Article

DNA surface exploration and operator bypassing during target search

https://doi.org/10.1038/s41586-020-2413-7

Received: 20 December 2018

Accepted: 7 April 2020

Published online: 24 June 2020

Check for updates

Emil Marklund[1,3], Brad van Oosten[1,3], Guanzhong Mao[1,3], Elias Amselem[1,3], Kalle Kipper[1], Anton Sabantsev[1], Andrew Emmerich[1], Daniel Globisch[2], Xuan Zheng[1], Laura C. Lehmann[1], Otto G. Berg[1], Magnus Johansson[1], Johan Elf[1]✉ & Sebastian Deindl[1]✉

Many proteins that bind specific DNA sequences search the genome by combining three-dimensional diffusion with one-dimensional sliding on nonspecific DNA[1–5]. Here we combine resonance energy transfer and fluorescence correlation measurements to characterize how individual *lac* repressor (LacI) molecules explore the DNA surface during the one-dimensional phase of target search. To track the rotation of sliding LacI molecules on the microsecond timescale, we use real-time single-molecule confocal laser tracking combined with fluorescence correlation spectroscopy (SMCT–FCS). The fluctuations in fluorescence signal are accurately described by rotation-coupled sliding, in which LacI traverses about 40 base pairs (bp) per revolution. This distance substantially exceeds the 10.5-bp helical pitch of DNA; this suggests that the sliding protein frequently hops out of the DNA groove, which would result in the frequent bypassing of target sequences. We directly observe such bypassing using single-molecule fluorescence resonance energy transfer (smFRET). A combined analysis of the smFRET and SMCT–FCS data shows that LacI hops one or two grooves (10–20 bp) every 200–700 μs. Our data suggest a trade-off between speed and accuracy during sliding: the weak nature of nonspecific protein–DNA interactions underlies operator bypassing, but also speeds up sliding. We anticipate that SMCT–FCS, which monitors rotational diffusion on the microsecond timescale while tracking individual molecules with millisecond resolution, will be applicable to the real-time investigation of many other biological interactions and will effectively extend the accessible time regime for observing these interactions by two orders of magnitude.

Sequence-specific binding and recognition of DNA target sites by proteins such as transcription factors, polymerases and DNA-modifying enzymes is at the core of cellular information processing. However, the 'target search problem' of how to rapidly—yet accurately—find a specific target sequence remains incompletely understood. One aspect of the search problem is addressed by facilitated diffusion[1], in which proteins search the genome by combining 3D diffusion with 1D sliding on DNA[1,2,5,6]. However, little is known about how sliding proteins explore the DNA surface. For example, it is unknown whether the sliding protein redundantly samples each base (as would be expected for faithful 1D diffusion) or whether the protein trades redundancy for speed, and occasionally skips bases. Here we shed light on the trade-off between sliding speed and accuracy in the recognition of target sites by measuring how a prototypical DNA-binding protein, the transcription factor LacI, explores the DNA surface during sliding.

LacI slides past operator sites

We first used smFRET[7,8] to directly monitor the kinetics with which LacI binds to its natural *lacO*$_1$ operator site (hereafter O_1). We generated a DNA construct that contains O_1 with a Cy5 acceptor dye located 5 bp from the O_1 edge. The DNA was surface-immobilized, and individual DNA molecules were monitored with a total-internal-reflection fluorescence microscope (Fig. 1a). Upon addition of LacI labelled with rhodamine in the DNA-binding domain (LacI–R) (Extended Data Fig. 1, Supplementary Table 1), the specific binding of LacI (Extended Data Fig. 2) to O_1 led to a sudden appearance of fluorescence signals and FRET (Fig. 1b). LacI containing a single donor (one-step photobleaching) on the distal or proximal subunit gave rise to a FRET distribution for binding events with peaks at FRET = 0.16 ± 0.001 and 0.89 ± 0.002, respectively (Fig. 1c, Supplementary Methods). The rate of operator binding ($k_{on,obs}$) depended on both LacI and Na^+ concentrations (Fig. 1d, Extended Data Fig. 2g) and the rate of LacI dissociation ($k_{off,obs}$) increased with increasing Na^+ concentration (Extended Data Fig. 2d, g). Importantly, LacI–R retained the ability to bind to both naturally occurring O_1 and the *lacO*$_3$ operator site (hereafter O_3) with binding affinities that were essentially identical to those that have previously been reported[9] (Fig. 1e), which indicates that the labelling of LacI–R did not substantially affect its operator binding. We also compared the binding affinity of LacI–R to that of a construct in which LacI was labelled further

[1]Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Uppsala, Sweden. [2]Department of Medicinal Chemistry, Science for Life Laboratory, Uppsala University, Uppsala, Sweden. [3]These authors contributed equally: Emil Marklund, Brad van Oosten, Guanzhong Mao, Elias Amselem. ✉e-mail: johan.elf@icm.uu.se; sebastian.deindl@icm.uu.se

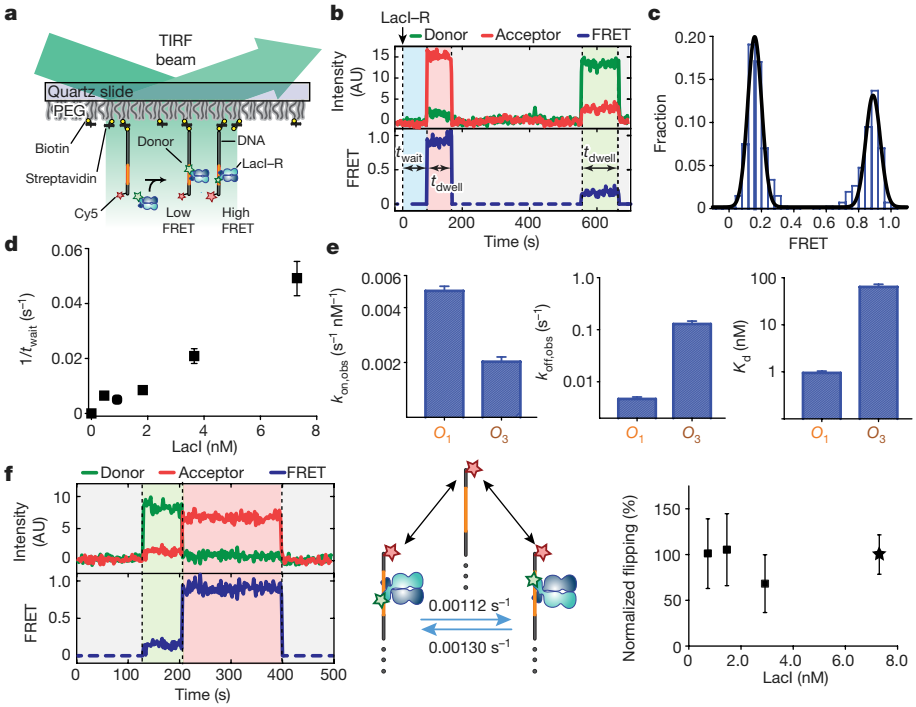


Fig. 1 | LacI–operator interaction dynamics. a, FRET detection for LacI dimer binding to its operator (orange). FRET donor, green; acceptor, red; PEG, polyethylene glycol; TIRF, total-internal-reflection fluorescence microscope. **b**–**e**, Binding to O_1. **b**, Traces showing binding of LacI containing a single donor on the proximal (left) or distal (right) monomer. t_{wait} and t_{dwell} denote the waiting and dwell times, respectively. AU, arbitrary units. **c**, Histogram from 6,500 binding events. Peaks at FRET = 0.16 ± 0.001 and 0.89 ± 0.002 (centroid values and standard errors from Gaussian fitting) are distal and proximal labelling configurations, respectively. **d**, Dependence of t_{wait} on LacI concentration (NaCl = 1 mM). Data points are mean ± s.e.m. from fitting a single-exponential decay to the distributions (n = 310 or 100 trajectories for LacI = 7.3 nM or otherwise, respectively). **e**, Rates of observed binding to (left) ($k_{on,obs} = t_{wait}^{-1}[LacI]^{-1}$), and dissociation from (middle) ($k_{off,obs}$), O_1 and O_3, and dissociation constants K_d (right). LacI–R = 7.3 nM, NaCl = 1 mM. Mean rate constants ± s.e.m. derived from t_{wait} (for $k_{on,obs}$ n = 248 and 109 for O_1 and O_3, respectively) or t_{dwell} (for $k_{off,obs}$ n = 367 and 91 for O_1 and O_3, respectively). $K_d = k_{off,obs}/k_{on,obs}$ with errors from propagating errors in the rate constants. **f**, Flipping. Left, LacI–R transitions from low to high FRET. Centre, three LacI binding states observed with a single O_1. The flipping rates, 0.0011 ± 0.0004 s^{-1} (distal to proximal, n = 28,000 s of observed time in each state) and 0.0013 ± 0.0004 s^{-1} (proximal to distal, n = 28,000 s of observed time in each state), are mean ± 95% confidence intervals. Right, flipping rates observed at various concentrations of LacI–R, normalized to LacI = 7.3 nM (asterisk). Values are mean ± 95% confidence intervals from n = 306, 189, 262 and 574 molecules (from left to right). In **c**–**f**, data are from at least three independent experiments at each condition.

away from the DNA binding domain (LacI–Far). The two labelling variants displayed identical sliding speeds (Extended Data Fig. 1c, d) and similar O_1 affinities (Extended Data Fig. 1e), further showing that the label impairs neither specific nor nonspecific DNA binding. FRET traces of LacI binding events exhibited instantaneous transitions between the proximal and distal binding orientations on the same operator (Fig. 1f). These 'flipping' transitions presumably arise from microscopic dissociation events, in which initially operator-bound LacI slides away from the operator and undergoes a spontaneous flipping transition before rebinding the operator in a flipped orientation.

To test for operator bypassing, we generated DNA constructs with two outer O_1 sites separated by 30 bp of DNA that was either random or contained a third site (O_1 or O_3) (Fig. 2b). To discriminate between binding to the one versus the other outer O_1, we used Cy5 and Alexa Fluor 750 as distinct acceptor dyes. In the presence of LacI–R, traces from individual 'O_1-random-O_1' molecules exhibited spontaneous 'switching' transitions due to LacI sliding from one outer operator site to the other (Fig. 2a). On the basis of the following observations, we conclude that these switching transitions involved a single LacI dimer. First, switching transitions were marked by fluorescence signals that appeared in the one acceptor channel and simultaneously disappeared in the other. In contrast to these rapid changes, the corresponding mean waiting time (t_{wait}) for a single binding event at the same LacI concentration (Fig. 1d) was substantially longer, with t_{wait} = 199 s (Supplementary Methods). Second, the frequency of switching transitions did not depend on the LacI concentration (Extended Data Fig. 3). We reasoned that a third, intervening operator should capture LacI sliding away from one of the outer O_1 sites—thereby sequestering it outside the FRET range (Supplementary Information section 3.2.2) of either acceptor—and abolish switching. Indeed, an intervening O_3 or O_1 reduced the switching rate (by a factor of 1.35 or 3.71, respectively) but did not completely abolish it (Fig. 2b). Thus, a single LacI dimer can bypass intervening sites and slide between the outer O_1 sites.

DNA sliding is coupled to rotation

To better understand bypassing, we sought to determine how sliding LacI explores the DNA surface. For this purpose, LacI–R was homogeneously (Extended Data Fig. 4a) labelled with rhodamine bifunctionally attached to two proximal cysteines to reduce rotation of the dye relative to the protein. We first characterized the orientation of individual LacI–R molecules by measuring their polarization of fluorescence while sliding on flow-stretched λ-DNA (49 kB) using single-molecule wide-field epifluorescence and camera-based polarization detection[10–14] (Fig. 3a, Extended Data Fig. 4b–j). These measurements showed a clear anisotropic polarization, which implies a nonrandom fluorophore orientation during LacI sliding (Fig. 3a, Extended Data Fig. 4b–j, Supplementary Information section 3.3). However, the limited temporal resolution (5 Hz) of these camera-based measurements could not resolve fast rotations of the protein around the DNA. Confocal detection can, however, be used to study interactions on a sub-millisecond timescale[15,16]. To more directly observe the sliding protein rotate around the DNA,

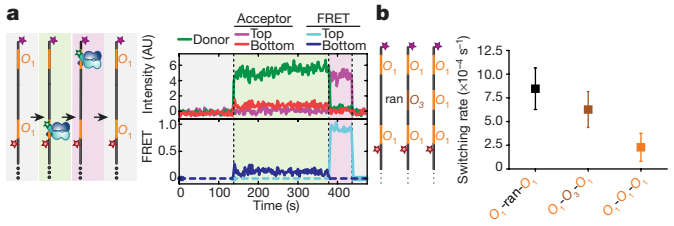


Fig. 2 | Observation of operator bypassing. a, Left, schematic of DNA constructs with two outer O_1 sites (orange), indicating the possible bound states of LacI–R. Top and bottom O_1, Alexa Fluor 750 (purple) and Cy5 (red) acceptor fluorophores, respectively. Right, donor fluorescence (green), bottom- (red) and top (purple)-site acceptor fluorescence, as well as bottom (dark blue) and top site (light blue) FRET traces showing the transition of a single LacI dimer initially bound at the bottom site (shaded green) switching to the top site (shaded purple). In the absence of acceptor signals, FRET is set to zero (dashed). **b**, Left, schematic of DNA constructs with two outer O_1 sites (orange) and intervening random (ran) DNA (left) or an additional O_3 (brown) (middle) or O_1 (right) site. Top and bottom O_1, Alexa Fluor 750 (purple) and Cy5 (red) acceptor, respectively. Right, hidden-Markov-model-derived switching rates (Extended Data Fig. 3) are mean ± 95% confidence intervals (n = 1,108, 1,137 and 699 individual molecules from left to right) from at least 10 independent experiments for each construct.

we used SMCT–FCS (Fig. 3b). This enabled us to monitor rotational diffusion on the microsecond timescale (Fig. 3b–f), at the same time as translational diffusion was tracked on the millisecond timescale. The translational movements—both parallel and perpendicular to the long axis of the DNA (Fig. 3d)—of individual LacI–R molecules were tracked and used to classify the molecules as sliders or nonsliders (that is, protein stuck on the glass surface) (Extended Data Table 1, Supplementary Methods, Supplementary Tables 2, 3). For FCS analysis[17], the photon emission was collected with nanosecond accuracy (Extended Data Fig. 5). We determined the autocorrelation function (ACF) of the fluorescence signal for molecules that bleached in a single step (Fig. 3e, f). If LacI sliding were coupled to its rotation around the DNA, the component of the ACF decay due to changes in the fluorophore orientation should be correlated with the rate of translational diffusion. A decrease in the translational diffusion rate is therefore expected to slow ACF decay in the relevant time regime. Indeed, when the experiment was repeated with a larger, maltose-binding-protein fusion of LacI–R (LacI–MBP–R), we measured slower translational diffusion (0.027 ± 0.002 μm^2 s^{-1} versus 0.035 ± 0.002 μm^2 s^{-1}) (Fig. 3d, Extended Data Fig. 6a) as well as slower ACF decay in the 20–100-μs range (Fig. 3e). No such difference between LacI–R and LacI–MBP–R was observed when they were immobilized on the glass surface (Fig. 3f, Extended Data Fig. 6b). The contributions to the ACF decay due to dye photophysics or flexibility in rotational attachment were therefore essentially identical for the two proteins. Thus, the difference in ACF between small and large proteins was due to 1D diffusion. Our data show that LacI sliding is coupled to its rotation around the DNA, with characteristic decay times on the order of 40 μs.

LacI slips out of the groove

To estimate the base pair distance that LacI translocates per revolution, we fit the difference in ACF to a model in which the pitch of the helical rotation is the only free parameter (Fig. 3g–j, Extended Data Fig. 7). The fitting method accurately returned the correct pitches when tested on theoretical rotational ACFs, convoluted with background noise processes obtained from the stationary molecules (Extended Data Fig. 7a, b). For the experimental data, fitting resulted in a pitch estimate of 39 ± 9 bp (Fig. 3g, Extended Data Fig. 7c). To explore the signal-to-noise ratio in our SMCT–FCS experiments, we carried out simulations of fluorophore rotation using the experimentally estimated pitch, as well as the same amount of data, shot noise and

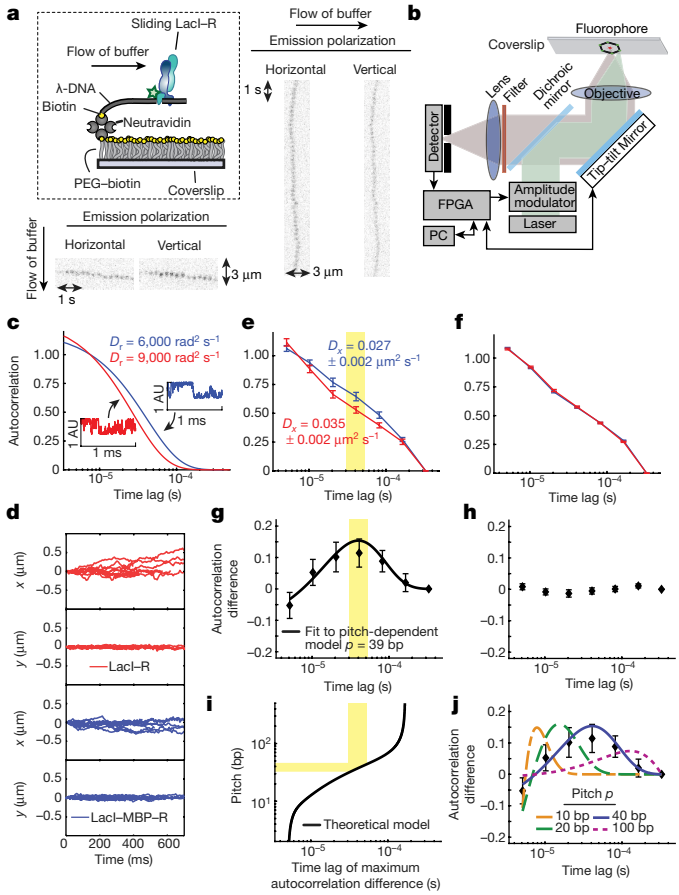


Fig. 3 | Determining pitch for rotation-coupled sliding. a, Schematic for flow-stretching (top left). Kinetic series of images in horizontal and vertical emission polarization, showing two representative sliding LacI–R molecules when DNA is stretched in the vertical (bottom left) or horizontal (right) direction. Fifty-two and 27 sliding molecules were captured for the vertical and horizontal direction, respectively. **b**, Schematic of SMCT–FCS setup. Black arrows, direction of communication; green dots, laser pattern during fluorophore (red dot) tracking. **c**, Simulated traces (inset) of faster (red) and slower (blue) dye rotation and the resulting autocorrelations. PC, personal computer; FPGA, field programmable gate array. **d**, x (DNA direction) and y coordinate of sliding LacI–MBP–R (blue, 151 molecules) and LacI–R (red, 90 molecules). Ten representative traces are shown (Extended Data Fig. 6a). **e**, **f**, Mean normalized autocorrelation of the fluorescence for sliding (**e**) (n = 86 and 54 informative autocorrelations for LacI–MBP–R and LacI–R, respectively) and stationary (**f**) (n = 2,273 and 1,064 informative autocorrelations for LacI–MBP–R and LacI–R, respectively) LacI–MBP–R (blue) and LacI–R (red). Diffusion constants are 20% trimmed mean ± s.e.m. **g**, **h**, Difference in autocorrelation between LacI–MBP–R and LacI–R for sliding (**g**) (n = 86 and 54 informative autocorrelations for LacI–MBP–R and LacI–R, respectively) and stationary (**h**) (n = 2,273 and 1,064 informative autocorrelations for LacI–MBP–R and LacI–R, respectively) molecules. Data in **e**–**h** are mean ± s.e.m. Black line in **g**, best fit to a rotation-coupled sliding model; bootstrapping the trajectories yields an average pitch (p) for the rotation of 39 ± 9 bp (mean ± s.e.m.). **i**, Theoretical dependence of the pitch on the time lag of maximum difference in the autocorrelations. Yellow area, time regime corresponding to pitch values within the s.e.m of the experimentally measured pitch. **j**, Best fits of the rotation-coupled sliding model when the pitch is constrained at different levels.

filtering steps as in the experiments. Notably, the resulting simulated differences in ACF were very similar to the experimentally determined ones (Extended Data Fig. 7d, Supplementary Methods), which confirms that our SMCT–FCS experiments yielded signal amplitudes and errors as expected from theory. We conclude that the sliding protein does not faithfully track the DNA helix, but instead slides with a longer pitch. A sliding mode in which LacI sometimes slips between grooves via

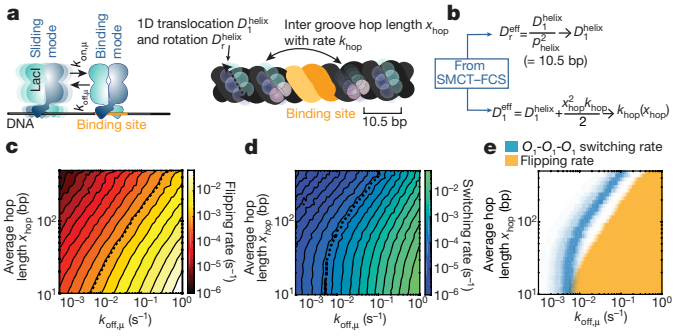


Fig. 4 | Determining hop length and frequency. a, Schematic of target search in which LacI combines faithful DNA-groove tracking (1D diffusion) with short and frequent intergroove hops that can bypass a specific binding site (orange). **b**, The effective rotational (D_r) and translational (D_1) diffusion coefficients from SMCT–FCS are used to set parameters ($k_{hop}(x_{hop})$) in the simulations of the FRET experiments. **c**, **d**, Flipping (**c**) and switching (**d**) rates calculated from stochastic simulations as a function of the model parameters. Dashed lines, experimentally observed rate of flipping on the same operator (**c**) and switching between two outer operators with an intervening operator (**d**). **e**, Parameter probability densities from experimental uncertainties. Parameters compatible with observed switching (blue) and flipping (orange) rates. Surface transparency scaled according to the probability density for the parameters given the experimental data. The flipping rate surface does not decrease with increasing values of $k_{off,\mu}$, as the simulations yield the maximum possible flipping rate. $k_{on,\mu}$ was infinite in these simulations; see Extended Data Fig. 8 for finite values of $k_{on,\mu}$.

microscopic hops would agree with these observations, and contribute to operator bypassing.

To determine which of the microscopic parameters for hopping (hop length and frequency) are consistent with the experimentally observable processes (switching rate, flipping rate and pitch), we simulated the processes for broad ranges of microscopic parameters (Fig. 4a,b). Flipping (Fig. 4c) and switching (Fig. 4d) rates were sampled in simulations as in the experiments (Supplementary Methods). Because both the overall 1D diffusion rate and the pitch of rotational sliding are known from the SMCT–FCS experiments, the hop frequency (k_{hop}) can be calculated for each hop length (x_{hop}) (Fig. 4b). The absolute hop and switching rates also depend on how often LacI dissociates from the operator (Fig. 4c, d), because LacI cannot hop if it is bound to the operator. The experimental flipping rate (dashed line in Fig. 4c) represents a lower bound for the absolute hop rate, as LacI cannot flip without hopping. At the same time, the absolute rate of the switching transitions defines a relation between the hop length and the operator dissociation rate, in which long and rare hops that do not frequently bypass the intervening operator are compensated for by more frequent dissociation (dashed line in Fig. 4d). By determining where the dashed lines in Fig. 4c, d overlap (Supplementary Methods) to satisfy the experimental constraints (Fig. 4e), we find that the average hop length cannot exceed 16 ± 8 bp, corresponding to a minimum hop frequency of 4 ± 1 ms^{-1}.

Discussion

Our data show that LacI rotates while sliding, and with a pitch that exceeds the 10.5-bp DNA pitch, as a consequence of frequent and short hops. Such hopping may result from the nonspecific binding that is sufficiently weak to optimize overall search speed, while at the same time inevitably leading to frequent operator bypassing. Bypassing does not necessarily reduce the rate of operator binding, as 1D diffusion involves many revisits to the same bases[1]. In fact, the observed hopping frequency allows for rapid scanning of the DNA and speeds up the first encounter with the specific site by about 100% in stochastic simulations of target search, despite the frequent bypassing (Supplementary Information section 3.5).

Finally, we anticipate that SMCT–FCS will lend itself to characterizing many molecular interactions at a scale hundreds of times faster than was previously accessible[18].

Online content

Any methods, additional references, Nature Research reporting summaries, source data, extended data, supplementary information, acknowledgements, peer review information; details of author contributions and competing interests; and statements of data and code availability are available at https://doi.org/10.1038/s41586-020-2413-7.

1. Berg, O. G., Winter, R. B. & von Hippel, P. H. Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. *Models and theory. Biochemistry* **20**, 6929–6948 (1981).
2. Hammar, P. et al. The *lac* repressor displays facilitated diffusion in living cells. *Science* **336**, 1595–1598 (2012).
3. Blainey, P. C., van Oijen, A. M., Banerjee, A., Verdine, G. L. & Xie, X. S. A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proc. Natl Acad. Sci. USA* **103**, 5752–5757 (2006).
4. Gowers, D. M., Wilson, G. G. & Halford, S. E. Measurement of the contributions of 1D and 3D pathways to the translocation of a protein along DNA. *Proc. Natl Acad. Sci. USA* **102**, 15883–15888 (2005).
5. von Hippel, P. H. & Berg, O. G. Facilitated target location in biological systems. *J. Biol. Chem.* **264**, 675–678 (1989).
6. Blainey, P. C. et al. Nonspecifically bound proteins spin while diffusing along DNA. *Nat. Struct. Mol. Biol.* **16**, 1224–1229 (2009).
7. Stryer, L. & Haugland, R. P. Energy transfer: a spectroscopic ruler. *Proc. Natl Acad. Sci. USA* **58**, 719–726 (1967).
8. Ha, T. et al. Probing the interaction between two single molecules: fluorescence resonance energy transfer between a single donor and a single acceptor. *Proc. Natl Acad. Sci. USA* **93**, 6264–6268 (1996).
9. Bondeson, K., Frostell-Karlsson, A., Fägerstam, L. & Magnusson, G. Lactose repressor-operator DNA interactions: kinetic analysis by a surface plasmon resonance biosensor. *Anal. Biochem.* **214**, 245–251 (1993).
10. Ha, T., Enderle, T., Chemla, S., Selvin, R. & Weiss, S. Single molecule dynamics studied by polarization modulation. *Phys. Rev. Lett.* **77**, 3979–3982 (1996).
11. Persson, F., Westerlund, F., Tegenfeldt, J. O. & Kristensen, A. Local conformation of confined DNA studied using emission polarization anisotropy. *Small* **5**, 190–193 (2009).
12. Backer, A. S., Lee, M. Y. & Moerner, W. E. Enhanced DNA imaging using super-resolution microscopy and simultaneous single-molecule orientation measurements. *Optica* **3**, 659–666 (2016).
13. Sosa, H., Peterman, E. J., Moerner, W. E. & Goldstein, L. S. ADP-induced rocking of the kinesin motor domain revealed by single-molecule fluorescence polarization microscopy. *Nat. Struct. Biol.* **8**, 540–544 (2001).
14. Forkey, J. N., Quinlan, M. E., Shaw, M. A., Corrie, J. E. T. & Goldman, Y. E. Three-dimensional structural dynamics of myosin V by single-molecule fluorescence polarization. *Nature* **422**, 399–404 (2003).
15. Verbrugge, S., Kapitein, L. C. & Peterman, E. J. G. Kinesin moving through the spotlight: single-motor fluorescence microscopy with submillisecond time resolution. *Biophys. J.* **92**, 2536–2545 (2007).
16. Verbrugge, S., Lansky, Z. & Peterman, E. J. G. Kinesin's step dissected with single-motor FRET. *Proc. Natl Acad. Sci. USA* **106**, 17741–17746 (2009).
17. Magde, D., Elson, E. & Webb, W. W. Thermodynamic fluctuations in a reacting system—measurement by fluorescence correlation spectroscopy. *Phys. Rev. Lett.* **29**, 705 (1972).
18. Persson, F., Lindén, M., Unoson, C. & Elf, J. Extracting intracellular diffusive states and transition rates from single-molecule tracking data. *Nat. Methods* **10**, 265–269 (2013).

Reporting summary

Further information on research design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

SMCT–FCS raw data are available at https://doi.org/10.17044/scilifelab.12401906. All other data are available from the corresponding authors upon reasonable request. Uncropped gels are provided in Supplementary Fig. 1.

Code availability

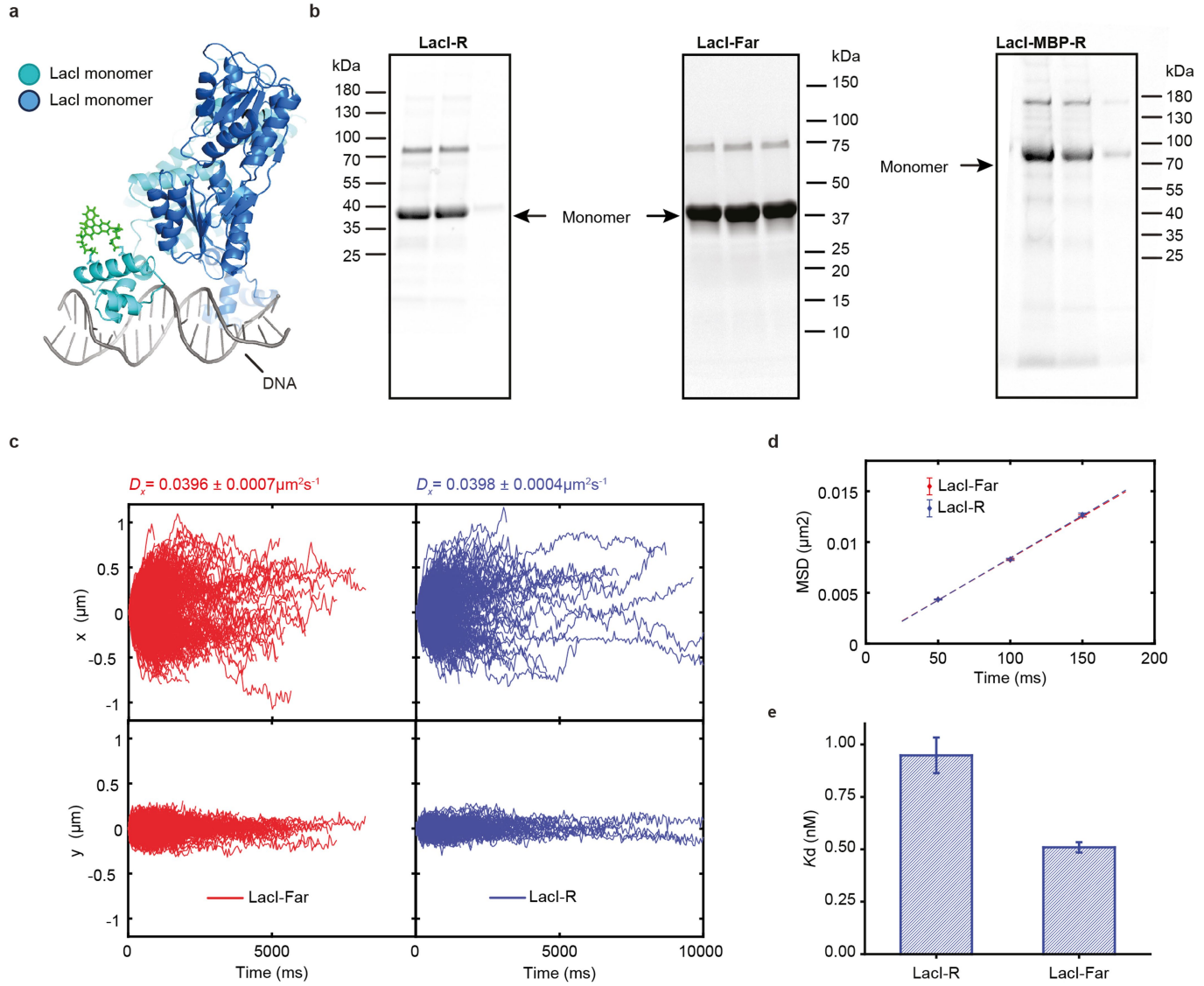
Code developed for SMCT–FCS analyses is available at GitHub (https://github.com/elfware/SMCT_FCS). All analysis software developed for this project is available from the corresponding authors upon reasonable request.

Acknowledgements We thank P. Blainey and K. Xiong for advice on flow-stretching; C. Joo for discussions; J. Laksman for initial code; and D. Fange, I. Barkefors and M. Ehrenberg for helpful comments on the manuscript. The project was funded by the European Research Council (ERC), the Swedish Research Council (VR) and the Knut and Alice Wallenberg Foundation (KAW).

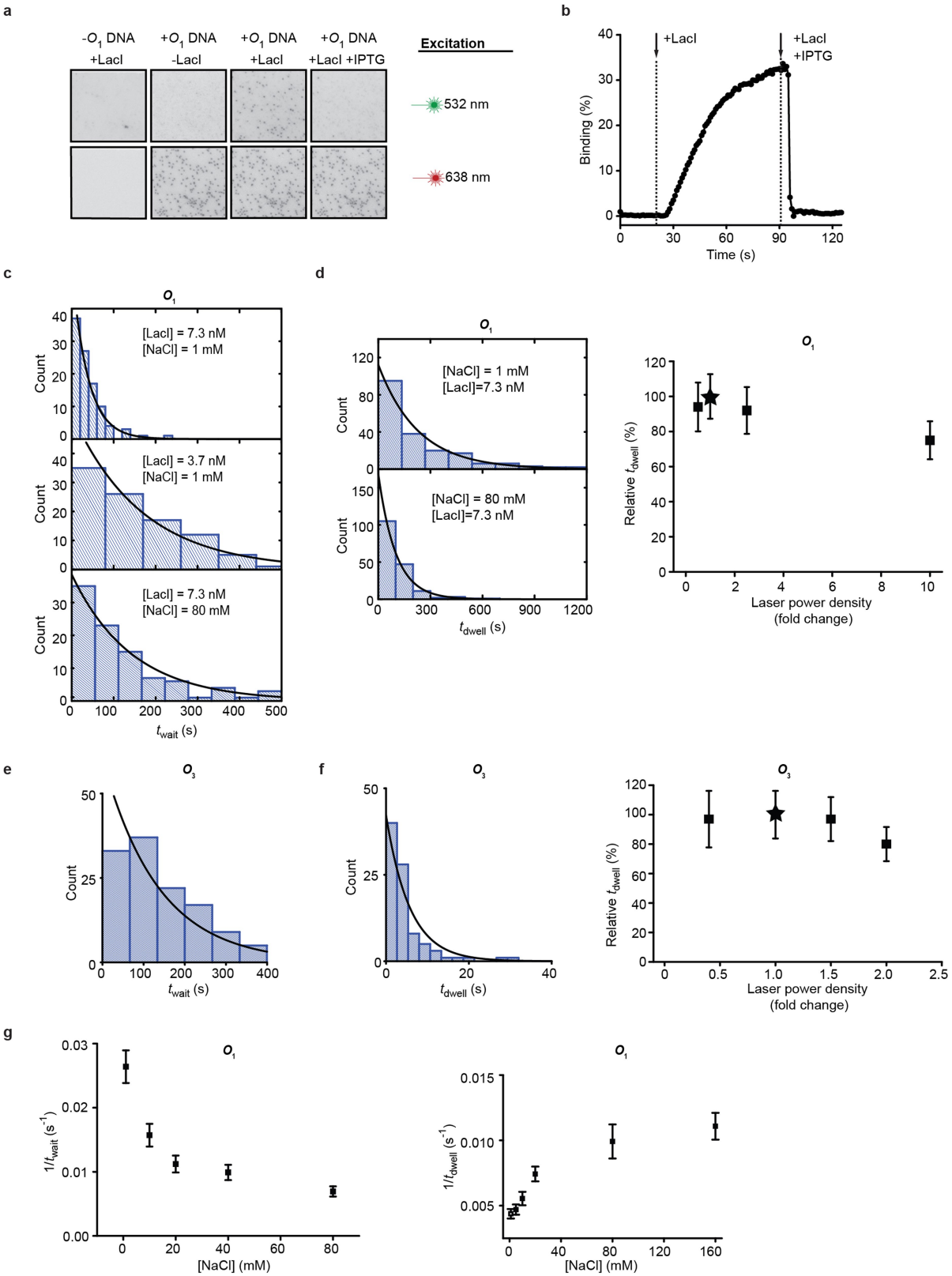
Author contributions E.M., B.v.O., G.M. and E.A. contributed equally to this manuscript. S.D. conceived of the smFRET approach to directly observe operator bypassing and J.E. conceived of the confocal tracking with polarization readout concept; S.D., B.v.O., G.M., J.E. and A.S. designed the smFRET study; J.E., E.M. and E.A. conceived of the SMCT–FCS implementation; E.A. developed and built the SMCT–FCS microscope; S.D., E.M. and K.K. developed the purification and labelling scheme; M.J. suggested the use of a bifunctional dye for SMCT–FCS tracking experiments; E.M. and L.C.L. purified and labelled LacI; E.M. implemented the fluidic assay, with input from M.J. E.M. carried out all tracking experiments, with assistance from E.A., K.K. and X.Z.; E.M. and E.A. developed theoretical models and data analysis methods for SMCT–FCS tracking experiments; E.M. carried out analysis of SMCT–FCS tracking experiments; G.M. generated all smFRET DNA constructs and collected smFRET data with B.v.O.; B.v.O. analysed FRET time traces with G.M.; A.E. provided initial code; D.G. supported high-performance liquid chromatography purification and CNBr-mediated cleavage experiments. E.M. designed and carried out stochastic simulations, with input from J.E, S.D. and O.G.B.; S.D., J.E., E.M. and E.A. wrote the paper, with input from all authors.

Competing interests The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41586-020-2413-7.
Correspondence and requests for materials should be addressed to J.E. or S.D.
Peer review information *Nature* thanks Peter von Hippel and Achillefs Kapanidisfor their contribution to the peer review of this work.
Reprints and permissions information is available at http://www.nature.com/reprints.

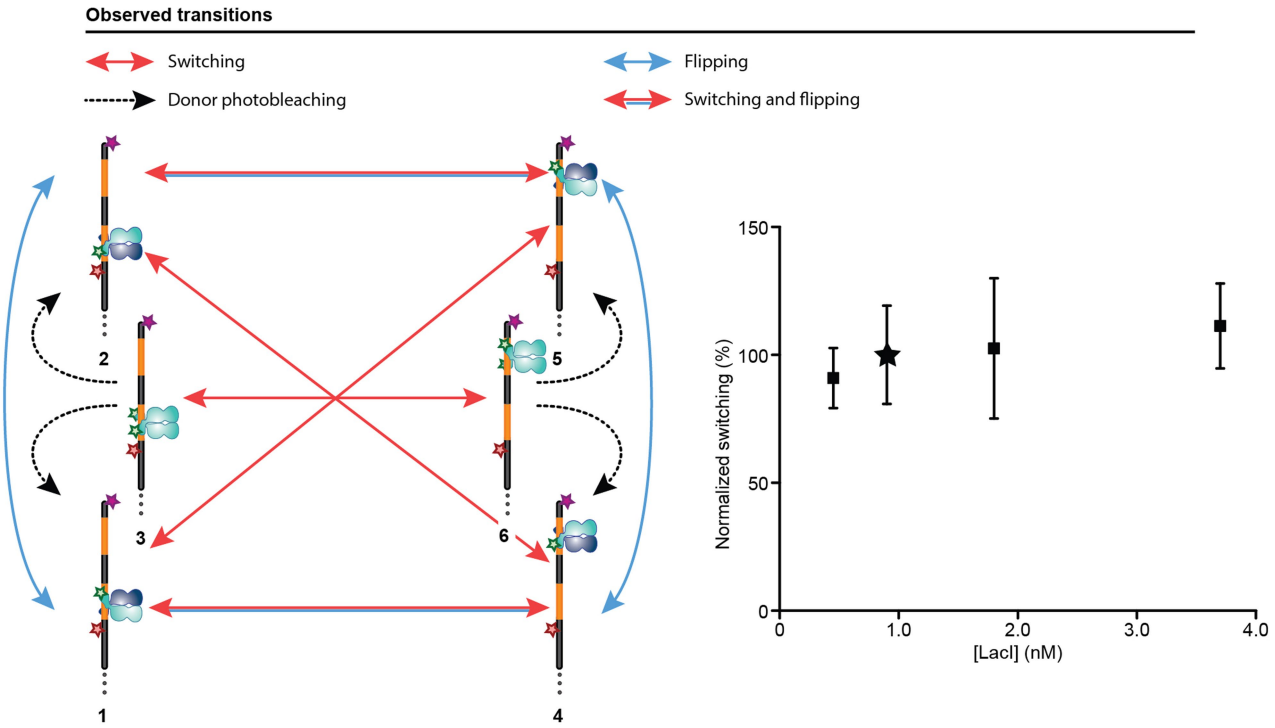


Extended Data Fig. 1 | Analysis of LacI labelling and binding. a, Structural model (based on Protein Data Bank (PDB) code 1OSL). LacI, blue and cyan; rhodamine, green; DNA, grey. **b**, SDS–PAGE of labelled LacI fractions after dye removal visualized using rhodamine (LacI–R and LacI–MBP–R) or Cy3 (LacI–Far) fluorescence. The bands corresponding to the monomeric sizes expected for LacI–R (left), LacI–Far (middle) and LacI–MBP–R (right) are indicated with arrows. The intensity of the monomeric band relative to the sum of the intensities of monomeric and dimeric bands is 77% and 86% for LacI–R and LacI–MBP–R, respectively. For gel source data, see Supplementary Fig. 1. **c**, *x* (DNA direction) and *y* coordinate of sliding LacI–Far (red) and LacI–R (blue) molecules obtained by electron multiplying charge-coupled device (EMCCD) tracking with a 50-ms frame rate. In total, 779 and 409 sliding molecules were captured for LacI–Far and LacI–R, respectively. Diffusion constants along the *x* coordinate (D_x) are mean ± s.e.m. **d**, Mean squared displacement (MSD) for different time steps for LacI–Far (red) (n = 779 sliding molecules) and LacI–R (blue) (n = 409 sliding molecules) ± s.e.m. **e**, Dissociation constants (K_d) for LacI–Far and LacI–R. LacI = 7.3 nM, NaCl = 1 mM. K_d values were calculated as $K_d = k_{off,obs}/k_{on,obs}$ with errors from propagating errors in the rate constants. Rate constants are mean ± s.e.m. ($k_{on,obs}$, n = 248 and 200 events for LacI–R and LacI–Far, respectively; $k_{off,obs}$, n = 367 and 322 events for LacI–R and LacI–Far, respectively).

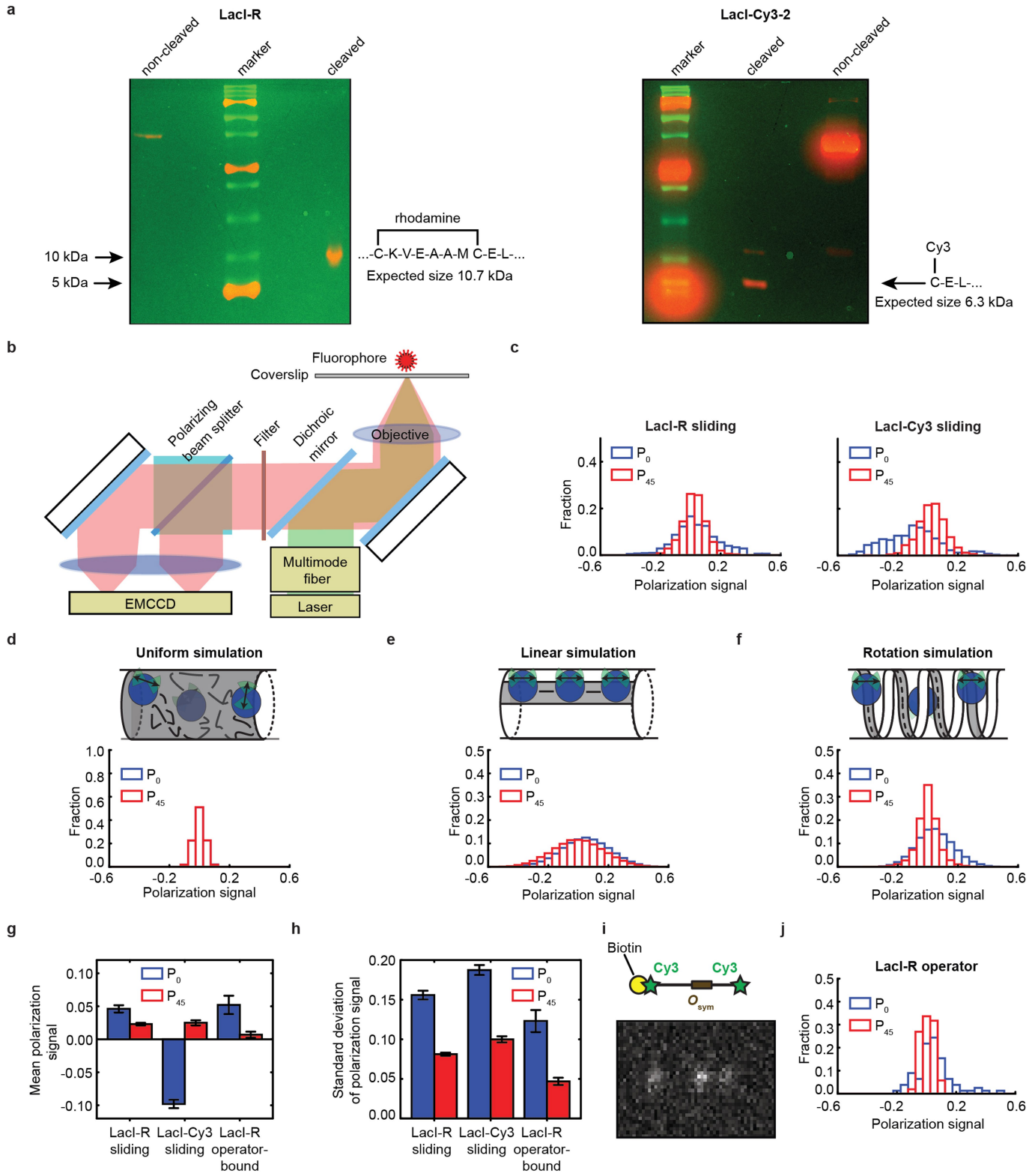


Extended Data Fig. 2 | See next page for caption.

Extended Data Fig. 2 | LacI binding kinetics. a, EMCCD images of acceptor channel under 532 nm (top) (donor excitation) and 638 nm (bottom) (acceptor excitation) illumination. Specific binding of individual donor-labelled LacI–R molecules (individual spots, top row) occurs in the presence (+O_1 DNA; individual spots, bottom row; DNA image was acquired before the addition of LacI, and the same DNA image is shown in columns 3 and 4 for reference) but not in the absence (–O_1 DNA) of acceptor-labelled DNA containing an operator site. Addition of isopropyl β-D-1-thiogalactopyranoside (+IPTG) displaces specifically bound LacI in the same field of view. This control experiment was carried out once. **b**, Time trace of LacI–R occupancy in one field of view (1,244 individual O_1 DNA molecules). The 7.3 nM LacI–R was supplied after 20 s (first dotted line), followed by the addition of 7.3 nM LacI–R together with 100 mM IPTG after 90 s. **c**, Histograms of t_{wait} values for specific binding to O_1 (n = 310 or 100 trajectories for LacI = 7.3 nM and NaCl = 1 mM or otherwise, respectively) at various LacI and NaCl concentrations as indicated. **d**, Left, histograms of t_{dwell} values for specific binding to O_1 (n = 190 or 184 events at LacI–R = 7.3 nM and NaCl = 1 mM (top) or 80 mM (bottom), respectively). Right, under standard imaging conditions (marked by the asterisk) measurements of t_{dwell} for binding to O_1 are not affected by photobleaching. Mean relative t_{dwell} for specific binding to O_1 (n = 106, 146, 92 and 243 individual molecules from left to right, with mean binding duration ± s.e.m.) observed at NaCl = 1 mM, and using various laser power densities, is shown. Dwell times and laser power densities were normalized to the standard imaging laser power density used for all other analyses (asterisk). **e**, Histogram of t_{wait} values for specific binding to O_3 (n = 109 individual molecules) at LacI = 7.3 nM and NaCl = 1 mM. **f**, Left, histogram of t_{dwell} values for specific binding to O_3 (n = 91 individual molecules at LacI–R = 7.3 nM and NaCl = 1 mM). Right, under standard imaging conditions (marked by the asterisk), measurements of t_{dwell} for binding to O_3 are not affected by photobleaching. The mean relative t_{dwell} for specific binding to O_3 (n = 79, 196, 83 and 107 individual molecules from left to right, with mean binding duration ± s.e.m.) observed at NaCl = 1 mM, and using various laser power densities, is shown. Dwell times and laser power densities were normalized to the standard imaging laser power density used for all other analyses (asterisk). **g**, Dependence of the mean t_{wait} (left) and t_{dwell} (right) value for binding on NaCl (LacI–R = 7.3 nM) concentrations. t_{wait} (left), n = 130 or 100 individual molecules for NaCl = 1 mM or for all other conditions, respectively; t_{dwell} (right), n = 190, 164, 164, 82, 184 and 153 individual molecules from left to right. Data are shown as mean ± s.e.m.

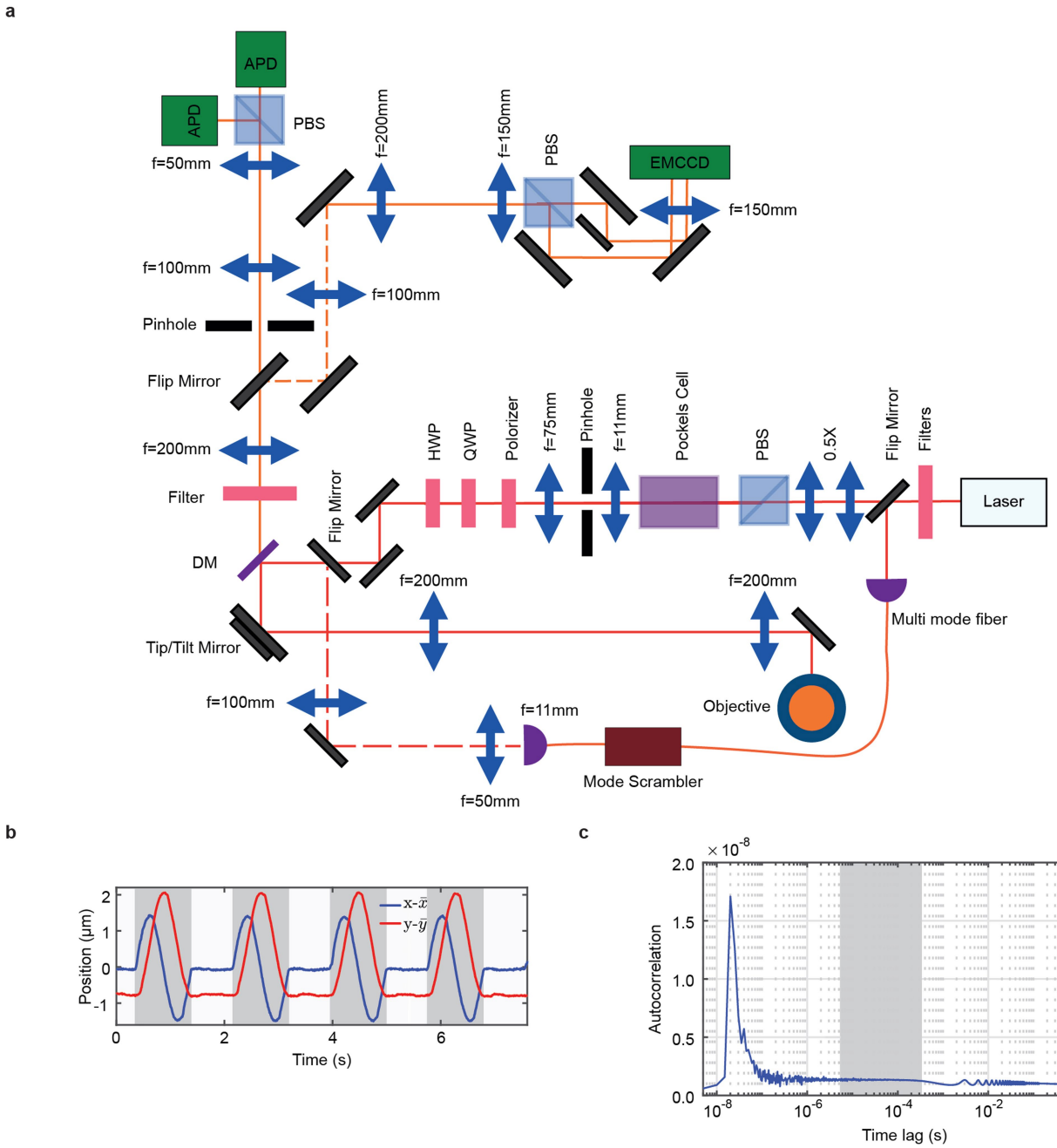


Extended Data Fig. 3 | Predominant switching and flipping transitions. Left, cartoon schematic of the six distinct LacI–R binding states (Supplementary Methods) observed with a construct featuring two outer O_1 sites. Transitions between distinct states are depicted by arrows. Right, switching rates observed at various LacI–R concentrations, normalized to the LacI–R concentration of 0.9 nM, (asterisk, used for the determination of all switching rates shown in Fig. 2), which is in the concentration regime in which switching is not affected by the occupancy of multiple LacI molecules. Error estimates represent s.e.m. (n = 22,720, 12,408, 13,714 and 3,601 individual molecules from left to right).

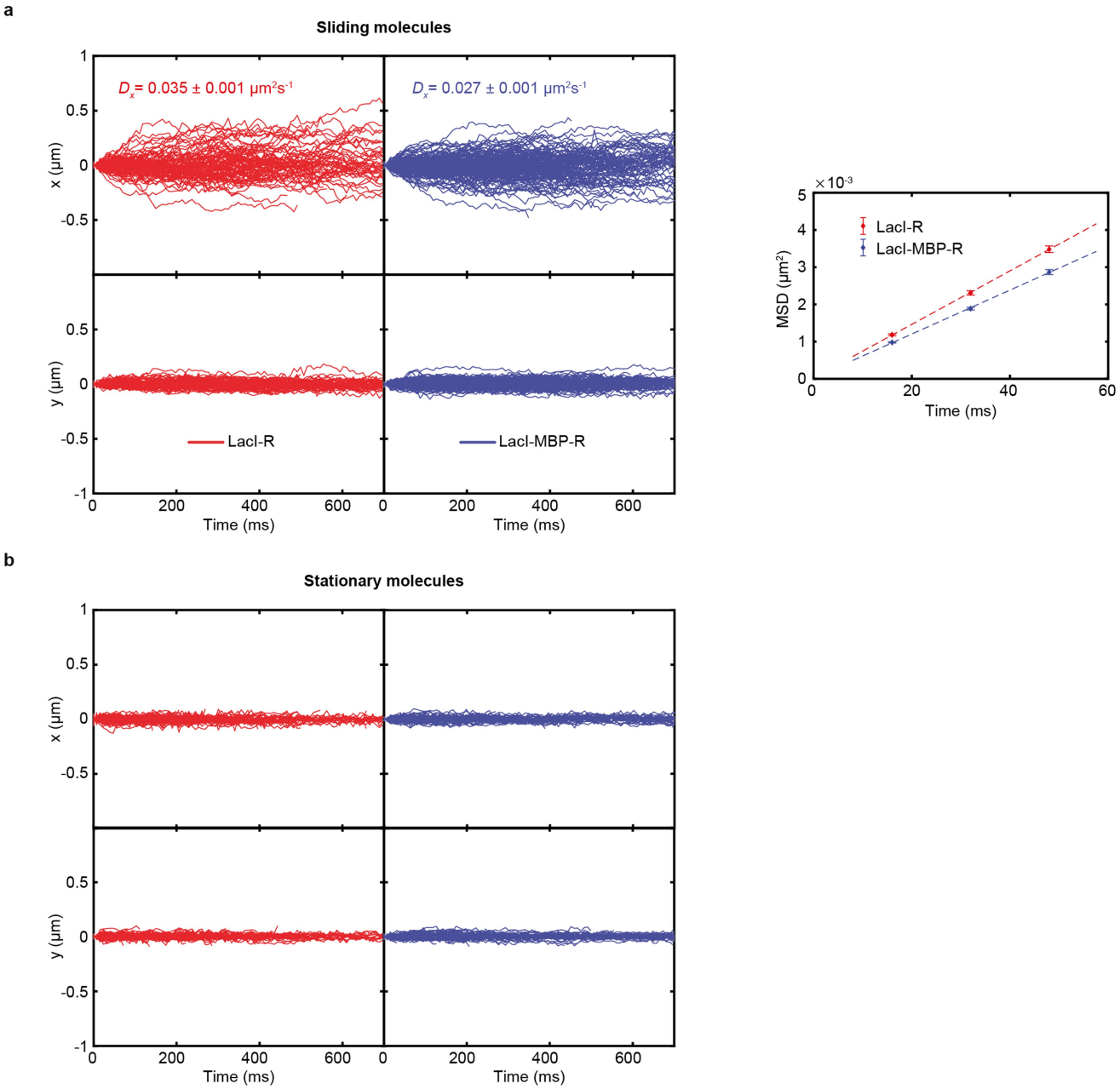


Extended Data Fig. 4 | See next page for caption.

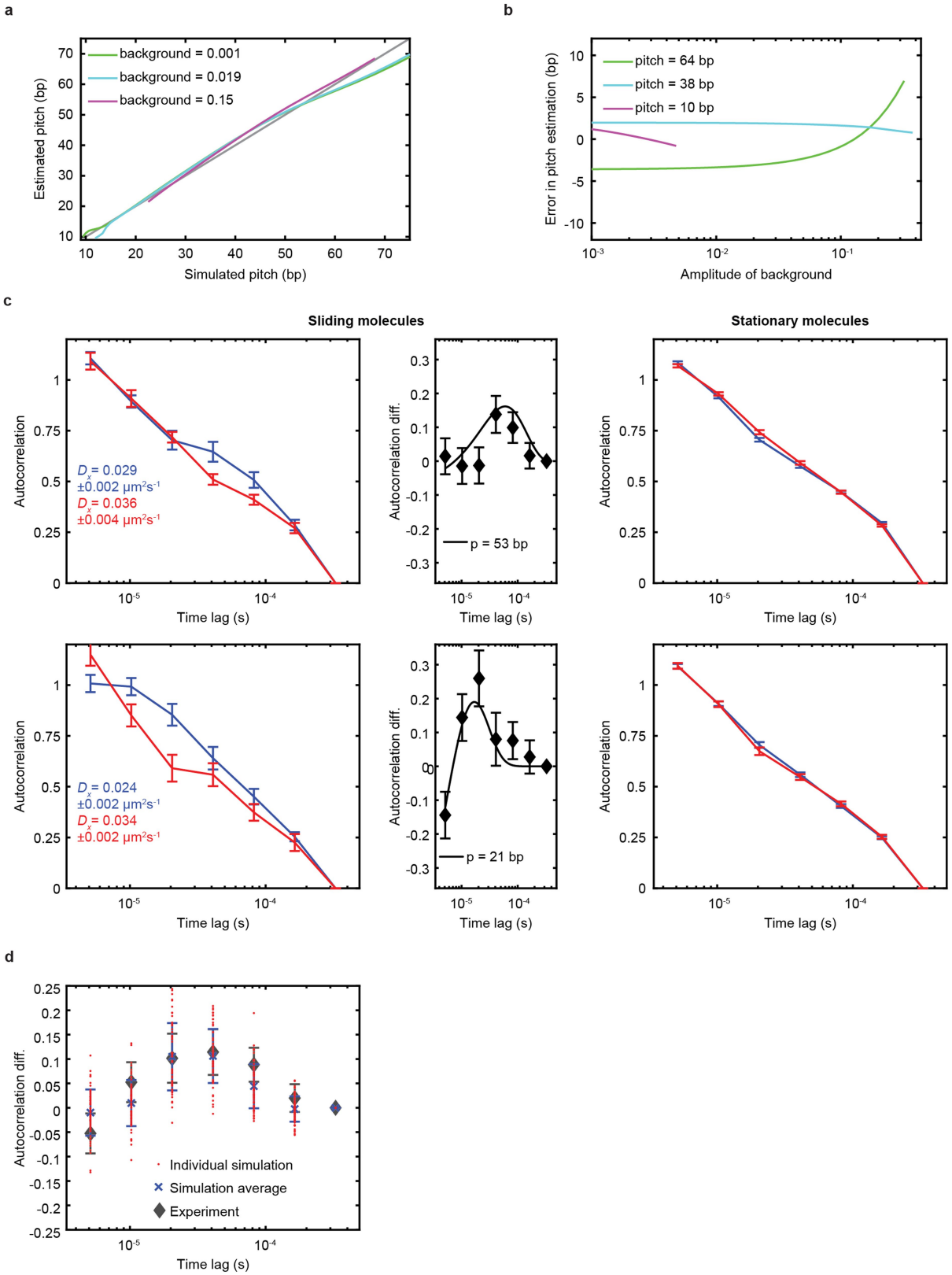
Extended Data Fig. 4 | Camera-based polarization measurements and characterization of dye labelling. a, SDS–PAGE analysis of LacI–R and LacI–Cy3-2 after peptide cleavage with CNBr. LacI–R was designed for bifunctional labelling (rhodamine attachment to both adjacent Cys residues) and LacI–Cy3-2 for monofunctional (using only one of the two Cys) labelling of the same α-helix. Precision Plus Protein Dual Xtra Standards (Bio-Rad) were used as a ladder for both gels. The experiments were repeated at least three times. For gel source data, see Supplementary Fig. 1. **b**, Schematic of the setup for camera-based polarization measurements. **c**, P_0 (blue) and P_{45} (red) polarization distributions averaged over 600 ms (3 frames) per count measured for sliding LacI–R (left) (79 and 172 sliding events for P_0 and P_{45}, respectively) and LacI-Cy3 (right) (61 and 53 sliding events for P_0 and P_{45}, respectively). For P_0 measurements, the horizontal polarization axis was aligned with the stretching direction of DNA; for P_{45} measurements, the horizontal polarization axis was rotated 45° away from the stretching direction of DNA. Details are provided in Supplementary Methods. **d**–**f**, Simulated polarization distributions for the uniform (**d**), linear (**e**) and rotation-coupled (**f**) sliding models. **g**, **h**, Mean ± s.e.m. (**g**) and s.d. ± s.e. (**h**) of the experimental polarization distributions in **c**, **j**. From left to right, n = 800, 1,606, 887, 673, 78 and 104 polarization signals averaged over 600-ms time steps. Error bars represent s.e.m. **i**, Schematic of the 7-kb DNA used in operator-binding polarization measurements (top) and EMCCD image of LacI–R bound to the artificial, strong O_{sym} operator on the flow-stretched DNA, in which 11 binding events were detected for measurements of P_0 (bottom). **j**, P_0 (blue) and P_{45} (red) polarization distributions averaged over 600 ms (3 frames) per count measured for operator-bound LacI–R. Eleven and 15 binding events were detected for measurements of P_0 and P_{45}, respectively.



Extended Data Fig. 5 | Optical layout and calibration data for SMCT–FCS. **a**, Optical layout of the combined system for fluorescence polarization microscopy and single-molecule confocal tracking. In the excitation path, a half-wave plate (HWP) and a quarter-wave plate (QWP) are placed behind a polarizer to create circularly polarized light at the objective. APD, avalanche photodiode; DM, dichroic mirror; HWP, half-wave plate; PBS, polarization beam splitter; QWP, quarter-wave plate. **b**, Immobilized fluorescent beads were used as a platform for testing the tracking capabilities of the real-time tracking system. A tracking trajectory is shown in which the stage is moving in circles with a diameter of 2.8 µm and the photon count rate is on average 24 kc s^{-1} (18 kc s^{-1} including laser off-time). Each greyed-out area represents a 1-s revolution followed by an 800-ms pause. In each paused section, the positioning s.d. is 20 nm and includes stage noise from its feedback loop. **c**, ACF for the bead shown in b. The plot depicts the high afterpulsing peak at 20 ns, which rapidly decreases and flattens out after 500 ns. The grey area corresponds to the time regime plotted for ACFs in Fig. 3. After 1 ms, an oscillation appears which is caused by the 4-ms instrument tracking period. The autocorrelation curve is here compensated for the 166-µs measurement off time between each 500-µs measurement.

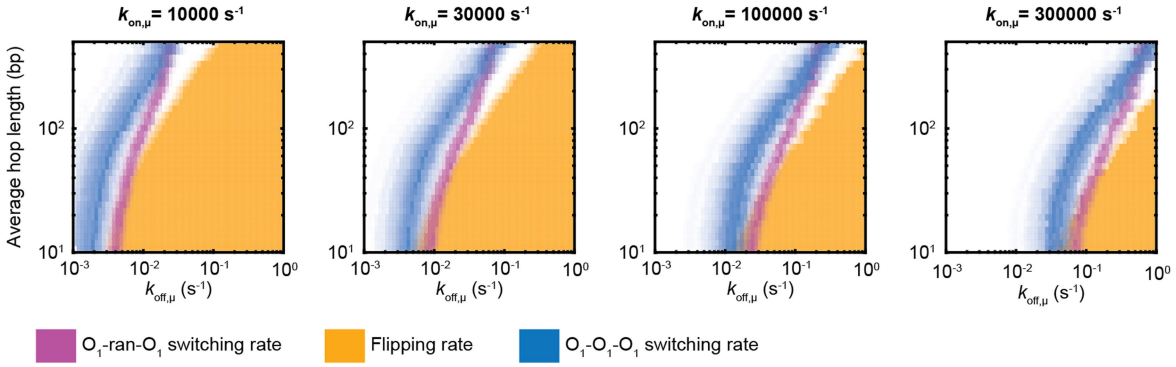


Extended Data Fig. 6 | Traces captured with confocal tracking. a, Left, *x* (DNA direction) and *y* coordinate of sliding LacI–MBP–R (blue) and LacI–R (red) molecules obtained by confocal tracking. In total, 151 and 90 sliding molecules were captured for LacI–MBP–R and LacI–R, respectively. The average diffusion constants along the *x* coordinate were 0.027 ± 0.001 $\mu m^2 s^{-1}$ and 0.035 ± 0.002 $\mu m^2 s^{-1}$. Diffusion constants are indicated as 20% trimmed mean ± s.e.m. Right, mean squared displacement for different time steps for LacI–MBP–R (blue) (n = 151 sliding molecules) and LacI–R (red) (n = 90 sliding molecules) ± s.e.m. **b**, *x* (DNA direction) and *y* coordinate of stationary LacI–MBP–R (blue) and LacI–R (red) molecules obtained by confocal tracking. For clarity, 100 representative trajectories are shown for each LacI species. In total, 3,773 and 1,594 stationary molecules were detected for LacI–MBP–R and LacI–R, respectively.



Extended Data Fig. 7 | See next page for caption.

Extended Data Fig. 7 | Fit of pitch-dependent autocorrelation model and repeats of SMCT–FCS experiments. a, Pitch estimates from simulated autocorrelations as a function of the pitch used in the simulation for different amplitudes of the background. **b**, Error in the pitch estimate as a function of background amplitude for different pitches. To consider only the simulated results relevant for the experimental results observed in Fig. 3, values are reported only for the estimated pitches between 9 and 75 bp and when the largest distance (amplitude) of the difference in the simulated autocorrelation is at least 50% of the amplitude of the observed experimental difference. **c**, The average normalized autocorrelation of the fluorescence signal for sliding (left) and stationary (right) LacI–MBP–R (blue) and LacI–R (red) molecules when the data have been pooled in two series (separated chronologically when the data were captured) for earlier (top) and later (bottom) recorded data. The reported diffusion constants are averages of all tracked molecules for each series. Error bars and error estimates represent s.e. Diffusion constants are indicated as 20% trimmed mean ± s.e.m. For the first experimental series, 69 sliding and 2,758 stationary molecules were analysed for LacI–MBP–R, and 36 sliding and 903 stationary molecules were analysed for LacI–R. For the second experimental series, 82 sliding and 1,015 stationary molecules were analysed for LacI–MBP–R, and 54 sliding and 691 stationary molecules were analysed for LacI–R. Middle, the difference in the mean autocorrelation between LacI–MBP–R and LacI–R ± s.e. Lines represent the best fit to a rotation-coupled sliding model. **d**, Difference in autocorrelation from simulations replicating the experimental datasets and filtering steps. The autocorrelation differences are indicated by the red dots. The number of traces, average trace length and average photon counts were the same as in the experiments, and the rotational diffusion constants were the fitted values from the experiments. The mean and expected s.e. (blue crosses) was calculated as the mean and s.d. of 60 individual simulations. The mean and s.e. of the experiments (black diamonds) were estimated as described in Fig. 3g (n = 86 and 54 informative autocorrelations for LacI–MBP–R and LacI–R, respectively).



Extended Data Fig. 8 | Probability densities of model parameters when taking into account the experimental FRET results. Model parameters from simulations that are compatible with the observed rates of switching between two outer O_1 operators with (blue) and without (purple) an intervening internal O_1 operator, and compatible with the flipping orientation on the same operator (orange). The transparency values of the surfaces have been scaled with the probability density for the parameters given the experimental data. The mean and standard error of the average hop length calculated from the joint probability densities for the four different $k_{on,\mu}$ values are—from left to right—22 ± 16 bp, 18 ± 11 bp, 16 ± 8 bp and 18 ± 12 bp. The flipping rate surface does not decrease with increasing values of $k_{off,\mu}$, as the simulations yield the maximum possible flipping rate. In other words, parameters that correspond to a higher maximum flipping rate could have also generated a lower actual flipping rate. Although $k_{on,\mu}$ was finite and fixed at four different levels in these simulations, $k_{on,\mu}$ was infinite in the simulations shown and used in Fig. 4c–e. The sample size was n = 1,108, 574 and 699 individual molecules for the O_1-random-O_1 switching rate, the LacI–R flipping rate and the O_1-O_1-O_1 switching rate, respectively.

Extended Data Table 1 | Sliding events in EMMCD and confocal tracking experiments

Experiment	Number of single-molecule detections	Number of sliding events	Sliding event per single-molecule detection
LacI-R, 300 pM λ-DNA, camera tracking	7782	251	0.032
LacI-R, (-) λ-DNA, camera tracking	968	2	0.0021
LacI-MBP-R, 300 pM λ-DNA, confocal tracking	8780	151	0.0172
LacI-MBP-R, (-) λ-DNA, confocal tracking	294	0	<0.0034

Molecules are classified as sliding by our data analysis pipeline; sliding events in negative control measurements without λ-DNA thus reflect molecules that are falsely classified as sliding. A single-molecule detection event is defined as a captured trajectory that lasts for at least 600 (camera-based) or 200 (confocal tracking) ms. Where no sliding events were detected, the upper limit for the final column in the table was calculated by assuming a single sliding event.

nature research

Corresponding author(s): Johan Elf and Sebastian Deindl

Last updated by author(s): Mar 25, 2020

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☒	☐	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	SMCT-FCS experiments: Custom-developed LabView (Version 2016 - 32 bit) code. smFRET experiments: For data collection, ImageJ (Version 1.48) as well as MicroManager (Version 1.4.19) were used (both open source)
Data analysis	SMCT-FCS experiments: Custom-developed MATLAB (R2018a) code. smFRET experiments: Data analysis was carried out with IDL (Version 8.7.0), ImageJ (Version 1.48), Excel for Mac 2011, Matlab (R2014b, R2017b), Python (Version 2.7.9) and HMMlearn (Version 0.2.0) using their standard functions/modules.

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw data and analysis software that were developed for this project are available from the corresponding authors upon reasonable request.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	SMCT-FCS experiments: Every experiment with SMCT-FCS yielded 10-20 sliding trajectories. Data were collected twice for LacI-R and three times for LacI-MBP-R until ~50 trajectories had been recorded for each protein. Analysis indicated that this sample size was sufficient large, since the resulting mean diffusion coefficients and autocorrelation functions exhibited sufficiently low s.e.m. to allow meaningful interpretation and comparison of the data. The entire experimental series was then repeated, so that approximately the same number of trajectories were collected again, in two experiments for LacI-R and three experiments for LacI-MBP-R. When compared with the first experimental series (Extended Data Fig. 7), the second set of experiments yielded consistent mean values for diffusion coefficients and autocorrelation functions. smFRET experiments: Sample sizes were not predetermined, but rather experiments were repeated until s.e.m. values were sufficiently low to facilitate meaningful comparison of the measured parameters under different experimental conditions. A minimum of three independent experiments were performed for each measurement.
Data exclusions	No data were excluded from analysis. Data filtering is described in the Methods section, Extended Data Table 2, and Supplementary Table 1-2.
Replication	For SMCT-FCS, the entire experimental series was repeated as shown in Extended Data Fig. 7. All values determined by single-molecule FRET measurements were obtained in at least three independent experiments with similar outcomes.
Randomization	The order of carrying out single-molecule experiments under different conditions was varied between distinct data collection sessions.
Blinding	Data analysis was not blinded. However, critical analysis steps that could be susceptible to human bias were automated.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☒	☐ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☒	☐ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging