2864–2877 *Nucleic Acids Research, 2006, Vol. 34, No. 10*
doi:10.1093/nar/gkl382

DNA looping by two-site restriction endonucleases: heterogeneous probability distributions for loop size and unbinding force

Gregory J. Gemmen, Rachel Millin and Douglas E. Smith*

Department of Physics, Mail Code 0379, University of California, San Diego, 9500 Gilman Drive, La Jolla, CA 92093, USA

Received February 22, 2006; Revised April 4, 2006; Accepted May 2, 2006

ABSTRACT

Proteins interacting at multiple sites on DNA via looping play an important role in many fundamental biochemical processes. Restriction endonucleases that must bind at two recognition sites for efficient activity are a useful model system for studying such interactions. Here we used single DNA manipulation to study sixteen known or suspected two-site endonucleases. In eleven cases (BpmI, BsgI, BspMI, Cfr10I, Eco57I, EcoRII, FokI, HpaII, NarI, Sau3AI and SgrAI) we found that substitution of Ca^{2+} for Mg^{2+} blocked cleavage and enabled us to observe stable DNA looping. Forced disruption of these loops allowed us to measure the frequency of looping and probability distributions for loop size and unbinding force for each enzyme. In four cases we observed bimodal unbinding force distributions, indicating conformational heterogeneity and/or complex binding energy landscapes. Measured unlooping events ranged in size from 7 to 7500 bp and the most probable size ranged from less than 75 bp to nearly 500 bp, depending on the enzyme. In most cases the size distributions were in much closer agreement with theoretical models that postulate sharp DNA kinking than with classical models of DNA elasticity. Our findings indicate that DNA looping is highly variable depending on the specific protein and does not depend solely on the mechanical properties of DNA.

INTRODUCTION

A wide range of important biological processes, including DNA transcription, replication, recombination and repair, involve formation of protein complexes that bind at multiple sites on the same DNA molecule (1–8). When such proteins are bound at distant sites along DNA they may interact via looping of the intervening DNA, and this phenomenon has been shown to have important implications for the regulation of DNA-directed biochemical processes (1,9–11). For example, DNA looping interactions can enhance or repress gene expression (12).

The dependence of the rate of loop formation on loop size has been considered in many theoretical studies (13–22). General considerations in polymer physics suggest there should be an optimum loop size. Very small loops are unfavorable due to DNA bending rigidity, whereas very large loops are entropically unfavorable. Considering only DNA mechanics, an optimum loop size of ~500 bp is predicted by models which treat DNA as a worm like chain (WLC) with homogeneous bending ridigity. However, a controversy has recently arisen over whether such models can properly describe data on the cyclization of very short DNA molecules (<100 bp) (23–25). While cyclization involves only DNA mechanics, theories for protein-induced DNA looping must account for additional effects such as the span and elasticity of the protein complex, loop rearrangement entropy and geometry of loop closure (Figure 1). These effects tend to shift the loop size distribution to slightly smaller values. In an effort to explain very small loops some investigators have also developed models that postulate spontaneous or protein-induced DNA kinking (15,17,19,21,25).

Two-site restriction endonucleases (REases) acting on long linear DNA molecules provide a model system for studying protein-mediated DNA looping (26). With several enzymes (NaeI, Cfr10I, EcoRII and Sau3AI), loops have been directly imaged by electron microscopy (27–30). In these experiments, Ca^{2+} was substituted for the normal enzyme cofactor Mg^{2+}. Under these conditions it was found that specific binding occurs but cleavage does not. Evidence of looping with SfiI and Cfr10I comes from DNA recombination and gel mobility shift measurements (31–33), with NgoMIV from fluorescence resonance energy transfer measurements (34), with BspMI from magnetic tweezers measurements (35), and with NarI and NaeI from tethered particle assays

*To whom correspondence should be addressed. Tel: +1 858 534 5241; Email: des@physics.ucsd.edu

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors

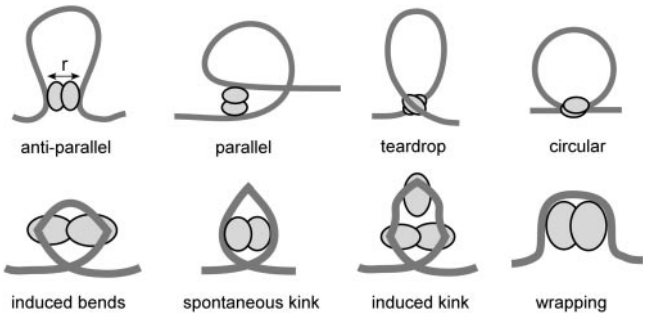

Figure 1. Schematic illustrations of various possible looping geometries. The parameter *r* indicates the span of the protein.

(36). In a number of other cases measurements show that efficient cleavage only occurs on templates containing two or more recognition sites (37–40), suggesting a DNA looping mechanism. Finally, in additional cases other enzymes cleavage has been found to be stimulated by short oligonucleotide duplexes containing the recognition sequence, suggesting that the enzyme complex can bind at two sites in trans (i.e. on two different molecules) (41–44).

Recently we used optical tweezers manipulation to study the DNA-tension dependence of cleavage by twenty-one different one-site and two-site REases. We found that the activity of two-site enzymes was universally 'switched off' by application of 5 pN of tension to the DNA, whereas that tension had virtually no effect on one-site enzymes (Gemmen, Millin, Smith, submitted manuscript). This finding is in accord with several recent theories that predict a strong tension dependence of DNA looping (19,20). Here, we report direct measurements of stable DNA looping with eleven two-site REases under conditions where cleavage was blocked by substituting Ca^{2+} for Mg^{2+}. These measurements allow us to characterize the rate of looping, distribution of loop sizes and binding strengths of loops. Our findings indicate that DNA looping depends strongly on the specific protein and not solely on the mechanical properties of DNA. We compare our measurements with many different theoretical predictions in the literature and find that in most cases our data agree more closely with models that postulate sharp DNA kinking than with classical models that assume homogeneous bending rigidity.

MATERIALS AND METHODS

Endonucleases

BamHI, BpmI, BsgI, BspMI, BstNI, EcoRI, EcoRV, FokI, HaeIII, HpaII, MboII, MspI, NarI, NaeI, SacII, Sau3AI, SfiI and SgrAI were obtained from New England Biolabs (NEB). BfiI, Cfr9I, Cfr10I and Eco57I were obtained from Fermentas. EcoRII and Ksp632I were obtained from Roche. The two-site enzymes and their properties are listed in Table 1.

Enzyme buffers

BfiI – Y + /Tango (33 mM Tris–acetate, 10 mM calcium acetate, 66 mM potassium acetate, 0.1 mg/ml BSA); BpmI – NEB3 + BSA [50 mM Tris–HCl, 100 mM NaCl, 10 mM $CaCl_2$, 1 mM DTT, 100 μg/ml BSA, (pH 7.9)]; BsgI – NEB4 + SAM [20 mM Tris–acetate, 50 mM potassium acetate, 10 mM calcium acetate, 1 mM DTT, 80 μM S-adenosylmethionine (pH 7.9)]; BspMI – NEB3 [50 mM Tris–HCl, 100 mM NaCl, 10 mM $CaCl_2$, 1 mM DTT (pH 7.9)]; Cfr9I – 10 mM Tris–HCl, 5 mM $CaCl_2$, 200 mM Na-glutamate, 100 μg/ml BSA (pH 7.2); Cfr10I – 10 mM Tris–HCl, 5 mM $CaCl_2$2, 100 mM NaCl, 0.02% Triton X-100, 100 μg/ml BSA (pH 8.0); Eco57I – 10 mM Tris–HCl, 10 mM $CaCl_2$, 50 mM NaCl, 10 mM S-adenosylmethionine, 100 μg/ml BSA (pH 7.5) EcoRII – SuREH [50 mM Tris–HCl, 100 mM NaCl, 10 mM $CaCl_2$, 1 mM DTT (pH = 7.5)]; FokI – NEB4 [20 mM Tris–acetate, 50 mM potassium acetate, 10 mM calcium acetate, 1 mM DTT (pH 7.9)]; HpaII – NEB1 [10 mM Bis–Tris–Propane–HCl, 10 mM $CaCl_2$, 1 mM DTT (pH 7.0)]; Ksp632I – 33 mM Tris–acetate, 66 mM potassium acetate, 10 mM calcium acetate, 0.5 mM DTT (pH 7.9); MboII – NEB2 [10 mM Tris–HCl, 50 mM NaCl, 10 mM $CaCl_2$, 1 mM DTT (pH 7.9)]; NarI – NEB1 [10 mM Bis–Tris–Propane–HCl, 10 mM $CaCl_2$, 1 mM DTT (pH 7.0)]; SacII – NEB4 [20 mM Tris–acetate, 50 mM potassium acetate, 10 mM calcium acetate, 1 mM DTT (pH 7.9)]; SgrAI – NEB4 (20 mM Tris–acetate, 50 mM potassium acetate, 10 mM calcium acetate, 1 mM DTT (pH 7.9)].

Table 1. Properties of the REases as reported in REBASE

Enzyme	Type	MW (kDa)	Form in solution	Active complex	Recognition sequence
BfiI	IIS	40	Dimer	Dimer	ACTGGG (5/4)
BpmI	IIE, G, S	117	*	*	$CTGGAG(N)_{16}\downarrow$
BsgI	IIE, G, S	121	*	*	$GTGCAG(N)_{16}\downarrow$
BspMI	IIE	222	Tetramer	Tetramer	$ACCTGC(N)_4\downarrow$
Cfr9I	IIE	37	*	*	C↓CCGGG
Cfr10I	IIF, P	320	Tetramer	Tetramer	R↓CCGGY
Eco57I	IIE, G	117	Monomer	*	$CTGAAG(N)_{16}\downarrow$
EcoRII	IIE, P	92	Dimer	Dimer	↓CCWGG
FokI	IIS	66	Monomer	Dimer	$GGATG(N)_9\downarrow$
HpaII	IIE, P	41	*	*	C↓CGG
Ksp632I	IIE	*	*	*	CTCTTC (1/4)
MboII	IIS	49	Monomer	Dimer	GAAGA (8/7)
NarI	IIE	*	*	*	GG↓CGCC
SacII	IIE	*	*	*	CCGC↓GG
Sau3AI	IIE, P	56	Monomer	Dimer	↓GATC
SgrAI	IIP	38	Dimer	Tetramer	C(A/G)↓CCGG(C/T)G

Type IIE REases bind at two sites, but only one is cleaved, whereas type IIF cleave coordinately at both binding sites. Type IIG have restriction and modification activities in the same subunit and Type IIS enzymes recognize asymmetric sequences and cleave at least one strand outside of the recognition sequence. Type IIP enzymes recognize symmetric sequences. Entries marked with an asterisk are those for which no information was available.

DNA constructs

'LBAC-A' was prepared by ligating a 4282 bp digoxygenin (DIG)-labeled PCR fragment to a 10 845 bp biotin-end-labeled restriction fragment of pBACe3.6. The PCR fragment was generated by amplification of a sequence from pFastBac HT-b (Invitrogen) using the primers d(GTGGTATGGCTGATTATGATC) and d(GCAGCCTGAATGGCGAATGG) and was labeled by incorporation of 20 μM of dUTP-11-DIG

(Roche) and 200 μM each of dATP, dCTP, dGTP, dTTP in the PCR. The multiple DIG labeling was used to provide a stronger attachment in some experiments. The 10 845 bp fragment was produced by digesting pBACe3.6 (Children's Hospital of Oakland Research Institute) with BsrGI (New England Biolabs, 'NEB') and end labeling by using the Klenow fragment of *Escherichia coli* DNA polymerase I, exo$^-$, (NEB) to incorporate dATP-14-biotin (Invitrogen). Both fragments were purified using the Qiagen PCR purification kit and digested with XhoI (NEB). To isolate the desired product the samples were run on a 1% agarose gel in 1× TAE buffer and purified using the Qiagen Gel Extraction kit. The two fragments were then ligated by use of T4 DNA ligase (NEB).

'LBAC-B' was prepared by labeling the aforementioned biotin-labeled, 10 845 bp XhoI fragment of pBACe3.6 by using the Klenow fragment of DNA polymerase I, exo-, to incorporate dUTP-11-DIG.

'½-λ-L' was prepared by using the Klenow fragment of DNA polymerase I, exo$^-$, to fill in the ends of methyladenine-free λ DNA (NEB) with biotin-dATP and dCTP (Invitrogen). The DNA was then digested by XbaI and purified using the Promega Wizard DNA clean up kit. A second fill-in was then done with DIG-labeled dUTP (Roche). The fragments were then digested with XhoI to select the left end (24 508 bp). '½-λ-R' was prepared in the same manner, except ApaI was used instead of XhoI to select the right end (23 994 bp).

PhiX174 DNA was purchased from NEB and was labeled by digesting with XhoI and end-labeling with dATP-14-biotin (Invitrogen). The DNA was then digested with StuI, purified using the Promega Wizard DNA clean up kit, and end-labeled with dUTP-11-DIG.

DNA tethering

Streptavidin coated microspheres (200 μl of 0.5% w/v, 2.2 μm diameter, Spherotech) were washed by twice centrifuging at 10 000 *g* and re-suspending in 200 μl of phosphate-buffered saline (PBS) pH 7.4 (Fisher Scientific) and 0.1 mg/ml BSA (New England Biolabs). A total of 5 μl of diluted DNA (~10–100 ng/μl) was mixed with 5 μl of microspheres and incubated for ~45 min at room temperature on a slowly rotating mixer. A total of 5–10 μl of these microspheres were diluted in 0.5 ml of PBS and loaded into a 1 ml tuberculin syringe for injection into the sample chamber.

Protein G coated microspheres (200 μl of 0.5% w/v, 2.8 μm diameter, Spherotech) were washed in the same manner and re-suspended in 20 μl PBS. Then, 5 μl of anti-DIG (200 μg/ml, Roche) was added. The microspheres were incubated on the mixer for ~45 min and then washed three more times and resuspended in 20 μl PBS. A total of 5 μl of the microspheres were loaded into a syringe for injection into the sample chamber.

Our optical tweezers instrument has been described previously (45). The anti-DIG coated microsphere was held by a micropipette while the microsphere carrying the DNA was trapped with the optical tweezers. The two microspheres were brought into proximity such that the DIG-labeled end of one DNA molecule bound to the anti-DIG coated bead, forming a DNA tether between them.

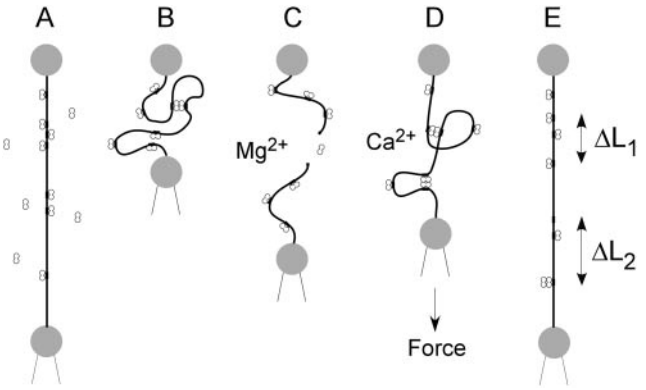

Figure 2. Schematic of the DNA looping measurements. The DNA is tethered between two microspheres, one manipulated by optical tweezers and the other manipulated by a piezoelectrically positioned micropipette. (**A**) The DNA is held taut while a solution containing enzyme is introduced. (**B**) The molecule is then relaxed for a prescribed incubation time, during which the active complex may form via DNA looping. (**C**) In the presence of Mg^{2+} ions, the DNA molecule is generally cleaved. (**D**) When Ca^{2+} is substituted for Mg^{2+} cleavage is blocked and stable DNA looping is detected. (**E**) Upon separating the microspheres, the looped complexes are disrupted, causing sudden increases in the tether length ΔL and drops in the DNA tension.

RESULTS

Detection of stable DNA loops

Sixteen known or suspected two-site enzymes were studied. These were selected based on the recognition sites on the DNA templates we had available (Materials and Methods; Table 1). Our measurement technique is illustrated in Figure 2. A single DNA molecule was held stretched (fractional extension 95%, corresponding to a tension of 5 pN) while the enzyme solution was flowed into the sample chamber. The enzymes were diluted in their standard reaction buffers except that Ca^{2+} was substituted for Mg^{2+}. The entire chamber was uniformly filled with the enzyme solution, the flow was stopped, and the DNA was then relaxed to a fractional extension of 35% (corresponding to a low tension of ~0.06 pN). The molecules were incubated for 2 min and the DNA was then stretched at a rate of 150 nm/s to assess loop formation. The tension in the DNA was measured at a rate of 100 Hz.

If the DNA remained tethered after reaching a tension of 60 pN it was relaxed back to 5 pN and incubation and stretching were repeated. If the tether detached from the microspheres, which typically occurred after 1 to 10 stretch cycles, the enzyme solution was drained from the sample chamber, a new DNA molecule was tethered, and a new aliquot of enzyme solution was introduced. Measurements were repeated ~70 to 300 times with each enzyme to accumulate statistics on loop formation.

Clear evidence for DNA looping was obtained with eleven out of sixteen putative two-site enzymes (BpmI, BsgI, BspMI, Cfr10I, Eco57I, EcoRII, FokI, HpaII, NarI, Sau3AI and SgrAI). Typical force-extension datasets are shown in Figure 3. Prior to introducing the enzyme the measured elasticity was as expected for a single, naked DNA molecule (46). After incubation with the enzyme the DNA tether was often shortened by a variable length, consistent with loop

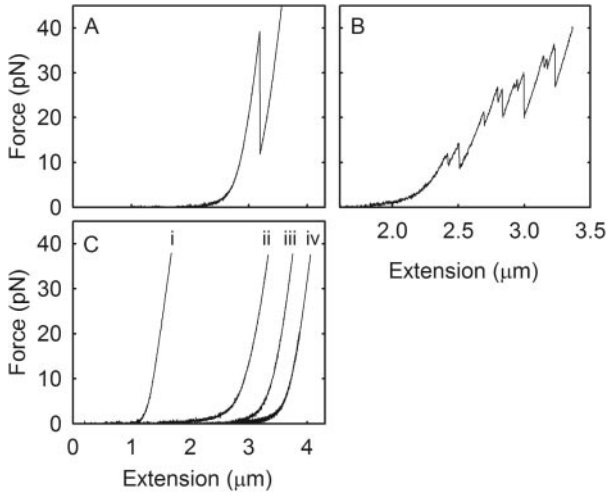

Figure 3. Typical DNA force-extension plots. (**A**) The two-site enzyme BsgI tested on the LBAC-B template, which has four recognition sites. The one detected loop had a measured length of 946 bp, in excellent agreement with the separation between the binding sites at positions 8342 and 9287 bp (a distance of 945 bp). (**B**) Sau3AI on the LBAC-A template, which has 55 recognition sites. (**C**) Control experiments: (i) the two-site enzyme Sau3AI was tested on bacteriophage phiX174 DNA, which contains no copies of its recognition sequence; (ii) the one-site enzyme HaeIII was tested on LBAC-B DNA containing 36 copies of its recognition sequence; (iii) the two-site enzyme SfiI was tested on LBAC-B DNA containing one copy of its recognition sequence and (iv) LBAC-B DNA, containing 26 FokI binding sites, incubated in the FokI reaction buffer, but no FokI added.

formation. Upon stretching we recorded sudden drops in the measured force, each followed by a steady increase in tension. These 'sawteeth' indicate events in which sequestered lengths of DNA are suddenly released, consistent with disruption of the individual DNA loops. Analysis of these events yields the number of loops formed, and the disruption force and DNA length change associated with each.

The observed length changes were consistent with the possible loop sizes given the separations of recognition sites on the DNA templates. With templates containing relatively few recognition sites, we typically observed zero or only one unlooping event per trial. An example is shown in Figure 3A, in which BsgI was tested on a template containing only four recognition sites. Only a single event was recorded in ~100 trials and yielded a length change of 946 bp, which corresponds closely to one of the possible separations between sites on the DNA template (the 945 bp distance between sites at 8342 and 9287 bp). In the other 99 trials there were no unlooping events. Therefore, all reported statistics on BsgI were obtained using a different template containing 31 sites. Multiple sawteeth were often recorded in cases where the DNA template contained many sites. An example is shown in Figure 3B, in which Sau3AI was tested on a template containing 55 recognition sites.

Four different types of control experiments were done (Figure 3C). First, DNA was stretched in the reaction buffers with no enzyme added to confirm that there were no sawteeth recorded. Second, several one-site REases (BstNI, HaeIII and MspI) were tested and no events were observed. Two of these, BstNI and MspI, were chosen for controls because they are isoschizomers of the two-site enzymes EcoRII and HpaII. Third, several two site-enzymes (NarI, SacII, SgrAI and Sau3AI) were tested with templates containing no recognition sites and, except with SgrAI, no events were observed. SgrAI was unusual in that it appeared to cause frequent non-specific looping, in accord with a previous report that SgrAI can bind non-cognate sites (47). Fourth, several two-site enzymes (NaeI, SacII and SfiI) were tested on a template containing only one site and no loop formation was detected.

Loops were not detected in five of the sixteen cases (BfiI, Cfr9I, Ksp632I, MboII and SacII) despite use of templates containing multiple sites. With BfiI, Cfr9I, MboII and SacII we previously observed inhibition of cleavage by tension in solutions containing Mg^{2+}, suggesting a looping mechanism. Here we detected no stable looping with Cfr9I or Ksp632I. In the case of Cfr9I, this may be attributable to the paucity of sites on the DNA template (six), but in the case of Ksp632I this was unexpected as the template had 20 sites and no loops were observed in 50 trials despite use of extended incubations of 5 min. With MboII, Ca^{2+}-dependent binding of DNA in trans was reported previously (39), but we did not detect DNA looping. With BfiI it was not possible to assess loop formation because we found, consistent with previous reports (48), that substitution of Ca^{2+} for Mg^{2+} did not block cleavage (in 21 out of 21 trials). Unexpectedly, cleavage was also observed in many trials with SacII (19 out of 33) and thus we did not attempt to collect statistics on looping with SacII.

Distributions of unbinding forces

With the eleven enzymes that exhibited looping in Ca^{2+}, loop disruption forces ranged from ~3 to 60 pN (Figure 4). The most probable force varied from ~10 to 30 pN (Figure 5). This range is similar to that measured for disruption of other protein–DNA complexes by optical tweezers, such as in our previous study of nucleosome unraveling (45). Statistics were collected with nine enzymes and are given in Table 2. Although looping with NarI and SgrAI was observed, statistics were not collected on these two enzymes as they have been reported to form non-specific loops (36,47). We note that the force distributions for different enzymes varied in shape. For example, those for EcoRII and BpmI were broad and roughly Gaussian whereas those for HpaII, Cfr10I and Eco57I were broad but skewed to low force, indicative of heterogeneous behavior dependent on enzyme.

Interestingly, the force distributions for FokI, Sau3AI, BspMI and BsgI were clearly bimodal. Such distributions may have several possible explanations. As DNA looping involves simultaneous protein–DNA and protein–protein interactions, one may consider whether the two force regimes corresponded to these two separate interactions. However, if each loop were bound by two links of unequal strengths, the unbinding of the weaker link would be more frequent. Only in the case of Eco57I was there evidence of a smaller peak at high forces, and the opposite was seen with BspMI, BsgI, FokI and Sau3AI. Thus, an alternative explanation is needed. As discussed below, one possibility is that the complexes could have heterogeneous binding modes; a second is that the binding energy landscapes contain multiple barriers (49).

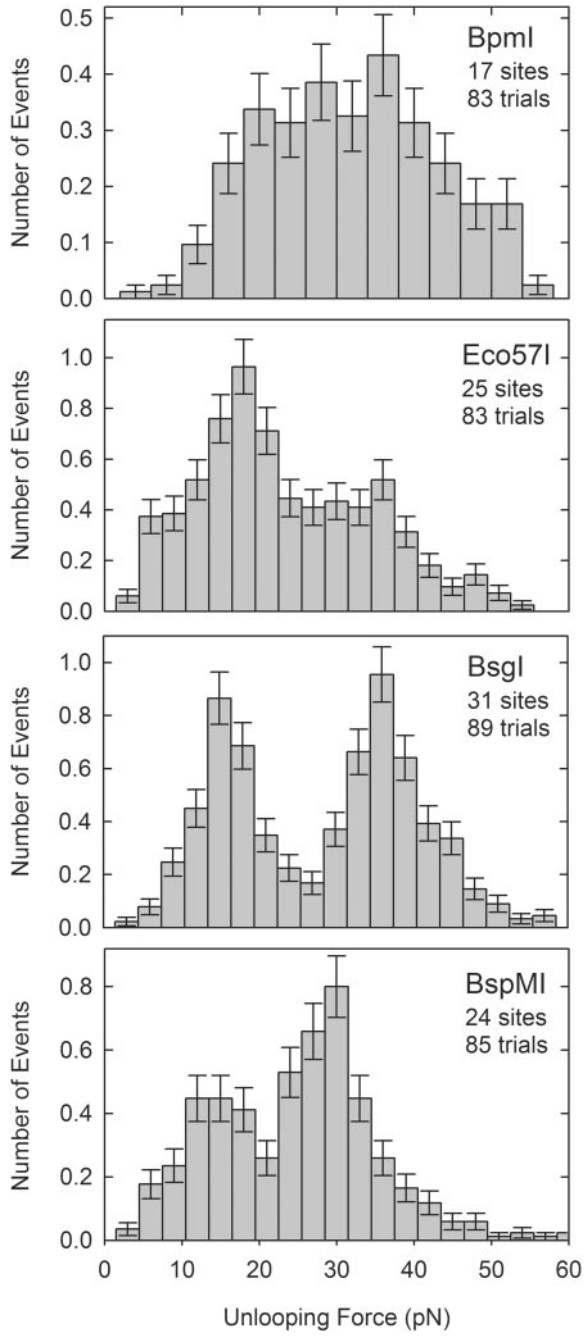

Figure 4. Histograms of the measured forces required to disrupt loops for different two-site enzymes. Number of events is reported as number per trial. So as not to low bias the force distributions, trials in which the tether broke before reaching 40 pN were not included. The number of recorded events ranged from 256 to 1330 depending on the enzyme.

Dependence on divalent cations

Many protein–DNA interactions are strongly dependent on divalent cations, and dependence on Mg^{2+} for cleavage is a hallmark of nearly all Type II REases. To investigate whether divalent ions were required for DNA looping, we conducted experiments in which 1 mM EDTA was substituted for Ca^{2+}. In the case of BpmI, BsgI, BspMI, Cfr10I, Eco57I and HpaII we did not detect any loop formation. However, we did detect

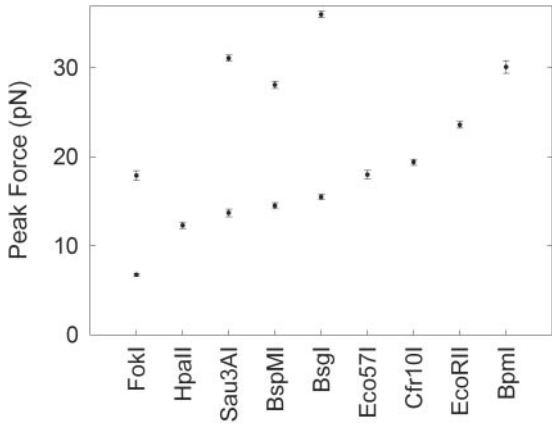

Figure 5. Most probable loop disruption force for each enzyme. For bimodal distributions, values for both the high and low peaks are plotted. The error bars report the standard error.

stable loop formation with EcoRII, FokI and Sau3AI (Figure 6). Consistent with previous reports, we found that BfiI cleaved DNA in the absence of metal ions (i.e. with 1 mM EDTA) (48). BfiI is considered very unusual in this regard since all other restriction enzymes appear to require Mg^{2+}.

Distributions of loop disruption forces measured with EcoRII, FokI and Sau3AI in Ca^{2+} versus EDTA are compared in Figure 6. In all three cases fewer loops were formed in EDTA. With Sau3AI and EcoRII the average unbinding force was lower, in accord with the expectation that divalent ions increase the binding specificity. With Sau3AI the force distribution was bimodal with Ca^{2+} but became monomodal in EDTA, which suggests qualitatively different binding modes dependent on ionic conditions. In the case of FokI, however, the unbinding force was unexpectedly higher with EDTA.

Measurements with HpaII were carried out with [Ca^{2+}] varying from 1 μM to 300 mM (Figure 7). No events were observed with EDTA and events were rarely observed with [Ca^{2+}] below 0.1 mM. The number of loops increased ~10-fold as [Ca^{2+}] was increased from 0.1 to 5 mM, but dropped ~10-fold as the calcium was further increased to 300 mM. The optimum point for loop formation occurred at ~5–10 mM, the same range reported to be optimal for cleavage activity with Mg^{2+}. Surprisingly, while the number of loops decreased with [Ca^{2+}] above 5 mM, suggesting nonspecific salt inhibition, the unbinding force increased steadily with [Ca^{2+}] up to 100 mM. This finding suggests that individual complexes can have different numbers of bound ions resulting in sub-populations with different binding strengths (50).

Distributions of loop sizes

Although many theoretical models have predicted the dependence of the probability of loop formation on loop size, little experimental data are available for comparison with these theories. Also, little is known about how loop sizes depend on the protein involved. An advantage of our method is that loop sizes are measured directly and loop size distributions are obtained from measurements repeated on an ensemble of complexes.

Table 2. Experimental conditions and results

Enzyme	C U/ml	DNA template	N_{sites}	N_P	$N/N_P \times 100$	<F>	<L>	L_S	L_L	ρ_S %
BpmI	20	½-λ-L	17	136	2.3 (0.42)	29.6 (11.2)	777 (208)	60	3634	18 (12)
BsgI	30	½-λ-L	31	465	1.7 (0.22)	26.9 (12.3)	645 (83)	45	3344	−3.2 (12)
BspMI	20	½-λ-L	24	276	0.91 (0.14)	23.2 (10.3)	613 (124)	25	4494	−16 (9.0)
Cfr10I	100	½-λ-L	56	1540	1.2 (0.052)	22.5 (12.4)	470 (89)	22	7444	−23 (5.9)
Eco57I	50	½-λ-L	25	300	2.7 (0.19)	21.8 (11.2)	573 (103)	19	7512	−6.7 (11)
EcoRII	50	½-λ-L	36	630	0.29 (0.026)	25.6 (12.2)	665 (176)	21	6940	−24 (10)
FokI	40	LBAC-B	26	325	0.79 (0.038)	14.8 (11.0)	301 (61)	18	4120	−22 (5.6)
HpaII	100	LBAC-B	49	1176	0.51 (0.041)	16.6 (9.92)	325 (65)	20	2500	−23 (6.9)
Sau3AI	40	LBAC-A	55	1485	0.66 (0.041)	25.0 (10.8)	423 (85)	7	2682	−11 (5.1)

N_{sites} is the number of recognition sites on the DNA template; C is the enzyme concentration in U/ml; N_P is the number of pairs of sites that can form loops on the DNA template; N/N_P is the number of observed loops per molecule normalized by N_P (and standard error); <F> is the mean loop disruption force (and SD) in pN; <L> is the mean observed loop size in base pairs (and SD) for the normalized distributions (see text); L_S is the shortest observed event size in the ensemble of data; L_L is the longest observed event size; ρ_S is the correlation coefficient between disruption force and event size for events <150 bp (and the SD expected for uncorrelated data).

The separations between recognition sites on the DNA template dictate the sizes of loops that can form in our experiment. Due to our use of long DNA templates with many binding sites these distributions are broad and extend out to ~10–20 kb (Supplementary Data). Comparisons between measured and possible loop sizes are shown in Figure 8. Here we have chosen a histogram bin size equal to the persistence length of DNA (150 bp), a characteristic measure of DNA rigidity (46). Several features are immediately evident. The distributions of possible sizes are continuous and nearly flat over the range from 0 to 3000 bp, while the measured distributions are strongly skewed towards shorter loops. This finding of few long loops is consistent with the theoretical expectation that such loops are entropically unfavorable.

While all datasets indicate strong suppression of large loops, the distributions of short loops varied dramatically for different enzymes. Among the four examples shown in Figure 8, the shortest loops were formed with EcoRII and Cfr10I, although the distribution is clearly broader with Cfr10I. The distribution is also broad with BsgI but is shifted to larger loop size. To determine the inherent probability distributions for loop size, correcting for the influence of DNA template, we normalized each distribution by dividing the number of measured events in each bin by the number of possible loop sizes (Figure 9). We also reduced the bin widths, which enabled us to identify optimum loop sizes (peak in the distribution) in most cases. The finding of a decrease in probability in the limit of small loop size is consistent with the theoretical expectation that small loops are unfavorable due to the bending rigidity of DNA. As shown in Figure 10A, the optimum loop size ranged from less than 75 bp to nearly 500 bp, dependent on enzyme. Interestingly, with BpmI and BspMI the size distributions had small secondary peaks. Such behavior may indicate multiple possible loop geometries for these enzymes. Statistics on measured loop sizes for all enzymes are given in Table 2.

Unlooping events ranged from as small as 7 bp to as large as ~7500 bp. Our resolution in detecting small loops was not limited by instrument resolution (~5 bp), but by the distribution of sites in the DNA template. Very few sites were separated by less than 10 bp. Minimum sizes measured with different enzymes ranged from 7 to 60 bp (Table 2). Our findings clearly show that loops substantially smaller than the persistence length are readily formed. In Figure 10B the rates of formation of short loops with different enzymes are compared.

Comparisons of the data with the predictions of various theoretical models are shown in Figures 9 and 10A. In all cases except with BsgI and BspMI, the measured loop sizes agree better with models that assume sharp DNA kinking (Figure 1) than with models that assume classical DNA elasticity. It is evident, however, that no single model was able to describe all of our data sets. Systematic comparisons of all of the datasets with all of the published models are given in the Supplementary Data section and discussed below.

Frequency of loops

The number of loops formed in a single DNA molecule with the enzyme solution can be directly counted in the force-extension datasets (Figure 3). We measure the number of loops formed after a 2 min incubation, which reports on the initial kinetics of loop formation. An incubation time of 2 min was used because a sufficient number of loops were formed in this interval to obtain statistics while also avoiding saturation [i.e. substantially fewer loops were formed than the maximum number possible ($N_{sites}/2$)]. This kinetic rate is a relevant quantity as the natural function of these enzymes is to cleave the DNA irreversibly after a loop is formed. Very high forces are required for unlooping and the loops are essentially irreversible on the time scale of our experiment. The maximum number of loops that can form in a given molecule is equal to $N_{sites}/2$, truncated to the nearest integer. In an ensemble of measurements, however, the number of different possible loops equals the number of combinations of pairs of sites $N_{pairs} = N_{sites}(N_{sites} - 1)/2$. To assess the relative rates of loop formation with different enzymes, we therefore calculated the mean number of observed loops per number of possible pairs of sites ($<N>/N_{pairs}$) (Table 2). Significant differences in looping frequency among enzymes are clearly evident.

DISCUSSION

Looping on multi-site templates

Our approach of using single DNA manipulation and templates containing a large number of recognition sites

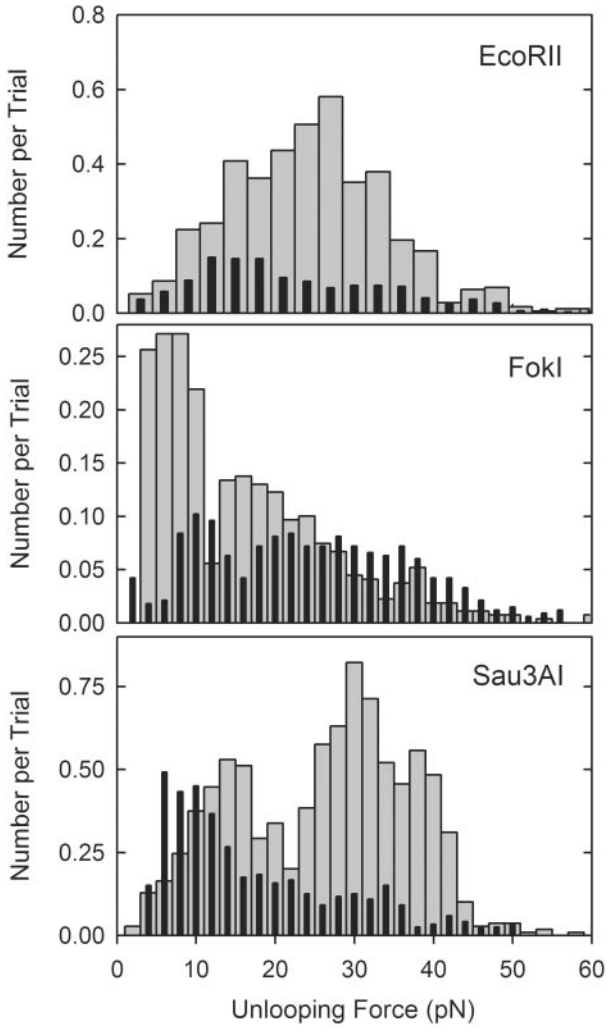

Figure 6. Distribution of forces at which loop disruption occurred in buffers containing Ca^{2+} (gray bars) versus no divalent cations and 1 mM EDTA added (black bars).

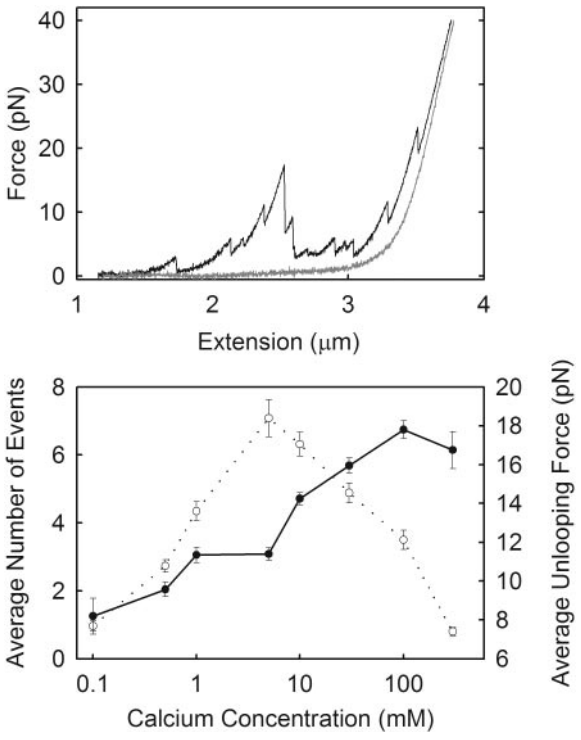

Figure 7. Dependence of DNA looping by HpaII on Ca^{2+}. (Upper) Examples of force-extension curves with 5 mM Ca^{2+} (black line, having many peaks) and with no divalent cations and 1 mM EDTA added (gray line, no peaks). (Lower) Mean number of loops (open circles, dashed line, left vertical axis) and mean loop disruption force (closed circles, solid line, right vertical axis) versus Ca^{2+} concentration.

allows for loop formation with a broad spectrum of possible loop sizes to be studied in a single experiment. However, a limitation of this technique is that in some cases (in particular, REases with nonpalindromic recognition sequences) individual recognition sites can have different possible orientations on the DNA. Because our templates contain a large number (hundreds to thousands) of possible pairs of sites, they allow for formation of many different loops of similar size, and we therefore cannot discern which two sites are involved in any one loop. Thus, we cannot dissect the effect of site orientation on looping. For those enzymes with non-symmetric recognition sites our templates contained many sites with both possible orientations. A previous study of BspMI observed an orientation effect with supercoiled DNA and it was speculated that such effects might occur for short loop sizes with linear DNA (51). On the other hand, two out of the nine cases we studied have symmetric recognition sequences (Sau3AI and HpaII) and in these cases the sites do not have distinct orientations. In another two cases (EcoRII and Cfr10I) the sequences are 'pseudo-symmetric' (CCWGG and RCCGGY) so they may also be insensitive to orientations. In the case of EcoRII, a previous study reported no effect of site orientation even for small site-separations (52).

The study of looping on a template with many sites has certain advantages. As our templates contained numerous sites, our experiment approximates a situation where there is a continuum of loop sizes. Thus our measurements report on which sizes occur given a near continuum of possible choices. In fact, this situation would correspond to the natural situation (e.g. Sau3AI, having only a four-letter recognition sequence, would be expected to encounter a long DNA with many sites).

Another benefit of having multiple sites is that this situation corresponds to that considered by a number of the published theories to which we can compare our data. A final advantage is that, in a set of experiments using only one DNA template, we obtain a great deal of information about the distribution of sizes and strengths of loops that can form. While certain additional information, such as on orientation effects, could be obtained by using templates with only two sites, such studies have only been done with a few enzymes and limited set of site separations (33,51–54). To systematically investigate the range of loop sizes we observed for nine different enzymes using templates with only two sites would require engineering hundreds to thousands of different DNA templates.

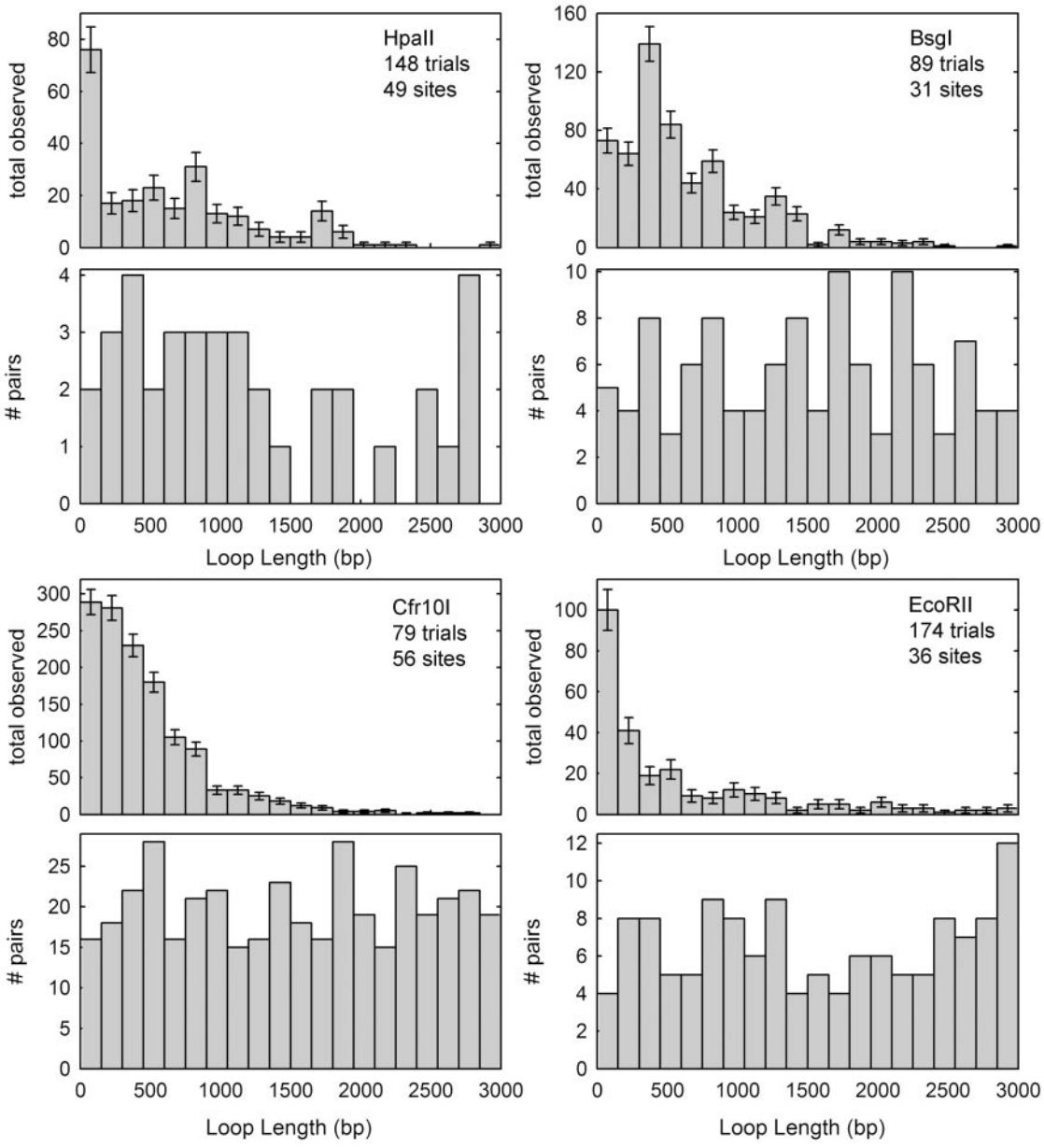

Figure 8. Histograms of observed loop sizes formed with various two-site enzymes. Below each graph is a histogram of the possible loop sizes plotted over the same range. The bin size is equal to 150 bp, which is the approximate persistence length of DNA.

Evidence for DNA looping

We detected DNA looping in buffers containing Ca^{2+} with eleven of the sixteen two-site enzymes that we studied. In five cases (with BspMI, Cfr10I, EcoRII, NarI and Sau3AI) our observations corroborate previous direct evidence for DNA looping (28–30,32,35,36). In four cases (with BpmI, BsgI, FokI and SgrAI) our observations confirm previous indirect evidence for DNA looping, which was postulated based on experiments showing that two sites are required for efficient cleavage (38,55). In two other cases (Eco57I and HpaII) we report the first evidence for DNA looping. These enzymes were previously suspected to act via a two-site mechanism based on reports that their activities on long templates are stimulated by addition of short oligonucleotide duplexes containing the recognition sequence (41–43).

Loop disruption forces

In our experiments loop opening is accelerated by application of force and the magnitudes of the disruption forces provide a measure of binding strength. An advantage of this approach is that we can rapidly probe high affinity complexes that would take an extremely long time to dissociate under equilibrium conditions (56). Average unlooping forces measured in our experiment ranged from ~15 to 30 pN. Forces of a similar magnitude were measured for unraveling of DNA bound in nucleosomes (45,57) and shown via force spectroscopy to correspond to a very slow zero force dissociation rate of

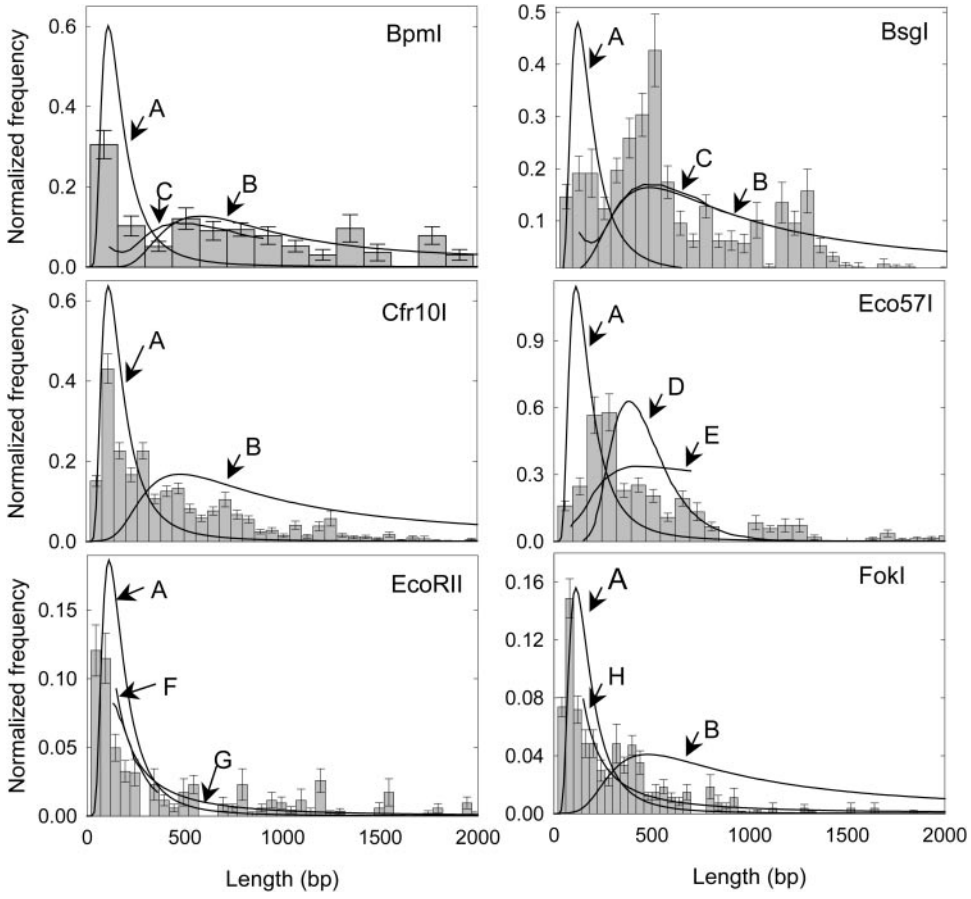

Figure 9. Normalized distributions of observed lengths of unlooping events following incubation of DNA with the enzyme solution for 2 min. To normalize for differences in DNA templates, histograms of the number of events (per molecule) in each length bin were divided by the number of pairs of sites on the DNA template in the bin. The solid lines are comparisons with theoretical distributions taken from the following references: (A) Ref. (19) 90° kink; (B) Ref. (16) with $r = 10$ nm; (C) Ref. (21) with hinge $\mu = 11$ and free-ends; (D) Ref. (18) with $\beta v = 0$ and $\beta\varepsilon = 15$; (E) Ref. (25) 90° kink with $P = 0.002$; (F) Ref. (17) 120° curvature with $r = 10$ nm; (G) Ref. (21) 90° kink and hinge with $\mu = 11$ and free-ends; (H) Ref. (17) 90° curvature with $r = 10$ nm. The theoretical predictions were normalized so that their integrated area over the range of the comparison was equal to the area under the corresponding data (see text and Supplementary Data).

$\sim 10^{-7}$ s^{-1}, assuming a single energy barrier model (57). In contrast, studies of Cfr10I via DNA recombination assays found a lifetime of loops of only ~90 s (32). Taken together with our force data, this finding would imply a very different binding energy landscape for DNA looped by Cfr10I than exists for DNA bound in nucleosomes.

Cfr10I, EcoRII and FokI are the only two-site enzymes for which association constants have been reported ($\sim 10^9$ M^{-1}, $\sim 2 \times 10^8$ M^{-1} and $\sim 10^9$ M^{-1}, respectively) (58–60). We measured a bimodal force distribution with FokI and only the higher of the two peak forces was comparable to the single peak force observed for Cfr10I (Figure 5), whereas EcoRII exhibited an even higher disruption force. Crystal structure data indicates that FokI has a small dimerization interface and, on this basis, it was predicted to have relatively weak association compared with other REases (61). Our measurements support this prediction as FokI had the lowest average disruption force (~15 pN). Another study published very recently provides further evidence that the inherent FokI monomer–dimer association in solution is very weak (K_D ~100 nM) but that DNA bound monomers have enhanced protein–protein association due to the fact that the DNA tethers them in close proximity (62).

We analyzed the unbinding force data for correlations with enzyme properties. One might expect larger proteins or those that recognize a longer sequence to bind more strongly. Little correlation was observed with molecular weight (correlation coefficient $= -0.04$, $P = 0.5$) but a positive correlation (coefficient $= 0.5$, $P = 0.08$) was observed between average force and recognition sequence length. This correlation was not universal, however. For example, although BspMI forms a tetramer that recognizes a 6 bp sequence and Sau3AI forms a dimer that recognizes only a 4 bp sequence, these two complexes exhibited similar unlooping forces.

We also analyzed individual datasets for correlations between unbinding force and loop size. Due to the bending rigidity of DNA, loops store mechanical stress and one might expect this stress to accelerate the opening of loops. When a molecule is stretched, however, the applied force could act in a direction to relieve this stress (for an antiparallel loop, as in Figure 1), or to increase it (for a parallel loop, which would be squeezed). Interestingly, for loops

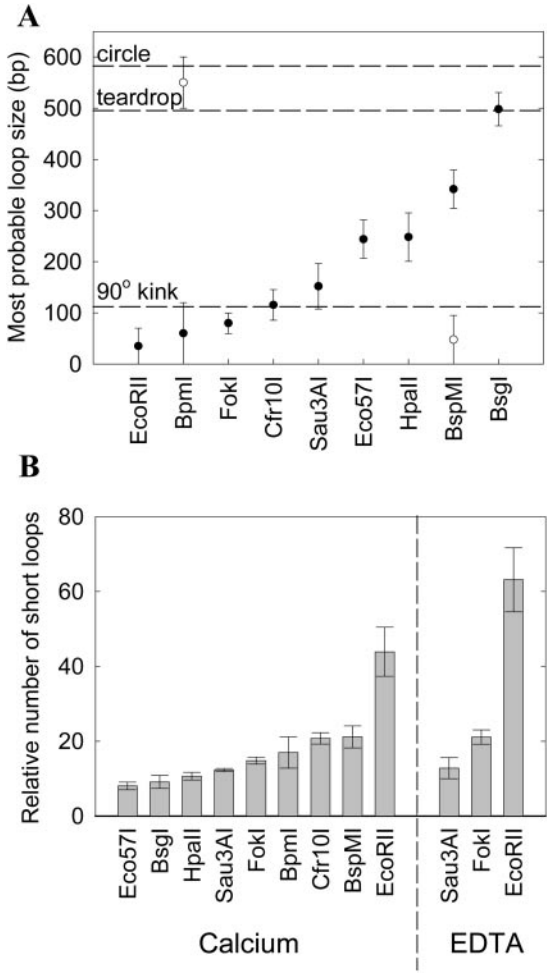

Figure 10. Data on loop sizes. (**A**) Most probable loop sizes for each enzyme. For those distributions which had two discernible peaks, the smaller peak is plotted as an open circle. The dashed lines in this plot indicate the most probable length predicted theoretically by the circle, teardrop and 90° kink models in Ref. (19). The error bars on EcoRII and BpmI extend to zero size because peaks occurred in the first bin of the histograms. In these cases our resolution was not limited by the resolution of our instrument, but by the available sites on the DNA. (**B**) Relative tendencies for short loop formation by different enzymes. The relative number of short loops was calculated as the number of loops shorter than one persistence length per number possible divided by total number observed per total possible.

shorter than one persistence length we found a negative correlation between unbinding force and loop size with eight out of nine enzymes (Table 2), suggesting that parallel loop geometries are favored.

Bimodal force distributions

Force distributions for FokI, Sau3AI, BspMI and BsgI were bimodal. Bimodal distributions have been reported previously in the unbinding of lipid molecules from a bilayer (63) and in the unbinding of RAN protein from nuclear importin receptor B1 (64). To our knowledge our findings constitute the first observations of bimodal unbinding force distributions for protein–DNA complexes. We note that the average sizes of loops within the two peaks in the force distribution were not significantly different. One possible explanation for bimodality is molecular heterogeneity. Individual complexes may have different binding modes with varying binding strengths, resulting from different loop geometries, binding surfaces or coordinated ions. Variable loop geometries have previously been postulated for the two-site REase SfiI based on varying mobilities observed in gel shift assays (33). When the recognition sequence is non-palindromic different loop geometries could arise due to different site orientations (i.e. repeating versus inverted). As many of these enzymes form dimers or pseudo-dimers, the weaker and stronger events could correspond to binding of complexes at one versus two surfaces. Heterogeneity could also stem from variable protein–DNA interactions. Some REases, such as NaeI, have been reported to contain two different DNA binding clefts (65) while others, such as SgrAI, have been reported to bind at secondary sites differing slightly in sequence (47). In addition, some Type IIE enzymes may have two effector sites in addition to a catalytic site and could possibly form loops between any two of these sites (66). Such cases would be difficult to discern in any experiment that measures binding or looping when cleavage is inhibited. However, none of these effects have been reported in the cases of FokI, Sau3AI, BspMI and BsgI. Alternatively, one need not postulate molecular heterogeneity to explain a bimodal force distribution. Recent theoretical calculations indicate that a bimodal distribution can arise during forced disruption of homogeneous complexes if the binding energy landscape has multiple energy barriers (49).

Dependence on ionic conditions

Most Type II REases require divalent cations as a cofactor. In certain cases the dependence is known to be dramatic. For example, the dissociation constant of the one-site REase PvuII is reported to be ~60 000 times lower in 10 mM Ca^{2+} than in 1 mM EDTA (67). Studies of the two-site enzyme SfiI by recombination assays showed that the looped complex had a lifetime of >7 h in Ca^{2+} versus only ~4 min in EDTA (32). Here, we report stable looping with EcoRII, FokI and Sau3AI in the absence of divalent ions. Previous studies of EcoRII and Sau3AI by electron microscopy did not report on whether Ca^{2+} was required for loop formation. A recent study of FokI complexes by analytical ultracentrifugation concluded that FokI dimerizes only in the presence of divalent cations (43). We observed clear looping with FokI in EDTA although these complexes exhibited the lowest average disruption force among all of the REases we studied. Apparently the looped complex we detected is not stable enough to be detected by analytical ultracentrifugation with the concentrations of FokI and DNA chosen. An advantage of our method is that it allows loop formation to be probed on short time scales.

Differences in frequency of loops

Significant differences in the frequency of looping with different enzymes are clearly evident (Table 2). For example, one may compare the results on EcoRII and Eco57I. Both were used at the same concentration and incubated with the DNA for the same amount of time, yet the frequency of looping was nearly an order of magnitude higher with Eco57I than with EcoRII. Surprisingly, however, a lower unbinding force was measured with Eco57I than with EcoRII (Table 2).

The major difference between these two enzymes was that Eco57I exhibited a significantly larger optimum loop size than EcoRII (~250 versus <75 bp) (Figure 10), suggesting an inhibitory effect of DNA bending rigidity. On the other hand, this difference in looping frequency might well be explained by a recent report by Tamulaitis *et al*. (68) that EcoRII requires interaction at three sites for full activity.

Overall, no universal trends were observed across different enzymes as there were no statistically significant correlations between frequency of looping and optimum loop size (correlation coefficient = 0.21, $P = 0.3$), between frequency and average unbinding force (coefficient = 0.37, $P = 0.16$), or between loop size and unbinding force (coefficient = 0.07, $P = 0.43$). These results further support the conclusion that loop formation is strongly dictated by the structural specifics of each complex.

Comparison of loop sizes with theories

Many studies of DNA mechanics have found good agreement with the predictions of the WLC model. For example, measurements of the elasticity of single DNA molecules are in excellent agreement with this model and indicate a persistence length of ~150 bp (69–72). In early work, Stockmayer and Yamakawa (73) calculated that a 'teardrop' shape, corresponding to a loop opening angle of ~81°, would minimize bending energy in the WLC model. Shimada and Yamakawa (13) subsequently considered thermal fluctuations and derived an expression for the dependence of DNA cyclization probability on molecular length. Although there are clearly differences between cyclization and looping, it has been proposed by a number of authors that the Shimada–Yamakawa expression may be used as an approximation for modeling DNA looping (14,19,20). Modifications to this model have also been proposed to account for the effect of the span of the protein (15–17), DNA tension (19,35), different elastic potentials (25) and loop rearrangement entropy (18). More detailed models have also been developed that can account for sequence-dependent DNA bending and flexibility and the geometry and flexibility of the protein complex (22,74).

Short-distance looping by some protein complexes, including the two-site REase SfiI, has been shown to be modulated due to the helical pitch of DNA (2,54). This effect has been modeled in some theoretical studies and causes the activity to be modulated up and down as the site separation is changed by ~5 bp (half the helical pitch) (13,22). Our measurements, however, examine the overall dependence on length, which is the primary dependence addressed by most of the theoretical studies, rather than this fine-scale modulation. Our template does not have site separations corresponding to every multiple of 5 bp. Thus, in binning our histograms at ~100 to 150 bp the effect of the helical modulation is expected to average out within these bins and therefore not strongly influence these distributions.

While the WLC model has been successfully applied in many cases, a recent study surprisingly reported much faster cyclization of short DNA molecules (<100 bp) than was anticipated based on this model (23,24). In an attempt to explain these findings, a number of new theoretical models were proposed that allow for spontaneous sharp kinks in DNA (25,75,76). However, based on further experiments and calculations some investigators have concluded that such spontaneous kinks are very unlikely (25). On the other hand, DNA kinking may be quite relevant in the case of protein-mediated DNA looping as many proteins, including some REases such as EcoRV, induce sharp bends in DNA (77). Several recent theoretical studies have considered the effect of such kinks on DNA looping (17,19,21).

In almost all cases the measured loop size distributions were in much better agreement with models that assume sharp DNA kinking than with models that assume classical DNA elasticity (Figures 9 and 10, and Supplementary Data). The Shimada–Yamakawa model predicts an optimum size of ~500 bp and very few loops <200 bp in size, whereas most of the experimental distributions indicated optimum values of <200 bp. A number of possible effects besides kinking could be considered in an attempt to reconcile the loop sizes with classical theory. First, the DNA persistence length could be shorter than the often-assumed value of 150 bp. Values as low as ~120 bp have been reported in solutions containing divalent cations, as used in our experiments (71). Second, the protein complex has a finite span, which can reduce the necessary bending of the DNA. Third, loop rearrangement entropy is predicted to cause compression of loops (18). Only in the cases of BsgI and BspMI (which exhibited optimum sizes of ~500 and ~330 bp, respectively) was the data in better agreement with classical WLC models including these corrections than with models that postulate DNA kinks. Even in these two cases there were a significant fraction of loops that were shorter than predicted, and with the seven other enzymes these corrections were insufficient to reconcile the data with the classical WLC theory.

On the other hand, the span of the protein complex leads to an underestimation in the measured loop sizes. When a loop is disrupted we measure the change in length ΔL of the DNA but this does not exactly correspond to the distance between the two sites. The initial length of the tethered complex is given by $L_{\text{initial}} = (L_{\text{DNA}} - L_{\text{loop}}) + r$, where L_{DNA} is the total length of the DNA, L_{loop} is the length of DNA inside the loop, and r is the span of the protein. Since the final length is $L_{\text{final}} = L_{\text{DNA}}$, the measured length change is $\Delta L = L_{\text{loop}} - r$. Thus, when comparing measured and predicted 'loop sizes,' one must consider the magnitude of r. Based on the molecular weights of the enzyme complexes we estimate that r is ~5 to 20 bp. Thus, the measured length changes most likely underestimate the loop lengths by approximately this amount. In the context of the comparison between theory and experiment, however, this effect is not of sufficient magnitude to reconcile the small loops we observe with the classical models.

Another experimental consideration is the effect of tension in the DNA. Two groups have recently modeled this effect theoretically and a shift to smaller loop size is predicted as the tension is increased. In our experiment the DNA was held an extension of 35%, which corresponds to a small tension of ~0.06 pN. However, theoretical calculations predict that this tension would have negligible effect on the shape of the distribution of loop sizes. Therefore, this effect also cannot reconcile our findings with the classical models.

In most cases (with BpmI, Cfr10I, Eco57I, EcoRII, FokI, HpaII and Sau3AI) the measured loop size distributions were in better agreement with models that assume sharp

DNA kinking. The model that fit many datasets best over the wide range from 0 to 1000 bp was the 90° kink model of Sankararaman and Marko (19). However, the expressions proposed by Rippe for loops with 30–120° kinks and with a 10 nm protein span fit considerably better over their range of validity (>150 bp) in many cases (17). While some models fit better than others, we emphasize that none fit perfectly and none could account for the excess of short loops observed in some cases (e.g. with EcoRII and FokI). Furthermore, secondary peaks in the loop size distribution were observed with BpmI and BspMI, suggesting heterogeneous loop conformations.

Yan *et al*. (75) have considered the possibility that short loops may be facilitated by kink formation in DNA by 'thermally activated hinges' due to localized strand-separation. However, such events were predicted to be rare, with an average separation >1 kb, and are thus unlikely to occur very often between closely spaced sites in our experiment. Further, on the basis of cyclization experiments and Monte Carlo simulations, Du *et al*. (25) conclude that the probability for spontaneous kinking of a 100 bp segment is <0.02%. In accord with these predictions, we find that models postulating spontaneous kinks do not agree with our data as well as models that postulate permanent kinks.

Protein binding is known in certain cases to induce sharp kinks in DNA, but in our experiment induced kinking inside a loop could only occur if additional binding sites existed between the two sites in question, and this is not very likely for sites which are closely spaced. Of the enzymes studied, FokI is the only one for which a structure of the DNA–protein complex has been determined and this shows that FokI does not induce significant bending of DNA (78). Thus, at least in this case we can seemingly rule out protein-induced DNA kinking as a mechanism for short loop formation. Crystal structures of BfiI, Crf10I and EcoRII have also been determined, but only for the free proteins (66,79,80).

In a case where induced kinks do occur, we imagine that such kinks at the closure point of the loop, rather than at the apex, could have a similar effect in facilitating short loops. Additionally, to explain very small looping events (e.g. of ~20 bp) we imagine that the DNA may be wrapped across the surface of the protein complex, akin to how DNA is wrapped in the nucleosome, rather than looping freely through the solution. It has been pointed out that flexibility of the protein complex probably plays an important role in facilitating the formation of such short loops (1,22,74).

Comparison of loop sizes with results of previous studies

Of the nine REases studied here, dependence of loop formation in linear DNA on loop size has only been studied for EcoRII and BspMI. The most extensive comparisons can be made with previous results on EcoRII. Reuter *et al*. (52,53) studied the dependence of cleavage rate on distance for site separations of 5, 10, 21, 31, 63, 73, 191 and 952 bp. Remarkably, they found that highest activity occurred on the 10 bp template and reported no activity with separations greater than 1000 bp. In our experiment the DNA template allowed for loops ranging in size from ~30 to 16 000 bp, and we measured loop sizes from 21 to 6940 bp with an optimum size smaller than 70 bp. This finding is consistent with that reported by Reuter *et al*. although with this enzyme our resolution was more limited due to a lack of short separations on our DNA template. On the other end of the spectrum, however, we detected large loops up to ~7 kb, significantly longer than that reported previously.

We may also compare our results with recent findings on BspMI. Looping in linear DNA has been studied using magnetic tweezers and loop sizes ranging from ~90 to 1500 bp were reported (35). However, statistics on loop sizes were not reported in this previous study. Here we report the distribution of loops, ranging from ~25 to 4500 bp, and report that the optimum size is ~340 ± 40 bp (Figure 10A). We also observed a second peak in the size distribution at <100 bp, suggesting an alternate binding mode.

Our measurements of loop size distributions with the seven other enzymes (BpmI, BsgI, Cfr10I, Eco57I, FokI, HpaII, Sau3AI) constitute novel results as such information has not been previously reported for these enzymes. Overall, our measurements reveal a large variability dependent on the protein. These findings clearly indicate that one may not fully understand protein-mediated DNA looping by considering only DNA mechanics. As emphasized by Zhang *et al*. (22) detailed models will have to consider specific structural details, such as loop geometry, sequence-dependent DNA bending, protein-induced DNA bending, protein span and protein elasticity.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank K. Hailey, K. Haushalter, A. Rajkumar, and R. Sim for assistance. This research was supported by a Burroughs Wellcome Fund Career Award, a Searle Scholars Award from the Kinship Foundation, and a Young Investigator Award from the Arnold and Mabel Beckman Foundation. Funding to pay the Open Access publication charges for this article was provided by University of California, San Diego.

Conflict of interest statement. None declared.

REFERENCES

1. Schleif,R. (1992) DNA looping. *Annu. Rev. Biochem.*, **61**, 199–223.
2. Dunn,T.M., Hahn,S., Ogden,S. and Schleif,R.F. (1984) An operator at −280 base pairs that is required for repression of araBAD operon promoter: addition of DNA helical turns between the operator and promoter cyclically hinders repression. *Proc. Natl Acad. Sci. USA*, **81**, 5017–5020.
3. Mukherjee,S., Erickson,H. and Bastia,D. (1988) Enhancer origin interaction in plasmid-R6K involves a DNA loop mediated by initiator protein. *Cell*, **52**, 375–383.
4. Su,W., Middleton,T., Sugden,B. and Echols,H. (1991) DNA looping between the origin of replication of Epstein–Barr virus and its enhancer site: stabilization of an origin complex with Epstein–Barr nuclear antigen 1. *Proc. Natl Acad. Sci. USA*, **88**, 10870–10874.
5. Oehler,S., Amouyal,M., Kolkhof,P., von Wilcken-Bergmann,B. and Muller-Hill,B. (1994) Quality and position of the three lac operators of *E.coli* define efficiency of repression. *EMBO J.*, **13**, 3348–3355.
6. Weiner,B.M. and Kleckner,N. (1994) Chromosome pairing via multiple interstitial interactions before and during meiosis in yeast. *Cell*, **77**, 977–991.

7. Blackwood,E.M. and Kadonaga,J.T. (1998) Going the distance: a current view of enhancer action. *Science*, **281**, 60–63.
8. Chen,Y. and Rice,P.A. (2003) New insight into site-specific recombination from Flp recombinase-DNA structures. *Annu. Rev. Biophys. Biomol. Struct.*, **32**, 135–159.
9. Vilar,J.M. and Leibler,S. (2003) DNA looping and physical constraints on transcription regulation. *J. Mol. Biol.*, **331**, 981–989.
10. Buchler,N.E., Gerland,U. and Hwa,T. (2003) On schemes of combinatorial transcription logic. *Proc. Natl Acad. Sci. USA*, **100**, 5136–5141.
11. Bintu,L., Buchler,N.E., Garcia,H.G., Gerland,U., Hwa,T., Kondev,J. and Phillips,R. (2005) Transcriptional regulation by the numbers: models. *Curr. Opin. Genet. Dev.*, **15**, 116–124.
12. Ogata,K., Sato,K. and Tahirov,T.H. (2003) Eukaryotic transcriptional regulatory complexes: cooperativity from near and afar. *Curr. Opin. Struct. Biol.*, **13**, 40–48.
13. Shimada,J. and Yamakawa,H. (1984) Ring-closure probabilities for twisted wormlike chains - application to DNA. *Macromolecules*, **17**, 689–698.
14. Rippe,K., von Hippel,P.H. and Langowski,J. (1995) Action at a distance: DNA-looping and initiation of transcription. *Trends Biochem. Sci.*, **20**, 500–506.
15. Merlitz,H., Rippe,K., Klenin,K.V. and Langowski,J. (1998) Looping dynamics of linear DNA molecules and the effect of DNA curvature: a study by Brownian dynamics simulation. *Biophys. J.*, **74**, 773–779.
16. Ringrose,L., Chabanis,S., Angrand,P.O., Woodroofe,C. and Stewart,A.F. (1999) Quantitative comparison of DNA looping *in vitro* and *in vivo*: chromatin increases effective DNA flexibility at short distances. *EMBO J.*, **18**, 6630–6641.
17. Rippe,K. (2001) Making contacts on a nucleic acid polymer. *Trends Biochem. Sci.*, **26**, 733–740.
18. Sankararaman,S. and Marko,J.F. (2005) Entropic compression of interacting DNA loops. *Phys. Rev. Lett.*, **95**, 078104.
19. Sankararaman,S. and Marko,J.F. (2005) Formation of loops in DNA under tension. *Phys. Rev. E.*, **71**, 021911.
20. Blumberg,S., Tkachenko,A.V. and Meiners,J.C. (2005) Disruption of protein-mediated DNA looping by tension in the substrate DNA. *Biophys. J.*, **88**, 1692–1701.
21. Yan,J., Kawamura,R. and Marko,J.F. (2005) Statistics of loop formation along double helix DNAs. *Phys. Rev. E.*, **71**, 061905.
22. Zhang,Y., McEwen,A.E., Crothers,D.M. and Levene,S.D. (2006) Statistical-mechanical theory of DNA looping. *Biophys. J.*, **90**, 1903–1912.
23. Cloutier,T.E. and Widom,J. (2004) Spontaneous sharp bending of double-stranded DNA. *Mol. Cell*, **14**, 355–362.
24. Cloutier,T.E. and Widom,J. (2005) DNA twisting flexibility and the formation of sharply looped protein–DNA complexes. *Proc. Natl Acad. Sci. USA*, **102**, 3645–3650.
25. Du,Q., Smith,C., Shiffeldrim,N., Vologodskaia,M. and Vologodskii,A. (2005) Cyclization of short DNA fragments and bending fluctuations of the double helix. *Proc. Natl Acad. Sci. USA*, **102**, 5397–5402.
26. Halford,S.E., Welsh,A.J. and Szczelkun,M.D. (2004) Enzyme-mediated DNA looping. *Annu. Rev. Biophys. Biomol. Struct.*, **33**, 1–24.
27. Topal,M.D., Thresher,R.J., Conrad,M. and Griffith,J. (1991) NaeI endonuclease binding to pBR322 DNA induces looping. *Biochemistry*, **30**, 2006–2010.
28. Siksnys,V., Skirgaila,R., Sasnauskas,G., Urbanke,C., Cherny,D., Grazulis,S. and Huber,R. (1999) The Cfr10I restriction enzyme is functional as a tetramer. *J. Mol. Biol.*, **291**, 1105–1118.
29. Mucke,M., Lurz,R., Mackeldanz,P., Behlke,J., Kruger,D.H. and Reuter,M. (2000) Imaging DNA loops induced by restriction endonuclease EcoRII. A single amino acid substitution uncouples target recognition from cooperative DNA interaction and cleavage. *J. Biol. Chem.*, **275**, 30631–30637.
30. Friedhoff,P., Lurz,R., Luder,G. and Pingoud,A. (2001) Sau3AI, a monomeric type II restriction endonuclease that dimerizes on the DNA and thereby induces DNA loops. *J. Biol. Chem.*, **276**, 23581–23588.
31. Szczelkun,M.D. and Halford,S.E. (1996) Recombination by resolvase to analyse DNA communications by the SfiI restriction endonuclease. *EMBO J.*, **15**, 1460–1469.
32. Milsom,S.E., Halford,S.E., Embleton,M.L. and Szczelkun,M.D. (2001) Analysis of DNA looping interactions by type II restriction enzymes that require two copies of their recognition sites. *J. Mol. Biol.*, **311**, 515–527.
33. Watson,M.A., Gowers,D.M. and Halford,S.E. (2000) Alternative geometries of DNA looping: an analysis using the SfiI endonuclease. *J. Mol. Biol.*, **298**, 461–475.
34. Katiliene,Z., Katilius,E. and Woodbury,N.W. (2003) Single molecule detection of DNA looping by NgoMIV restriction endonuclease. *Biophys. J.*, **84**, 4053–4061.
35. Yan,J., Skoko,D. and Marko,J.F. (2004) Near-field-magnetic-tweezer manipulation of single DNA molecules. *Phys. Rev. E*, **70**, 011905.
36. van den Broek,B., Vanzi,F., Normanno,D., Pavone,F.S. and Wuite,G.J. (2006) Real-time observation of DNA looping dynamics of Type IIE restriction enzymes NaeI and NarI. *Nucleic Acids Res.*, **34**, 167–174.
37. Kruger,D.H., Barcak,G.J., Reuter,M. and Smith,H.O. (1988) EcoRII can be activated to cleave refractory DNA recognition sites. *Nucleic Acids Res.*, **16**, 3997–4008.
38. Bath,A.J., Milsom,S.E., Gormley,N.A. and Halford,S.E. (2002) Many type IIs restriction endonucleases interact with two recognition sites before cleaving DNA. *J. Biol. Chem.*, **277**, 4024–4033.
39. Soundararajan,M., Chang,Z., Morgan,R.D., Heslop,P. and Connolly,B.A. (2002) DNA binding and recognition by the IIs restriction endonuclease MboII. *J. Biol. Chem.*, **277**, 887–895.
40. Gowers,D.M., Bellamy,S.R. and Halford,S.E. (2004) One recognition sequence, seven restriction enzymes, five reaction mechanisms. *Nucleic Acids Res.*, **32**, 3469–3479.
41. Reuter,M., Kupper,D., Pein,C.D., Petrusyte,M., Siksnys,V., Frey,B. and Kruger,D.H. (1993) Use of specific oligonucleotide duplexes to stimulate cleavage of refractory DNA sites by restriction endonucleases. *Anal. Biochem.*, **209**, 232–237.
42. Kupper,D., Moncke-Buchner,E., Reuter,M. and Kruger,D.H. (1999) Oligonucleotide stimulators allow complete cleavage of agarose-embedded DNA by particular type II restriction endonucleases. *Anal. Biochem.*, **272**, 275–277.
43. Vanamee,E.S., Santagata,S. and Aggarwal,A.K. (2001) FokI requires two specific DNA sites for cleavage. *J. Mol. Biol.*, **309**, 69–78.
44. Roberts,R.J., Belfort,M., Bestor,T., Bhagwat,A.S., Bickle,T.A., Bitinaite,J., Blumenthal,R.M., Degtyarev,S., Dryden,D.T., Dybvig,K. *et al.* (2003) A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.*, **31**, 1805–1812.
45. Gemmen,G.J., Sim,R., Haushalter,K.A., Ke,P.C., Kadonaga,J.T. and Smith,D.E. (2005) Forced unraveling of nucleosomes assembled on heterogeneous DNA using core histones, NAP-1, and ACF. *J. Mol. Biol.*, **351**, 89–99.
46. Bustamante,C., Smith,S.B., Liphardt,J. and Smith,D. (2000) Single-molecule studies of DNA mechanics. *Curr. Opin. Struct. Biol.*, **10**, 279–285.
47. Wood,K.M., Daniels,L.E. and Halford,S.E. (2005) Long-range communications between DNA sites by the dimeric restriction endonuclease SgrAI. *J. Mol. Biol.*, **350**, 240–253.
48. Sapranauskas,R., Sasnauskas,G., Lagunavicius,A., Vilkaitis,G., Lubys,A. and Siksnys,V. (2000) Novel subtype of type IIs restriction enzymes. BfiI endonuclease exhibits similarities to the EDTA-resistant nuclease Nuc of *Salmonella typhimurium*. *J. Biol. Chem.*, **275**, 30878–30885.
49. Derenyi,I., Bartolo,D. and Ajdari,A. (2004) Effects of intermediate bound states in dynamic force spectroscopy. *Biophys. J.*, **86**, 1263–1269.
50. Pingoud,A., Fuxreiter,M., Pingoud,V. and Wende,W. (2005) Type II restriction endonucleases: structure and mechanism. *Cell Mol. Life Sci.*, **62**, 685–707.
51. Kingston,I.J., Gormley,N.A. and Halford,S.E. (2003) DNA supercoiling enables the type IIS restriction enzyme BspMI to recognise the relative orientation of two DNA sequences. *Nucleic Acids Res.*, **31**, 5221–5228.
52. Reuter,M., Kupper,D., Meisel,A., Schroeder,C. and Kruger,D.H. (1998) Cooperative binding properties of restriction endonuclease EcoRII with DNA recognition sites. *J. Biol. Chem.*, **273**, 8294–8300.
53. Kruger,D.H., Kupper,D., Meisel,A., Tierlich,M., Reuter,M. and Schroeder,C. (1995) Restriction endonucleases functionally interacting with two DNA sites. *Gene*, **157**, 165.
54. Wentzell,L.M. and Halford,S.E. (1998) DNA looping by the Sfi I restriction endonuclease. *J. Mol. Biol.*, **281**, 433–444.

55. Bilcock,D.T., Daniels,L.E., Bath,A.J. and Halford,S.E. (1999) Reactions of type II restriction endonucleases with 8-base pair recognition sites. *J. Biol. Chem.*, **274**, 36379–36386.
56. Evans,E. (2001) Probing the relation between force--lifetime--and chemistry in single molecular bonds. *Annu. Rev. Biophys. Biomol. Struct.*, **30**, 105–128.
57. Brower-Toland,B.D., Smith,C.L., Yeh,R.C., Lis,J.T., Peterson,C.L. and Wang,M.D. (2002) Mechanical disruption of individual nucleosomes reveals a reversible multistage release of DNA. *Proc. Natl Acad. Sci. USA*, **99**, 1960–1965.
58. Waugh,D.S. and Sauer,R.T. (1993) Single amino acid substitutions uncouple the DNA binding and strand scission activities of Fok I endonuclease. *Proc. Natl Acad. Sci. USA*, **90**, 9596–9600.
59. Skirgaila,R. and Siksnys,V. (1998) Ca2+-ions stimulate DNA binding specificity of Cfr10I restriction enzyme. *Biol. Chem.*, **379**, 595–598.
60. Reuter,M., Schneider-Mergener,J., Kupper,D., Meisel,A., Mackeldanz,P., Kruger,D.H. and Schroeder,C. (1999) Regions of endonuclease EcoRII involved in DNA target recognition identified by membrane-bound peptide repertoires. *J. Biol. Chem.*, **274**, 5213–5221.
61. Wah,D.A., Bitinaite,J., Schildkraut,I. and Aggarwal,A.K. (1998) Structure of FokI has implications for DNA cleavage. *Proc. Natl Acad. Sci. USA*, **95**, 10564–10569.
62. Catto,L.E., Ganguly,S., Milsom,S.E., Welsh,A.J. and Halford,S.E. (2006) Protein assembly and DNA looping by the FokI restriction endonuclease. *Nucleic Acids Res.*, **34**, 1711–1720.
63. Evans,E. and Williams,P. (2004) *Dynamic Force Spectroscopy In Physics of Bio-Molecules and Cells, Ecoles des HOUCHES d Ete LXXV. EDP Sciences - Springer-Verlag, pp. 145 - 185*. Springer Verlag, Berlin.
64. Nevo,R., Stroh,C., Kienberger,F., Kaftan,D., Brumfeld,V., Elbaum,M., Reich,Z. and Hinterdorfer,P. (2003) A molecular switch between alternative conformational states in the complex of Ran and importin beta 1. *Nature Struct. Biol.*, **10**, 553–557.
65. Huai,Q., Colandene,J.D., Topal,M.D. and Ke,H.M. (2001) Structure of NaeI-DNA complex reveals dual-mode DNA recognition and complete dimer rearrangement. *Nature Struct. Biol.*, **8**, 665–669.
66. Zhou,X.E., Wang,Y., Reuter,M., Mucke,M., Kruger,D.H., Meehan,E.J. and Chen,L. (2004) Crystal structure of type IIE restriction endonuclease EcoRII reveals an autoinhibition mechanism by a novel effector-binding fold. *J. Mol. Biol.*, **335**, 307–319.
67. Bowen,L.M. and Dupureur,C.M. (2003) Investigation of restriction enzyme cofactor requirements: a relationship between metal ion properties and sequence specificity. *Biochemistry*, **42**, 12643–12653.
68. Tamulaitis,G., Sasnauskas,G., Mucke,M. and Siksnys,V. (2006) Simultaneous binding of three recognition sites is necessary for a concerted plasmid DNA cleavage by EcoRII restriction endonuclease. *J. Mol. Biol.*, **358**, 406-419.
69. Bustamante,C., Marko,J.F., Siggia,E.D. and Smith,S. (1994) Entropic elasticity of lambda-phage DNA. *Science*, **265**, 1599–1600.
70. Vologodskii,A. (1994) DNA extension under the action of an external force. *Macromolecules*, **27**, 5623–5625.
71. Wang,M.D., Yin,H., Landick,R., Gelles,J. and Block,S.M. (1997) Stretching DNA with optical tweezers. *Biophys. J.*, **72**, 1335–1346.
72. Bouchiat,C., Wang,M.D., Allemand,J.F., Strick,T., Block,S.M. and Croquette,V. (1999) Estimating the persistence length of a worm-like chain molecule from force-extension measurements. *Biophys. J.*, **76**, 409–413.
73. Yamakawa,H. and Stockmayer,W. (1972) Statistical mechanics of wormlike chains.2. excluded volume effects. *J. Chem. Phys.*, **57**, 2843.
74. Zhang,Y.L. and Crothers,D.M. (2003) Statistical mechanics of sequence-dependent circular DNA and its application for DNA cyclization. *Biophys. J.*, **84**, 136–153.
75. Yan,J. and Marko,J.F. (2004) Localized single-stranded bubble mechanism for cyclization of short double helix DNA. *Phys. Rev. Lett.*, **93**, 108108.
76. Wiggins,P.A., Phillips,R. and Nelson,P.C. (2005) Exact theory of kinkable elastic polymers. *Phys. Rev. E*, **71**, 021909.
77. Winkler,F.K., Banner,D.W., Oefner,C., Tsernoglou,D., Brown,R.S., Heathman,S.P., Bryan,R.K., Martin,P.D., Petratos,K. and Wilson,K.S. (1993) The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.*, **12**, 1781–1795.
78. Wah,D.A., Hirsch,J.A., Dorner,L.F., Schildkraut,I. and Aggarwal,A.K. (1997) Structure of the multimodular endonuclease FokI bound to DNA. *Nature*, **388**, 97–100.
79. Grazulis,S., Manakova,E., Roessle,M., Bochtler,M., Tamulaitiene,G., Huber,R. and Siksnys,V. (2005) Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc. Natl Acad. Sci. USA*, **102**, 15797–15802.
80. Bozic,D., Grazulis,S., Siksnys,V. and Huber,R. (1996) Crystal structure of *Citrobacter freundii* restriction endonuclease Cfr10I at 2.15 A resolution. *J. Mol. Biol.*, **255**, 176–186.