Tethered fluorophore motion: studying large DNA conformational changes by single-fluorophore imaging

Supporting Material

Peter F. J. May[1,*], Justin N. M. Pinkney[1,*], Pawel Zawadzki[2], Geraint W. Evans[1], David J. Sherratt[2], Achillefs N. Kapanidis[1]

[1]Biological Physics Research Group, Clarendon Laboratory, Department of Physics, University of Oxford, Parks road, Oxford, OX1 3PU, UK.

[2]Department of Biochemistry, University of Oxford, South Parks road, Oxford, OX1 3QU, UK.

Corresponding author: Achillefs N. Kapanidis, Clarendon Laboratory, University of Oxford, Parks Road, Oxford, OX1 3PU, UK. Phone: +44 (0) 1865-272-401; Fax: +44 (0) 1865-272-400. E-mail: a.kapanidis1@physics.ox.ac.uk

* These authors contributed equally to this work

Supplementary Figures

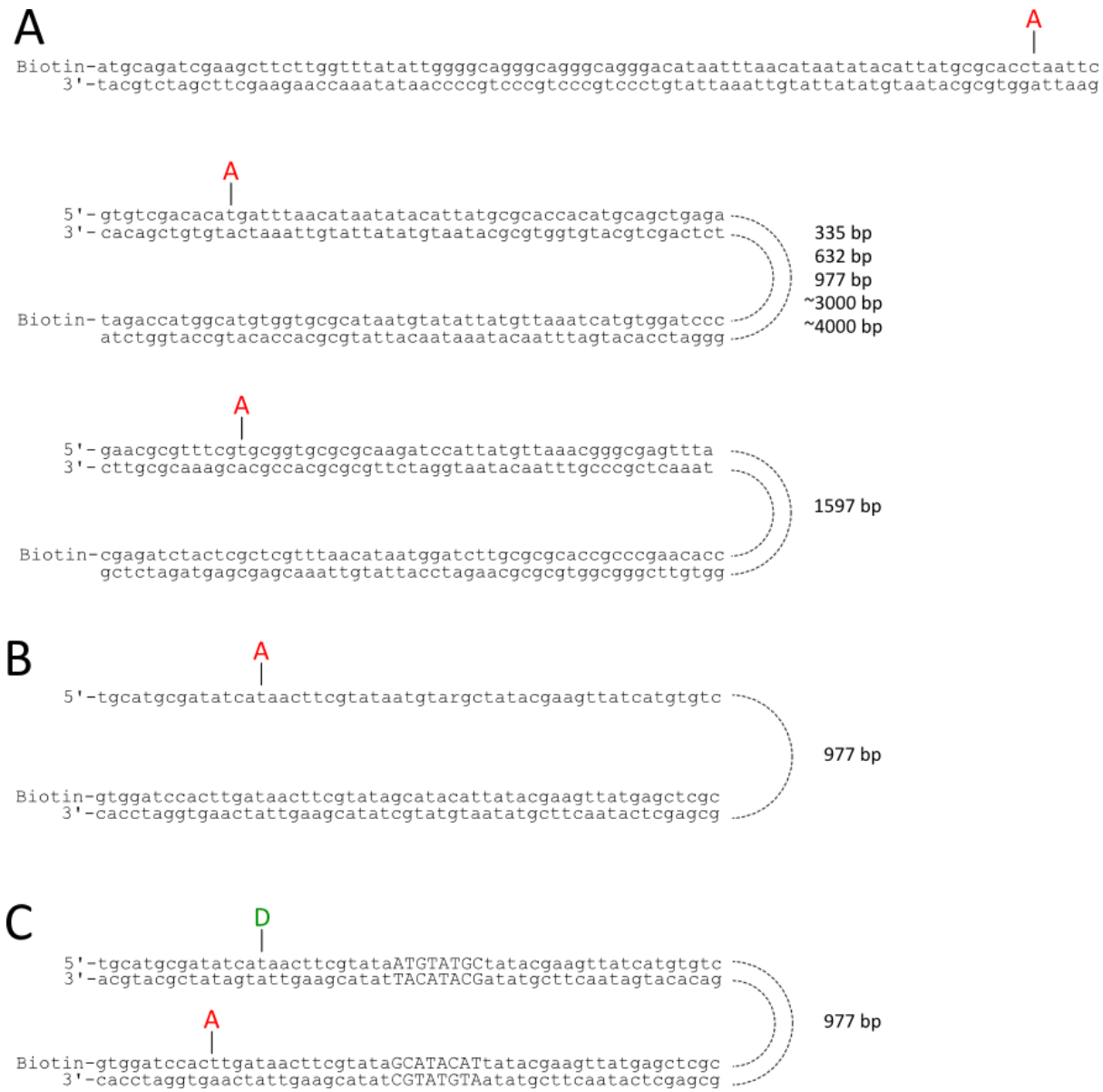

Figure S1. The sequences of DNAs used throughout this study. Primer sequences are shown, with labelling locations of Cy3B and Cy5 indicated by the D and A respectively. The long stretches of intervening DNA are not shown, but are indicated by dashed lines with their lengths indicated. DNAs were prepared as outlined in the Materials and Methods. (A) Singly labelled DNA used to characterise the TFM observables. (B) Partially double-stranded DNA used in KF experiments. (C) Doubly labelled DNA used in Cre experiments.

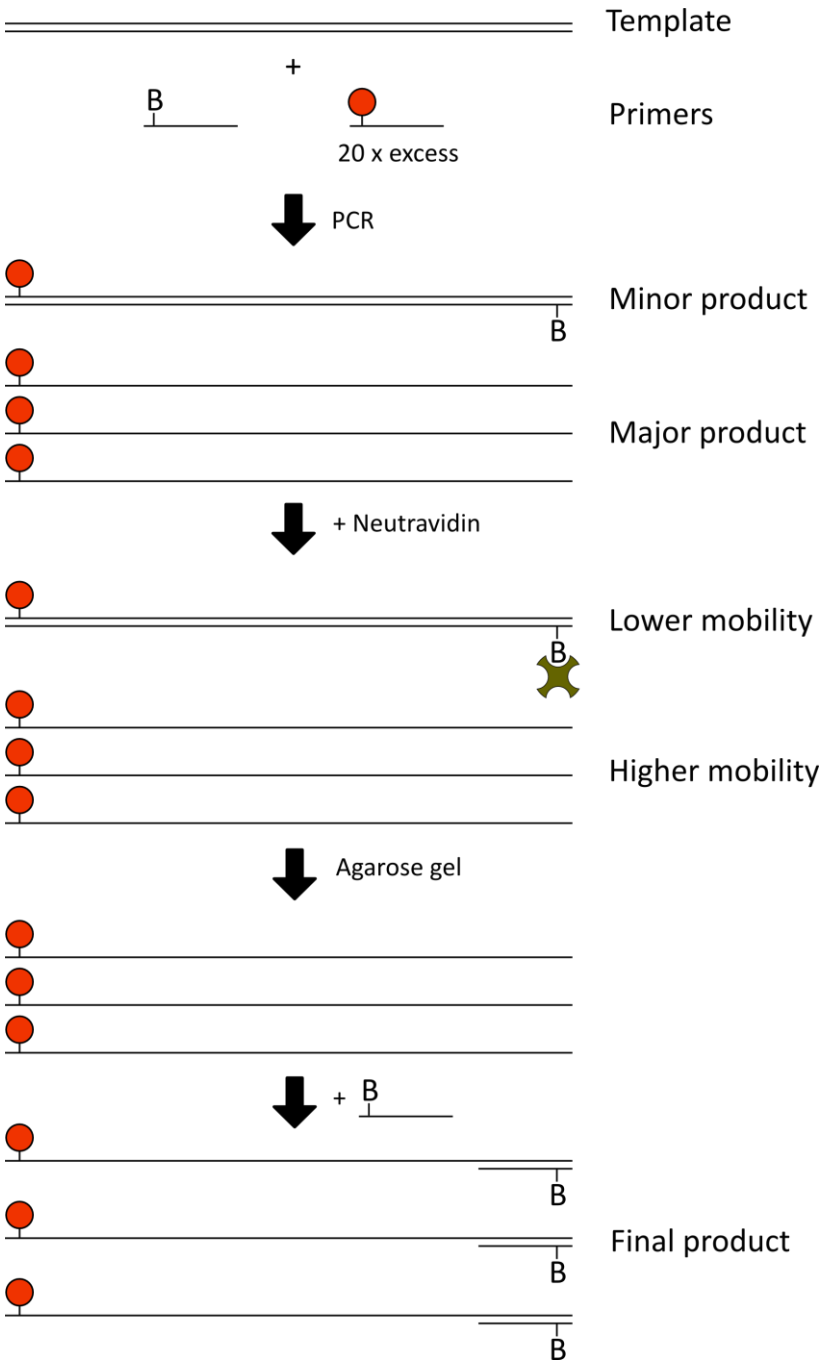

Figure S2. Schematic of KF substrate production. A PCR with a 20-fold excess of labelled primer produced mostly labelled single stranded DNA, along with double stranded biotinylated DNA. Addition of Neutravidin before gel purification ensured good separation between ssDNA and dsDNA of the same backbone length. The single stranded, labelled DNA was purified from the gel fragment and annealed to an excess of the biotinylated primer. The partially double stranded product was be purified on an affinity-binding column.

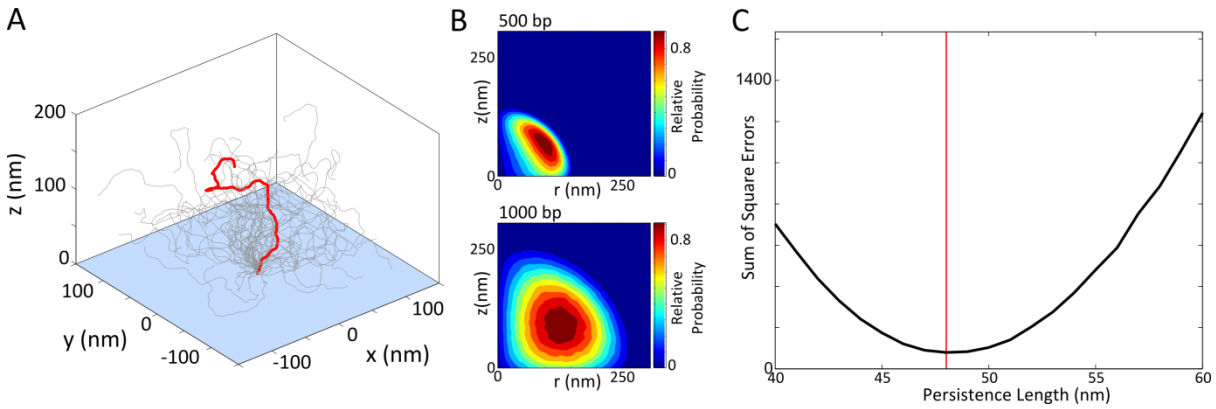

Figure S3. (A) Simulated tethered DNA configurations of length 1000 bp. A single conformation is highlighted in red. (B) Fluorophore probability distributions using WLC simulation, showing the relative probability of finding the fluorophore at a radius r and depth z from the tether point, for 500 bp (top) and 1000 bp (bottom) chains. Cylindrical symmetry has been used to show the distributions as 2D contour plots. (C) Fitting the persistence length of the WLC simulation to experimental data. The optimal fit (see Materials and Methods) occurs at a value of P = 48 nm (red line), and the results of simulations using P = 48 nm are shown in Fig. 3B.

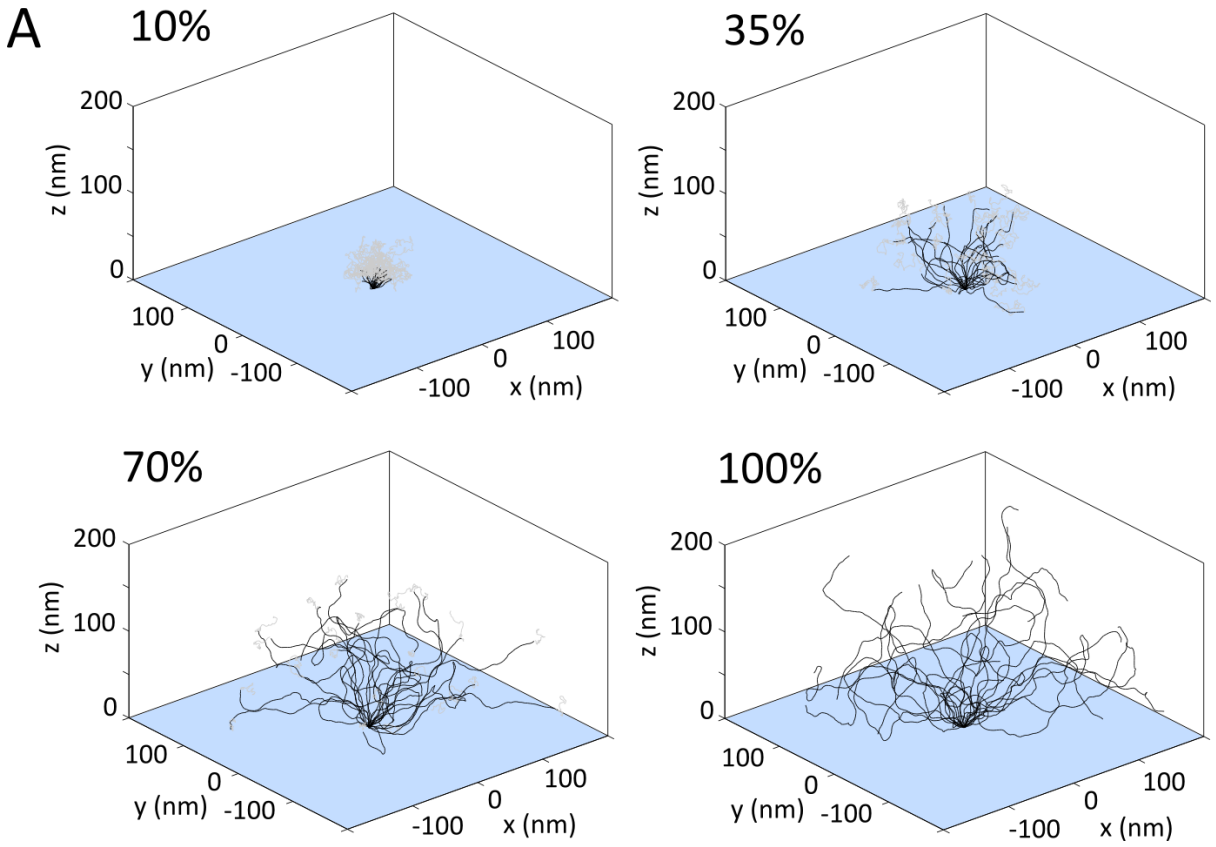

Figure S4. KF polymerisation (A) Simulations of 1087 bp DNA with increasing double stranded fractions. Single stranded DNA is shown in grey, double stranded in black. As the DNA is polymerised the FIW increases.

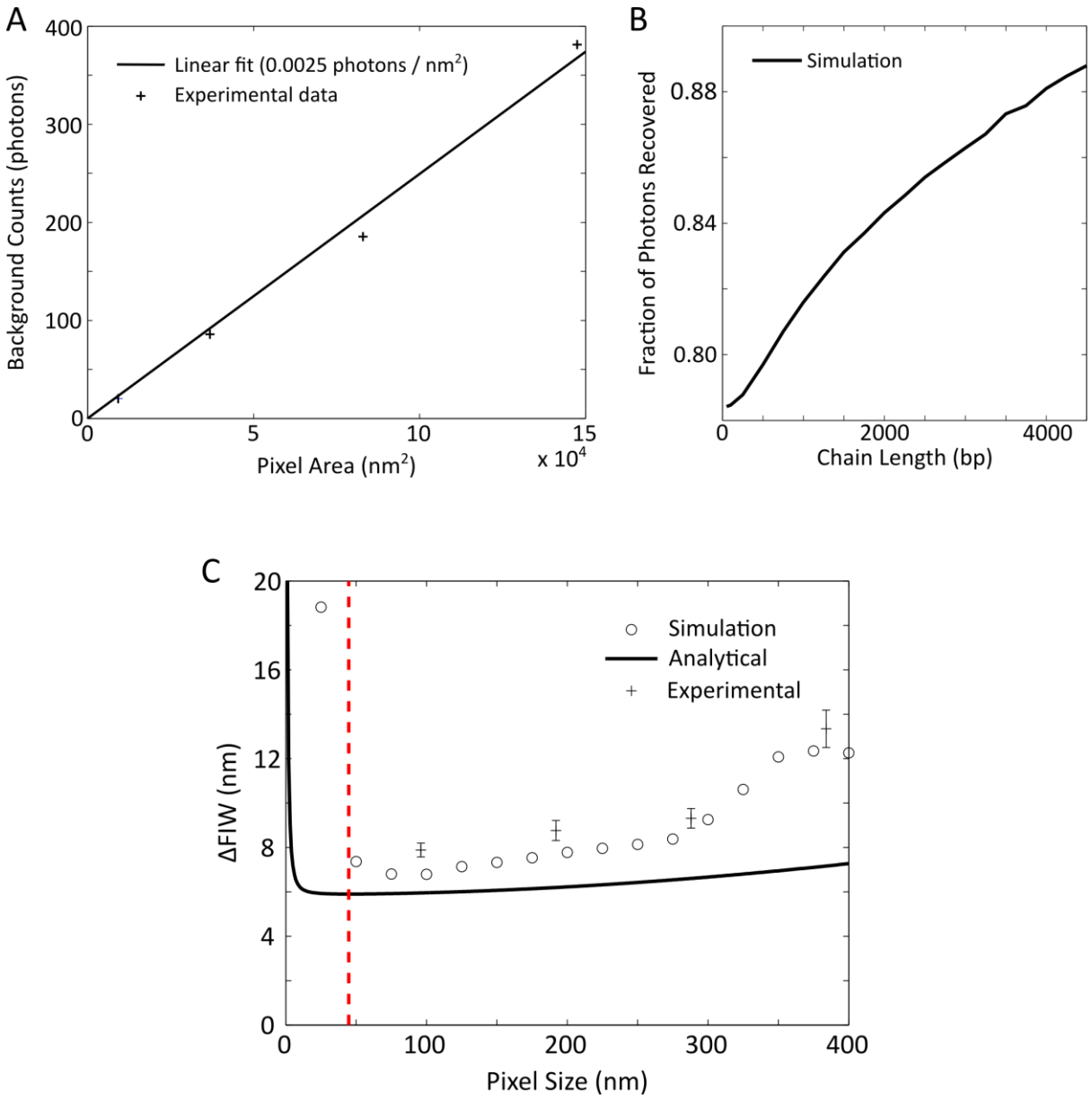

Fig S5. (A) Dependence of background photon count on the area covered by each pixel in the image plane. Four different pixel areas were achieved using binning on our EMCCD camera. A linear fit was constrained to go through the origin. This dependence was taken into account in our simulation and evaluation of our analytical expression when deducing an optimum pixel size. (B) Underestimation of photon count by OLS Gaussian fitting to simulated images. We simulated 80,000 images per chain length using a persistence length of 48 nm, pixel size of 96 nm, $N = 1000$, and a gain, read noise and background count chosen to give $\sigma_b = 5$. As the chain length is increased, the underestimation becomes less significant, which coupled with intensity changes due to the evanescent excitation field in TIRF makes recovered photon count an unreliable reporter of DNA length. (C) Optimum pixel size, using $N = 2250$, and 1000 bp DNA with a FIW of 188 nm. The analytical expression has a minimum at a pixel size of 45 nm (red dashed line), which is smaller than the simulation minimum at around 100 nm. As with the dependence on intensity, the analytical expression appears to underestimate the amount of noise present; the underestimation at larger pixel sizes can be explained by our approximation of sums as integrals in the derivation. We have taken the underestimation of photon count, by our image analysis algorithm, into account by matching recovered counts with our

simulation. Error bars are the standard error in the mean from > 15 molecules. Data were taken at 2 Hz.