]. Two types of devices were used in this study:

(1) Dense arrays of devices were used to develop and test the nanowell-confined chemistry method. Gold alignment marks were defined after nanotube growth in an 8 mm x 8 mm grid with a pitch of 50 μm. High-resolution scanning electron microscopy was used to scan the substrate and identify the coordinates of long nanotubes. One nanotube was chosen and isolated using e-beam lithography and O_2 plasma reactive-ion etching (RIE). Using e-beam lithography, a dense array of electrodes was defined along the nanotube, comprising 20 blocks of 35 electrode pairs separated by 4 μm each. The electrodes were made by evaporating Ti (0.5 nm), Pd (20 nm) and Au (50 nm), with Pd ensuring good electrical contacts to the nanotube[3].

(2) For real-time electrical measurements, smFET devices were fabricated. Thirty-two parallel Ti-electrode pairs with source-drain separation of 4 μm were patterned perpendicular to the direction of nanotube growth, after which the substrate was annealed in vacuum at 350 °C. Titanium forms low-resistance electrical contacts with nanotubes[3], as well as a native oxide layer after air exposure that passivates the electrodes. Scanning electron microscopy, Raman spectroscopy and local electrical probing were used to select a nanotube candidate to form the device channels, and the remaining nanotubes were removed using O_2 plasma reactive-ion etching (RIE). A microfluidic flow cell (800 μm large x 7 mm long) was made in polydimethylsiloxane (PDMS) and stamped over the device array to allow measurement in aqueous conditions. Devices were gated using either the substrate back gate (V_G) or electrolytically in saline buffer using on-chip Pt bars as pseudo-reference electrodes (V_L).

Single-point functionalization. Nanowells were patterned in a thin layer of poly(methyl methacrylate) (PMMA A2 950k), spincoated at 5000 rpm for 60s. The PMMA thickness was measured as approximately 70 nm using AFM. Writing was done using a high-resolution electron beam lithography writer (NanoBeam nB4). To maximize the resolution, writing was done a low current (1 nA) and patterns were developed in a 4 °C solution of isopropyl alcohol and deionized water (3:1 IPA:H_2O). Nanowells size and full-depth development were assessed using AFM and by evaporating a thin metal layer (Ti, 8 nm) inside the nanowells followed by lift-off. Single-point functionalization was obtained inside the smallest 20 nm wide nanowells, using aryldiazonium chemistry[4]. The 4-carboxybenzenediazonium tetrafluoroborate (CBDT) reagent was synthesized according to McNab et al.[5] Functionalization was done by dissolving 10mM CBDT in aqueous phosphate buffer (100 mM, pH 8) and immediately immersing mask-covered devices in the solution. Devices were incubated in the solution at room temperature for 24 h to saturate the yield of attached functional groups[4]. Although this chemistry can usually be done in a variety of solvents, aqueous conditions are necessary here to prevent dissolution of the PMMA mask.

DNA secondary functionalization. Secondary DNA functionalization was obtained on the carboxyphenyl-nanotube adduct after dissolution of the nanowells mask. The DNA oligomer was purchased HPLC-grade from Integrated DNA Technologies (Integrated DNA Technologies, HPLC grade), with the following sequence: 5'-$NH_2(CH_2)_6$ AGA GGG TTA GGG TTA GGG TTA GGG-3' with the alkyl amine functionalization placed at the 5' terminus. Once received, the DNA was suspended in sodium phosphate buffer (100mM pH 8.4) to obtain a DNA concentration of 100 μM. Several aliquots were prepared and kept in the freezer until the experiment. Tethering the DNA to the carboxyl group was done using a two-step EDC/NHS coupling protocol[6]. In the first reaction, the carboxyl termination of the single-molecule probe was activated using 1 mM EDC (Thermo Scientific) and 1 mM N-hydroxysuccinimide (NHS, Sigma-Aldrich 98%) in

dimethylformamide (DMF) for 10 min[7]. In the second reaction, we introduced 100 μM of the 5'-amino-modified DNA oligomer in a sodium phosphate buffer solution (100 mM, pH 8.4) for 30 min. The 5' amino moiety reacts with the NHS ester[6] to form an amide bond between the 5' terminus of the DNA oligomer and the carboxyphenyl single-molecule probe on the nanotube. After the reaction, the flow cell was flushed with Tris-HCl buffer (10 mM, pH 7.5) for 10 min at 45 °C to rinse away any residual unconjugated DNA oligomer, EDC and NHS traces and the Na^+ ions that were introduced from the sodium phosphate buffer.

Electrical measurements and analysis. Large arrays of carbon nanotube devices as illustrated in Supplementary Fig. 4 were probed using an automatic probe station (Summit 12k, Cascade Microtech). Current voltage characteristics were measured, before and after functionalization, by sweeping the back gate voltage V_G from -10 V to 10 V while holding the source and drain potentials at V_S = 50 mV and V_D = 0 mV. Voltages were applied and currents were read using a semiconductor parameter analyzer (Agilent 4155C).

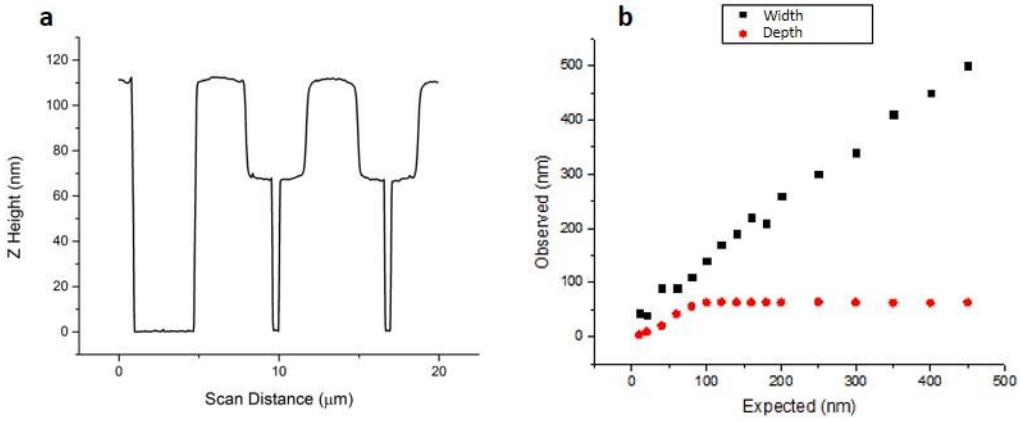
Real-time electrical measurements on smFET devices were realized using a custom printed circuit board, designed to operate, gate and measure up to 11 devices simultaneously. The board is composed of 11 measurement channels, each connected to a dedicated transimpedance amplifier with 1M front-end gain. An adjustable gain stage follows the transimpedance amplifier, with selectable values of 2x, 10x, 20x, 100x, and 200x. The signal is filtered by a second- order 5 kHz anti-aliasing filter and then sampled at a rate of 20 kS/s per channel. For real-time measurements, devices were typically kept at 100 mV DC source-drain bias (V_{SD}) and at – 300 mV (V_L) electrolytic gate bias using on-chip Pt bars immersed in saline buffer.

Real-time data was analyzed using a hidden Markov model (HMM) to simultaneously obtain the most probable sequence of low-G and high-G states (idealized path) and construct the baseline-corrected trajectories, as previously described[8–10]. In the case of DNA experiments, visual inspection of the raw data revealed the absence of transitions in the millisecond timescale, allowing us to filter and decimate the data before HMM analysis using a built-in *resample()* MATLAB function using a 100x down-sampling factor with a 2nd order FIR filter. Each idealized path was visually verified and corrections were applied manually when necessary (5 in the K^+ trace and 3 in the Na^+ trace). For each conductance state, all dwell times were compiled in a histogram representing the survival probability of this state (Supplementary Fig. 13). The obtained distributions were fitted with a double exponential decay function, indicating the presence of two distinct characteristic lifetimes for each state, as compiled in Supplementary Table 2. We evaluated the stability of the estimation of the rates with respect to incorrect classification of a dwell from each time series by cross validation: a new time series was formed through absorbing a randomly drawn dwell into its previous

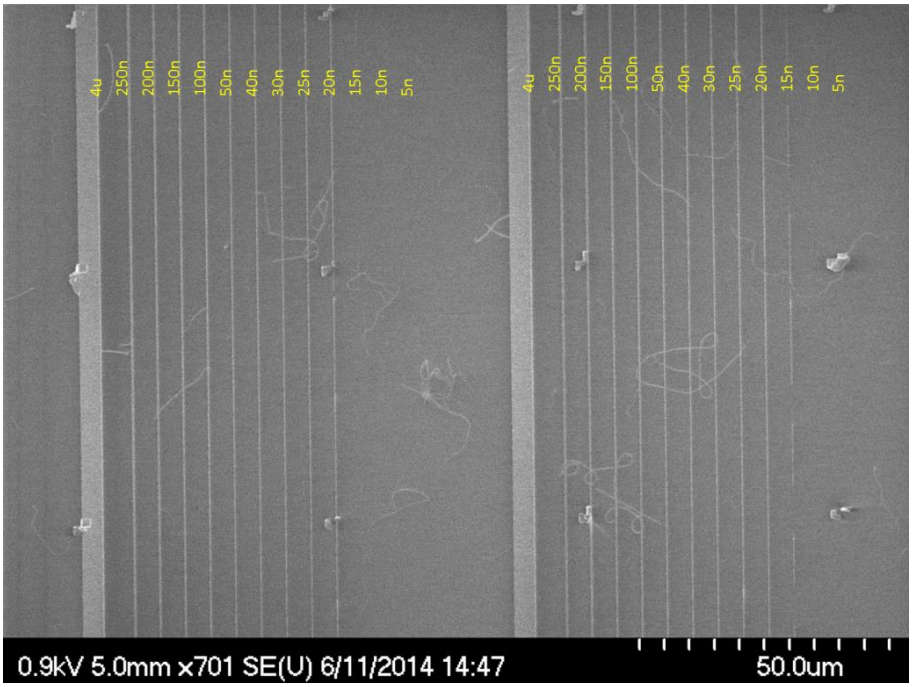
and next dwell by flipping its class, and the rates were re-estimated with the new time series. The results of this analysis repeated 1000 times are shown in Supplementary Fig. 14 and reveal that the fast lifetimes are mostly insensitive to misclassifications while the slow dwell times are not, because faster lifetimes are sampled more often than slow lifetimes. Finally, error bars were generated for each lifetime by bootstrapping the idealized traces, *i.e.* using the obtained lifetimes to re-simulate the Markov chain a hundred times (with the same number of points) and taking the variance of the obtained distribution as the error.

References:
(1) Hilty, S.; Webb, B.; Frankel, R. B.; Watt, G. D. *J. Inorg. Biochem.* **1994**, *56*, 173.
(2) Li, Y.; Kim, W.; Zhang, Y.; Rolandi, M.; Wang, D.; Dai, H. *J. Phys. Chem. B* **2001**, *105*, 11424.
(3) Chen, Z.; Appenzeller, J.; Knoch, J.; Lin, Y.; Avouris, P. *Nano Lett.* **2005**, *5*, 1497.
(4) Schmidt, G.; Gallon, S.; Esnouf, S.; Bourgoin, J. P.; Chenevier, P. *Chem. Eur. J.* **2009**, *15*, 2101.
(5) McNab, H.; Monahan, L. C. *J. Chem. Soc. Perkin Trans. 1* **1989**, *53*, 419.
(6) Hermanson, G. T. In *Bioconjugate Techniques*; Elsevier, 2013; pp 229–258.
(7) Fujisaki, F.; Oishi, M.; Sumoto, K. *Chem. Pharm. Bull. (Tokyo).* **2007**, *55*, 124.
(8) Bruno, W. J.; Ullah, G.; Daniel Mak, D. O.; Pearson, J. E. *Biophys. J.* **2013**, *105*, 68.
(9) Bronson, J. E.; Fei, J.; Hofman, J. M.; Gonzalez, R. L.; Wiggins, C. H. *arXiv:0907.3156 [q-bio.QM]* **2009**.
(10) Sorgenfrei, S.; Chiu, C. Y.; Gonzalez, R. L.; Yu, Y.-J. J.; Kim, P.; Nuckolls, C.; Shepard, K. L.; Gonzalez Jr., R. L.; Yu, Y.-J. J.; Kim, P.; Nuckolls, C.; Shepard, K. L.; Gonzalez Jr., R. L.; Yu, Y.-J. J.; Kim, P.; Nuckolls, C.; Shepard, K. L. *Nat. Nanotechnol.* **2011**, *6*, 126.

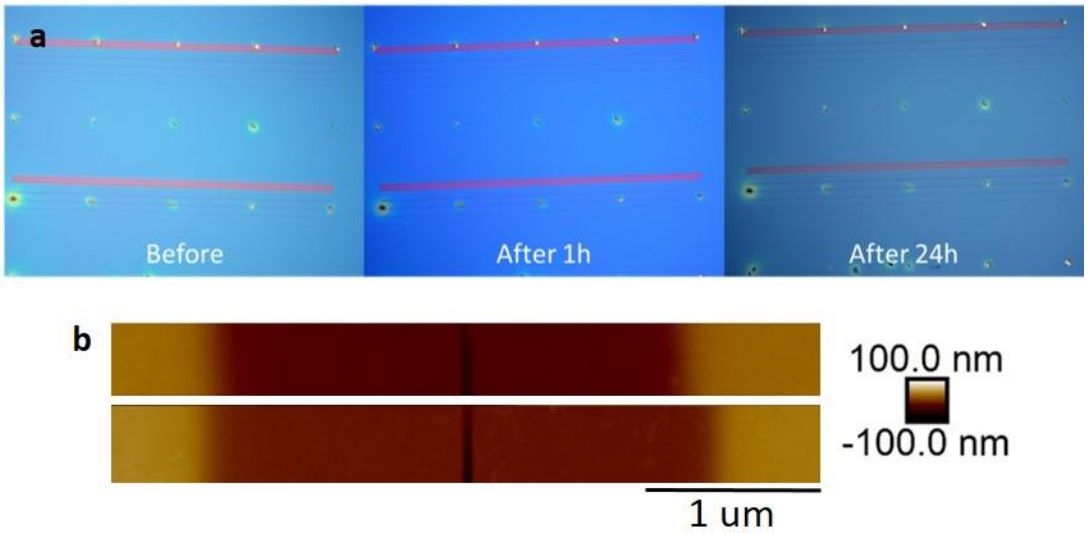
Mask characterization



Supplementary Figure 1. AFM characterization of PMMA masks. (a) Profile height for three masks with nanowells of different widths, separated by electrodes. (b) Measured width and depth of various nanowells as a function of nominal width.

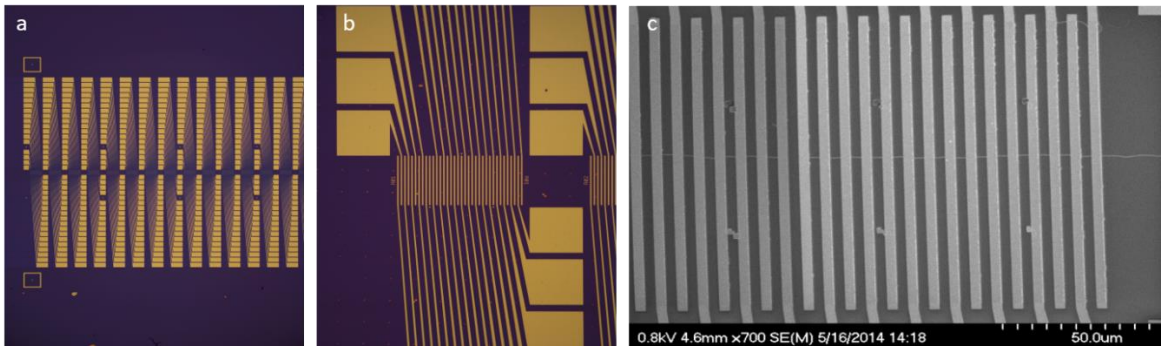


Supplementary Figure 2. SEM characterization of metal patterns produced with PMMA masks. Metal lines (Ti, 8 nm thick) evaporated through masks with different nanowell widths and imaged with SEM. We see that nanowells down to 20 nm form reproducible metallic lines indicating that the mask was developed all the way down to the substrate.

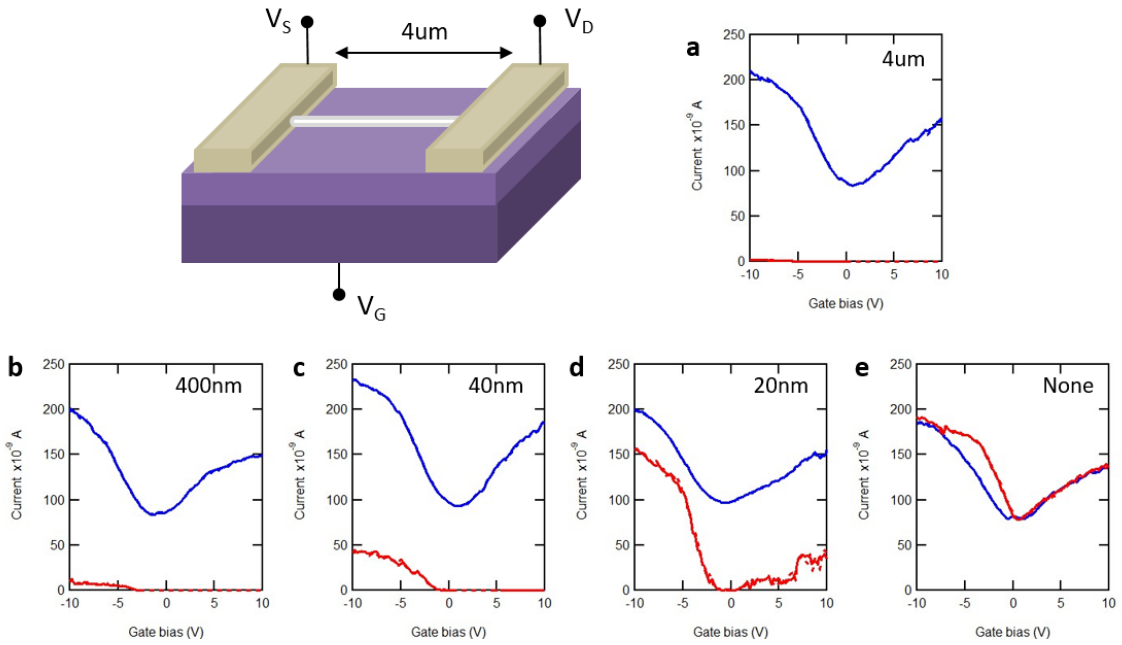


Supplementary Figure 3. Stability of PMMA masks. (a) Optical images before and after immersion for 24h in phosphate buffer solution (0.1M, pH 8). (b) AFM images of a mask before and after reaction in CBDT solution.

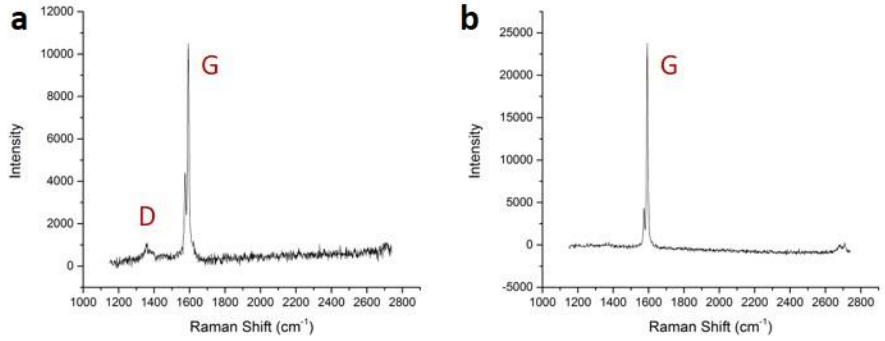
2. Device and functionalization characterization



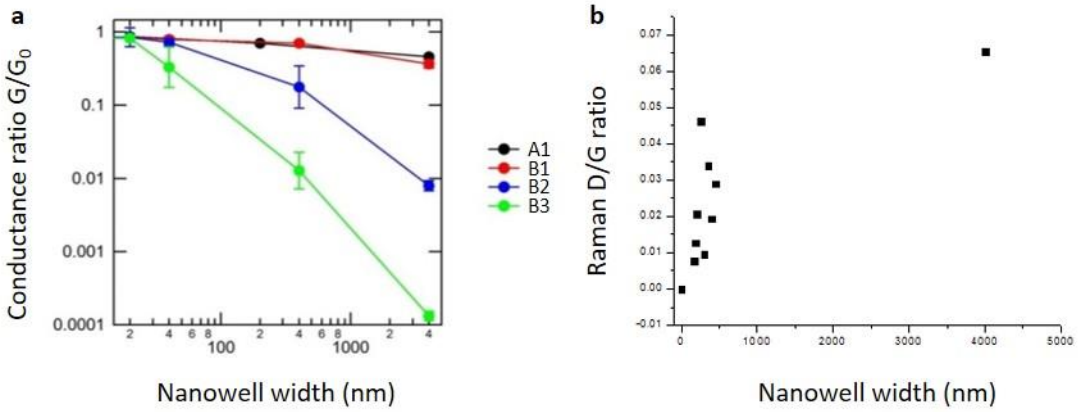
Supplementary Figure 4. Layout for large arrays of carbon nanotube devices. Up to 700 devices are fabricated over a 1cm-long carbon nanotube by defining densely-packed arrays of electrodes. (a) Optical image, zoom 2.5x, (b) Optical image, zoom 20x, (c) Scanning electron microscope image, zoom 700x. The horizontal carbon nanotube is visible in the SEM image.



Supplementary Figure 5. Device layout and back-gated current-voltage curves. Device illustration showing source (S), drain (D) and gate (G) electrodes and channel length of 4um. Examples of I-V curves acquired with V_S=50mV, V_D=0mV and sweeping the back-gate potential V_G. Blue is in pristine state and red is after functionalization through a nanowell. Nanowell widths tested are (a) 4um (i.e. no mask), (b) 400nm, (c) 40nm, (d) 20nm, and (e) no exposure (i.e. control).

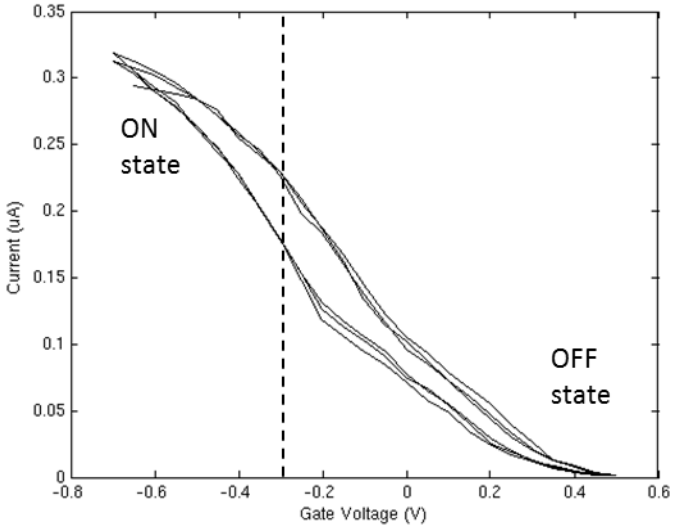


Supplementary Figure 6. Raman spectroscopy on controls. Raman spectra of fully-exposed (a) and unexposed (b) nanotube, for positive and negative controls. Exposed nanotubes show a D peak around 1300cm^{-1} indicating functionalization with CBDT.



Supplementary Figure 7. Effect of nanowell width. (a) Conductance change G/G_0 after functionalization as a function of nanowell width in several ensembles of devices. Each color represents a set of devices made from a different carbon nanotube, named A1-B1-B2-B3. Each set was divided in subsets in which devices were functionalized using a different nanowell width. Each data point in the figure represents the conductance change averaged over all devices in a subset (between N=5 and N=90) and the error bar represents the standard deviation of that ensemble. (b) Raman D/G ratio, which decreases as nanowell width decreases. Each data point is acquired from a separate device exposed using a different nanowell width.

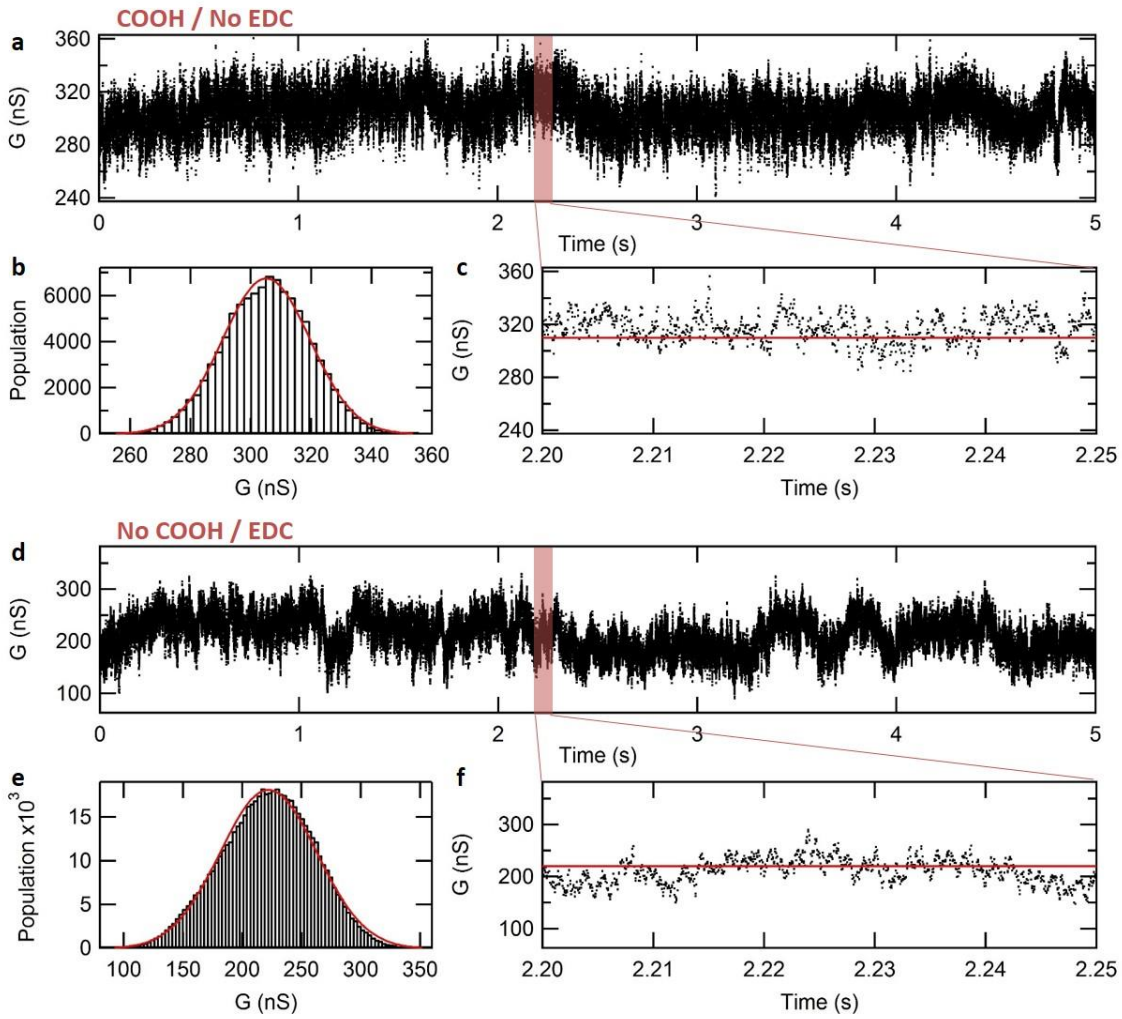
3. Real-time data analysis



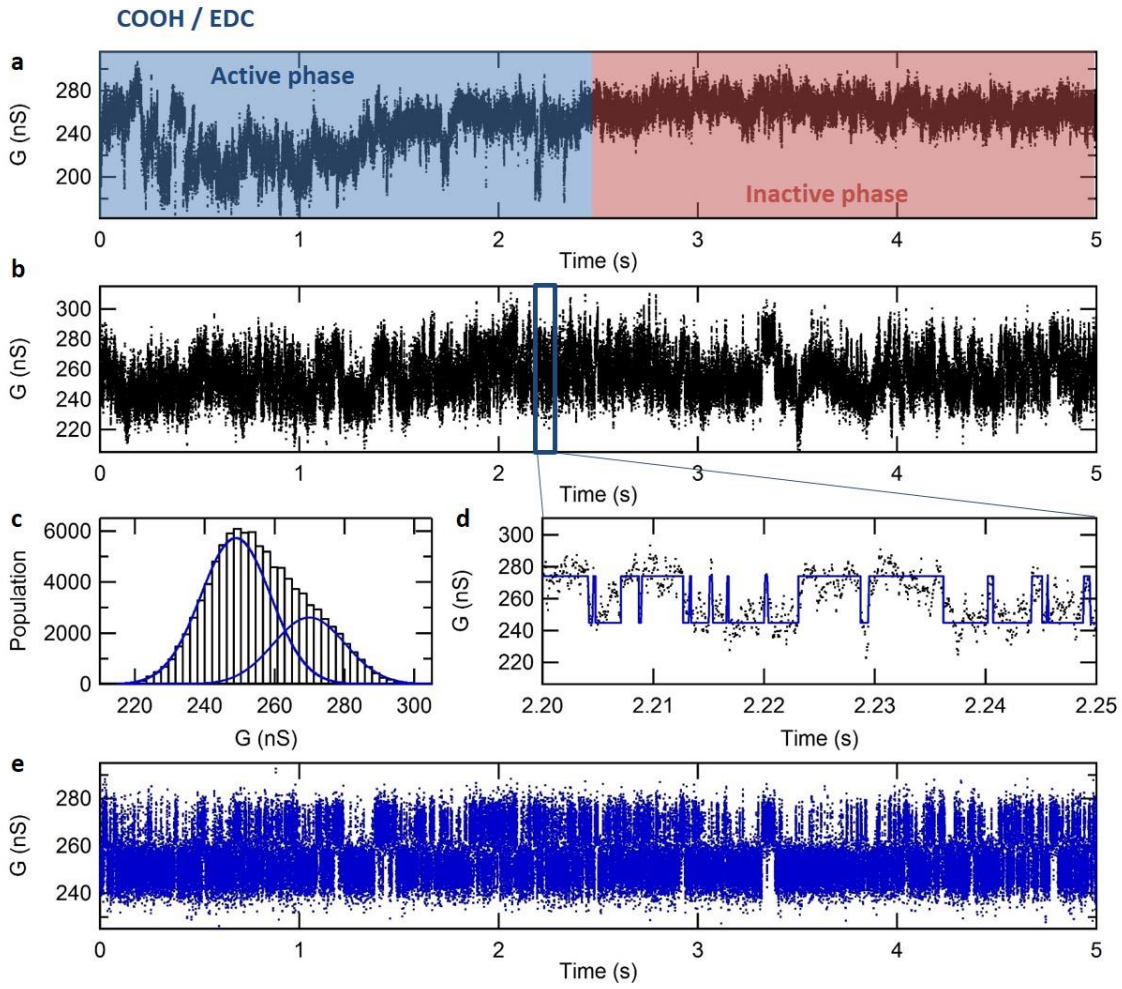
Supplementary Figure 8. Typical transfer curve of an smFET in aqueous conditions. Conductance of the nanotube switches from an ON-state at high negative gate bias to an OFF-state at positive bias. Real-time measurements are performed at -300 mV gate bias to have the nanotube conductance between ON and OFF states. Drain-source potential is fixed at 100 mV and gate potential is applied through Pt pseudoreference electrodes in contact with the electrolytic solution (see Fig. 3A). This specific transfer curve is from device B3 (see Supplementary Table 1) before DNA conjugation.

Supplementary Table 1. Summary of device status and activity on the smFET platform

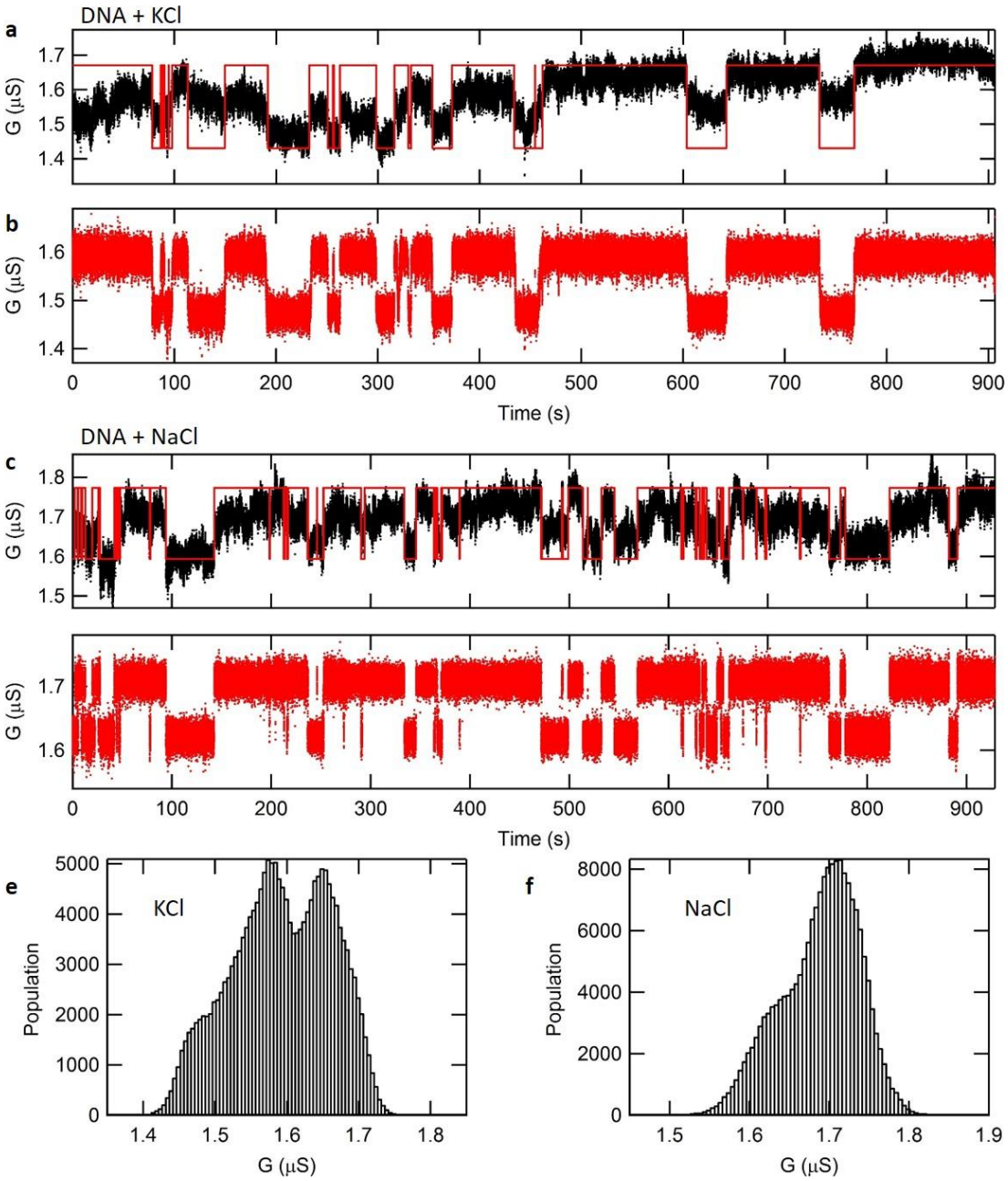
Device	ID	Type	Conductance drop	Functionalization status	Single-molecule EDC activity	Single-molecule DNA activity
1	A1	Unexposed	No	--	No	No
2	A3	Exposed	0.75	Functionalized	No	No
3	A4	Exposed	0.81	Functionalized	No	No
4	A6	Exposed	0.91	Functionalized	No	No
5	A8	Exposed	No	Unfunctionalized	No	No
6	B1	Unexposed	No	--	No	No
7	B3	Exposed	0.6	Functionalized	Yes	Yes
8	B6	Exposed	No	Unfunctionalized	No	No
9	B8	Exposed	0.82	Functionalized	Yes	No
10	C1	Unexposed	No	--	No	No
Yield				5/7 = 71% functionalized 2/7 = 29% unfunctionalized	2/7 = 29% EDC-active	1/2 = 50% DNA-active



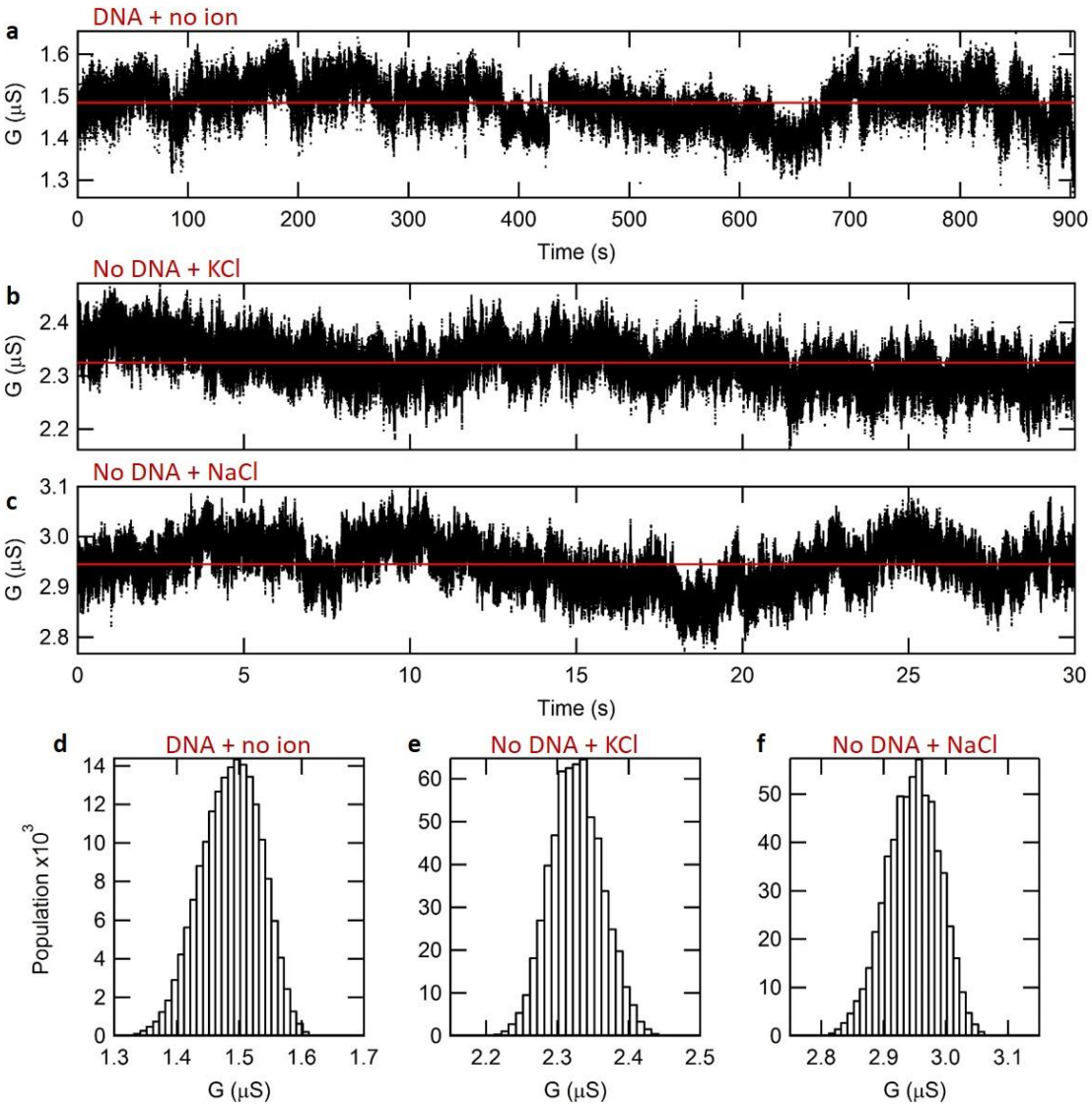
Supplementary Figure 9. Control experiments for EDC activity. (a) Control trace from a functionalized device in absence of EDC. (b) Distribution of conductance values in the above trace, with a single Gaussian fit indicative of single-state dynamics. (c) Zoom on a 50 ms region and idealized trace (red) for single-state dynamics. (d) Control trace from an unexposed device in presence of 50 μM EDC. (e) Distribution of conductance values in the above trace, with a single Gaussian fit indicative of single-state dynamics. (f) Zoom on a 50 ms region and idealized trace (red) for single-state dynamics.



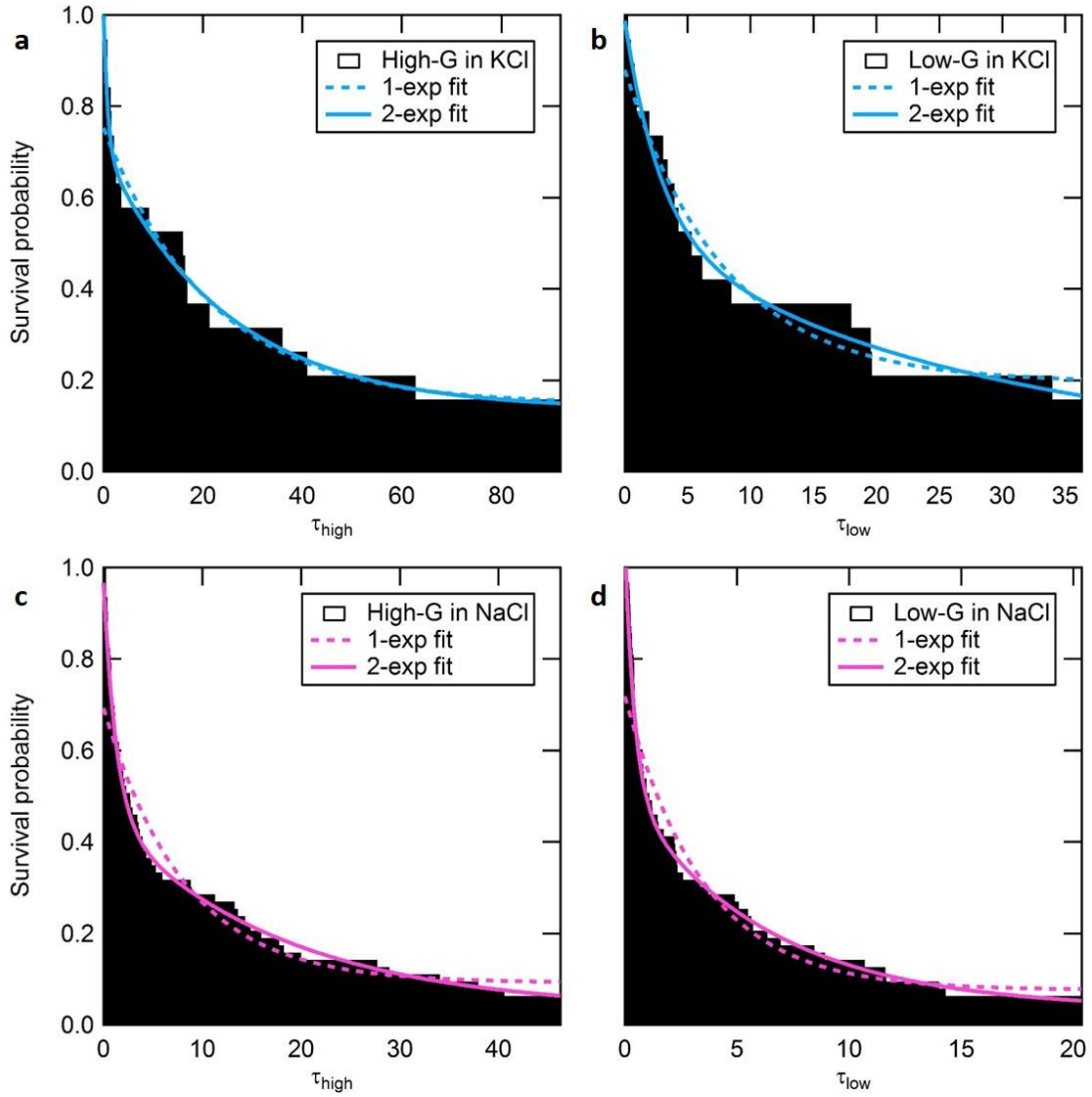
Supplementary Figure 10. Real-time COOH/EDC activity. (a) Example of transition between an active and inactive phase in the COOH/EDC dynamics. Typical lifetime of each phase is > 5 s. (b) Real-time trajectory of a functionalized device in presence of 50 μM EDC. (c) Distribution of conductance values in the above trace, fitted with two Gaussians as indicative of two-state dynamics. (d) Zoom on a 50 ms segment, superimposed with the idealized two-state trajectory obtained from a Hidden-Markov model (HMM) fit. (e) Baseline-corrected trajectory obtained from the HMM fit.



Supplementary Figure 11. Real-time trajectories for DNA activity. (a) Full raw trajectory in presence of 10mM KCl, superimposed with the idealized trajectory (red) obtained from the HMM fit. (b) Baseline-corrected trajectory for KCl environment, according to the HMM fit. (c) Full raw trajectory in presence of 10mM NaCl, superimposed with the idealized trajectory (red) obtained from the HMM fit. (b) Baseline-corrected trajectory for NaCl environment, according to the HMM fit. (e) Distribution of conductance values for the KCl trace in **a**, showing two peaks as indicative of two-state dynamics. The shoulder seen on the left side is attributed to the slow drift of the conductance baseline. (f) Distribution of conductance values for the NaCl trace in **a**, showing two peaks as indicative of two-state dynamics.



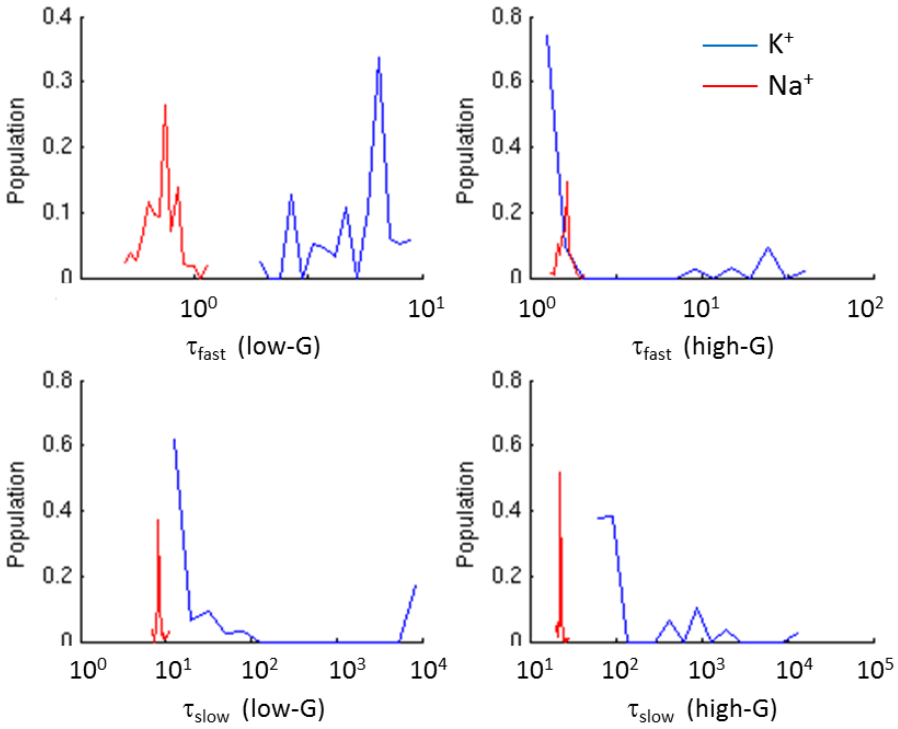
Supplementary Figure 12. Real-time trajectories for DNA activity. (a) Full raw trajectory for DNA in buffer in absence of any monovalent cation like K+ or Na+. Superimposed is the average value of the conductance. (b/c) Raw trajectories for a control device in presence of 10mM KCl (b) and 10mM NaCl (c). Superimposed is the average value of the conductance. (d/e/f) Distribution of conductance values in the above traces, showing a single Gaussian shape indicative of single-state dynamics.



Supplementary Figure 13. Survival probability. Survival probability for transitions from high-G to low-G state (left) and low-G to high-G state (right) as a function of dwell time. Data obtained in the presence of K+ ions (top) or Na+ ions (bottom). 18 distinct dwells were identified in the K+ idealized trace, and 53 dwells in the Na+ idealized trace. Dotted lines show fits with a simple exponential model and plain lines show fits with a double-exponential model. The double-exponential model is a better fit in each case, as it captures both the short-lived and long-lived dwells.

Supplementary Table 2. Characteristic lifetimes of low-G and high-G states, in presence of K^+ or Na^+ ions. Time constants are obtained from the double-exponential fits in Fig. S9 and S10. Errors are obtained from standard bootstrapping (see section 1.5)

		Lifetimes (s)		Error on lifetimes (s)		DNA state
		τ_{fast}	τ_{slow}	$\Delta\tau_{fast}$	$\Delta\tau_{slow}$	
High-G	K^+	**1.354**	**75.036**	0.42876	38.282	**Unfolded**
	Na^+	**1.0832**	**21.252**	0.17153	2.6706	
Low-G	K^+	**3.5682**	**34.191**	1.0119	7.6984	**Folded**
	Na^+	**0.44885**	**10.042**	0.13827	2.0055	



Supplementary Figure 14. Stability of τ estimates with respect to dwell assignation. Lifetime estimates for each state were calculated 1000x with the introduction of a random dwell inversion. Fast lifetimes form well-delimited distributions, indicating good stability in regard to dwell assignation errors; slow lifetime distributions are on the contrary very sensitive to dwell assignation, due to the fact that they are much less sampled in the trajectories.