Quantitative Transcription Factor Binding Kinetics at the Single-Molecule Level

Yufang Wang,[†] Ling Guo,[‡] Ido Golding,[‡] Edward C. Cox,[‡]* and N. P. Ong[†]
[†]Department of Physics and [‡]Department of Molecular Biology, Princeton University, Princeton, New Jersey

ABSTRACT   We investigated the binding interaction between the bacteriophage λ-repressor CI and its target DNA using total internal reflection fluorescence microscopy. Large stepwise changes in the intensity of the red fluorescent protein fused to CI were observed as it associated with and dissociated from individually labeled single-molecule DNA targets. The stochastic association and dissociation were characterized by Poisson statistics. Dark and bright intervals were measured for thousands of individual events. The exponential distribution of the intervals allowed direct determination of the association and dissociation rate constants (k_a and k_d, respectively). We resolved in detail how k_a and k_d varied as a function of three control parameters: the DNA length L, the CI dimer concentration, and the binding affinity. Our results show that although interactions with nonoperator DNA sequences are observable, CI binding to the operator site is not dependent on the length of flanking nonoperator DNA.

INTRODUCTION

The CI protein, the product of the *cI* gene of bacteriophage λ, has long served as an important model system for understanding protein-DNA interactions and gene regulation (1,2). In λ-phage, CI monomers assemble into active dimers (CI_2s) that bind to three DNA sites (O_R1, O_R2, and O_R3) on the right operator O_R, which determines genetic switching between the lysogenic and lytic lifestyles. Adjacent dimers interact cooperatively, and these higher-order complexes can interact with similar complexes bound to three related tandem operators (O_L1, O_L2, and O_L3) ~2.4 kilo basepairs (bp) away (1). Understanding the dynamics of binding is important for unraveling the relevant physical mechanism of genetic control, as well as broader problems such as macromolecular recognition.

The interactions between CI and its targets may be the best-studied repressor-operator system in molecular biology (3–7). However, we know nothing about how individual stochastic binding events sum to yield the ensemble averages revealed by nitrocellulose filter binding assays (8), DNase footprinting (9), electrophoretic mobility shifts (10), and isothermal titration calorimetry (6). In recent years, several single-molecule techniques have been introduced to study protein-DNA interactions (11–14). The tethered-particle motion technique (15) provides dynamic information about single-molecule length at 100 nm scale. Fluorescence-aided single-molecule sorting (16,17) provides the fraction of molecules bound in the reaction at the single-molecule level. A very promising technique is single-molecule fluorescence microscopy (SMFM) (18), which is capable of imaging the same molecule for long periods. To study target specificity

and binding on a molecule-by-molecule basis, we used the SMFM method, tethering the target DNA individually labeled with single fluorophore to a surface and labeling the CI protein by fusion to a red fluorescent protein, tdimer RFP (19).

As noted above, because each DNA molecule can interact with several CI_2 molecules, this interaction may have multiple outcomes, with variable numbers of CI_2 molecules interacting with their targets and each other—the regulatory circuit is complex. To investigate these interactions at the most basic level, we restricted our studies to the binding of CI_2 to a single operator sequence, O_R1, thus avoiding the complexities of these higher-order interactions.

In these kinds of experiments, large uncertainties often arise from difficulties in comparing substrate and protein concentrations [CI_2] from sample to sample, and experiment to experiment. We addressed this problem by arranging different DNA populations in an array geometry on the same substrate and immersing the DNA in the same protein solutions. This allowed one variable (e.g., the DNA length L) to be varied while holding [CI_2] fixed. With these DNA arrays, we observed reproducible variations of the dissociation and association rate constants with respect to both L and [CI_2].

Results from DNA molecules with and without O_R1 sites, with O_R1 flanked by a range of different nonspecific DNA lengths, and with CI repressors from both wild-type (WT) and a tight-binder mutant were compared. Large stepwise changes in fluorescence intensity enabled us to resolve clearly the association and dissociation of single CI_2 molecules from their targets. This proved valuable in the stochastic analysis of the intensity versus time traces. We measured dissociation as well as association rate constants k_d and k_a. The values we measured agree qualitatively, although not quantitatively, with those obtained in bulk assays, which provide ensemble averages. We found that operator-specific rate constants had

very little dependence on DNA length. With increasing affinity of a mutant repressor for its target, the dissociation rate constant becomes smaller, whereas the association rate constant increases. We also found that although CI binding to nonspecific DNA sequences is easily observed, nonspecific DNA flanking the O_R1 sequence has no influence on specific rate constants, which suggests that one-dimensional (1D) searching along nonspecific sequences is not an important aspect of repressor targeting over the DNA length L studied here.

MATERIALS AND METHODS

DNA and CI

The O_R1 sequence was inserted into the pGEM-7Zf(+) vector (Promega, Madison WI) at the ApaI site. The local sequence is 5′-…GGG CCC TTA TCA CCG CCA GAG GTA GGG CCC…-3′, where GGG CCC is the ApaI recognition site. The construct was verified by sequencing. Polymerase chain reaction (PCR) was used to amplify the O_R1 sequence together with the flanking sequences using the Expand HiFi PCR system (Roche Diagnostics, Indianapolis, IN). The PCR product was purified with Qiagen QIA PCR Kit (Qiagen, Valencia, CA). As shown in Fig. 1 A, a common left primer (5′-CAG TGA ATT TGT AAT ACG ACT CAC TAT AGG G-3′) was used with different right primers (Table 1) to create DNA of varying lengths. The common left primer has a biotin group attached to the 5′ end to immobilize the DNA. A green fluorescent Alexa Fluor 488 was attached

at the base to the ninth T from the 5′ end of the left primer. All primers, including the modifications, were ordered from Integrated DNA Technologies (Coralville, IA). In particular, the internal Alexa Fluor 488 dT modification is available through off-catalog order. We call the DNA with the O_R1 sequence operator (op) DNA. We also made DNA molecules without the operator sequence using the original vector as the PCR template. We call those molecules nonoperator (nop) DNA. DNA lengths were verified using electrophoresis.

All CI protein preparations used in our single-molecule investigations were fusions with tdimer (a red fluorescent protein, tdimer(2)12 (19)). HN-tagged WT CI constructs, Lys-34 CI, Cys-88 CI, and their respective tdimer fusions were cloned into the PROTet 6 × HN cloning vector by standard methods (20). All plasmid inserts were verified by sequencing. CI function was established by demonstrating that strains carrying each construct in the PROTet vector (CLONTECH Laboratories, Mountain View, CA) displayed the expected sensitivity to λvir and i^{434}, and resistance to λ-PaPa superinfection, hallmarks of functional CI. CI and the tdimer:CI fusions were then purified by Talon column chromatography (CLONTECH Laboratories). The molarities of purified fractions were established by quantitative densitometry of each preparation by comparison with known quantities of bovine serum albumin (BSA) run in adjacent lanes of sodium dodecyl sulfate (SDS)-polyacrylamide gels. Molecules of tdimer:Cys-88 CI, whose threonine residue at position 88 had been replaced by a cysteine, were shown to be dimers by electrophoresis in nonreducing SDS-polyacrylamide gels, as expected (21). Purified fractions were tested for in vitro activity by standard gel-shift assays using a 200 bp DNA fragment containing the λ-O_R operator consensus sequence and nonspecific competitor DNA as a control for sequence specificity (22). The tdimer:CI fusions were as active or more so than respective CI proteins on a molar basis.
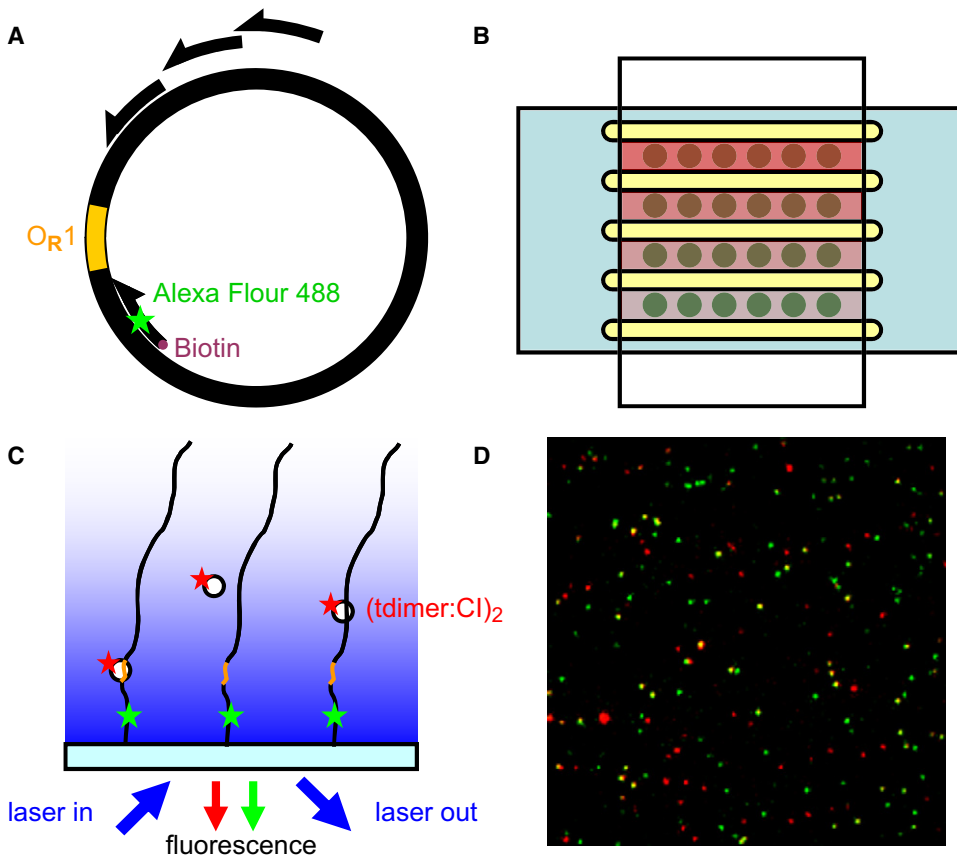

FIGURE 1 Sample preparation and single-molecule imaging of CI-DNA interaction. (A) Schematic description of the PCR primers used here. The O_R1 sequence was inserted into a vector. A common PCR left primer with a 5′ biotin and an internal Alexa Fluor 488 was used in each PCR reaction. The right primers were selected so that different primers would produce different DNA lengths. nop DNA samples were synthesized using the vector lacking the O_R1 sequence as the PCR template. (B) Layout of 24 DNA patches in four channels defined by vacuum grease (*yellow stripes*) between two rectangular coverslips. Each channel had a different [CI$_2$] (shown as different *red opacities*), which allowed us to compare different L DNA at the same CI concentration. (C) Schematic drawing of the reaction in one patch. Green fluorescently labeled DNA molecules were immobilized on the coverslip surface. In solution, and also bound on DNA were (tdimer:CI)$_2$ molecules (CI binds as a dimer). Both molecules were visualized by TIRF microscopy. (D) Superposition of red and green fluorescence images. Green spots are individual DNA molecules fixed to the coverslip surface. Red spots are tdimer:CI molecules. Colocalized DNA and CI (*yellow spots*) indicate CI-DNA binding.

TABLE 1 Right primers used for each DNA sample

DNA	Primer sequence
nop 78 bp	5′-CCGAATTCCTCGAGTCTAGAGG-3′
nop 100 bp	5′-TATCGATTTCGAACCCGGGG-3′
nop 200 bp	5′-GGAAACAGCTATGACCATGATTAC G-3′
nop 300 bp	5′-TAATGTGAGTTAGCTCACTCATTAGGC-3′
nop 400 bp	5′-CAATACGCAAACCGCCTCTC-3′
nop 500 bp	5′-GTGGATAACCGTATTACCGCCT-3′
op 78 bp	5′-CATGCGACGTCGGGC-3′
op 100 bp	5′-GAATTCCTCGAGTCTAGAGGAGC-3′
op 200 bp	5′-GCCAAGCTATTTAGGTGACACTATAGAA-3′
op 300 bp	5′-GGCACCCCAGGCTTTAC-3′
op 400 bp	5′-CGCGTTGGCCGATTCAT-3′
op 500 bp	5′-GAGTGAGCTGATACCGCTCG-5′

The tdimer:CI preparations always contained a contaminating protein with a molecular mass of ~25 kD. Because this is close to the mass of native CI, which would act as a nonfluorescent competitor in our experimental system, we cut out both the tdimer:CI band and the ~25 kD band from a denaturing polyacrylamide gel and identified them by mass spectrometry. The fusion product exhibited the expected peptides from the tdimer and CI. There was no CI signal in the contaminating band. The identified peptides were entirely from the tdimer. We suspect that this contaminant arises either by proteolysis (unlikely) or premature chain termination of the tdimer as it is synthesized.

Preparation of DNA arrays in channels

DNA molecules were immobilized on coverslip surfaces. Glass surfaces were treated as previously described (23). Coverslips (22 mm × 40 mm) were cleaned by first soaking in 5% Contrad 70 (Decon Laboratories, King of Prussia, PA) for 2 h, followed by rinsing and immersion in deionized water for 1 h. They were then rinsed with and immersed in acetone for 5 min, followed by 5 min in Vectabond (Vektor Laboratories, Burlingame, CA) solution (2% v/v in acetone). They were then washed with acetone and deionized water, and blow-dried with clean compressed air. The coverslip surfaces were exposed to polyethylene glycol (PEG) coating solution, which was prepared as 25% (w/v) mPEG-SPA (Nektar Therapeutics, Huntsville, AL) and 0.25% (w/v) biotin-PEG-NHS (Nektar Therapeutics) in 50 mM NaHCO$_3$ (pH 8.5). After incubation in a humid chamber for ≥ 3 h, the coverslip surfaces were rinsed with deionized water and kept in deionized water for an additional 2 h. After blow-drying, the coverslips were exposed to 0.1 mg/ml streptavidin solution for 15 min, rinsed with deionized water, and blow-dried again.

Solutions of biotin-labeled DNA were pipetted onto a coverslip to form a typically 4 × 6 array of patches (Fig. 1 B). After a 15-min incubation in a humid chamber, DNA molecules were immobilized by biotin-streptavidin interaction. To avoid mixing between different DNA populations, excess solution on the coverslip was blotted dry with bibulous paper. Binding buffer (10 mM PIPES, 50 mM KCl, 5% DMSO, 0.1 mM EDTA) was pipetted to the same spots and blotted dry repeatedly to remove unbound DNA. We divided the array into four channels with walls of vacuum grease. The channels were capped with a second coverslip coated with PEG and filled with tdimer:CI solution in binding buffer before they were sealed with beeswax.

Protein absorption on the glass surfaces was minimized by the PEG coating. Nonspecific absorption could change the actual protein concentration at low protein concentration. We observed that protein absorption was at most one molecule per μm^2. The liquid layer was ~200 μm thick. At the lowest CI$_2$ concentration of 70 pM, we estimate that nonspecific binding could decrease the concentration by at most 20%, which would lower the measured association rate constants by only a small amount.

To minimize tdimer:CI denaturation, protein solutions filled into the channels were freshly diluted ~1000-fold into binding buffer from a small amount of the μM stock solution just before each experiment. Concentra-

tions of CI$_2$ were calculated assuming a monomer-dimer dissociation constant of 20 nM (24). By following this protocol, each patch of DNA in the same channel in Fig. 1 B was exposed to the same [CI$_2$]. From one channel to the next, [CI$_2$] increased by a factor of ~3. This allowed us to significantly reduce problems associated with differences in surface preparation and protein concentration. Hence, the 24 patches with varying L and [CI$_2$] were investigated on the same substrate in one experimental run. The patches of DNA molecules were incubated with CI solution for ~60 min before fluorescence microscopy images were acquired.

Time-lapse series were captured for all DNA lengths immersed in a selected CI concentration for each prepared sample. A too-low CI concentration resulted in very few binding events, whereas a too-high concentration resulted in multiple CI$_2$ bindings to a single DNA. The CI concentration that maximized single-CI$_2$ DNA interactions was selected. An experimental run was typically 5–6 h for each prepared array. During each run, six 40- or 60-min time-lapse image series were acquired. The CI protein degradation was insignificant, as judged by fluorescence intensity during the experiment.

SMFM

DNA was immobilized on the coverslip surface and immersed in CI solution. Snapshots and time-lapse images were captured at room temperature by total internal reflection fluorescence (TIRF) microscopy (Fig. 1 C) with a Nikon TE2000-U inverted microscope. Automated image acquisition was achieved with an HF204 emission filter wheel and H112 focus drive (Prior Scientific, Rockland, MA). The sample was illuminated with the argon 488 nm laser line incident near the critical angle.

The evanescent wave depth δ equals $\lambda/2n\pi\sqrt{\sin\theta_i - \sin\theta_c}$, where λ is the light wavelength in vacuum, 488 nm; n is the refractive index of glass, 1.52; and θ_c is the critical angle from glass to water, 61.4°. The excitation intensity at the glass-water interface is also a monotonically decreasing function of the incident angle that can be theoretically calculated (25). We calibrated the incident angle by measuring the intensities of 20 nm fluorescent spheres immobilized on a glass surface. We estimated that the evanescent wave depth was 300 nm or deeper in our experiments. Thus the entire length of the DNA was illuminated. We also note that the longer DNAs, especially those longer than the 200 bp persistence length, describe a volume that is on average closer to the surface, and hence on average the bound CI molecules were in the evanescent field.

Both Alexa Fluor 488 and tdimer were excited at 488 nm. The emission of Alexa Fluor 488 peaks around 517 nm (green), whereas the emission of tdimer peaks around 579 nm (red). Dichroic mirror and green and red emission filters were used to separate excitation light and fluorescence from different molecules (Z488rdc, HQ525/50m, HQ645/75m; Chroma Technology, Rochingham, VT). Images were recorded with an EM-CCD camera (DV877-DCS-BV, Andor Technology, South Windsor, CT). To minimize fluorophore bleaching, the laser power density was kept at ~19 W/cm^2 over an illuminated area of ~60 × 40 μm^2.

For still images, a green frame was acquired, immediately followed by the acquisition of a red frame, each with 1 s exposure time. For time-lapse images, a single green frame was only taken once at the beginning of each time-lapse run with 1 s exposure time, since the DNA molecules were immobilized. Red time-lapse images were acquired with 0.1 s exposure time every 10 or 15 s. The sample is only illuminated while a picture is being taken, to minimize tdimer bleaching and hence maximize event recording. In our experiment, the time it took for a bound CI$_2$ to dissociate from DNA was on the order of minutes. We chose [CI$_2$] so that it would take about the same amount of time for an unbound CI$_2$ to associate with DNA. In Movie S1 in the Supporting Material, the DNA molecules (marked with *blue circles*) can be seen dynamically interacting with CI$_2$.

Image analysis

First, the background fluorescence was subtracted from every image. The background fluorescence comes mainly from the buffer autofluorescence, which is proportional to the illumination intensity. The laser illumination

has a two-dimensional (2D) Gaussian profile. The image background was determined by fitting the image, most of whose pixels were background pixels, to a 2D Gaussian function. After background subtraction, the image was also normalized by the same 2D Gaussian profile to compensate for the nonuniform illumination. Only the center of the image, the region with the highest signal/noise ratio, was used for further analysis.

In each image, fluorescent dye or protein molecules appear as diffraction-limited bright spots. These spots were identified by thresholding. The intensity of a bright spot was calculated as the average intensity of the 3×3 pixels around the local maxima above background. In the still images, the red intensity was normalized by the mean intensity of tdimer:Cys-88 CI molecules, which form covalent dimers independently of concentration (21), because dimers are the smallest active units that interact with DNA. From now on, we will refer to tdimer:CI as CI in this article since all our CI proteins are tdimer fusions.

Cross-correlation was calculated between the green DNA image and the red CI image. Colocalization of CI_2 and DNA resulted in a strong cross-correlation peak, which was also used to align the green and red images. By overlaying the aligned green image on the red one, we could easily identify the binding events of CI_2 to DNA as the colocalization of green and red spots, which acquired a yellow hue.

Time-lapse images were also aligned to correct horizontal drift. For time-lapse images, it was often observed that at the locations of green DNA molecules, the red intensity showed stepwise switches between zero and nonzero levels. We identify these changes with the association and dissociation of red fluorescent CI_2 to and from DNA.

A semiautomated computer program was developed to divide the red intensity versus time curve ($I(t)$) into a series of relative flat segments connected by sharp transitions. Before analyzing the raw trace of red intensity $I(t)$, we measured two intensities I_0 and I_1 in several manually selected traces. I_0 is the average quiescent intensity when no CI_2 proteins are bound to the DNA. I_1 is the average intensity when only one CI_2 is bound. The values of I_0 and I_1 are reproducible from run to run and are well defined throughout the experiment. Assuming that the intensity is linearly scaled with number of CI_2 bounded, the average intensity of more than one CI_2 bound to the DNA can be estimated. Given the raw trace of the intensity $I(t)$ versus t, the program achieves automatic segmentation in two steps. With the average intensities I_n ($n = 0, 1, 2...$) known, the program readily identifies the portions of $I(t)$ belonging to no CI_2, one CI_2, and multiple CI_2 states. However, at this stage, the actual transition time between these states may only be estimated roughly.

To fix the transition time accurately, we employ a minimum-fluctuation method. At such a low fluorescence level of a few CI_2 molecules, the recorded intensity fluctuation is dominated by the dark current of the camera. The intensity root mean-square (rms) amplitude σ_0 is constant regardless of the CI_2 number, as long as the CI_2 number remains small and constant. In a raw trace $I(t)$, when a transition occurs (CI_2 number changes), the rms amplitude increases significantly. We consider $I(t)$ in a time interval $[t_1, t_2]$. If the rms amplitude σ within $[t_1, t_2]$ is $<2\sigma_0$, this interval is treated as a continuation of the previous state. However, if $\sigma \geq 2\sigma_0$, the program recognizes that a transition has occurred. It then picks an arbitrary point t_3 within $[t_1, t_2]$, and calculates the rms amplitude σ_1 within $[t_1, t_3]$ and σ_2 within $[t_3, t_2]$. By minimizing the quantity $(t_3 - t_1)\sigma_1^2 + (t_2 - t_3)\sigma_2^2/t_2 - t_1$ with respect to t_3, the program can then accurately determine the actual transition time. In addition, we also recognize small clusters of $I(t)$ data points with significantly different mean values from the surrounding data points as very short segments.

Because a low signal/noise level can confuse the program, the segmented curves were visually checked and manually corrected if necessary. Only segments that begin and end within the complete trace of $I(t)$ were used for our histograms.
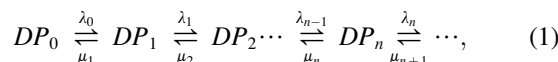
Gel-shift assay

The same fluorescent DNA and CI used in the single-molecule experiments were used for the gel-shift assays. DNA and CI were mixed in binding buffer

supplemented with 100 μg/mL BSA. The binding buffer is also the buffer used in the single-molecule measurements. The concentration of the DNA was 20 nM in every reaction, whereas the concentrations of CI_2 varied (5, 36.5, or 452 nM). The reaction mixtures were incubated at room temperature. After ~1 h, 20 μL reaction mixtures were electrophoresed in Bio-Rad 4–20% Gradient TBE gels (Bio-Rad, Hercules, CA) in a cold (4°C) room. The gels were stained with ethidium bromide for DNA. The gel images were taken under ultraviolet illumination.

Stochastic series

To analyze how the intensity time trace $I(t)$ is informed by the transition rates, and to define some terms, we describe transitions between states DP_n (defined as a DNA bound by n CI_2s) as the time-dependent stochastic series

$$DP_0 \underset{\mu_1}{\overset{\lambda_0}{\rightleftharpoons}} DP_1 \underset{\mu_2}{\overset{\lambda_1}{\rightleftharpoons}} DP_2 \cdots \underset{\mu_n}{\overset{\lambda_{n-1}}{\rightleftharpoons}} DP_n \underset{\mu_{n+1}}{\overset{\lambda_n}{\rightleftharpoons}} \cdots, \qquad (1)$$

where λ_n and μ_n are the forward and backward transition rates governing the time evolution of the probability of finding a complex DP_n at time t.

A powerful feature of dynamic SMFM is that it allows a single transition in the series Eq. 1 to be studied in isolation. In general, the time trace $I(t)$ at a colocalized site shows both single- (see Fig. 3 A) and multiple-step transitions (see Fig. 3 B) relative to the background intensity. We used a semiautomated computer program that efficiently distinguishes large, stepwise transitions from small-amplitude fluctuations.

By collecting the set of traces exhibiting single-step transitions, we accumulated statistics on the transition $DP_0 \rightleftharpoons DP_1$. The "up" and "down" transitions separate the bright durations t_{bright} and dark durations t_{dark}, whose probability distribution can be described by $\mathcal{P}(t_{dark}) = \lambda_0 e^{-\lambda_0 t_{dark}}$ and $\mathcal{P}(t_{bright}) = \mu_1 e^{-(\mu_1 + \lambda_1)t_{dark}}$, respectively (see Fig. 3, C and D). From the state DP_1, the molecule complex goes to DP_0 with rate constant μ_1. DP_1 also goes to DP_2 with rate constant λ_1. During the same period of time, the number of DP_1 to DP_0 transitions was counted as N_{10}, whereas the number of DP_1 to DP_2 transitions was counted as N_{12}. We then have $N_{10}/N_{12} = \mu_1/\lambda_1$. From the individual traces of $I(t)$ at each DNA bound to CI_2 (200 ~300 in the field of view), we accumulated hundreds of measurements of t_{dark} and t_{bright} in a single time-lapse data set. Assuming Poisson processes, λ_0 and $\mu_1 + \lambda_1$ was obtained by fitting the distributions of t_{dark} and t_{bright} to exponential distributions according to the unbiased estimation in Koster et al. (26). We obtain μ_1 by multiplying the second exponential constant with $N_{10}/N_{10} + N_{12}$. Hence, the histograms of t_{dark} and t_{bright} in the filtered set allow λ_0 and μ_1 to be measured directly (see Fig. 3, C and D).

If the transitions to $DP_2, DP_3, \ldots$ are negligible, the association and dissociation rate constants k_a and k_d, of the reaction

$$DP_0 + CI_2 \underset{k_d}{\overset{k_a}{\rightleftharpoons}} DP_1 \qquad (2)$$
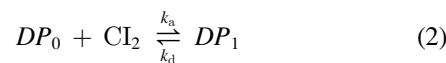
are just $k_a = \lambda_0/[CI_2]$, $k_d = \mu_1$, with the dissociation constant $K_D = k_d/k_a$.

RESULTS

Snapshots and time-lapse images of CI interacting with DNA

In a typical snapshot with the red channel image overlaid on the green channel image, there are many bright spots (Fig. 1 D). The colors of the bright spots fall into three easily separable classes: green, red, and yellow (identified as DNA, CI, and DNA-CI, respectively).

The TIRF technique allows us to capture simultaneously the transitions in each of the DNA molecules within the field of view. When observing single fluorophore-labeled

molecules, the blinking and bleaching of fluorophores might interfere with the measurements.

Molecule bleaching was moderate in our experiments. The average intensity of all fluorescent molecules in the field of view decreased by 10–30% after 100 s of continuous illumination (Fig. 2 A). The intensities of single Alexa Fluor 488 molecules were very stable, except for the occasional characteristic single-step bleaching events (Fig. 2, B and C). The fluorescent intensity of tdimer is more critical. Blinking and bleaching of tdimer might be confused with binding and unbinding of CI_2 from DNA. To control for this possible problem, we investigated tdimer:Cys-88 CI molecules nonspecifically bound to a glass surface. We could identify single tdimer:Cys-88 CI dimer molecules as individual bright spots with similar intensities. With time, the intensities of these spots decreased gradually (Fig. 2, D–G). It is worth emphasizing that the molecules were only illuminated when the image was acquired during the process of time-lapse image acquisition. The accumulated exposure time during a 40- or 60-min movie was ~100 s. With tdimer bleaching slowly and gradually during this period of time, the binding and unbinding of CI_2 can be easily resolved.

In the dynamic experiments, we captured 40- or 60-min movies at each of the six patches in the same channel. Fig. 3 A shows a typical time trace of the fluorescence signal from one molecule over a 40-min interval. The baseline (minimum signal intensity) corresponds to the state with no CI_2 bound to DNA, whereas the plateaus at finite intensity represent the state with one CI_2 bound to DNA. At very low $[CI_2]$, most DNA has not bound CI_2 over the duration of data acquisition. However, at high $[CI_2]$, several finite intensity states were observed in a single time trace, resulting from multiple CI_2 interacting with a single DNA molecule. We adjusted the $[CI_2]$ to optimize the number of single CI_2 binding events. In each field of view, there are time traces for ~200 molecules on average, each displaying one to 10 bright states. As explained above, the "filtered set" contains only the traces with single-step transitions. The histograms of the sets of t_{dark} and t_{bright} in Fig. 3, C and D, are well fitted by an exponential distribution, consistent with Poisson processes and implying that association and dissociation are best modeled as single-step processes. This then allows λ_0 and μ_1 in each of the populations to be determined.

In some data sets, similar dark-bright red intensity switching patterns were observed at locations where there was no green fluorescence. Some of these locations may have harbored a bleached DNA molecule, whereas others may simply have been CI molecules nonspecifically bound to the surface. To make sure we were looking at the interaction between CI_2 and DNA, we counted events only at locations where green fluorescence was high.

CI binds DNA both with and without the operator sequence, showing different binding profiles

We first discuss measurements of binding profiles for single DNA molecules by looking at still images.
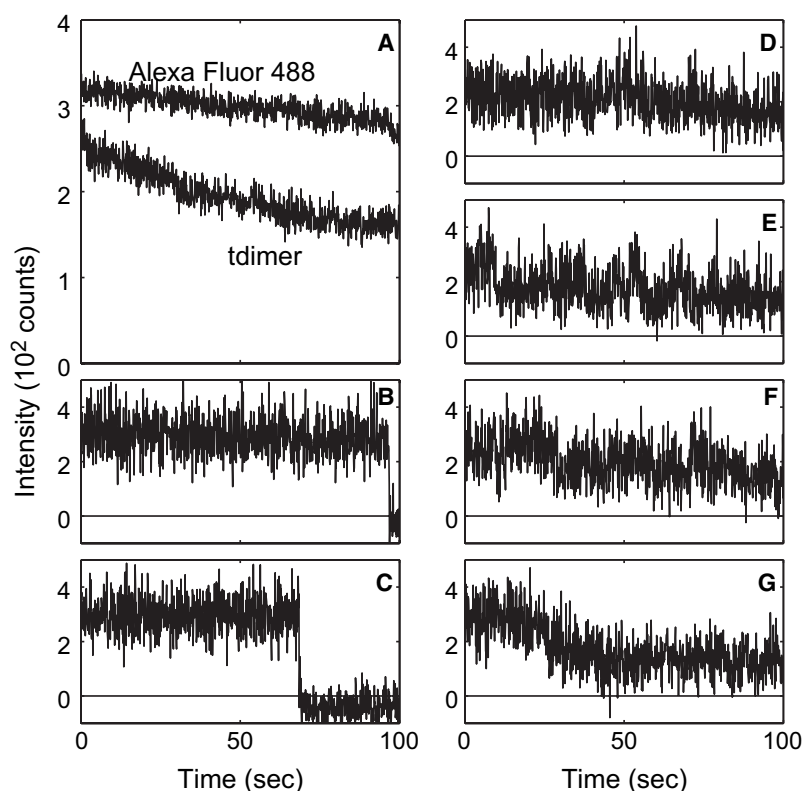

FIGURE 2   Molecular bleaching. (A) Average bleaching curve of all molecules in a field of view as a function of time. The Alexa Fluor 488 fluorescence decreased 10% and the tdimer fluorescence decreased 30% after 100 s of continuous illumination. (B and C) Intensity-time traces of two typical Alexa Fluor 488 molecules. The intensity abruptly dropped to the background level. (D–G) Intensity-time traces of four typical tdimer:Cys-88 CI molecules. Compared with single Alexa Fluor 488 molecules, there was no abrupt, large drop in intensity. With time, the intensities of these molecules decreased gradually.
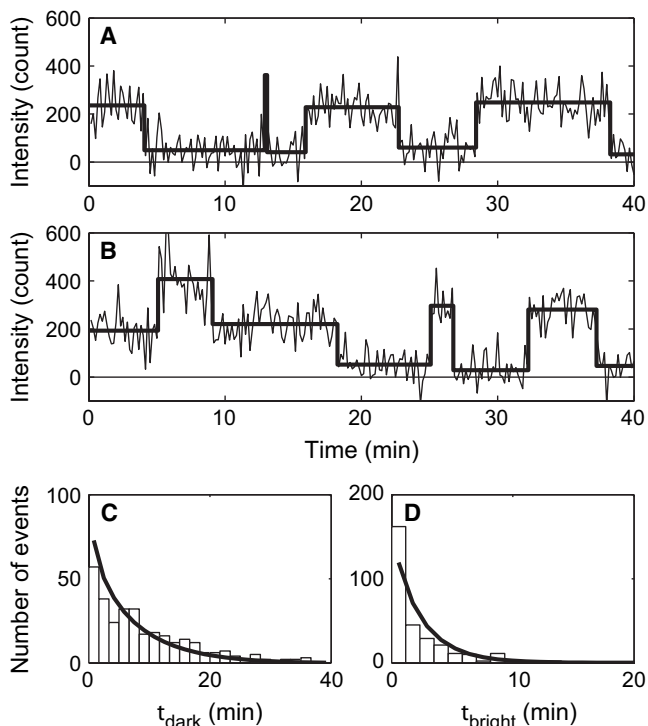
FIGURE 3 Dynamics of CI_2 interacting with DNA (WT CI and op 300 bp DNA are shown here). (*A*) Time trace of red fluorescent signal I versus t from a single op DNA molecule displaying the transitions $DP_0 \rightleftharpoons DP_1$. Bright intensities correspond to binding of a single (tdimer:$CI)_2$. The bold line segments are generated by the transition-recognition program that counts the events. (*B*) A single DNA molecule could be bound by more than one CI. During 5–9 min, this DNA was bound by two CIs. We filter out multiple binding events and accumulate statistics of dark intervals t_{dark} and the bright intervals t_{bright} using traces as in panel *A*. The histograms of the distribution of t_{dark} and t_{bright} are shown in *C* and *D*, respectively. Fits to exponential distributions $\mathcal{P}(t_{dark})$ and $\mathcal{P}(t_{bright})$ (*bold curves*) assuming Poisson processes yields $\lambda_0 = [CI_2]k_a$ and $\mu_1 = k_d$. (Also see main text.)

Recall that the red intensities were normalized using the intensity of single CI_2 molecules. A series of intensity histograms is shown in Fig. 4 for CI_2 interacting with op DNA. The histogram consists of two peaks: a narrow one centered around the background intensity (*dashed line*) identifies the population of DNA without CI_2 bound to it. For DNA from 78 to 200 bp, the position of the second peak is consistent with one CI_2 bound to DNA. These peaks overlap with the histogram of single Cys-88 CI dimer molecules, confirming that CI binds to DNA as a dimer. With a further increase in L, the peak broadens and shifts, implying that two or more CI_2s bind to each DNA molecule as the CI_2 concentration increases.

As mentioned above, the array design allows the binding probability to be investigated as both L and $[CI_2]$ are varied. A revealing way to view the trends is to display in the (L, I) plane the intensity histogram as a heat map, with blue indicating zero counts and red indicating large numbers of counts (Fig. 5). The six vertical strips within each panel display the
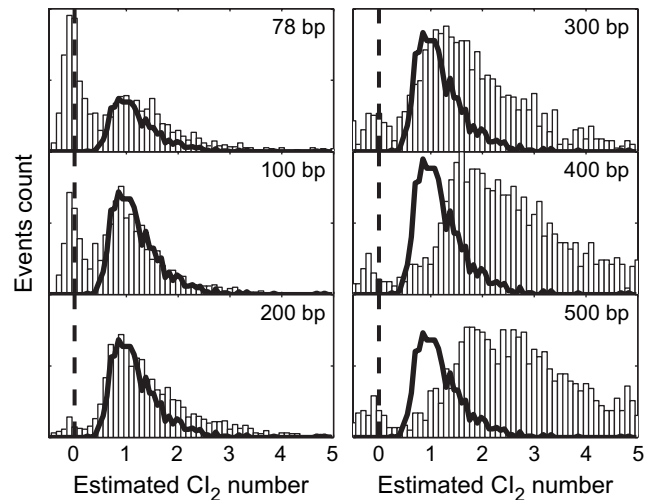
FIGURE 4 Histograms of red fluorescence intensities from (tdimer:$CI)_2$ (Lys-34 CI in this figure) per green fluorescent spot (op DNA in this figure), indicating the number of CI_2s per DNA. DNA lengths L are labeled in each panel. These DNA molecules were exposed to the same $[CI_2]$. Two peaks are apparent in each panel. The narrow peak near zero intensity (aligned with the *dashed line*) is the DNA population with no bound CI_2. The second peak is the DNA population bound with one molecule of CI_2 ($n \geq 1$). As L increases, the peak broadens and shifts to higher intensity, reflecting the increase in the number of CI_2s per DNA. A histogram of tdimer:Cys-88 CI intensities is plotted as a thick line in each panel as the dimer standard.

normalized event-count distribution in the six patches (with successive increasing L) all exposed to the same $[CI_2]$. The four upper panels are for nop DNA, and the lower panels are for op DNA. With $[CI_2]$ held fixed, more CI_2 was observed to bind to op DNA (*lower panels*), as expected from the increased affinity. Clearly, in each panel the high-intensity peak moves to larger values as L increases from 78 to 500 bp. However, in very dilute $[CI_2]$ (*right-most panels*), short (78 ~300 bp) nop DNA molecules mostly had no CI_2 bound to them, whereas the large fraction of short op DNA had one CI_2 bound. It was very likely that this CI_2 molecule bound to DNA at the O_R1 sequence.

To quantify these trends, we calculated the mean number $\langle n \rangle$ of bound CI_2s per DNA molecule in each frame as the mean red intensity normalized by the intensity of single CI_2 molecules. Fig. 6 summarizes how $\langle n \rangle$ varies with L and concentration in the following four cases: WT CI binding to nop (Fig. 6 *A*) and op DNA (Fig. 6 *B*), and the tight-binder Lys-34 CI mutant binding to nop (Fig. 6 *C*) and op DNA (Fig. 6 *D*). In each panel, $\langle n \rangle$ is plotted against L for each of the three or four $[CI_2]$ (nominal values stated in Fig. 6). The measurements were repeated 2–3 times to check reproducibility. For a given $[CI_2]$, we found that $\langle n \rangle$ increased linearly with L, and may be fit by a straight line. With increasing $[CI_2]$, the slope s of the linear fits increased, initially in proportion to CI, but tending toward saturation at high $[CI_2]$ (Fig. 6, *A*, *C*, and *D*).

In general, the L-linear increase in $\langle n \rangle$ reflects the increased availability of nonoperator sites for binding. Significantly, if
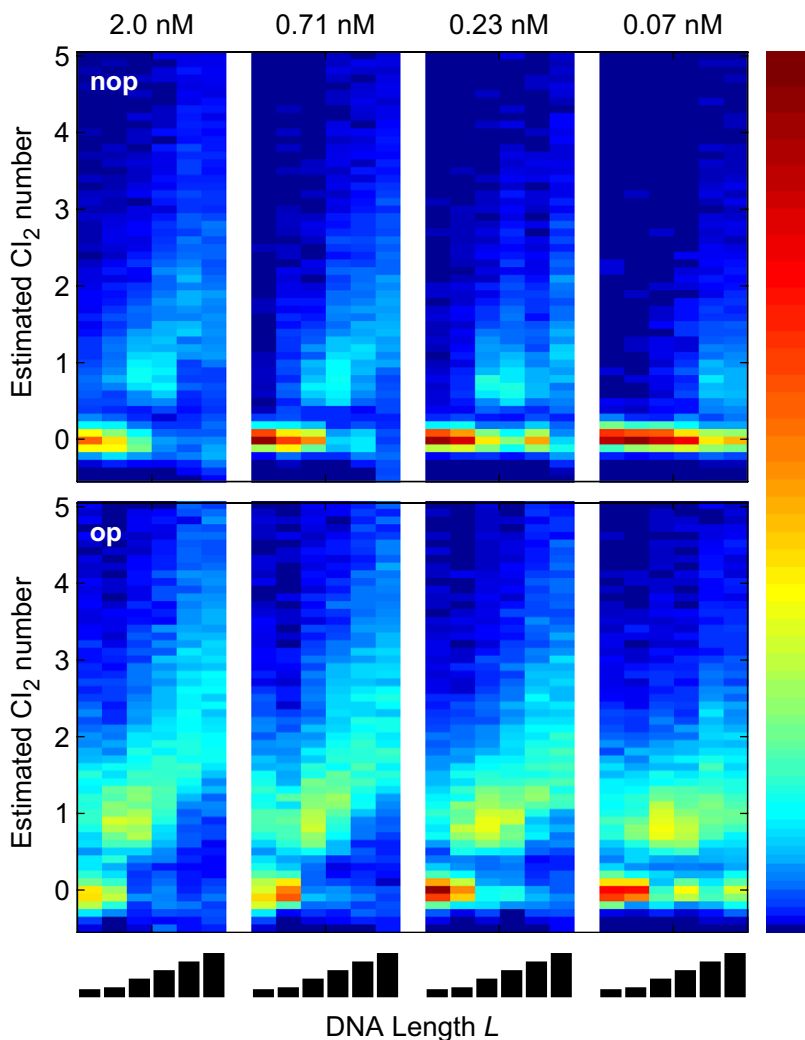
FIGURE 5 Evolution of the fluorescence intensities of Lys-34 CI bound to DNA versus DNA length L and [CI_2]. Normalized frequencies of fluorescence in each intensity bin are shown as a heat map. The DNA is nop (op) in the upper (lower) panels. Each panel shows how the histograms change as L steps through the values 78, 100, 200, 300, 400, 500 bp, with [CI_2] fixed at the value shown on the top. Blue represents zero frequency, and red represents high frequency. In each panel, as L increases, the center of the second peak shifts to a larger CI_2 number as the peak also gets broader, reflected in increasing $\langle n \rangle$. Increasing [CI_2] at fixed L also increases $\langle n \rangle$. At the same L and [CI_2], op DNA captures more CI_2 than nop DNA.

we look closely at the runs with op DNA (Fig. 6, *B* and *D*), we find that $\langle n \rangle$ displays a finite intercept n_0, viz. $\langle n \rangle = n_0 + sL$. At low [$CI_2$], the strong dominance of binding to the O_R1 site over nonoperator sites renders the association rate nearly insensitive to L.

Dissociation and association rate constants

We first compare the binding strengths of WT CI_2 to op and nop DNA embedded in two length extremes: 78 and 500 bp. Fig. 7 *A* shows the measured dissociation rate constant k_d. Values of k_d for nop DNA (*open symbols*) were found to be larger than the op case (*solid symbols*) by a factor of 2–4. As expected, WT CI_2 binds more strongly to op DNA compared to nop. Within the uncertainty of our experiments, k_d was not significantly different between the two lengths, i.e., a sixfold increase in L had no observable effect on how long the CI_2 remained bound. Three runs with nominally similar [CI_2] are shown.

The corresponding association rate constant $\lambda_0 = k_a[CI_2]$ showed the opposite relationship between op and nop

DNA. The values for λ_0 in op DNA (*solid symbols*) were larger than nop (*open symbols*) by a factor of 2–4 (Fig. 7 *B*), consistent with the higher affinity for op DNA. Moreover, the measured λ_0 for op DNA showed no significant dependence on L. This is expected because binding of CI_2 to the single O_R1 site in op DNA dominates the op binding processes, and is independent of L. By contrast, λ_0 should increase with L for nop DNA. We indeed observed that λ_0 in WT+nop (*open symbols*) increased ~4-fold between the 78-bp and 500-bp populations. We revisit this point later.

To bolster these conclusions, we extended measurements to several DNA populations of intermediate L (100, 200, 300, and 400 bp). In addition, we investigated the Lys-34 CI mutant, which binds more tightly than WT CI (27). Fig. 8 summarizes the dissociation rates k_d measured with WT CI + nop DNA (*solid symbols*) and WT CI + op DNA (*open symbols*), and for the tight-binder Lys-34 CI + op DNA (*asterisks*). For each DNA construct, op or nop, six patches of different lengths were investigated. The measured k_d values were well separated in the three populations. In each case, the gray band is centered at the mean value of k_d, and its
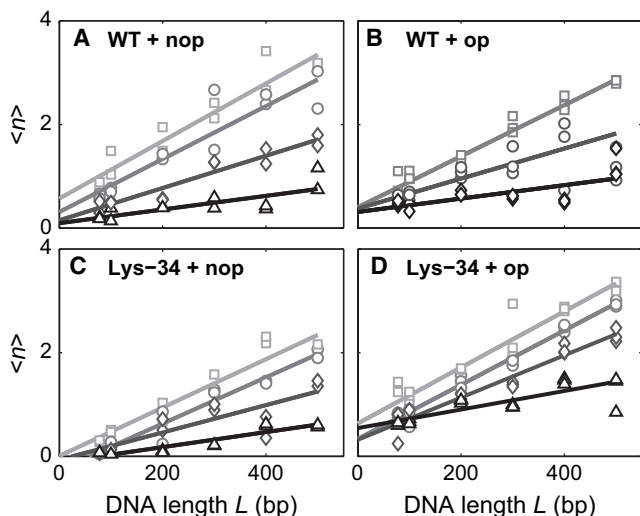
FIGURE 6 Mean number $\langle n \rangle$ of CI_2s per DNA molecule plotted versus L. (A) WT CI and nop DNA, (B) WT CI and op DNA, (C) Lys-34 CI and nop DNA, (D) Lys-34 CI and op DNA. In each panel, $\langle n \rangle$ is calculated as the mean red fluorescent intensity per DNA site, normalized by the intensity of single CI_2. Straight lines are the best linear fits to $\langle n \rangle$. Within each panel, $[CI_2]$ decreases ~3-fold between data sets with successively smaller slopes. The values (in nM) are 25, 11, 4.4, 1.7 (A); 4.4, 1.7, 0.6 (B); 2.0, 0.71, 0.23, 0.07 (C); and 2.0, 0.71, 0.23, 0.07 (D).

halfwidth gives the standard deviation. As expected, Lys-34 CI + op DNA showed the strongest binding (smallest k_d), WT CI + op DNA was intermediate, and WT CI + nop DNA had the weakest binding (~10 times weaker than Lys-34 CI + op DNA). In addition, the results strongly reinforce the conclusion that the k_d in each case is unchanged by a sixfold increase in L.
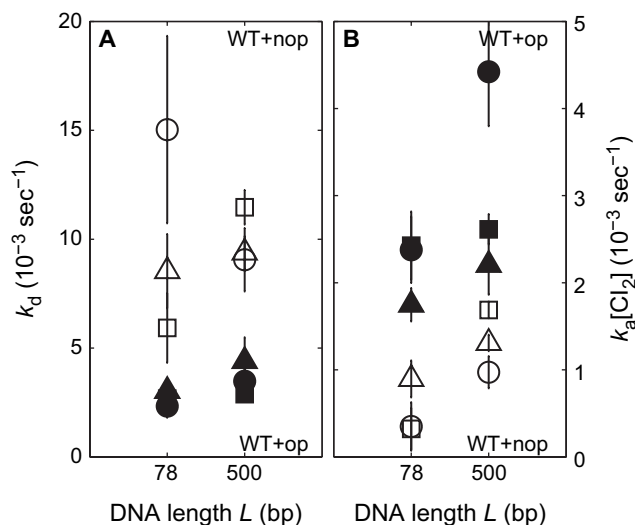

FIGURE 7 Direct comparisons of rate constants for binding of WT CI_2 to nop versus op DNA. Panel A compares the dissociation rate constants k_d for op DNA (solid symbols) with that for nop DNA (open symbols). For each DNA population, two lengths were investigated in three runs at the same nominal $[CI_2]$ of 11 nM (different symbols). Panel B shows the association rate constants k_a extracted from the same experiments.
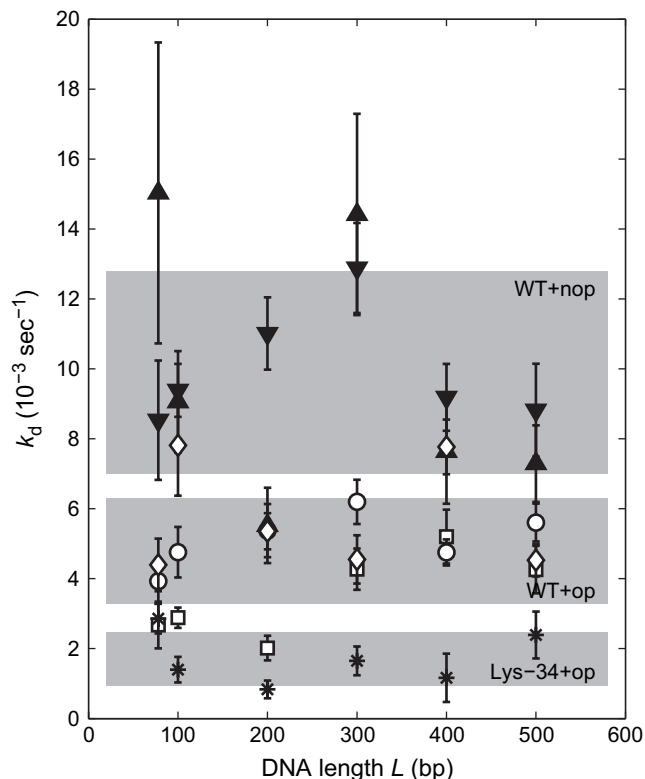
FIGURE 8 Comparison of the dissociation rate constants k_d measured for binding of WT CI to nop DNA (solid symbols) and op DNA (open symbols), and binding of the Lys-34 CI to op DNA (asterisks). Six DNA lengths were investigated (at fixed $[CI_2]$) in each case. For WT CI + nop DNA, $[CI_2]$ was 11 nM in both runs (solid triangles). For WT CI + op DNA, $[CI_2]$ was 11 nM in two runs (open squares and circles) and 1.7 nM in a third run (open diamonds). In Lys-34 CI + op DNA, $[CI_2]$ was 0.07 nM. In each case, a gray band is centered at the mean value of k_d and its halfwidth gives the variance. The observed values of k_d increase significantly as we go from Lys-34+op to WT+op, to WT+nop (successively weaker binding of CI_2). However, k_d does not display a resolvable L or $[CI_2]$ dependence.

The experiments for the WT CI + nop DNA and WT CI + op DNA were run two and three times, respectively, with different sample preparations. The runs are shown as different symbols. Within the error limits, the values of k_d were insensitive to changes in $[CI_2]$, consistent with the expectation that dissociation should not be influenced by $[CI_2]$. This insensitivity to $[CI_2]$ allows the data for k_d from different runs to be meaningfully compared, as in Fig. 8. However, λ_0, which we discuss next, is obviously L dependent.

Fig. 9 plots $k_a[CI_2] = \lambda_0$ obtained in these experiments. For nop DNA (Fig. 9 A), λ_0 displayed a clear increase with L, which confirms the trend hinted at in Fig. 7 B, and is consistent with the results of Wang et al. (28). The nominally L-linear increase is consistent with the expectation that a longer DNA target has a larger probability of capturing CI_2. However, this is valid only for nop DNA. With op DNA, λ_0 showed no resolvable L dependence for both WT CI (Fig. 9 B) and Lys-34 CI (Fig. 9 C). Apparently, the O_R1 binding site, when present, dominates any advantage
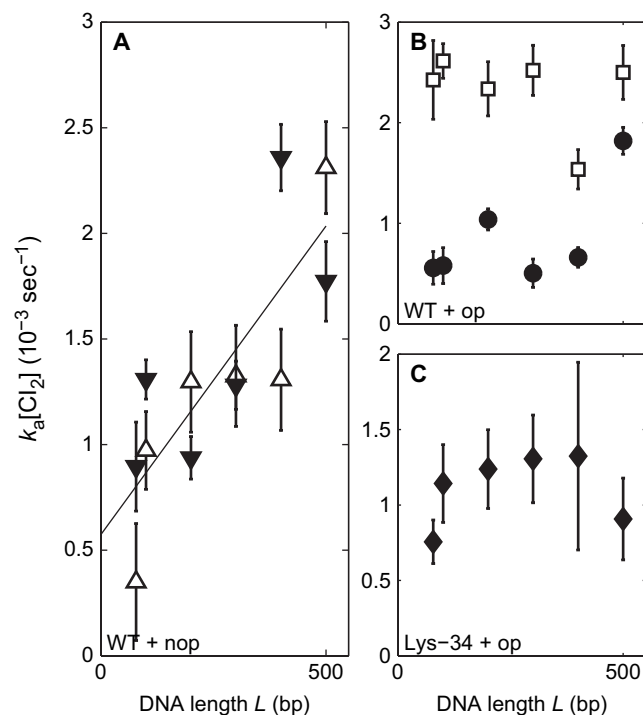
FIGURE 9 Corresponding values of k_a extracted from the series of experiments involving the CI-DNA populations. (*A*) WT CI + nop DNA. The two different symbols are for two different experimental runs. (*B*) WT CI + op DNA. (*C*) Lys-34 CI + op DNA shown in Fig. 8. Panel *A* shows that k_a in nop DNA increases linearly with *L*, whereas in the two op DNA populations, k_a is virtually *L*-independent. In panel *A*, the nominal [CI$_2$] is ~11 nM (*open* and *solid triangles*). In *B*, [CI$_2$] is 11 nM (*open square*) and 1.7 nM (*solid circles*). In *C*, [(Lys-34 CI)$_2$] is 0.07 nM.

that increasing *L* might confer on repressor capture at the operator site. In contrast with k_d, λ_0 increases by a factor of 4–5 for a ~6-fold increase in [CI$_2$] (Fig. 9), as we expect that a higher concentration leads to more binding events.

Control gel-shift assays

The association rates measured using the single-molecule microscopy method are two orders of magnitude smaller than the values in the literature (27). To test whether this discrepancy was caused by the experimental method or the experimental materials, we performed gel-shift assays using the same DNA and CI, as well as a similar buffer, as used in the single-molecule experiments. The reaction of 20 nM DNA and three CI$_2$ concentrations (5 nM, 36.5 nM, and 452 nM) were investigated. The results are shown in Fig. 10.

The nop DNA captured CI$_2$, which was also observed in the single-molecule measurements. About half of the DNA molecules were bound to CI$_2$ at the CI$_2$ concentration of 452 nM. The equilibrium constant K_D for nop and CI$_2$ is estimated to be about hundreds of nanomolars, similar to the single-molecule measurement results (150 ± 50 nM for 78 bp nop DNA, 64 ± 9 nM for 500 bp nop DNA).

For op DNA and WT CI, about half of the DNA molecules were bound to CI$_2$ at a lower CI$_2$ concentration (36.5 nM), in
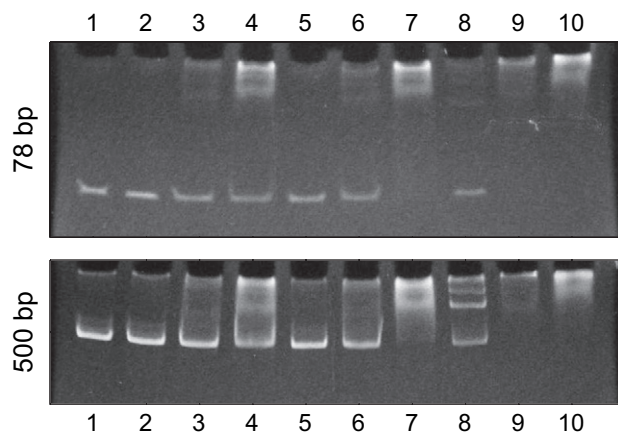

FIGURE 10 Gel-shift assays of DNA-CI reactions. Upper panel: 78 bp DNA. Lower panel: 500 bp DNA. DNA concentration in each lane was 20 nM. Lane 1: nop DNA control with no CI$_2$. Lanes 2–4: nop DNA interacting with increasing concentrations of WT CI$_2$. Lanes 5–7: op DNA interacting with increasing concentrations of WT CI$_2$. Lanes 8–10: op DNA interacting with increasing concentrations of Lys-34 CI$_2$. The three CI$_2$ concentrations are 5 nM (*lanes 2*, *5*, and *8*), 36.5 nM (*lanes 3*, *6*, and *9*), and 452 nM (*lanes 4*, *7*, and *10*).

agreement with our finding that op DNA has a higher affinity for CI$_2$. The equilibrium constant K_D is estimated from the gel-shift assays result to be about tens of nanomolars, similar to our single-molecule results of 22 ± 3 nM.

Compared with WT CI, the Lys-34 mutant of CI has a higher affinity for op DNA. When the CI$_2$ concentration was larger than that of DNA, almost all of the DNA molecules were bound. Only at the lowest CI$_2$ concentration, 5 nM, did we observe some free DNA. We estimate that the dissociation constant K_D is much smaller than 20 nM. Our single-molecule measurements suggest that it is 0.11 ± 0.02 nM.

DISCUSSION AND CONCLUSIONS

Blinking and bleaching of the tdimer

To investigate the behavior of tdimer:CI, we immobilized Cys-88 CI on glass surfaces without a PEG coating by nonspecific binding. Recall that Cys-88 CI forms covalent dimers, which is the same as the smallest active units that binds DNA at the molarities explored here. When the protein concentration was low enough, we observed distinct bright spots with similar intensities in the red channel. With time, the intensities of these spots decreased gradually.

It is known that green fluorescent protein exhibits "on-off" blinking behavior on the timescale of seconds (29). For GFP blinking, the fractional "on" time is defined as $F_{on} = \langle \Delta t_{on} \rangle / (\langle \Delta t_{on} \rangle + \langle \Delta t_{off} \rangle)$, where Δt_{on} (Δt_{off}) are the intervals when the fluorescence is on (off). The F_{on} value depends on excitation intensity I_{ex} such that $F_{on} = I_s/(I_s + I_{ex})$. The saturation excitation intensity I_s equals 1.5 kW/cm^2 for S65T-GFP (30). In our experiments, the excitation intensity is as low as 19 W/cm^2. If the monomers of tdimer have similar I_s, F_{on} would

be nearly one, which means that the monomers would be in the "on" state most of the time. If the mean "off" time $\langle \Delta t_{\mathrm{off}} \rangle$ is on the order of seconds, like S65T-GFP, the mean "on" time $\langle \Delta t_{\mathrm{on}} \rangle$ would be as long as the total exposure time in our experiment runs.

Further studies with the red fluorescent protein DsRed showed that this closely packed tetramer protein has four distinctive intensity levels for each single molecule. Each intensity level has a lifetime of seconds [31]. However, little is known about the tdimer, which is a tandem dimer of two RFP monomers. Each CI_2 has two independent tdimers. One would therefore expect that (tdimer:$CI)_2$ would bleach in at least two steps and possibly up to four steps. However, the intensity change in each step would be small compared to the photon number fluctuations in our experiments. In addition, the DNA-CI_2 complex has considerable rotation freedom in solution, which causes intensity fluctuation with the linear polarized excitation illumination used here. For these reasons, then, we believe that the intensity loss of the (tdimer:$CI)_2$ fusion due to bleaching appears as a smooth gradual process. Hence, it is reasonable to identify the rapid and stepwise changes in intensity with the unbinding of CI_2 from DNA.

Single-molecule versus ensemble measurements

Several kinds of bulk assays have been used to investigate DNA-protein interactions. The nitrocellulose filter binding assay [8] is based on the fact that protein binds to nitrocellulose, trapping any DNA bound to it. In the DNase footprinting assay [9], proteins protect DNA sequences from enzymatic cleavage when the proteins are specifically bound to the sequences. In gel-shift assays [10], protein-bound DNA migrates more slowly during electrophoresis. Calorimetry [6] is based on heat generation/absorption when proteins bind to DNA. Surface plasmon resonance [32] has also been used to investigate the interaction between molecules.

Some of these bulk assays can provide binding profiles, whereas others provide binding dynamics. All but the calorimetry experiments are nonequilibrium measurements. In contrast to bulk assays, an advantage of SMFM is that it can provide both the binding profile and the binding dynamics in great detail. When analyzing binding dynamics, we can filter out multiple molecule interaction, leaving only the one-on-one DNA-protein interactions. We can also study association and dissociation in isolation with the reaction system in equilibrium.

The observed rate constants versus values in the literature

The association rate constant we measured are considerably smaller than those measured with bulk assays [27,33]. Could this be because of our relatively slow sampling rate, since events that occur faster than the time interval between two consecutive frames would not be detected? First, we note

that most of the binding events last for many sampling frames. It is not easy to see how we could miss fast on and off rates at the same site. Second, we were careful to control for this possibility by increasing our sampling rate in preliminary experiments, where we did not observe any significant changes in the observed rate constants. Increasing the sampling rate caused the tdimer to bleach faster, resulting only in fewer events being recorded.

A second possible explanation for the discrepancy between our rates and values in the literature is that the fusion of the bulky tdimer prevents CI from dimerization or binding to DNA. CI molecules bind to DNA via the N-terminus and dimerize through the C-terminus. Without the C-terminus, the N-terminal domain alone is known to have much lower affinity ($K_D \sim 1 \ \mu M$) than the intact CI protein [34]. We reason that, because the tdimer was fused to the N-terminus, it should not affect CI dimerization. Moreover, according to the crystal structure [35], the N-terminus points away from DNA, so the tdimer should not affect the DNA binding either. As noted in Materials and Methods, these constructs confer immunity to superinfection. They have also been used to study oligomerization by fluorescence correlation spectroscopy [36]. We compared the intensity of CI bound to DNA with the intensity of the Cys-88 covalent CI dimer. By comparing the molecular mass in both reducing and nonreducing SDS gels, we confirmed that the Cys-88 mutant CI forms a dimer. As shown in Fig. 4, the Cys-88 covalent CI dimer and tdimer:CI bound on short DNA molecules have similar fluorescence levels. In summary, the evidence suggests that these fusion products function as dimers, both in vivo and in vitro, and at concentrations in the nanomolar range.

Yet another possibility is that the glass surface may somehow severely restrict the accessibility of CI to the DNA molecules, despite the long distance (40 bp, ~14 nm) between the operator sequence and the immobilization point.

To address these questions, we conducted control gel-shift assays with the same DNA and CI molecules reacting in buffer solutions similar to that used in the single-molecule experiment. The results show that the affinity of the tdimer:CI is not as high as reported in the literature (e.g., $K_D \sim (2.9 \pm 0.9) \times 10^{-13}$ M for op DNA and WT CI_2 at 50 mM KCl) [33], but rather similar to our single-molecule results ($K_D = (2.2 \pm 0.3) \times 10^{-8}$ M for op DNA and WT CI_2). Moreover, the affinity difference between CI binding to op and nop DNA, as well as between WT and the tight-binding Lys-34 mutant, was confirmed in these fusion proteins.

Binding dynamics suggested very short 1D search lengths

It is widely believed that transcription factors find their targets by first associating in the nonoperator sequence, and then diffusing along DNA using sliding and hopping to find their targets. Since hops are assumed to cover very short distances, they can be combined with sliding by scaling up the sliding

length to the length of a 1D search s_e (37). The interaction of CI_2 with DNA can then be summarized with a 1D-3D model: CI_2 spends its time either by diffusing three-dimensionally in the space or by sliding one-dimensionally along the DNA. When CI_2 jumps on(off) the DNA molecule, we will see an abrupt red intensity increase(decrease) at the location of DNA. The dynamics should allow effective 1D searching lengths s_e of 10^3 bp to be distinguished from those of ~20 bp by comparing 78–500 bp DNA molecules.

We first consider k_a for nop DNA. We expect it to scale with the length of DNA since doubling the DNA lengths doubles the cross section for capturing a CI_2. A linear fit to our data shows that k_a increases linearly with L, but with an intercept $k_a \sim L + 200$ bp. The reason for a nonzero intercept could be that immobilizing the DNA on a surface makes the two ends of DNA nonequivalent for CI_2 capture. If so, the positive value suggests that the immobilized end captures CI_2 more often than the free end.

For op DNA, one would expect to measure a length-dependent k_a if CI_2 associates with any sequence and then finds its target through 1D diffusion. Instead, we observed that k_a in op DNA exhibits little length dependence. This implies that CI_2 associates with the operator sequence with much higher probability than with a nonoperator sequence, which is expected. It also implies that the effective 1D search length is shorter than 78 bp, the shortest DNA length in our experiments. Even so, one might still expect some length dependence since the CI_2's first interaction with DNA might be nonspecific. However, the operator sequence is closer to the immobilized end and is always at the same distance from the immobilization point. The result with nop DNA suggests that the immobilized end and the free end behave differently. Because the immobilized end captures more CI_2, we expect to see more CI_2 binding to the operator sequence than CI_2 binding with the nonspecific sequence farther from the immobilizing point. Thus, the length dependence of k_a associated with nonspecific interaction may be weakened.

As for k_d, we observed clearly that CI_2 bound to op DNA dissociates more slowly than CI_2 bound to nop DNA. This confirms that the operator sequence has a higher affinity. We also see that Lys-34 CI has a smaller k_d than WT CI when binding to op DNA. This confirms that the interaction is protein specific. As shown in Fig. 8, when affinity increases, k_d decreases ($1 \times 10^{-2}\,\mathrm{s}^{-1}$ for WT CI + nop DNA compared with $5 \times 10^{-3}\,\mathrm{s}^{-1}$ for WT CI + op DNA), by a factor of 2. For the tight-binder Lys-34 mutant CI + op, k_d further decreases to $1.6 \times 10^{-3}\,\mathrm{s}^{-1}$ by a factor of 3.

Apart from dissociation from locations along the length of the DNA, CI_2 may also dissociate at one of the ends of the DNA molecules. However, immobilization of one end via a biotin-streptavidin link to the surface likely prevents dissociation from that end. Instead, reflection is possible, and the CI_2 returns to 1D diffusion. For nop DNA, we observed no length dependence of k_d. It is possible that the free DNA end is totally reflective, or 1D diffusion covers only a small range of DNA, so that CI_2 dissociates from a local sequence. For the first case, we would expect to see some k_d length dependence for CI_2 and op DNA, since CI_2 on shorter DNA encounters more reflections. In the second case (1D diffusion covers a very short distance), there will be no length dependence for op DNA either, because CI_2 falls off the operator sequence before it can interact with the nonoperator sequences. Our observation supports the second case. We also observed little length dependence of the tight-binder mutant Lys-34 CI_2 to op DNA.

Finally, as noted above, if k_a involves 1D diffusion—and the evidence summarized here suggests that it does not—then L-dependence should also appear in k_d measured as a function of L, since for simple Poisson events (Fig. 3, C and D) the law of microscopic reversibility holds. The fact that k_d shows no such dependence for op sequences (Figs. 7 A and 8) strengthens our view that 1D diffusion plays little if any role in target-finding over a range of 78–500 bp.

Extensive 1D diffusion on nonspecific DNA can slow the search to a rate that is even slower than 3D searching alone. However, in bacteria, target sequences for most repressors, including CI, lie very close to their coding sequences (37). Moreover, the motion of single messenger RNA molecules in living cells was recently reported to follow a subdiffusion process (38). This implies that the local concentration of repressor is higher—and the search volume smaller—than is often assumed. Although k_a is seen to increase with L for binding to nop DNA (Fig. 9 A), we could not resolve any L-dependence for op DNA (Fig. 9, B and C). This suggests to us that, although nonspecific binding of CI scaling with L is readily seen, 1D diffusion is unimportant once an operator is present.

We note that proteins that bind to and diffuse on DNA are often designed to clamp onto the DNA (39) as they slide along. For those transcription factors with multiple distant targets, they often have high molarities in the cell. In the case of CI, once it is bound to O_R1 and O_R2, which lie adjacent to the cI coding region, cI expression is upregulated, resulting in increasing concentrations of CI_2. The expression is not repressed until the low-affinity binding site O_R3 is also bound (4,40). Increasing concentration then increases the probability that CI_2 will bind to O_L sequences 2.4 kbp away, as well as the operator sites on superinfecting λ-DNA. Recent results show that CI transcription is further regulated by DNA looping induced by CI_2 binding at O_L1, O_L2, O_R1, and O_R2 (7). DNA looping also pulls the distant binding site O_L to the vicinity of the cI coding region, increasing the probability that CI_2 will find the two lower-affinity sites O_R3 and O_L3.

In summary, we used quantitative SMFM to investigate the interaction of CI_2 with DNA molecules. We have shown that CI_2 can bind DNA at both operator and nonoperator sequences. However, CI_2's interaction with the operator sequence is not affected by the length of flanking nonoperator sequences. Our data suggest that CI_2 finds its target operator sequence mainly by 3D diffusion.

SUPPORTING MATERIAL

A movie is available at http://www.biophysj.org/biophysj/supplemental/S0006-3495(08)00050-7.

REFERENCES

1. Ptashne, M., A. Jeffrey, A. D. Johnson, R. Maurer, B. J. Meyer, et al. 1980. How the λ repressor and cro work. *Cell*. 19:1–11.

2. Johnson, A. D., A. R. Poteete, G. Lauer, R. T. Sauer, G. K. Ackers, et al. 1981. λ Repressor and cro–components of an efficient molecular switch. *Nature*. 294:217–223.

3. Reichardt, L., and A. D. Kaiser. 1971. Control of λ repressor synthesis. *Proc. Natl. Acad. Sci. USA*. 68:2185–2189.

4. Meyer, B. J., D. G. Kleid, and M. Ptashne. 1975. λ Repressor turns off transcription of its own gene. *Proc. Natl. Acad. Sci. USA*. 72:4785–4789.

5. Strahs, D., and M. Brenowitz. 1994. DNA conformational changes associated with the cooperative binding of cI-repressor of bacteriophage λ to OR. *J. Mol. Biol*. 244:494–510.

6. Merabet, E., and G. K. Ackers. 1995. Calorimetric analysis of λ cI repressor binding to DNA operator sites. *Biochemistry*. 34:8554–8563.

7. Anderson, L., and H. Yang. 2008. DNA looping can enhance lysogenic CI transcription in phage λ. *Proc. Natl. Acad. Sci. USA*. 105:5827–5832.

8. Kim, J. G., Y. Takeda, B. W. Matthews, and W. F. Anderson. 1987. Kinetic studies on cro repressor-operator DNA interaction. *J. Mol. Biol*. 196:149–158.

9. Tullius, T. D., and B. A. Dombroski. 1986. Hydroxyl radical "footprinting": high-resolution information about DNA-protein contacts and application to λ repressor and cro protein. *Proc. Natl. Acad. Sci. USA*. 83:5469–5473.

10. Vipond, I. B., and S. E. Halford. 1995. Specific DNA recognition by EcoRV restriction endonuclease induced by calcium ions. *Biochemistry*. 34:1113–1119.

11. Nie, S., and R. N. Zare. 1997. Optical detection of single molecules. *Annu. Rev. Biophys. Biomol. Struct*. 26:567–596.

12. Schütz, G. J., W. Trabesinger, and T. Schmidt. 1998. Direct observation of ligand colocalization on individual receptor molecules. *Biophys. J*. 74:2223–2226.

13. Xie, X. S., and H. P. Lu. 1999. Single-molecule enzymology. *J. Biol. Chem*. 274:15967–15970.

14. Weiss, S. 2000. Measuring conformational dynamics of biomolecules by single molecule fluorescence spectroscopy. *Nat. Struct. Biol*. 7:724–729.

15. Nelson, P. C., C. Zurla, D. Brogioli, J. F. Beausang, L. Finzi, et al. 2006. Tethered particle motion as a diagnostic of DNA tether length. *J. Phys. Chem. B*. 110:17260–17267.

16. Kapanidis, A. N., N. K. Lee, T. A. Laurence, S. Doose, E. Margeat, et al. 2004. Fluorescence-aided molecule sorting: analysis of structure and interactions by alternating-laser excitation of single molecules. *Proc. Natl. Acad. Sci. USA*. 101:8936–8941.

17. Nalefski, E. A., E. Nebelitsky, J. A. Lloyd, and S. R. Gullans. 2006. Single-molecule detection of transcription factor binding to DNA in real time: specificity, equilibrium, and kinetic parameters. *Biochemistry*. 45:13794–13806.

18. Taguchi, H., T. Ueno, H. Tadakuma, M. Yoshida, and T. Funatsu. 2001. Single-molecule observation of protein-protein interactions in the chaperonin system. *Nat. Biotechnol*. 19:861–865.

19. Campbell, R. E., O. Tour, A. E. Palmer, P. A. Steinbach, G. S. Baird, et al. 2002. A monomeric red fluorescent protein. *Proc. Natl. Acad. Sci. USA*. 99:7877–7882.

20. Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

21. Sauer, R. T., K. Hehir, R. S. Stearman, M. A. Weiss, A. Jeitler-Nilsson, et al. 1986. An engineered intersubunit disulfide enhances the stability and DNA binding of the N-terminal domain of λ repressor. *Biochemistry*. 25:5992–5998.

22. Chen, Y., I. Golding, S. Sawai, L. Guo, and E. C. Cox. 2005. Population fitness and the regulation of *Escherichia coli* genes by bacterial viruses. *PLoS Biol*. 3:e229.

23. Pal, P., J. F. Lesoine, M. A. Lieb, L. Novotny, and P. A. Knauf. 2005. A novel immobilization method for single protein spFRET studies. *Biophys. J*. 89:L11–L13.

24. Burz, D. S., D. Beckett, N. Benson, and G. K. Ackers. 1994. Self-assembly of bacteriophage λ cI repressor: effects of single-site mutations on the monomer-dimer equilibrium. *Biochemistry*. 33:8399–8405.

25. Jackson, J. D. 1999. Classical Electrodynamics, 3rd ed. Wiley, New York.

26. Koster, D., C. Wiggins, and N. Dekker. 2006. Multiple events on single molecules: unbiased estimation in single-molecule biophysics. *Proc. Natl. Acad. Sci. USA*. 103:1750–1755.

27. Nelson, H. C., and R. T. Sauer. 1985. λ Repressor mutations that increase the affinity and specificity of operator binding. *Cell*. 42:549–558.

28. Wang, Y. M., J. O. Tegenfeldt, W. Reisner, R. Riehn, X. J. Guan, et al. 2005. Single-molecule studies of repressor-DNA interactions show long-range interactions. *Proc. Natl. Acad. Sci. USA*. 102:9796–9801.

29. Dickson, R. M., A. B. Cubitt, R. Y. Tsien, and W. E. Moerner. 1997. On/off blinking and switching behaviour of single molecules of green fluorescent protein. *Nature*. 388:355–358.

30. Garcia-Parajo, M., G. Segers-Nolten, J. Veerman, J. Greve, and N. van Hulst. 2000. Real-time light-driven dynamics of the fluorescence emission in single green fluorescent protein molecules. *Proc. Natl. Acad. Sci. USA*. 97:7237–7242.

31. Garcia-Parajo, M., M. Koopman, E. van Dijk, V. Subramaniam, and N. van Hulst. 2001. The nature of fluorescence emission in the red fluorescent protein DsRed, revealed by single-molecule detection. *Proc. Natl. Acad. Sci. USA*. 98:14392–14397.

32. Smith, E. A., M. G. Erickson, A. T. Ulijasz, B. Weisblum, and R. M. Corn. 2003. Surface plasmon resonance imaging of transcription factor proteins: interactions of bacterial response regulators with DNA arrays on gold films. *Langmuir*. 19:1486–1492.

33. Johnson, A. D., C. O. Pabo, and R. T. Sauer. 1980. Bacteriophage λ repressor and cro protein: interactions with operator DNA. *Methods Enzymol*. 65:839–856.

34. Sauer, R. T., C. O. Pabo, B. J. Meyer, M. Ptashne, and K. C. Backman. 1979. Regulatory functions of the λ repressor reside in the amino-terminal domain. *Nature*. 279:396–400.

35. Stayrook, S., P. Jaru-Ampornpan, J. Ni, A. Hochschild, and M. Lewis. 2008. Crystal structure of the λ repressor and a model for pairwise cooperative operator binding. *Nature*. 452:1022–1025.

36. Samiee, K. T., M. Foquet, L. Guo, E. C. Cox, and H. G. Craighead. 2005. λ-Repressor oligomerization kinetics at high concentrations using fluorescence correlation spectroscopy in zero-mode waveguides. *Biophys. J*. 88:2145–2153.

37. Kolesov, G., Z. Wunderlich, O. N. Laikova, M. S. Gelfand, and L. A. Mirny. 2007. How gene order is influenced by the biophysics of transcription regulation. *Proc. Natl. Acad. Sci. USA*. 104:13948–13953.

38. Golding, I., and E. C. Cox. 2006. Physical nature of bacterial cytoplasm. *Phys. Rev. Lett*. 96:098102.

39. López de Saro, F. J., M. G. Marinus, P. Modrich, and M. O'Donnell. 2006. The β sliding clamp binds to multiple sites within MutL and MutS. *J. Biol. Chem*. 281:14340–14349.

40. Dodd, I. B., K. E. Shearwin, A. J. Perkins, T. Burr, A. Hochschild, et al. 2004. Cooperativity in long-range gene regulation by the λ CI repressor. *Genes Dev*. 18:344–354.