Restriction enzyme Ecl18kI–induced DNA looping dynamics by single-molecule FRET

ARTICLE *in* THE JOURNAL OF PHYSICAL CHEMISTRY B · JUNE 2014

Impact Factor: 3.3 · DOI: 10.1021/jp504546v · Source: PubMed

THE JOURNAL OF PHYSICAL CHEMISTRY B

Article

pubs.acs.org/JPCB

Restriction Enzyme Ecl18kI-Induced DNA Looping Dynamics by Single-Molecule FRET

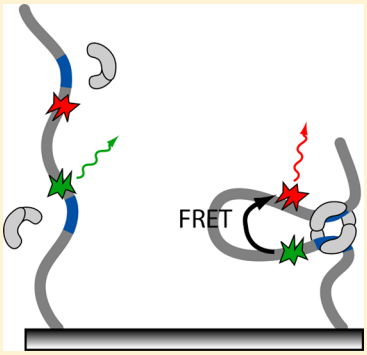
Danielis Rutkauskas,*[,†] Milda Petkelyte,[†] Paulius Naujalis,[†] Giedrius Sasnauskas,[‡] Gintautas Tamulaitis,[‡] Mindaugas Zaremba,[‡] and Virginijus Siksnys[‡]

[†]Institute of Physics, Center for Physical Sciences and Technology, Savanoriu 231, LT-02300, Vilnius, Lithuania

[‡]Institute of Biotechnology, Vilnius University, Graiciuno 8, LT-02241, Vilnius, Lithuania

Supporting Information

ABSTRACT: Many type II restriction endonucleases require binding of two copies of a recognition site for efficient DNA cleavage. Simultaneous interaction of the enzyme with two DNA sites results in DNA loop formation. It was demonstrated with the tethered particle motion technique that such looping is a dynamic process where a DNA loop is repeatedly formed and disrupted. Here we use a better and in the context of protein-induced DNA looping virtually unexploited strategy of single-molecule Förster resonance energy transfer of surface immobilized biomolecules to quantitatively study the dynamics of Ecl18kI endonuclease-induced DNA looping and determine the rate constants of loop formation and disruption. We show that two DNA-bound Ecl18kI dimers efficiently form a bridging tetramer looping out intervening DNA with a rate that is only a few orders of magnitude lower than the diffusion limited rate. On the other hand, the existence of Ecl18kI tetramer is only transient, and the loop is rapidly disrupted within about 1 s.



INTRODUCTION

Interaction of distant DNA sites via DNA looping is a recurring motif in various biochemical processes such as regulation of gene expression, DNA replication, and recombination.[1,2] In the past decade, it has become evident that DNA looping is also crucial for the function of many type II restriction endonucleases (REases). These enzymes recognize and cut short (4−8 base pairs) DNA sequences, and usually require only Mg^{2+} ions as a cofactor.[3] In vivo REases protect their host bacteria from viral attacks by cleaving foreign DNA; in vitro, they are widely used as a molecular tool for various DNA manipulations. Despite their similar function, REases vary in their subunit composition and DNA cleavage mechanism; for example, orthodox enzymes are homodimeric proteins that bind and cleave solitary copies of cognate DNA, type IIF enzymes are homotetramers capable of simultaneous binding and cleavage of two DNA sites,[4,5] while type IIE enzymes interact with several copies of the recognition sequence but cleave only one, the other sites acting only as allosteric activators.[6] The requirement for several sites of cognate DNA in this context is rationalized to act as a double check against the host DNA cleavage, as both IIF and IIE enzymes display only residual activity when bound to a single recognition site.[7,8] Concomitant interaction with two sites by a single REase molecule inevitably leads to loop formation, which was demonstrated for a number of type IIF and type IIE enzymes using various biochemical assays, atomic force microscopy, and single-molecule (SM) techniques.[9−11]

At present, we aim to study the phenomenon of protein-induced DNA looping exemplified by Ecl18kI, a well characterized[12] type IIF REase requiring two interrupted palindromic target sites 5′-CCNGG-3′ (here N is any nucleotide) and Mg^{2+} ions as a cofactor for its optimal activity. Ecl18kI is a dimer in solution,[13] but a combination of structural, biochemical, and SM work using the tethered particle motion (TPM) technique suggests that upon DNA binding two primary dimers of Ecl18kI associate into a catalytically active tetramer.[9] The goal here is a refined quantitative study of the dynamics of loop formation by Ecl18kI exploiting the fact that substitution of the cofactor Mg^{2+} ions with Ca^{2+} precludes DNA cleavage but at the same time retains the specific affinity of the REase to DNA. In the past, among the SM techniques, it was TPM that was used to study the real-time dynamics of dsDNA looping induced by restriction enzymes[10,14,15] and other molecular systems.[16−18] As in the previous SM work on Ecl18kI,[9] the temporal resolution of such measurements is limited by the necessity to average over the tethered particle motions. Also, the presence of the particle distorts the studied process. Here we use a different experimental strategy based on the Förster resonance energy transfer (FRET) of surface immobilized biomolecules that in the context of nucleic acid research has been successfully applied to study, for example, Holliday junction branch migration,[19] structural fluctuations of hairpin ribozyme,[20] and nucleosomal structural dynamics and positioning[21] but has not been applied to the problem of protein-mediated DNA looping. Compared to TPM, FRET measurement has a better time resolution, is more specific, and is free from the bead-associated distortions.

Received: May 8, 2014
Revised: June 25, 2014
Published: June 27, 2014

   dx.doi.org/10.1021/jp504546v | J. Phys. Chem. B 2014, 118, 8575−8582

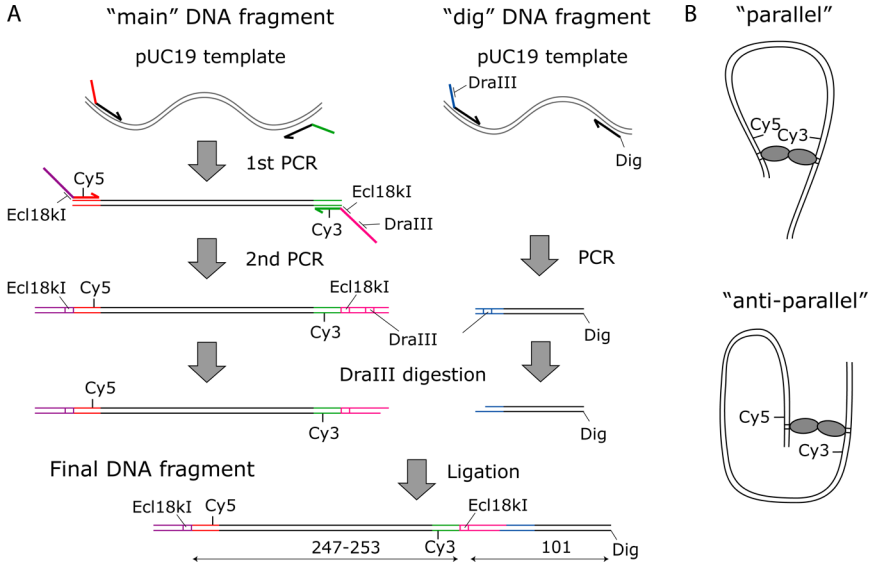


Figure 1. (A) Fluorescently labeled DNA construct and its synthesis. Ecl18kI and DraIII target sequences are represented by rectangular boxes; the digoxigenin is marked by "Dig". (B) Two possible Ecl18kI-associated loop conformations. An oval represents the Ecl18kI dimer. Only the "parallel" DNA loop conformation results in FRET between the Cy3 and Cy5 dye labels.

MATERIALS AND METHODS

Protein. Protein was purified as described previously.[13] The protein concentration was determined using an extinction coefficient of 77 660 M^{-1} cm^{-1} calculated for the dimer by the ProtParam tool at http://web.expasy.org/protparam/. Protein concentrations are expressed in terms of dimer if not stated otherwise.

DNA Fragments. The general design of the DNA construct used in the SM measurements is presented in Figure 1A.

The proper placement of the Cy3 and Cy5 dye labels, resulting in appreciable FRET upon "parallel" loop formation by the enzyme, was determined by extending the short nine bp DNA fragments in the crystal structure of Ecl18kI[12] with polyA/polyT duplexes (generated using the 3D-DART server[22]) and searching for a pair of thymine bases on the opposing DNA double strands with a suitable distance between their C5 atoms. For the 19th nucleotides from the target sequences in the 3′ direction, this distance is about 45 Å, which is less than the Förster radius for the Cy3/Cy5 pair (54 Å). Moreover, at this position, the dye labels are sufficiently distant from the binding sites of the enzyme so that the possible protein influence on the labels' fluorescence is precluded.

We have prepared six DNA fragments with intersite distances of 247, 249, 250, 251, 252, and 253 bp, named "0", "2", "3", "4", "5", and "6" DNA, respectively. The following set of primers was used in the fragment synthesis: R, 5′-CGTTTACGAACGAAATAATAAATTATCGCCACTGGCAGCAGC-3′; M0, 5′-CGTAATGCAGCCAATAATTATTAAACTACCAGCGGTGGTTTGTTT-3′; M2, 5′-CGTAATGCAGCCAATAATTATTAAACGCTACCAGCGGTGGTTTG-3′; M3, 5′-CGTAATGCAGCCAATAATTATTAAACCGCTACCAGCGGTGGTTTG-3′; M4, 5′-CGTAATGCAGCCAATAATTATTAAAACCGCTACCAGCGGTGG-3′; M5, 5′-CGTAATGCAGCCAATAATTATTAAACACCGCTACCAGCGGTG-3′; M6, 5′-CGTAATGCAGCCAATAATTATTAAACCACCGCTACCAGCGGT-3′; MCy5, 5′-GCACTGGACGCCAGGCGTTTACGAACGAAATAAMAAA-3′; MCy3, 5′-GAGCACCGCGTGCGCCAGGCGTAATGCAGCCAATAATMATT-3′; Dig1, 5′-TAGTGCACGCGGTGTTATCGCCACTGGCAGCAGC-3′; Dig2, 5′-Dig-CGTATGTCGTACCGGTAAGAACTCTGTAGCACCGCC-3′. The recognition sequences of Ecl18kI (5′-CCNGG-3′) and DraIII (5′-CACNNNGTG-3′) are underlined, "M" denotes the positions of the Cy3 or Cy5-labeled thymine derivatives, and "Dig" denotes digoxigenin.

The primers with internal dye labels (attached to the 5-C6-amino-2′-deoxythymidine via a six-carbon linker using NHS chemistry) were obtained from IBA, while the rest were purchased from Metabion.

The full DNA construct was synthesized by ligating the "main" and "dig" fragments, as depicted in Figure 1A. The "main" fragment was synthesized in two PCR steps. In the first step, we synthesized a series of 281−287 bp auxiliary unlabeled DNA fragments using primers "R" and "M0−6". The 5′-terminal flaps of primers used in the first PCR (red and green) are complementary to the 3′-terminal sequences of the dye-labeled primers MCy3 and MCy5 used in the second PCR. The 5′-flaps of the latter primers (dark purple and magenta) contain recognition sequences of Ecl18kI and DraIII restriction enzymes (see above). In the second PCR step, using primers MCy3 and MCy5, we synthesized a series of dye-labeled fragments using as templates the auxiliary fragments produced in the first PCR step. The "dig" fragment was obtained in a single PCR step, using a digoxigenin-labeled primer "Dig2" and a primer with a DraIII site "Dig1" (blue). The final DNA fragment was assembled by DraIII digestion and ligation of the "main" and "dig" fragments. The total length of the DNA construct is ∼370 bp; the lengths of different segments are 247−253 bp between the Ecl18kI target sites, 101 bp between the digoxigenin label and the closest Ecl18kI site, and 18 bp between the Cy3 or Cy5-labeled nucleotides and the Ecl18kI sites.

The PCRs were performed using the DreamTaq PCR Master Mix (K1071, Fermentas, ThermoFisher Scientific), and PCR products were purified using the GeneJet PCR purification kit (K0701, Fermentas, ThermoFisher Scientific). The "main" and "dig" fragments were digested with the FastDigest DraIII (FD1234, Fermentas, ThermoFisher Scientific), purified with the GeneJet PCR purification kit, and ligated with the T4 DNA

dx.doi.org/10.1021/jp504546v | J. Phys. Chem. B 2014, 118, 8575−8582

ligase (EL0011, Fermentas, ThermoFisher Scientific). The ligation product was separated from unreacted fragments by electrophoresis through agarose, recovered using the Freeze N' Squeeze centrifugal device (Biorad), and additionally purified with the GeneJet PCR purification kit. The details of the PCR, digestion, and ligation reactions are presented in the Supporting Information.

Control DNA Duplexes. A number of short control DNA duplexes were used for FRET image calibration and acquisition of Cy3 and Cy5 fluorescence spectra. Duplexes were annealed by heating 1 μM mixtures of complementary single-stranded oligonucleotides in the "yellow buffer" (YB): 33 mM Tris (pH 7.9 at 20 °C), 66 mM K-acetate to 90 °C and then permitting it to slowly cool at room temperature. Sequences of these Cy3-, Cy5-, and digoxigenin-labeled oligonucleotides are analogous to those used by Joo et al.[23] The sequence of Cy3-labeled oligonucleotide is 5′-TGGCGACGGCAGCGAGGCTT-TTTTTTTTTTTTTTTTTTTT-Cy3-3′, and that of Cy5- and digoxigenin-labeled oligonucleotide is 3′-Dig-ACCGCTGC-CGTCGCTCCG-Cy5-5′. According to Joo, in the Cy3- and Cy5-labeled duplex, the FRET value is around 0.55. Oligonucleotides were obtained from IBA.

Sample Cell Preparation. The flow cell for microscopy was prepared by melting a piece of Parafilm with a cut-out rectangular channel in between a coverslip and a coverslide with drilled inlet and outlet holes. Tefzel tubes with clamped-on ferrules (Upchurch Scientific) were pressed against the coverslide holes by means of a custom-made cell holder. The volume of the cell including the inlet and outlet bits of tubing was about 30 μL.

Microscopy coverslips (Menzel Glaser) were cleaned by first sonicating in acetone for 10 min, then rinsing in 2-propanol, and finally exposing to the air plasma (PDC-002, Harrick) for 5 min. Subsequently, they were stored in a closed staining dish in the air.

After drilling the holes, coverslides (Menzel Glaser) were sonicated in 1% Alconox type III deionized water solution for 10 min and rinsed with type III deionized water. Then, they were additionally cleaned using the same procedure as that for the coverslips.

SM Assay. The glass surface for the digoxigenin-labeled DNA fragment immobilization and its interaction with Ecl18kI was prepared as follows: the flow cell was incubated with 10 μg/mL of antidigoxigenin (11333089001, Roche) in YB for 10 min, washed with YB, infused with 10 mg/mL α-casein (C6780, Sigma-Aldrich) in YB and incubated for 0.5 h, washed with YB, infused with ∼20 pM DNA in YB, and washed without incubating with YB. For the measurement of DNA–protein interactions in the final step, the cell was infused with a 6 nM solution of Ecl18kI in YB supplemented with 10 mM Ca^{2+}, 3 U/mL of glucose oxidase (G6125, Sigma-Aldrich), 1% β-D-glucose (G0047, TCI Europe), and 2.5 mM Trolox (238813, Sigma). Catalase is omitted from the recipe, as it causes significant fluorescence background and its absence has no effect on Ecl18kI activity (see Figure S1 in the Supporting Information). The inlet and outlet tubes were sealed with Parafilm. Under these conditions, the fluorescence half-time of the control DNA duplex was about 0.5 min for Cy5 and more than 2 min for Cy3. Also, the fluorescence intensity of both Cy3 and Cy5 dyes was very constant, exhibiting no blinking (see Figure S2 in the Supporting Information).

SM Fluorescence Microscopy. The principle of the total internal reflection fluorescence (TIRF) setup that we used for the study of the DNA–protein interactions is described, for example, by Roy et al.[24] In short, FRET pair-labeled DNA molecules in the buffer solution are immobilized on the surface of a glass coverslip. The laser beam that is totally internally reflected (TIR) from the glass–water interface in the sample cell excites the donor fluorophores. The image of the excited sample surface is spectrally split into the "donor" and "acceptor" images. An uninterrupted sequence of these two-channel images (a movie) is acquired. Then, the evolution of the ratio of intensities of the bright spots in the two channels originating from the same DNA molecule reports on the development of FRET in that particular DNA molecule.

In our implementation, it is a home-built SM fluorescence microscope based on a biological inverted Nikon Eclipse Ti-U equipped with a 100× 1.4 Oil Plan Apo VC objective. The excitation source for the FRET donor is a 532 nm, 25 mW CW DPSS laser (Crystalaser). A 10× expanded laser beam is focused onto the back focal plane (BFP) of the objective, and to achieve TIR off the glass–water interface, the beam is translated to the periphery of the BFP by a mirror preceding the objective. The resulting excitation spot is about 80 μm in diameter on the sample surface. The laser power after the objective is 3.3 mW. Inside the microscope, the laser beam is reflected and the resulting fluorescence is transmitted by a dichroic mirror (Di01-R532, Semrock). The fluorescence image is focused onto an adjustable mechanical slit at the side port of Eclipse, filtered of the excitation light by a high-pass interference filter (HQ545LP, Chroma), and further split into the donor and acceptor image components by a dichroic mirror (645dcxr, Chroma). The two image components are focused onto the same EMCCD chip (DU-897E-CS0-UVB, Andor) side by side, and fluorescence movies are acquired with a 0.1 s integration time using a home-written program coded in Labview 2010 (National Instruments). An example of a fluorescence image is presented in Figure S3 in the Supporting Information.

The objective is focused on the sample surface by monitoring the transverse position of TIR of a reference infrared (IR) laser diode (L850P010, Thorlabs) guided into the objective by a dichroic mirror (t800dcsp, Chroma) in the second stratum of the microscope. Also, during the movie acquisition, the axial focal drift is actively compensated by a negative feedback loop between the objective piezo translator (P-725.1CD, E-712.3CDA controller, PI) and the position of the IR reflection on the video camera (WAT-902H3, Watec). The IR reflection is filtered out from the fluorescence detection with a low-pass interference filter (FF01-775/SP-25, Semrock).

Ensemble Fluorescence Spectroscopy. In addition to two-channel spectral detection, the microscope is also equipped with a Shamrock SR-303i polychromator (Andor) for full spectral acquisition that was used to measure bulk fluorescence spectra of DNA fragments with different intertarget distances. DNA solution was excited in a regular wide-field mode through a CFI S Plan Fluor ELWD 40× objective with a 532 nm laser. Fluorescence spectra were acquired by dispersing the signal with a 53-310R grating (Newport) onto the same EMCCD camera that was used for the SM fluorescence measurements. Fluorescence spectra were background-subtracted and corrected for the spectral detection sensitivity.

FRET Movie Analysis. To obtain FRET trajectories of individual DNA fragments from a FRET movie, the following algorithm was used: a number of frames at the beginning of the movie are averaged, the resulting image thresholded, and bright particles in the donor part of the image identified and framed

dx.doi.org/10.1021/jp504546v | J. Phys. Chem. B 2014, 118, 8575–8582

by a bounding rectangle region of interest (ROI). The positions of ROIs of corresponding acceptor particles are calculated from the ROIs of the donor particles by applying the translation and scaling transformations whose parameters were determined from a calibration FRET image acquired on a sample of Cy3-, Cy5-, and digoxigenin-labeled control DNA duplex. Then, a point in a trajectory of particle fluorescence intensity is calculated by summing all pixel intensities in that particle's ROI. The full trajectory is obtained by doing this for all frames in the movie. Donor and acceptor signals are subtracted respective offsets calculated on pieces of trajectories with zero FRET and before donor and acceptor photobleach. We noticed that the baseline of the donor signal in the experiment with Ecl18kI generally is much more dynamic than that in the absence of the protein (see Figure S4 in the Supporting Information). This could probably be due to the known enhancement of Cy3 fluorescence intensity by the proximity of protein.[25] To avoid transfer of this noise into the calculation of FRET, the positive values of the donor signal and the negative values in the acceptor signal are coerced to zero. FRET is calculated as $f = f_a - f_d$ and normalized to unity, where f_a is the acceptor intensity and f_d is the donor intensity. Such a definition (rather than the conventional $f = f_a/(f_a + f_d)$) yields a better contrast between the different FRET states. No additional corrections for the bleed-through of the donor signal into the acceptor channel or direct excitation of the acceptor label were applied, since we are not interested in the absolute FRET value but rather in its change due to DNA looping.

FRET Trajectory Analysis. The analysis of FRET trajectories is performed using freely available QuB software,[26] designed to analyze single-ion channel conductance data which essentially is very similar to that of SM fluorescence. First, all FRET trajectories are "idealized" by a segmental K-means (SKM) routine[27] using a two-state hidden Markov model. In SKM, the amplitude and the standard deviation of the low conductance class are fixed. This prevents the routine from identifying false high FRET dwells in the noisy baseline. The resulting sequence of dwells is then used by the maximum interval likelihood (MIL) procedure[28,29] for the maximum likelihood optimization of the kinetic rate constants in the reaction scheme. Importantly, MIL corrects for the missed events that are shorter than a set dead time. We set this dead time to 0.5 of the sampling time (0.05 s).

Ensemble FRET Efficiencies. The measurements were performed with 10 nM DNA without enzyme or with 100 nM Ecl18kI.

The extent of FRET due to DNA–Ecl18kI interactions was evaluated by comparing the fluorescence spectrum measured on the bulk sample of the "main" DNA fragment alone and then in the presence of Ecl18kI. The intensity increase upon the addition of Ecl18kI of the Cy5 spectral component in the spectrum of Cy3- and Cy5-labeled DNA fragment indicates FRET. Quantitatively, this can be expressed as

$$F_3 \sim \varepsilon_3 Q_3 \tag{1}$$

$$F_5 \sim \varepsilon_5 Q_5 \tag{2}$$

$$F_3' \sim (1 - f)\varepsilon_3 Q_3 \tag{3}$$

$$F_5' \sim (\varepsilon_5 + f\varepsilon_3) Q_5 \tag{4}$$

Here F are the observed integral intensities of the two spectral components of DNA alone and F' are those intensities of DNA reacting with Ecl18kI; ε are the absorption extinction coefficients at the excitation wavelength, Q are the fluorescence quantum efficiencies, and f is by definition the FRET efficiency. Assuming 100% labeling efficiencies of both dye labels, the concentration-dependent proportionality coefficients in eqs 1 and 2 and eqs 3 and 4 are the same. This then leads to

$$f = (p' - p)/(p' + s), \quad p' = F_5'/F_3', \quad p = F_5/F_3, \quad s = Q_5/Q_3 \tag{5}$$

The F ratios are obtained by decomposing the corresponding spectra into the Cy3 and Cy5 spectral components whose spectra are measured on the Cy3- or Cy5-labeled control DNA duplexes. $Q_5 = 0.3$ and $Q_3 = 0.15$ are taken from ATDBio.

RESULTS

The time resolution of our measurement is 0.1 s, and the duration of a fluorescence trajectory is limited to at most a few minutes due to fluorophore bleaching. Therefore, we can monitor processes occurring on a time scale from 0.1 s to a few tens of seconds. The loop formation time calculated for Ecl18kI from the earlier TPM measurements[9] was of the order of 1 s, which is an optimal time scale for our experimental setup. To maintain the same experimental time scale, we had to consider that the looping probability strongly depends on the loop size, helical phasing, and loop conformation.

To some extent, the likelihood of DNA looping is determined by the mechanical properties of DNA. Even though its validity has been recently challenged for very short DNA fragments,[30,31] a generally accurate description of DNA mechanics is provided by the wormlike chain (WLC) model.[32] A very useful quantity to characterize the looping probability is the J factor or the relative concentration of two sites on DNA that are brought together by looping.[33] The WLC model predicts different regimes of J dependence on the length of the DNA tether between the sites.[34] For short DNA tether lengths, the larger the intertarget spacing, the larger the J factor. At about 500 bp, the J factor reaches the maximum and then gradually drops. Including the requirement of rotational alignment of the sites causes the J dependence on DNA length to oscillate with a period of a helical repeat. Furthermore, this dependence is also affected by the criterion of when the sites are considered to be brought together: this will be different, for example, for DNA cyclization and protein-mediated DNA looping. Also, in the case of DNA interaction with a protein (as opposed to cyclization). the problem of looping is further complicated by the fact that the protein may trap different loop conformations.

Probabilities of different loop shapes were extensively modeled for lactose repressor (LacI).[35,36] Since it is a stable tetramer, a V-shaped LacI can form four different loop shapes. In the case of Ecl18kI, because it assembles on DNA, only two different loop conformations can be thought of. One is reminiscent of the "parallel" loop with extended LacI conformation and the other "anti-parallel" with extended instead of V-shaped LacI (Figure 1B). From calculations on LacI, it appears that the "parallel" loop with extended LacI is more probable than "anti-parallel" loops. If this held true also for Ecl18kI, then its "parallel" loop would be more probable. This is the reason the FRET pair labels are placed in such a way as to report on the formation of the "parallel" loop. Also, to simplify the kinetic analysis, we tried to minimize the likelihood of the alternative loop conformation. The difference in J factors between different

dx.doi.org/10.1021/jp504546v | J. Phys. Chem. B 2014, 118, 8575–8582

loops increases with decreasing intertarget spacing. We therefore chose an intertarget distance shorter than the optimal 500 bp with the aim of favoring the formation of one looping conformation.

To adjust the helical phasing, we searched for the exact intertarget spacing bringing the two targets in phase upon their juxtaposition. To do so, we considered six intertarget distances spanning more than half of the DNA helical repeat, since this should be sufficient to find one of the phasing extremes. The relative ease of DNA looping at a specific intertarget spacing was assessed from ensemble fluorescence measurements (Figure 2).

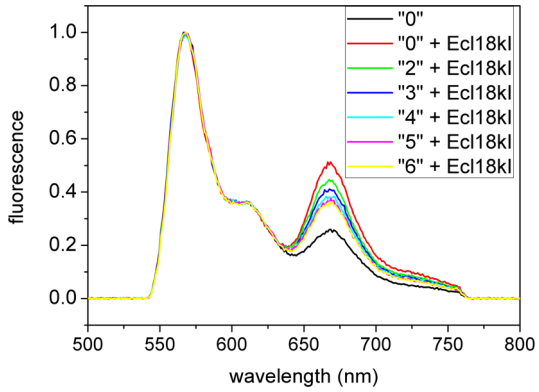


Figure 2. Bulk fluorescence spectra of fragments with different intertarget distances reacting with Ecl18kI. For comparison, presented is also the spectrum of the "0" DNA fragment without Ecl18kI. The abrupt spectral intensity drop at ~760 nm is due to the low-pass filter used to block the reference infrared laser reflection.

The extent of FRET was evaluated by comparing the fluorescence spectra of DNA samples with and without Ecl18kI as described in the Materials and Methods. Since the geometry of juxtaposition of fluorescent labels upon DNA looping does not depend on the intertarget distance, different ensemble FRET values are not associated with differences in intramolecular FRET efficiencies but with the variation of loop formation probability due to the requirement of more or less DNA twisting. The calculated values of FRET efficiency for different DNA fragments in the order of increasing intertarget distance are 0.10, 0.08, 0.06, 0.04, 0.04, and 0.02. It appears that the bulk FRET efficiency quasi-monotonously decreases from about 0.1 for "0" DNA to 0.02 for "6" DNA, and in theory, this value should start increasing if the intertarget distance was lengthened further. It can thus be concluded that coincidentally the starting intertarget distance allows for the fastest, torsional strain-free loop formation. Therefore, the "0" DNA fragment was chosen for the SM measurements to probe the DNA–Ecl18kI interactions.

To investigate the real-time dynamics of "0" DNA fragment–Ecl18kI interactions, we acquired two spectral component fluorescence movies of surface-immobilized DNA in the buffer solution with Ecl18kI. FRET trajectories of individual DNA molecules were extracted from these movies as described in the Materials and Methods.

An example of raw donor and acceptor trajectories together with a calculated FRET trajectory is shown in Figure 3 (more examples of traces are presented in the Supporting Information in Figure S5).

Changes in the donor and acceptor fluorescence intensity are clearly anticorrelated, as is expected in the presence of the

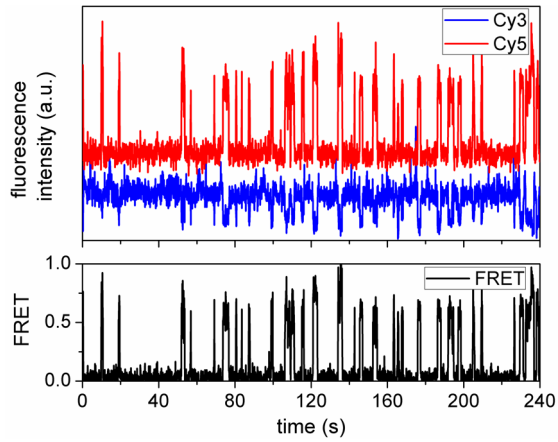


Figure 3. Single-molecule raw donor and acceptor fluorescence intensity traces (upper panel) and corresponding FRET trace (lower panel) calculated as described in the Materials and Methods.

excitation energy transfer. Inequality of Cy5 and Cy3 fluorescence signal change is probably due to the difference of the quantum efficiencies of the two dyes. The FRET value changes between two levels of distinctly different magnitude: 0 and 0.5–1. Zero FRET corresponds to the unlooped DNA, where the donor and the acceptor labels are separated by about 70 nm if the DNA was fully stretched. This distance is significantly larger than the Förster radius of the Cy3/Cy5 pair, and consequently, FRET should be zero. On the other hand, FRET of 0.5–1 corresponds to DNA looped by Ecl18kI, where the Cy3 and Cy5 labels are brought to close proximity.

The FRET trajectory provides a reduced view of the chemical kinetics in the system because it only reveals whether DNA is looped or unlooped at any moment of time. The unlooped state, however, consists of a number of different chemical substates indistinguishable in the time record, as shown in the proposed kinetic scheme of Ecl18kI interaction with a two-target-site DNA fragment (Figure 4).

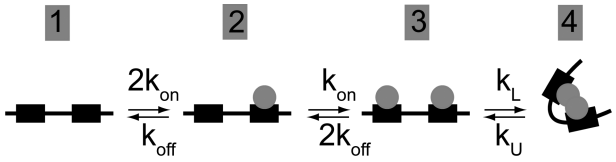


Figure 4. Reaction scheme of Ecl18kI interacting with two-target-site DNA. The target sites are represented by black rectangular boxes, and the Ecl18kI dimers, by gray circles. The different chemical states are numbered to ease the notation of the rate constants in the text.

The reaction model consists of four interconverting states: DNA, DNA with one bound Ecl18kI dimer, DNA with two bound dimers, and looped DNA. The first three states collectively form the unlooped state. The dimer association rate constant in the reaction scheme is $k_{on} = k_a C$; here, k_a is the second order association rate constant (dimension M^{-1} s^{-1}) and C is the molar dimer concentration. k_{off} is the dimer dissociation rate constant. The dimer associates with bare DNA with a double rate, since DNA contains two unoccupied target sites. Likewise, the doubly occupied DNA converts to DNA with one dimer bound with double dissociation rate, since both dimers have a chance to dissociate.

To analyze FRET trajectories, we follow along the lines of analysis originally developed for the measurements of single-ion

dx.doi.org/10.1021/jp504546v | J. Phys. Chem. B 2014, 118, 8575–8582

channel currents. The data of SM fluorescence is very similar to that derived from the patch clamp: in both cases, the time course of the signal changing between different levels of magnitude is measured. The underlying kinetic reaction mechanisms in both cases contain aggregated states: ion channels—zero current state consisting of a number of closed ion channel conformations—and SM fluorescence—zero-FRET state corresponding to different chemical species of unlooped DNA. The analogy extends even further in that the sequential model we assumed for Ecl18kI is identical to a well-studied model of nicotinic acetylcholine receptor.[28]

The first step in the analysis of FRET trajectory required for further computation is its idealization: each sample point is attributed to either the looped or unlooped state of DNA. This way the original FRET trajectory is converted into a sequence of discrete states and transitions between them. The simplest way to do this is via a threshold-crossing algorithm where the transition between the states is affirmed each time the signal value changes across the threshold. However, the method lacks a way to distinguish between the noise-induced transitions and the genuine short-lived dwells. It also requires some sort of essentially arbitrary criterion to set the threshold value.

An alternative approach free from the deficiencies of the thresholding method is the so-called hidden Markov modeling (HMM), where the idealized trajectory is built by maximizing the probability that a proposed sequence of states generates the observed data. This algorithm does not require user-defined parameters and (even though it is not relevant in our case) can be successfully applied to systems containing a greater number of states.

Whatever the method of idealization, the end result is a trajectory of a few levels of different but constant magnitude. This straightforwardly leads to the histograms or probability density functions (pdf's) of states' dwell times. A pdf can be represented by a mixture of exponential components whose number is equal to the number of kinetic states underlying the observed state.[37] Thereby, it is a common practice to fit pdf's with a sum of exponents. However, each exponential parameter, in general, is dependent on all the rates associated with the aggregated state and there is no trivial relationship between the kinetic rate constants and the exponential amplitudes and the time constants. In other words, exponential fits cannot be interpreted simply as kinetic rate constants.

In fact, for some simple models, the relationship between the exponential fits and the kinetic rates can be expressed analytically.[37] However, even for such relatively simple models, in general, these expressions are rather complex. Alternatively, model pdf's can be formulated using the Q-matrix formalism[38] and the model rates computed via the maximum likelihood fitting to the measured pdf's.[39] Finally, it is also possible to estimate the model parameters with the maximum likelihood fitting of the dwell time sequence directly without building the pdf's.[28] This approach is also implemented in freely available QuB software[26] that we used in our data analysis as described in the Materials and Methods.

First, calculated FRET traces are idealized using the SKM procedure in QuB. An example of an outcome of this operation is demonstrated in Figure 5: the idealized time dependence faithfully follows the apparent changes in the analyzed time trajectory.

The MIL procedure is then run on the whole set of idealized FRET traces. The total number of dwells in both FRET states in all trajectories is about 4400. Optimized are only the looping

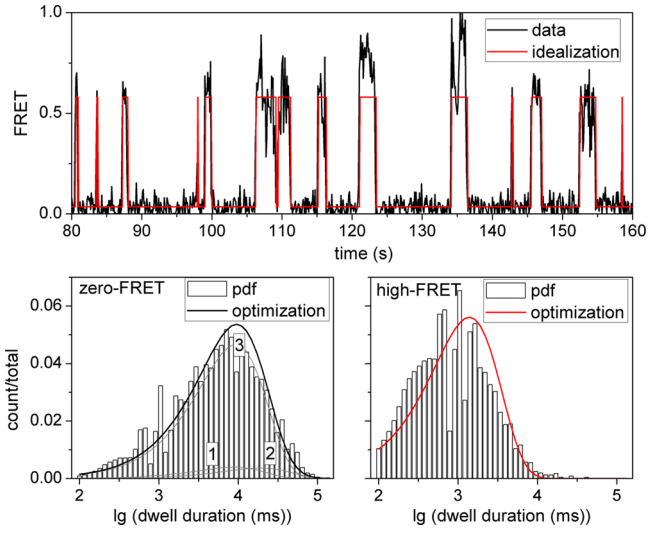


Figure 5. Outcomes of sequential K-means and maximum likelihood interval procedures. The upper panel contains a fragment of FRET trace with overlaid idealization. In the lower two panels are pdf's of dwell durations (bars) overlaid with calculated optimized pdf's (solid black and red curves). The zero-FRET state pdf is shown together with contributions from different substates (gray lines) enumerated as in the kinetic scheme in Figure 4.

and unlooping rates with the rest fixed to the values found in Zaremba et al.:[9] $k_{off} = 0.001\ s^{-1}$, $k_{on} = 10^7\ M^{-1}\ s^{-1} \times 6 \times 10^{-9}\ M = 0.06\ s^{-1}$. Even though optimization is done directly on idealized traces and not on the distributions of the dwell times, the goodness of the fit is visualized by overlaying the data histograms with pdf's calculated with optimized model parameters (Figure 5). The result of optimization is $k_L = 0.12 \pm 0.01\ s^{-1}$ and $k_U = 0.78 \pm 0.02\ s^{-1}$.

DISCUSSION

It was thought that, given a sufficient amount of data, it should, in principle, be possible to extract all the parameters of the kinetic scheme with one optimization. Qin et al.[28] have demonstrated that their maximum likelihood algorithm with good precision retrieves the kinetic parameters of various kinetic models from the synthetic data. The rates in the nicotinic acetylcholine receptor model (our kinetic scheme) were extracted from a data set with 15 000 events. We verified this by simulating in QuB with the same rate parameters as in the literature[28] a single segment of data with a similar number of events and then running MIL with the starting values of the rate constants very different from the actual ones. The result was that optimization converged to the values of the rate parameters rather close to those used to generate the data.

To verify if this could be successful with the model parameters of Ecl18kI, we simulated a single segment data set with our kinetic model and $k_{off} = 0.001\ s^{-1}$, $k_{on} = 0.06\ s^{-1}$, $k_L = 0.1\ s^{-1}$, and $k_U = 0.8\ s^{-1}$. Again, the trace contained about 15 000 events and consisted of 960 000 points, which corresponds to 400 experimental traces of 2400 points. The rate constants were pairwise scaled: $k_{12} = 2k_{23}$ and $k_{32} = 2k_{21}$ (Figure 4). Even in spite of this constraint, the MIL procedure did not converge even when run with the starting values equal to those used to generate the data. Similarly, MIL did not converge on the measured data. If we suppose that the actual rates underlying the data are of the order of those found in Zaremba et al., failure

dx.doi.org/10.1021/jp504546v | J. Phys. Chem. B 2014, 118, 8575−8582

to converge is not surprising, considering that the measured data set is smaller and more segmented.

From the comparison of the two calculations, it appears that optimization cannot be successful for an arbitrary combination of the rate constants. One of the reasons is that a rate constant will be estimated poorly or it will not be possible to estimate it at all if it leads from a state that is rarely visited.[40] Such is the case of states 1 and 2 in the kinetic scheme (Figure 4), since the dissociation rate constant leading to these states is orders of magnitude smaller than the rest of the rates in the scheme. This can also be seen from the contributions of different states to the overall pdf shown in Figure 5. The components of states 1 and 2 are small compared to that of state 3. The result of this imbalance of the rate values is that optimization of the whole kinetic scheme could not be accomplished. Thus, we compromised for a feasible task—found the looping and unlooping rates only, assuming the values of the rest of the rates from the literature.

Admittedly, the values of association and dissociation rates borrowed from the literature[9] are such that the weakly occupied states 1 and 2 introduce very little non-mono-exponentiality into the pdf of the zero-FRET state. We tried to optimize the rates of a simple two-state model on the data generated by a full Ecl18kI model (see above), and the resulting looping rate constant differs from the true value only by 3%. That is, assuming the values of dissociation and association constants such as we did, to find the looping rate constant, optimization could be simply substituted for fitting of the zero-FRET pdf with a single exponential. That such simplification is legitimate, however, cannot be known beforehand, and it is proper to do optimization on the trajectory rather than fit a pdf with single or multiple exponents and attribute to them physical meaning.

As explained in the Results, Ecl18kI hypothetically could form a loop with an alternative conformation that is not detected in our experimental scheme. Taking it into account would require adding of one more zero-FRET substate branching off state 3 in our kinetic scheme (Figure 4). Again, we wished to determine what would be the effect on the optimization results of this added loop. To this end, we simulated a trajectory with the Ecl18kI model with one more loop added. Even though this alternative loop is less likely, we set its looping rates to the optimized values calculated on the measured data with the one-loop model. Then, the looping and unlooping rates of the Ecl18kI model with one loop are optimized on this simulated data. The resulting looping rate constant is only 10% off the true value. This means that even if an alternative and unobserved loop was formed the rate constant of the measured looping calculated with the one-loop model would be rather close to the true value.

CONCLUSIONS

In conclusion, we successfully applied the SM FRET to study the dynamics of protein-induced DNA looping and from acquired data extracted the values of the rate constants governing the dynamics of looping and unlooping. The unlooping rate constant ~0.78 s^{-1} is in a good agreement with the value of 0.86 s^{-1} obtained previously.[9] This is not surprising, since this rate constant primarily depends on the affinity of the DNA-bound Ecl18kI dimers which should not be too different in the two experiments.

The looping rate constant 0.12 s^{-1} is smaller than the value 0.6 s^{-1} obtained by Zaremba et al.,[9] and this could be due to the fact that, even though we took care to find the exact intertarget distance for maximum looping frequency, this distance is ~250 bp, whereas in ref 9 it is ~520 bp. Since shorter DNA is more rigid, the looping in our case is less frequent. At the same time, in ref 9, the value of the looping rate constant calculated from the data is 0.15 s^{-1}. The presence of the bead creates a stretching force on the DNA, reducing the observed looping rate.[41] The true value of the looping rate is then obtained by multiplying the observed rate by a factor of 4 to correct for the effect of the bead-associated entropic force. Since this factor is an estimate, the end result is also to a certain degree uncertain. On the other hand, because in our experiment there is no distorting influence of the bead, the looping rate constant extracted from the measurement is the sought end result. At 250 bp, the predicted J factor for loop formation depending on the nature of interaction between the two DNA sites is between 10^{-7} and 10^{-8} M.[34] With the looping rate constant of 0.15 s^{-1}, this translates to the second order association rate constant between the target sites equal to $1.5 \times (10^6-10^7)$ M^{-1} s^{-1} which is only a few orders of magnitude slower than the diffusion-limited rate. Then, the overall picture is such that DNA-bound Ecl18kI dimers are rather efficient in forming a tetramer but their interaction is transient: it takes just more than a second for their association to be disrupted. It has to be noted, however, that even such relatively short time scale is not the limit for our setup. It could be used to study even faster DNA looping dynamics.

ASSOCIATED CONTENT

Supporting Information

The details of the biochemical preparations, a fluorescence image, and additional fluorescence and FRET traces. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*Phone: +370 64355421. Fax: +370 5 264 9774. E-mail: danielis@ar.fi.lt.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

This research was funded by a grant (No. MIP-43/2013) from the Research Council of Lithuania. We would like to thank Georgij Kostiuk for participating in biochemical preparations.

REFERENCES

(1) Schleif, R. DNA Looping. *Annu. Rev. Biochem.* **1992**, *61*, 199–223.

(2) Matthews, K. S. DNA Looping. *Microbiol. Rev.* **1992**, *56*, 123–136.

(3) Pingoud, A.; Fuxreiter, M.; Pingoud, V.; Wende, W. Type II Restriction Endonucleases: Structure and Mechanism. *Cell. Mol. Life Sci.* **2005**, *62*, 685–707.

(4) Wentzell, L. M.; Halford, S. E. DNA Looping by the Sfi I Restriction Endonuclease. *J. Mol. Biol.* **1998**, *281*, 433–444.

(5) Zaremba, M.; Sasnauskas, G.; Urbanke, C.; Siksnys, V. Conversion of the Tetrameric Restriction Endonuclease Bse634I into a Dimer: Oligomeric Structure-Stability-Function Correlations. *J. Mol. Biol.* **2005**, *348*, 459–478.

(6) Tamulaitis, G.; Sasnauskas, G.; Mucke, M.; Siksnys, V. Simultaneous Binding of Three Recognition Sites Is Necessary for a Concerted Plasmid DNA Cleavage by EcoRII Restriction Endonuclease. *J. Mol. Biol.* **2006**, *358*, 406–419.

(7) Halford, S. E.; Welsh, A. J.; Szczelkun, M. D. Enzyme-Mediated DNA Looping. *Annu. Rev. Biophys. Biomol. Struct.* **2004**, *33*, 1−24.
(8) Zaremba, M.; Sasnauskas, G.; Siksnys, V. The Link between Restriction Endonuclease Fidelity and Oligomeric State: A Study with Bse634I. *FEBS Lett.* **2012**, *586*, 3324−3329.
(9) Zaremba, M.; Owsicka, A.; Tamulaitis, G.; Sasnauskas, G.; Shlyakhtenko, L. S.; Lushnikov, A. Y.; Lyubchenko, Y. L.; Laurens, N.; van den Broek, B.; Wuite, G. J. L.; et al. DNA Synapsis through Transient Tetramerization Triggers Cleavage by Ecl18kI Restriction Enzyme. *Nucleic Acids Res.* **2010**, *38*, 7142−7154.
(10) Van den Broek, B.; Vanzi, F.; Normanno, D.; Pavone, F. S.; Wuite, G. J. L. Real-Time Observation of DNA Looping Dynamics of Type IIE Restriction Enzymes NaeI and NarI. *Nucleic Acids Res.* **2006**, *34*, 167−174.
(11) Shlyakhtenko, L. S.; Gilmore, J.; Portillo, A.; Tamulaitis, G.; Siksnys, V.; Lyubchenko, Y. L. Direct Visualization of the EcoRII - DNA Triple Synaptic Complex by Atomic Force Microscopy. *Biochemistry* **2007**, *46*, 11128−11136.
(12) Bochtler, M.; Szczepanowski, R. H.; Tamulaitis, G.; Grazulis, S.; Czapinska, H.; Manakova, E.; Siksnys, V. Nucleotide Flips Determine the Specificity of the Ecl18kI Restriction Endonuclease. *EMBO J.* **2006**, *25*, 2219−2229.
(13) Tamulaitis, G.; Solonin, A. S.; Siksnys, V. Alternative Arrangements of Catalytic Residues at the Active Sites of Restriction Enzymes. *FEBS Lett.* **2002**, *518*, 17−22.
(14) Laurens, N.; Bellamy, S. R. W.; Harms, A. F.; Kovacheva, Y. S.; Halford, S. E.; Wuite, G. J. L. Dissecting Protein-Induced DNA Looping Dynamics in Real Time. *Nucleic Acids Res.* **2009**, *37*, 5454−5464.
(15) Laurens, N.; Rusling, D. a; Pernstich, C.; Brouwer, I.; Halford, S. E.; Wuite, G. J. L. DNA Looping by FokI: The Impact of Twisting and Bending Rigidity on Protein-Induced Looping Dynamics. *Nucleic Acids Res.* **2012**, *40*, 4988−4997.
(16) Finzi, L.; Gelles, J. Measurement of Lactose Repressor-Mediated Loop Formation and Breakdown in Single DNA Molecules. *Science* **1995**, *267*, 378−380.
(17) Han, L.; Garcia, H. G.; Blumberg, S.; Towles, K. B.; Beausang, J. F.; Nelson, P. C.; Phillips, R. Concentration and Length Dependence of DNA Looping in Transcriptional Regulation. *PLoS One* **2009**, *4*, e5621.
(18) Vanzi, F.; Broggio, C.; Sacconi, L.; Pavone, F. S. Lac Repressor Hinge Flexibility and DNA Looping: Single Molecule Kinetics by Tethered Particle Motion. *Nucleic Acids Res.* **2006**, *4*, 3409−3420.
(19) Karymov, M.; Daniel, D.; Sankey, O. F.; Lyubchenko, Y. L. Holliday Junction Dynamics and Branch Migration: Single-Molecule Analysis. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 8186−8191.
(20) Tan, E.; Wilson, T. J.; Nahas, M. K.; Clegg, R. M.; Lilley, D. M. J.; Ha, T. A Four-Way Junction Accelerates Hairpin Ribozyme Folding via a Discrete Intermediate. *Proc. Natl. Acad. Sci. U. S. A.* **2003**, *100*, 9308−9313.
(21) Killian, J. L.; Li, M.; Sheinin, M. Y.; Wang, M. D. Recent Advances in Single Molecule Studies of Nucleosomes. *Curr. Opin. Struct. Biol.* **2012**, *22*, 80−87.
(22) Van Dijk, M.; Bonvin, A. M. J. J. 3D-DART: A DNA Structure Modelling Server. *Nucleic Acids Res.* **2009**, *37*, W235−W239.
(23) Joo, C.; McKinney, S. a; Nakamura, M.; Rasnik, I.; Myong, S.; Ha, T. Real-Time Observation of RecA Filament Dynamics with Single Monomer Resolution. *Cell* **2006**, *126*, 515−527.
(24) Roy, R.; Hohng, S.; Ha, T. REVIEW A Practical Guide to Single-Molecule FRET. *Nat. Methods* **2008**, *5*, 507−516.
(25) Hwang, H.; Kim, H.; Myong, S. Protein Induced Fluorescence Enhancement as a Single Molecule Assay with Short Distance Sensitivity. *Proc. Natl. Acad. Sci. U. S. A.* **2011**, *108*, 7414−7418.
(26) Milescu, L. S.; Nicolai, C.; Bannen, J. *QuB Software*; 2000.
(27) Qin, F. Restoration of Single-Channel Currents Using the Segmental K-Means Method Based on Hidden Markov Modeling. *Biophys. J.* **2004**, *86*, 1488−1501.
(28) Qin, F.; Auerbach, A.; Sachs, F. Estimating Single-Channel Kinetic Parameters from Idealized Patch-Clamp Data Containing Missed Events. *Biophys. J.* **1996**, *70*, 264−280.
(29) Qin, F.; Auerbach, A.; Sachs, F. Maximum Likelihood Estimation of Aggregated Markov Processes. *Proc. Biol. Sci.* **1997**, *264*, 375−383.
(30) Du, Q.; Smith, C.; Shiffeldrim, N.; Vologodskaia, M.; Vologodskii, A. Cyclization of Short DNA Fragments and Bending Fluctuations of the Double Helix. *Proc. Natl. Acad. Sci. U. S. A.* **2005**, *102*, 5397−5402.
(31) Cloutier, T. E.; Widom, J. Spontaneous Sharp Bending of Double-Stranded DNA. *Mol. Cell* **2004**, *14*, 355−362.
(32) Shimada, J.; Yamakawa, H. Ring-Closure Probabilities for Twisted Wormlike Chains. Application to DNA. *Macromolecules* **1984**, *698*, 689−698.
(33) Levene, S. D.; Giovan, S. M.; Hanke, A.; Shoura, M. J. The Thermodynamics of DNA Loop Formation, from J to Z. *Biochem. Soc. Trans.* **2013**, *41*, 513−518.
(34) Peters, J. P.; Maher, L. J. DNA Curvature and Flexibility in Vitro and in Vivo. *Q. Rev. Biophys.* **2010**, *43*, 23−63.
(35) Swigon, D.; Coleman, B. D.; Olson, W. K. Modeling the Lac Repressor-Operator Assembly: The Influence of DNA Looping on Lac Repressor Conformation. *Proc. Natl. Acad. Sci. U. S. A.* **2006**, *103*, 9879−9884.
(36) Zhang, Y.; McEwen, A. E.; Crothers, D. M.; Levene, S. D. Analysis of in-Vivo LacR-Mediated Gene Repression Based on the Mechanics of DNA Looping. *PLoS One* **2006**, *1*, e136.
(37) Shelley, C.; Magleby, K. L. Linking Exponential Components to Kinetic States in Markov Models for Single-Channel Gating. *J. Gen. Physiol.* **2008**, *132*, 295−312.
(38) Colquhoun, D.; Hawkes, A. G. Fitting and Statistical Analysis of Single-Channel Records. In *Single-channel recording*; Sakmann, B., Neher, E., Eds.; Springer: New York, 2009; pp 589−633.
(39) Qin, F.; Li, L. Model-Based Fitting of Single-Channel Dwell-Time Distributions. *Biophys. J.* **2004**, *87*, 1657−1671.
(40) Colquhoun, D.; Hawkes, A. G. The Principles of the Stochastic Interpretation of Ion-Channel Mechanisms. In *Single-channel recording*; Sakmann, B., Neher, E., Eds.; Springer: New York, 2009; pp 397−479.
(41) Segall, D. E.; Nelson, P. C.; Phillips, R. Excluded-Volume Effects in Tethered-Particle Experiments: Bead Size Matters. *Phys. Rev. Lett.* **2006**, *96*, 088306.