http://www.diva-portal.org

## Postprint

This is the accepted version of a paper published in Nature. This paper has been peerreviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):
Marklund, E., van Oosten, B., Mao, G., Amselem, E., Kipper, K. et al. (2020)
DNA surface exploration and operator bypassing during target search
Nature, 583(7818): 858-+
https://doi.org/10.1038/s41586-020-2413-7

Access to the published version may require subscription.
N.B. When citing this work, cite the original published paper.

Permanent link to this version:
http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-439327

# DNA surface exploration and operator bypassing during target search 


#### Abstract

E. Marklund ${ }^{1^{*}}$, B. van Oosten ${ }^{1^{*}}$, G. Mao ${ }^{1^{*}}$, E. Amselem ${ }^{1^{*}}$, K. Kipper ${ }^{1}$, A. Sabantsev ${ }^{1}$, A. Emmerich $^{1}$, D. Globisch ${ }^{2}$, X. Zheng ${ }^{1}$, L. C. Lehmann ${ }^{1}$, O. Berg ${ }^{1}$, M. Johansson ${ }^{1}$, J. Elf ${ }^{1}$ \& S. Deindl ${ }^{1}$ ${ }^{1}$ Department of Cell and Molecular Biology, Science for Life Laboratory, Uppsala University, Box 596, 75124, Uppsala, Sweden, ${ }^{2}$ Department of Medicinal Chemistry, Science for Life Laboratory, Uppsala University, Box 574, 75123, Uppsala, Sweden, *These authors contributed equally to this work.

Correspondence: J.E. (johan.elf@icm.uu.se), S.D. (sebastian.deindl@icm.uu.se). Many proteins that bind specific DNA sequences search the genome by combining three dimensional (3D) diffusion with one dimensional (1D) sliding on non-specific 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 1D phase of target search. To track the rotation of sliding LacI molecules on the microsecond time scale, we use real-time single-molecule confocal laser tracking combined with fluorescence correlation spectroscopy (SMCT-FCS). The fluorescence signal fluctuations are accurately described by rotation-coupled sliding, where LacI traverses $\sim \mathbf{4 0}$ base pairs (bp) per revolution. This distance substantially exceeds the 10.5 -bp helical pitch of DNA, suggesting that the sliding protein frequently hops out of the DNA groove, which would result in frequent bypassing of target sequences. Indeed, we directly observe such bypassing by single-molecule fluorescence resonance energy transfer (smFRET). A combined analysis of the smFRET and SMCTFCS data shows that LacI hops one to two grooves (10-20 bp) every 200-700 $\mu \mathrm{s}$. Our data suggest a speed-accuracy trade-off during sliding; the weak nature of non-specific protein-DNA interactions underlies operator bypassing but also speeds up sliding. We anticipate that the SMCT-FCS method to monitor rotational diffusion on the microsecond time scale while tracking individual molecules with millisecond time resolution will be applicable to the real-time investigation of many other biological interactions and effectively extend the accessible time regime 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 ${ }^{1}$, whereby proteins search the genome by combining 3D diffusion with 1D sliding on $\mathrm{DNA}^{1,2,5,6}$. Little is however known about how sliding proteins explore the DNA surface. For example, it is unknown if the sliding protein redundantly samples each base, as would be expected for faithful 1D diffusion, or if the protein trades redundancy for speed and occasionally skips bases. Here, we shed light on the trade-off between sliding speed and accuracy in target site recognition 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 single-molecule fluorescence resonance energy transfer (FRET) ${ }^{7,8}$ to directly monitor the kinetics with which LacI binds to its natural lac $O_{1}$ operator site $\left(O_{1}\right)$. We generated a DNA construct containing $O_{1}$ with a Cy5 acceptor dye 5 bp from the $O_{1}$ edge. The DNA was surface-immobilized, and individual DNA molecules were monitored with a total-internal-reflection fluorescence (TIRF) microscope (Fig. 1a). Upon addition of LacI labelled with rhodamine in the DNA-binding domain (LacI-R, Extended Data Fig. 1, Supplementary Table 1), its specific binding (Extended Data Fig. 2) to $O_{1}$ led to a sudden appearance of fluorescence signals and FRET (Fig. 1b). LacI containing a single donor (onestep photobleaching) on the distal or proximal subunit gave rise to a FRET distribution for binding events with peaks at $\operatorname{FRET}=0.16 \pm 0.001$ and $0.89 \pm 0.002$, respectively (Fig. 1c and Methods). The rate of operator binding ( $k_{\text {on,obs }}$ ) depended on both LacI and $\mathrm{Na}^{+}$ concentrations (Fig. 1d and Extended Data Fig. 2g) and the rate of LacI dissociation ( $k_{\text {off,obs }}$ ) increased with increasing $\mathrm{Na}^{+}$concentration (Extended Data Fig. 2d,g). Importantly, LacI-R retained the ability to bind to both naturally occurring $O_{1}$ and $O_{3}$ with binding affinities essentially identical to the ones previously reported ${ }^{9}$ (Fig. 1e), indicating that labelling of LacI-R did not substantially impact its operator binding. The binding affinity of LacI-R was also compared to a construct where LacI was labelled further 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 the specific nor the non-specific DNA binding. Interestingly, FRET traces of Lacl 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, where 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 (ran) 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 Alexa750 as distinct acceptor dyes. In the presence of LacI-R, traces from individual $O_{1-}$ ran- $O_{1}$ molecules exhibited spontaneous 'switching' transitions due to LacI sliding from one outer operator site to the other (Fig. 2a). These switching transitions involved a single LacI dimer, based on the following observations. First, switching transitions were marked by fluorescence signals appearing in the one acceptor channel and simultaneously disappearing in the other. In stark contrast to these rapid changes, the corresponding mean waiting time for a single binding event at the same LacI concentration (Fig. 1d) was substantially longer with $t_{\text {wait }}=199 \mathrm{~s}$ (see also 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 (see also Supplementary Information 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), yet 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 next 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 $\lambda$ 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, implying a non-random fluorophore orientation during LacI sliding (Fig. 3a, Extended Data Fig. 4b-j, Supplementary Information 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 time scale ${ }^{15,16}$. In order to more directly observe the
sliding protein rotate around the DNA, we therefore combined real-time single-molecule confocal laser tracking with fluorescence correlation spectroscopy (SMCT-FCS) (Fig. 3b). This allowed us to monitor rotational diffusion on the microsecond time scale (Fig. 3b-f), at the same time as translational diffusion was tracked on the millisecond time scale. 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 them as sliders or nonsliders (i.e., protein stuck on the glass surface) (Methods and Supplementary Tables 2 and 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 (LacI-MBP-R), we measured slower translational diffusion ( $0.027 \pm 0.002$ $\mu \mathrm{m}^{2} \mathrm{~s}^{-1}$ versus $0.035 \pm 0.002 \mu \mathrm{~m}^{2} \mathrm{~s}^{-1}$; Fig. 3d, Extended Data Fig. 6a) as well as slower ACF decay in the 20 to $100 \mu$ s range (Fig 3e). No such difference between LacI-R and LacI-MBPR 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 therefore show that LacI sliding is coupled to its rotation around the DNA with characteristic decay times on the order of $40 \mu \mathrm{~s}$.

## LacI slips out of the groove

To estimate the bp distance that LacI translocates per revolution, we fit the difference in ACF to a model where 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 the background noise processes obtained from the stationary molecules (Extended Data Fig. 7a,b). For the experimental data, fitting resulted in a pitch estimate of $39 \pm 9 \mathrm{bp}$ (Fig. 3g, see also 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 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 and Methods), confirming 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 where LacI sometimes slips between grooves via microscopic hops would agree with these observations and contribute to operator bypassing.

To determine which microscopic parameters for hopping, i.e., hop length and frequency, are consistent with the experimental observables, i.e., 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 (see Methods). Since both the overall 1D diffusion rate and the pitch of rotational sliding are known from the SMCT-FCS experiments, the hop frequency $k$ can be calculated for each hop length $x$ (see Fig. 4b). The absolute hop and switching rates also depend on how often LacI dissociates from the operator (Fig. 4c,d), since LacI cannot hop if it is bound to the operator. The experimental flipping rate (dashed line, Fig. 4c) represents a lower bound for the absolute hop rate, since 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, where long and rare hops that do not frequently bypass the intervening operator are compensated for by more frequent dissociation (dashed line, Fig. 4d). By determining where the dashed lines in Fig. 4c,d overlap (see Methods) to satisfy the experimental constraints (Fig. 4e), we find that the average hop length cannot exceed $16 \pm 8$ bp , corresponding to a minimum hop frequency of $4 \pm 1 \mathrm{~ms}^{-1}$.

## Discussion

Our data show that LacI rotates while sliding, and with a pitch that exceeds the $10.5-\mathrm{bp}$ 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, since 1 D 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 $\sim 100 \%$ in stochastic simulations of target search, despite the frequent bypassing (Supplementary Information 3.5).

Finally, we anticipate that SMCT-FCS will lend itself to characterising many molecular interactions hundreds of times faster than what has been accessible hitherto ${ }^{18}$.

## References

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. U. S. A. 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. U. S. A. 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. U. S. A. 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. U. S. A. 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, 3-6 (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. Threedimensional 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. U. S. A. 106, 17741-17746 (2009).
17. Magde, D., Elson, E. \& Webb, W. W. Thermodynamic fluctuations in a reacting systemmeasurement 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).

Figure 1 LacI-operator interaction dynamics. a, FRET detection for LacI dimer binding to its operator (orange). FRET donor: green; acceptor: red. b-e, Binding to $O_{1}$. b, Traces showing binding of LacI containing a single donor on the proximal (left) or distal (right) monomer. $t_{\text {wait }}$ and $t_{\text {dwell }}$ : waiting and dwell times, respectively. a.u., arbitrary units. c, Histogram from 6,500 binding events. Peaks at FRET $=0.16 \pm 0.001$ and $0.89 \pm 0.002$ (centroid values and standard errors from Gaussian fitting): proximal and distal labelling configurations. d, Dependence of $t_{\text {wait }}$ on [LacI] ([ NaCl$]=1$ mM ). Data points are mean $\pm$ standard error of mean (s.e.m). from fitting a single-exponential decay to the distributions ( $n=310$ or 100 trajectories for [LacI] $=7.3 \mathrm{nM}$ or otherwise, respectively). e, Rates of observed binding to (left: $k_{\mathrm{on}, \mathrm{obs}}=t_{\text {wait }}{ }^{-1}[\mathrm{LacI}]^{-1}$ ) and dissociation from (middle: $k_{\mathrm{off}, \mathrm{obs}}$ ) $O_{1}$ and $O_{3}$ and dissociation constants $K_{\mathrm{d}}$ (right). [LacI-R] $=7.3 \mathrm{nM},[\mathrm{NaCl}]=1 \mathrm{mM}$. Mean rate constants $\pm$
s.e.m. derived from $\mathrm{t}_{\text {wait }}$ (for $k_{\text {on,obs }}: n=248$ and 109 for $O_{1}$ and $O_{3}$, respectively) or $\mathrm{t}_{\text {dwell }}$ (for $k_{\text {off,obs }}: n=$ 367 and 91 for $O_{1}$ and $O_{3}$, respectively). $K_{\mathrm{d}}=k_{\mathrm{off}, \text { obs }} / k_{\mathrm{on}, \mathrm{obs}}$ with errors from propagating errors in the rate constants. f, Flipping. Left: LacI-R transitions from low $\rightarrow$ high FRET. Centre: three LacI binding states observed with a single $O_{1}$. The flipping rates, $0.0011 \pm 0.0004 \mathrm{~s}^{-1}$ (distal $\rightarrow$ proximal; $n=28,000$ s of observed time in each state) and $0.0013 \pm 0.0004 \mathrm{~s}^{-1}$ (proximal $\rightarrow$ distal; $n=28,000 \mathrm{~s}$ of observed time in each state), are mean $\pm 95 \%$ confidence intervals. Right: Flipping rates observed at various [LacI-R], normalized [LacI] $=7.3 \mathrm{nM}$ (asterisk). Values are mean $\pm 95 \%$ confidence intervals from $n$ $=306,189$, 262, and 574 molecules (left to right). c-f, Data from at least 3 independent experiments at each condition.

Figure $2 \mid$ 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 $O_{1}$ : Alexa750 (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). $\mathbf{b}$, Left: Schematic of DNA constructs with two outer $O_{1}$ sites (orange) and intervening random (ran) DNA (left panel) or an additional $O_{3}$ (brown, middle panel) or $O_{1}$ (right panel) site. Top and bottom $O_{1}$ : Alexa750 (purple) and Cy 5 (red) acceptor, respectively. Right: Hidden Markov model (HMM)-derived switching rates (Extended Data Fig. 3) are mean $\pm 95 \%$ confidence intervals ( $n=1108,1137$, and 699 individual molecules from left to right) from at least 10 independent experiments for each construct.

Figure 3 | Pitch determination 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. 52 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. a.u., arbitrary units. d, x(DNA direction) and y-coordinate of sliding LacI-MBP-R (blue, 151 molecules) and LacI-R (red, 90 molecules). Shown: 10 representative traces (see also 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) ( $n=2273$ and 1064 informative autocorrelations for LacI-MBP-R and LacI-R, respectively) LacI-MBP-R (blue) and LacI-R (red). Diffusion constants are $20 \%$ trimmed mean $\pm$ 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) ( $n=2273$ and 1064 informative autocorrelations for LacI-MBP-R and LacI-R, respectively) molecules. Data in (e-h) are mean $\pm$ s.e.m. Black line in (g): best fit to a rotation-coupled sliding model; bootstrapping the trajectories yields an average pitch for the rotation of $39 \pm 9 \mathrm{bp}$ (mean $\pm$ s.e.m). $\mathbf{i}$, 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. $\mathbf{j}$, Best fits of the rotation-coupled sliding model when the pitch is constrained at different levels.

Figure $4 \mid$ Determination of hop length and frequency. a, Schematic of target search where LacI combines faithful DNA groove tracking (1D diffusion) with short and frequent inter-groove hops that can bypass a specific binding site (orange). $\mathbf{b}$, The effective rotational $\left(D_{r}\right)$ and translational $\left(D_{l}\right)$ diffusion coefficients from SMCT-FCS are used to set parameters $\left(k_{h o p}\left(x_{h o p}\right)\right)$ 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, 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 $k_{\text {off }, \mu}$, since the simulations yield the maximum possible flipping rate. See also Extended Data Fig. 8.

Supplementary Information is linked to the online version of the paper.

Methods can be found in the Supplementary Information.

Author Contributions E.M., B.v.O, M.G., 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 HPLC purification and $\mathrm{CNBr}-$ mediate cleavage experiments.; E.M. designed and carried out stochastic simulations, with input from J.E, S.D., and O.B.; S.D., J.E., E.M., and E.A. wrote the paper, with input from all authors.

Author Information The authors declare no competing financial interests. Reprints and permissions information is available at www.nature.com/reprints. Correspondence and requests for materials should be addressed to J.E. (johan.elf@icm.uu.se) and S.D. (sebastian.deindl@icm.uu.se).

Data availability. The data (individual single-molecule fluorescence trajectories from FRET experiments and SMCT-FCS raw data as well as individual SMCT-FCS trajectories) 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).

Acknowledgements We thank Paul Blainey and Kan Xiong for advice on flow-stretching, Chirlmin Joo for discussions, Joakim Laksman for initial code, and David Fange, Irmeli Barkefors and Måns 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). All raw data and analysis software that were developed for this project are available from the corresponding authors (J.E. and S.D.) upon reasonable request.

Extended Data Figure 1: Analysis of LacI labelling and binding. (a) Structural model (based on PDB code: 1OSL). LacI: blue/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 $77 \%$ or $86 \%$ for LacI-R or LacI-MBP-R, respectively. For gel source data, see Supplementary Figure 1. (c) x (DNA direction) and y coordinate of sliding LacI-Far (red) and LacI-R (blue) molecules obtained by EMCCD tracking with a $50-\mathrm{ms}$ frame rate. In total, 779 and 409 sliding molecules were captured for LacI-Far and LacIR , respectively. Diffusion constants along the x -coordinate $\left(D_{\mathrm{x}}\right)$ are mean $\pm$ s.e.m. (d) Mean squared displacement for different time steps for LacI-Far (red; $n=779$ sliding molecules) and LacI-R (blue; $n=409$ sliding molecules) $\pm$ s.e.m. (e) Dissociation constants $\left(K_{\mathrm{d}}\right)$ for LacI-Far and LacI-R. [LacI] = $7.3 \mathrm{nM},[\mathrm{NaCl}]=1 \mathrm{mM} . K_{\mathrm{d}}$ values were calculated as $K_{\mathrm{d}}=k_{\text {off,obs }} / k_{\mathrm{on}, \mathrm{obs}}$ with errors from propagating errors in the rate constants. Rate constants are mean $\pm$ s.e.m. $\left(k_{\text {on,obs }}: n=248\right.$ and 200 events for LacIR and LacI-Far, respectively; $k_{\text {off,obs }}: n=367$ and 322 events for LacI-R and LacI-Far, respectively).

Extended Data Figure 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 ( $-O_{1} \mathrm{DNA}$ ) of acceptor-labelled DNA containing an operator site. Addition of IPTG (+ 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 ( 1244 individual $O_{1}$ DNA molecules). 7.3 nM LacI-R were 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 $t_{\text {wait }}$ values for specific binding to $O_{1}$ ( $n=310$ or 100 trajectories for $[\mathrm{LacI}]=7.3 \mathrm{nM},[\mathrm{NaCl}]=1 \mathrm{mM}$ or otherwise, respectively) at various LacI and NaCl concentrations as indicated. (d) Left: Histograms of $t_{\text {dwell }}$ values for specific binding to $O_{1}(n=190$ and 184 events at [LacI-R] $=7.3 \mathrm{nM}$ and $[\mathrm{NaCl}]=1 \mathrm{mM}$ (top) or 80 mM (bottom)). Right: Under standard imaging conditions (marked by the asterisk), measurements of $t_{\mathrm{dwell}}$ for binding to $O_{1}$ are not affected by photobleaching. Shown is the mean relative $t_{\text {dwell }}$ for specific binding to $O_{1}(n=106,146$, 92 , and 243 individual molecules from left to right, with mean binding duration $\pm$ s.e.m.) observed at $[\mathrm{NaCl}]=1 \mathrm{mM}$ and using various laser power densities. Dwell times and laser power densities were normalized to the standard imaging laser power density used for all other analyses (asterisk). (e) Histogram of $t_{\text {wait }}$ values for specific binding to $O_{3}(n=109$ individual molecules) at [LacI] $=7.3 \mathrm{nM}$ and $[\mathrm{NaCl}]=1 \mathrm{mM}$. (f) Left: Histogram of $t_{\text {dwell }}$ values for specific binding to $O_{3}(n=91$ individual molecules at $[$ LacI-R] $=7.3 \mathrm{nM}$ and $[\mathrm{NaCl}]=1 \mathrm{mM})$. Right: Under standard imaging conditions (marked by the asterisk), measurements of $t_{\mathrm{dwell}}$ for binding to $O_{3}$ are not affected by photobleaching. Shown is the mean relative $t_{\text {dwell }}$ for specific binding to $O_{3}(n=79,196,83$, and 107 individual molecules from left to right, with mean binding duration $\pm$ s.e.m.) observed at $[\mathrm{NaCl}]=1 \mathrm{mM}$ and using various laser power densities. 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 $t_{\text {wait }}($ left $)$ and $t_{\text {dwell }}($ right $)$ value for binding on $\mathrm{NaCl}([L a c I-R]=7.3 \mathrm{nM})$ concentrations. $t_{\text {wait }}$ (left): $n=130$ or 100 individual molecules for $[\mathrm{NaCl}]=1 \mathrm{mM}$ or for all other conditions, respectively; $t_{\text {dwell }}$ (right): $n=190,164,164,82,184$, and 153 individual molecules from left to right. Data are shown as mean $\pm$ s.e.m.

Extended Data Figure 3. Predominant switching and flipping transitions. Left: Cartoon schematic of the six distinct LacI-R binding states (see 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 where switching is not affected by the occupancy of multiple LacI molecules. Error estimates represent standard errors of the mean ( $n=22720,12408,13714$, and 3601 individual molecules from left to right).

Extended Data Figure 4. Camera-based polarization measurements and characterisation 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 $\alpha$-helix. Precision Plus Protein ${ }^{\text {TM }}$ Dual Xtra Standards (Bio-Rad) were used as a ladder for both gels. The experiments have been repeated at least three times. For gel source data, see Supplementary Figure 1. (b) Schematic illustration 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, whereas for $P_{45}$ measurements the horizontal polarization axis was rotated $45^{\circ}$ away from the stretching direction of DNA (see Methods for details). (d)-(f) Simulated polarization distributions for the uniform (d), linear (e), and rotationcoupled (f) sliding models. (g)-(h) Mean $\pm$ s.e.m (g) and standard deviations $\pm$ their standard errors (h) of the experimental polarization distributions in (c,j). From left to right, $n=800,1606,887,673$, 78 , and 104 polarization signals averaged over $600-\mathrm{ms}$ time steps. Error bars represent s.e.m. (i) Schematic illustration of the 7 kb DNA used in operator binding polarization measurements (top) and EMCCD image of LacI-R bound to the artificial, strong $O_{\text {sym }}$ operator on the flow-stretched DNA, where 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. 11 and 15 binding events were detected for measurements of $P_{0}$ and $P_{45}$, respectively.

## Extended Data Figure 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. Abbreviations: polarization beam splitter (PBS), avalanche photodiode (APD), electron multiplying charge-coupled device (EMCCD), dichroic mirror (DM). (b) As a platform for testing the tracking capabilities of the realtime tracking system, immobilized fluorescent beads were used. Here, a tracking trajectory is shownwhere the stage is moving in circles with a diameter of $2.8 \mu \mathrm{~m}$ and the photon count rate is on average $24 \mathrm{kc} / \mathrm{s}$ ( $18 \mathrm{kc} / \mathrm{s}$ including laser off-time). Each greyed-out area represents a $1-\mathrm{s}$ revolution followed by an 800 ms pause. In each paused section, the positioning standard deviation is 20 nm and includes stage noise from its feedback loop. (c) Autocorrelation function for the bead shown in (b). Extraction of photon time tagging information and calculation of the autocorrelation. The plot depicts the high after pulsing peak at 20 ns , which rapidly decreases and flattens out after 500 ns . The grey area corresponds to the time regime plotted for autocorrelation functions 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 $166 \mu \mathrm{~s}$ measurement off time between each $500 \mu \mathrm{~s}$ measurement.

Extended Data Figure 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 $0.027 \pm 0.001 \mu^{2} \mathrm{~s}^{-1}$ and $0.035 \pm 0.002$ $\mu \mathrm{m}^{2} \mathrm{~s}^{-1}$. Diffusion constants are indicated as $20 \%$ trimmed means $\pm$ 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) $\pm$ 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, 3773 and 1594 stationary molecules were detected.

Extended Data Figure 7. Fit of pitch-dependent autocorrelation model and repeats of SMCTFCS 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 only consider simulated results relevant for the experimental results observed in Fig. 3, values are only reported for 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 has been pooled in two series (separated chronologically when the data was 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 standard errors. Diffusion constants are indicated as $20 \%$ trimmed means $\pm$ s.e.m.. For the first experimental series 69 sliding and 2758 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 1015 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 $\pm$ standard error. Lines represent the best fit to a rotation-coupled sliding model. (d) Difference in autocorrelation from simulations replicating the experimental data sets 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 standard error (blue crosses) was calculated as the mean and standard deviation of 60 individual simulations. The mean and standard error of the experiments (black diamonds) were estimated as described in Fig. 3g of the main text ( $n=86$ and 54 informative autocorrelations for LacI-MBP-R and LacI-R, respectively)

## Extended Data Figure 8. Probability densities of model parameters when taking into account

 the experimental FRET results. Model parameters from simulations compatible with the observed rates for switching between two outer $O_{l}$ operators with (blue) and without (purple) an intervening internal $O_{l}$ 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 $k_{\text {on }, \mu}$ values are, from left to right: $22 \pm 16 \mathrm{bp}, 18 \pm$ $11 \mathrm{bp}, 16 \pm 8 \mathrm{bp}$ and $18 \pm 12 \mathrm{bp}$. The flipping rate surface does not decrease with increasing values of $k_{o f f ;}$, , since the simulations yield the maximum possible flipping rate. In other words, parameters corresponding to a higher maximum flipping rate could have also generated a lower actual flipping rate. While $k_{o n, \mu}$ was finite and fixed at four different levels in these simulations, $k_{o n, \mu}$ was infinite in the simulations shown and used in Figure $4 \mathrm{c}-\mathrm{e}$ of the main-text. The sample size was $n=1108,574$, and 699 individual molecules for the $O_{1}$-ran- $O_{1}$ 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.

 Molecules are classified as sliding by our data analysis pipeline; sliding events in negative control measurements without $\lambda$-DNA thus reflect molecules that are falsely classified as sliding. A singlemolecule detection event is defined as a captured trajectory that lasts for at least 600 (camera-based) or 200 (confocal tracking) ms. When no sliding events were detected, the upper limit for the last columns in the table have been calculated by assuming a single sliding event.




