Supplementary Section to "When a helicase is not a helicase: dsDNA tracking by the motor protein EcoR124I"

Louise K. Stanley, Ralf Seidel, Carsten van der Scheer, Nynke H. Dekker, Mark D. Szczelkun and Cees Dekker

Supplementary Methods

Construction of pRSgap

pRSgap (Supplementary Figure 3B) was constructed from pBluescriptII SK+ (Stratagene) by inserting 4 PCR fragments derived from λ-DNA into the multiple cloning site of the vector. PCR fragment 1 (forward primer: AAAATCTAGAAGTTCAGGAAGCGGTGATGCTG, reverse primer AAAAGAGCTCTTGGGCGGTTGTGTACATCGAC) copies the 4236 to 6137 bp region from λ-DNA. After digestion with the corresponding restriction enzymes it was cloned between the XbaI and the SacI site of the vector. PCR fragment 2 (forward primer: AAAAGAATTCGGTGACCCTTACGCGAATCC, reverse primer AAAATCTAGAGGCTTCAGCGACCTTGTCC) copies the 9043 to 11201 bp region from λ-DNA. After digestion with the corresponding restriction enzymes it was cloned between the EcoRI and the XbaI site of the vector. PCR fragment 3 (forward primer: AAAACTCGAGCGGTCGTGCTGTATTGTCTC, reverse primer AAAAGAATT**CCTCAGC**CAGAGACACAGCC**CCTCAGC**AGAACTTCCATTA TTCTCCTG) copies the 21425 to 22661 bp region from λ-DNA. It contains a single EcoR124I site. The two 20 bp spaced, directly-repeated BbvCI sites for construction of the ssDNA gaps were added to the reverse primer (underlined in bold). After digestion with the corresponding restriction enzymes the PCR product was cloned between the XhoI and the EcoRI site of the vector. PCR fragment 4 (forward primer: AAAAGGTACCAGTTCAGGAAGCGGTGATGCTG, reverse primer AAAACTCGAGCAGCAACCGCAAGAATGC) copies the 4236 to 5925 bp region from λ-DNA (same as fragment 1). After digestion with the corresponding restriction enzymes it was cloned between the KpnI and the XhoI site of the vector.

Production of gapped DNA substrates for magnetic tweezers experiments

pRSgap was digested with XhoI and KpnI. To either end of the ~8 kbp fragment, two ~700 bp PCR fragments were ligated[1], containing: multiple biotin-modified dUTP bases at the XhoI end; and, multiple digoxigenin-modified dUTP bases at the KpnI end. The construct was purified by phenol/chloroform and chloroform extraction followed by ethanol precipitation. To create 20 nt gaps the construct was treated with either Nt- or Nb.BbvCI (Supplementary Figure 3, Supplementary Table 3) for 1 hr at 37 ºC. Added to this was 0.5 vol of a solution containing 0.2 M EDTA, 0.2 M Tris-Cl, pH 8.0 and 2 μM of the oligonucleotide complementary to the DNA stretch to be removed. The samples were incubated for 10 min at 75 ºC and for 10 min at 37 ºC and subsequently purified by phenol/chloroform and chloroform extraction followed by ethanol precipitation. Successful gap production was tested by digesting constructs with BbvCI, which did not produce any detectable digestion products separated in agarose gels. To produce gaps smaller than 20 bp, 1 μM of the relevant HPLC-purified oligonucleotide (Supplementary Table 1) was hybridised into the 20 bp gap in 10 mM Tris-HCl, pH 8.0, 500 mM NaCl by heating the solution to 80 ºC and cooling it slowly at –1 K min^{-1}. Excess oligonucleotide was removed with a spin-column-purification kit (MoBio). BbvCI digestion of the substrates was restored, showing that the gaps were successfully rescued.

Construction of pMDS41 derivatives

pMDS41.1 and pMDS41.21 (Supplementary Figure 3B) were constructed from pMDS23 (ref. 2) by inserting into the NcoI site in one or other orientation a 73 bp DNA segment containing a pair of directly repeated BbvCI sites.

Production of Gapped DNA substrates for triplex displacement

CsCl-purified pMDS41 constructs were treated with Nt- or Nb.BbvCI (Supplementary Figure 3, Supplementary Tables 2 & 3) for 40 min at 37 ºC. Product DNA were initially purified by phenol/chloroform and chloroform extraction followed by isopropanol precipitation. The DNA were then re-suspended in NEB1 [10 mM Bis Tris Propane-Cl, 10 mM $MgCl_2$, 1 mM DTT (pH 7.0)] containing a 50-fold molar excess of an oligonucleotide complementary to the DNA stretch to be removed. 40 μl samples were heated to 85 ºC for 5 minutes in a thermal cycler, cooled at maximum

ramp to 20 ºC, and then transferred rapidly onto ice. Excess oligonucleotide was removed on a 1 ml MonoQ column (GE BioScience) by elution with a 20 ml linear gradient of NaCl (0.7 – 1.0 M). Plasmid-length DNA samples were dialysed and precipitated. Non-gapped DNA was removed by cleavage with EcoRV which cuts only dsDNA in the region between the BbvCI sites. On non-gapped substrates this separates the EcoR124I binding site from the triplex site so the triplex displacement signal only comes from full-length gapped DNA. Linear substrates were generated by cleavage with HindIII. To generate substrates with shorter gaps or modifications, a 4-fold molar excess of the relevant HPLC-purified oligonucleotide (Supplementary Tables 2 & 3) was incubated with the gapped DNA in NEB1, heated to 80 ºC for 5 minutes, and allowed to cool slowly to ambient temperature. Excess oligonucleotide was removed by ultrafiltration (100 kDa cut-off, Amicon MicroSpin). Success of the annealing was checked by Type II restriction digest.

To generate substrates with long gaps, 21 nM DNA nicked with the BbvCI mutants (see above) was treated with 200 U of exonuclease III (Promega) in the supplied reaction buffer, at 37 ºC for 20 s (150 nt gaps) or 45 s (300 nt gaps). Because of the polarity of Exonuclease III, the gaps on the 5′-3′ strand will be closer to the EcoR124I binding site than those on the opposite strand. The DNA was then purified and linearised as above. Type II restriction mapping was used to estimate the size of the gaps to within ± 50 nt (data not shown).

Construction of DNA substrates for crosslinking studies

For the bulk solution crosslinking studies, pLKS3 (Figure 2A) was generated from pNEO (Pharmacia) by cloning an oligoduplex containing an EcoR124I binding site into the HindIII site. For the single molecule crosslinking studies (Figure 2C), pSFVI (Invitrogen) carrying a single EcoR124I site was cleaved with SpeI and XhoI. To either end of the ~8 kbp fragment, two ~700 bp PCR fragments were ligated containing: multiple biotin-modified dUTP bases at the XhoI end; and, multiple digoxigenin-modified dUTP bases at the SpeI end (ref. 1). The construct was gel purified and crosslinked as described below.

Determining the number of ICLs per DNA molecule

To check the efficiency of linear DNA crosslinking, aliquots (~500 ng DNA) were removed from crosslinking reactions and DMSO added to a final concentration of

30% (v/v). The samples were heated to 90 ºC for 2 minutes and then rapidly cooled to 4 ºC. This treatment denatures the non-crosslinked DNA into separate strands whilst the crosslinked DNA remains as dsDNA. 0.5 volumes of STEB were then added to each sample and the DNA separated by agarose gel electrophoresis in TAE buffer (40 mM Tris-acetate, 1 mM EDTA, pH 8.0). On the basis of the difference in electrophoretic mobility between ssDNA and dsDNA, we could then estimate the fraction of crosslinked DNA (f_{ds}) from the amount of dsDNA relative to ssDNA. For each DNA substrate and HTMP concentration used here, we increased the UV irradiation time from t=0 until the DNA was fully crosslinked ($f_{ds} \rightarrow 1$). An exponential relationship between f_{ds} and irradiation time was observed in each case (data not shown). The average number of crosslinks per DNA molecule was then estimated assuming a random Poisson distribution of HTMP by $\ln[(1-f_{ds})^{-1}]$ (for $f_{ds} < 0.9$). From these values a linear relationship between irradiation time and the average number of ICLs per DNA could be obtained (data not shown). By extrapolating this linear relationship to infinite time (and thus an infinite number of ICLs), we could then estimate the irradiation time required to obtain a particular number of ICLs. For the circular DNA, we used the same conditions as calculated for the linear equivalent.

DNA unwinding and cleavage measurements

RecBCD unwinding reactions were carried out at 37 ºC for 30 minutes in 30 µl volumes containing 5 nM DNA, 0.4 nM RecBCD and 2 µM SSB in Unwinding Buffer (25 mM Tris-acetate, 3 mM $MgCl_2$, 1 mM Ca^{2+}, 1 mM DTT, 5 mM ATP, pH 8.0). Reactions were stopped by the addition of 10 µl of DMSO followed by 10 µl of 0.1 M EDTA (pH 8.0), 2.5% (w/v) SDS, 40% (v/v) glycerol and 0.4% (w/v) bromophenol blue. EcoR124I cleavage reactions on 5 nM crosslinked DNA were carried out at 37 ºC for 30 minutes in 30 µl volumes containing 5 nM DNA, 30 nM MTase, 90 nM HsdR, 100 µM *S*-adenosyl methionine and 4 mM ATP in Buffer R. Reactions were stopped by the addition of 15 µl of STEB [0.1 M EDTA, 0.1 M Tris-Cl (pH 8.0), 40% (w/v) sucrose, 0.4% (w/v) bromophenol blue]. EcoR124I cleavage reaction timecourses were carried out in 210 µl volumes of Buffer R containing 5 nM DNA, 30 nM MTase, 90 nM HsdR, and either: (for the gapped DNA at 20 ºC), 25 µM ATP, 1.5 U/ml creatine phosphokinase, 14.5 mM phosphocreatine and 100 µM *S*-adenosyl methionine; or, (for the crosslinked DNA at 37 ºC), 4 mM ATP and 100 µM

S-adenosyl methionine. Aliquots (20 µl) were removed at the timepoints indicated and were stopped by the addition of 10 µl of STEB. In all cases DNA substrates and products were separated by agarose gel electrophoresis.

Deriving the probabilities for bypass, stalling and dissociation from the single-molecule time traces

In order to calculate the probabilities for bypass, stalling and dissociation we first defined the position of the gap as a range of translocation distances based on the position of stalling events and dissociation events at the gap (see peaks in the histograms for the translocated distance, Supplementary Fig. 4). Typically a range of 1100 to 1400 bp was taken for the gap position. Based on this, all translocation events that were longer than 1400 bp were considered to bypass the gap and all translocation events that terminated between 1100 and 1400 bp were classified as having either dissociated or stalled, depending if visible stalling could be observed or not. The number of dissociation events was corrected for dissociation events that would have anyway occurred in the absence of the gap. The number of those "naturally" occurring, i.e. not gap induced, dissociation events was estimated from the mean event time of the translocation events. This was done using maximum likelihood estimation as published before (ref. 1), where the number of events shorter than the Brownian noise was determined. The individual probabilities were then calculated by dividing the number of bypass, stalling or the corrected number of dissociation events by the sum of the three values. For probabilities larger than zero the standard error was calculated from the standard error of a Poissonian distribution ($\Delta N = \sqrt{N}$). For probabilities (P) equal to zero, the error was calculated by $\Delta P = 1 - (1/3)^{1/N}$, where N is the total number of measured events.

Supplementary Discussion

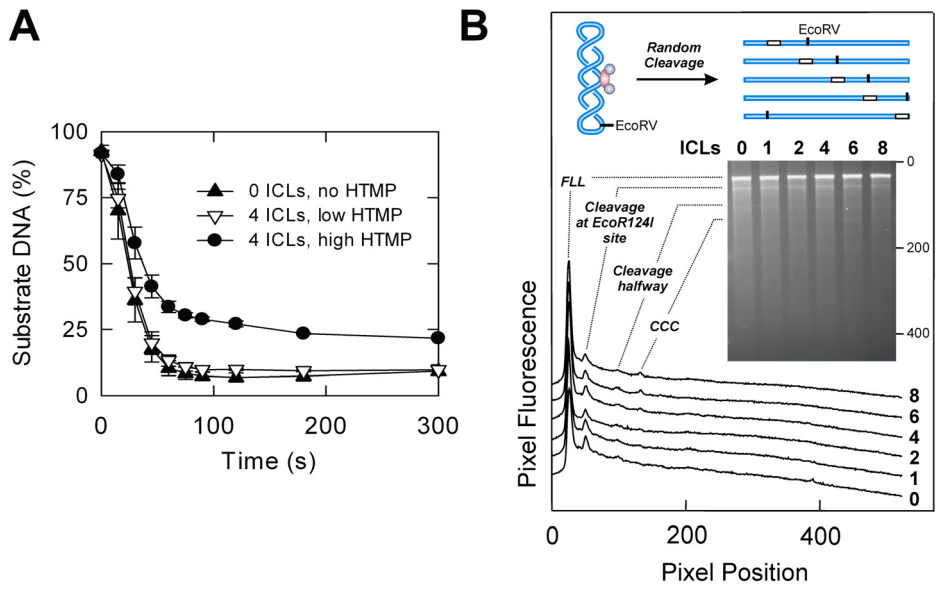
On the force dependence of protein stepping versus ssDNA looping

Within a model in which an enzyme steps over a gap or adduct, the protein has to make a conformational change/step which equals the size of the gap (s) against the load (i.e., the applied force, F). It therefore has to make an energetic transition from the beginning to the end of the gap over an energetic barrier. This maximum distance between the beginning of the gap and the barrier is s. Using a force-dependent Arrhenius equation, one can estimate the upper limit of the force dependency of the gap bypass rate (k_{over}) within a stepping model as (ref. 3): $k_{over}(F) = k_{over}(F = 0)\exp(-F \cdot s/k_B T)$. With F = 1.6 pN, s = 20 nt · 0.2 nm/nt (estimated extension of ssDNA at that force) one gets a reduction of ~5-fold in $k_{over}(F = 1.6$ pN) compared to $k_{over}(F = 0)$. Even assuming 0.3 nm/nt one gets a reduction of only 11-fold, but certainly not a value of >90-fold as observed in the tweezers. Furthermore, at 0.4 pN the reduction in k_{over} within the stepping model would drop to below 2-fold (still assuming 0.3 nm/nt, although this would be now much shorter). However, from our measurements a greater reduction (~10-fold) is still obtained. As the gap bypass probability in the bulk experiments is affected by the gap to an even lesser degree, the large reduction in the successful bypass events seen in the single-molecule experiments must be due exclusively to the applied force. Our results are not consistent however with gap bypass by protein stepping/conformational changes. They agree much better with a ssDNA looping model (Figure 7, main section), for which a force-induced reduction of ssDNA looping of up to 1000-fold can be estimated[4].

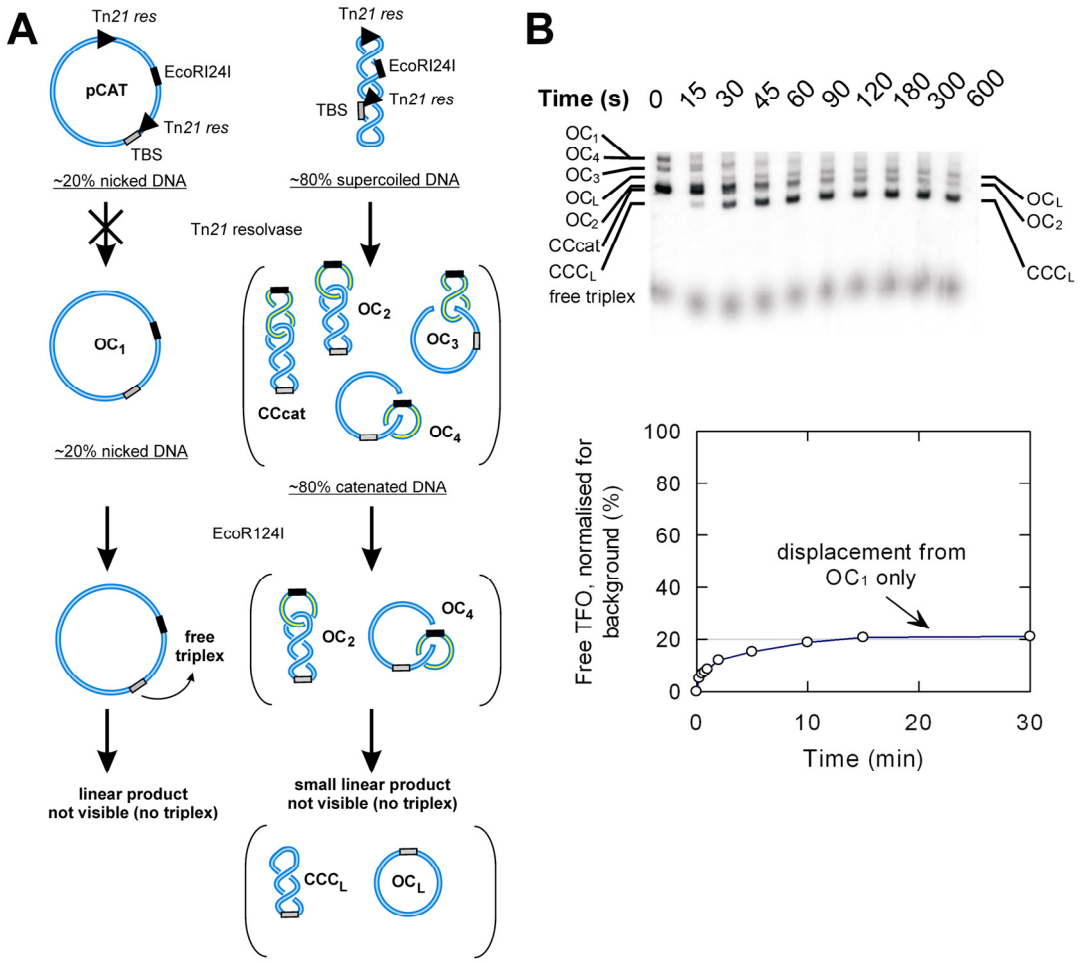
References to the Supplementary Section

1. Seidel, R. *et al.* Dynamics of initiation, termination and reinitiation of DNA translocation by the motor protein EcoR124I. *EMBO J.* **24**, 4188-4197 (2005).
2. Firman, K. & Szczelkun, M.D. Measuring motion on DNA by the type I restriction endonuclease EcoR124I using triplex displacement. *EMBO J.*, **19**, 2094-2102 (2000).
3. Howard, J. *Mechanics of Motor Protein and the Cytoskeleton*. Sinauer Associates, Sunderland, MA, USA (2001)

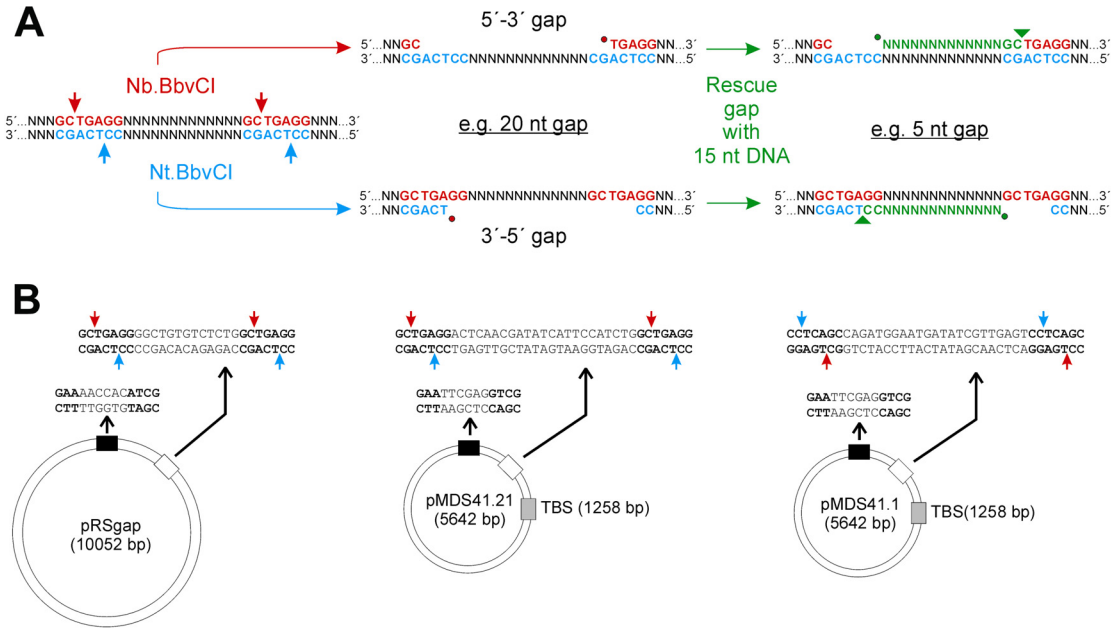
4. Sankararaman, S. & Marko, J.F. Formation of loops in DNA under tension. *Phys. Rev. E. Stat. Nonlin. Soft Matter Phys.* **71**, 021911 (2005).
5. Szczelkun, M.D., Dillingham, M.S., Janscak, P., Firman, K. & Halford, S.E. Repercussions of DNA tracking by the type IC restriction endonuclease EcoR124I on linear, circular and catenated substrates. *EMBO J.* **15**, 6335-6347 (1996)
6. Szczelkun, M.D., Janscak, P., Firman, K. & Halford, S.E. Selection of non-specific DNA cleavage sites by the type IC restriction endonuclease EcoR124I. *J. Mol. Biol*. **271**, 112-123 (1997).
7. Wang, H. & Hays, J.B. Simple and rapid preparation of gapped plasmid DNA for incorporation of oligomers containing specific DNA lesions. *Mol. Biotechnol.* **19**, 133-140 (2001).



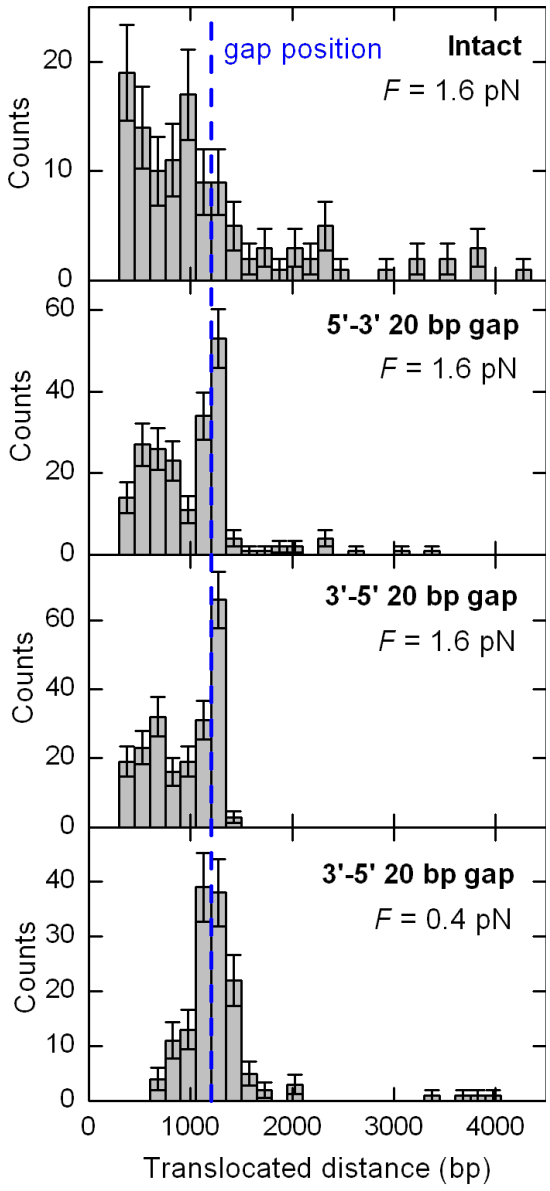
Supplementary Figure 1 DNA cleavage of crosslinked plasmid substrates by EcoR124I. (**A**) ~4 ICLs per pLKS3 molecule were introduced at two HTMP concentrations: <0.033 HTMP/dT (low) or >1 HTMP/dT (high). EcoR124I cleavage reactions were carried out at 37 ºC in 210 µl volumes containing 5 nM DNA (as shown), 30 nM MTase, 90 nM HsdR, 100 µM *S*-adenosyl methionine and 4 mM ATP in Buffer R. Aliquots (20 µl) were removed at the time points indicated and were stopped by the addition of 10 µl of STEB. DNA substrates and products were separated by agarose gel electrophoresis and the relative amounts of intact substrate (shown), nicked circle and full length linear DNA quantified[5]. The initial cleavage rate is unaffected by the ICLs. The reduced amplitude at >1 HTMP/dT indicates that monoadduct formation can, at high concentration, block EcoR124I activity. (**B**) Mapping of the EcoR124I cleavage loci on crosslinked pLKS3. Cleavage of circular DNA most likely results from collision between two HsdR subunits originating from the same endonuclease complex, one travelling around each DNA arc from the core MTase (S. McClelland and M.D.S., unpublished data). Because of the natural topoisomer distribution in our DNA preparations, different individual initiation times for translocation, and the collapse of motor events, cleavage loci on circular substrates are broadly distributed with only a small increase in activity at the locus 180º from the site and a relatively high proportion of cleavage events occurring close to the site[6]. To identify the EcoR124I cleavage sites, the product DNA can be cut with a site-specific Type II RM enzyme to generate a ladder of characteristic fragments[6]. We generated a series of pLKS3 substrates with different levels of ICLs (0-8 per DNA) using <0.033 HTMP/dT. 7 nM of each substrate was treated with 42 nM MTase, 126 nM HsdR, 100 µM *S*-adenosyl methionine and 4 mM ATP in Buffer R at 37 ºC for 10 min. The DNA was then cut with EcoRV and the products separated by agarose gel electrophoresis. The gels were stained with 500 ng/ml ethidium bromide for 2 hours followed by destaining for 40 minutes. Digital images were captured on a Gene Genius Bio-Imaging System (Syngene) without editing and using a linear intensity scale. Scans of ethidium bromide fluorescence intensity in each lane are shown. Apart from an increase in uncut covalently closed DNA (CCC) with increasing ICLs (due to the corresponding increase in monoadduct formation), the modified DNA show the same mapping pattern as unmodified DNA. Thus ICLs do not influence the distribution of cleavage loci.



Supplementary Figure 2 DNA translocation on a catenane substrate. (**A**) pCAT was produced by cloning a triplex binding site (TBS) into the HindIII of pMDS4a (ref. 5). Recombination of supercoiled pCAT by Tn*21* resolvase will produce a single interlinked (-2) catenane with the EcoR124I and triplex binding sites on separate rings. Nicked DNA (OC_1) will not be recombined. Catenanes were produced as described (ref. 5). Our preparation of pCAT contained ~20% nicked DNA and 80% supercoiled DNA; the resulting catenane preparation therefore contained ~80% CCcat (with some post-recombination nicked species OC_2, OC_3 and OC_4) and ~20% unrecombined OC_1 (data not shown). Cleavage of OC_1 by EcoR124I is preceded by triplex displacement so the linear product is not seen on the gel. EcoR124I will cleave CCcat (and nicked species) via nicked intermediates (OC_2 or OC_4) before liberating the linearised small DNA and the large ring with triplex still bound (CCC_L or OC_L). (**B**) Triplex displacement measured from the catenane/nicked DNA mixture using the radioactive assay (ref. 2). The final reaction conditions were 5 nM DNA (2.5 nM ^{32}P-labelled triplex), 30 nM MTase, 90 nM HsdR and 4 mM ATP at 20 ºC. Only ~20% free triplex was liberated in 30 minutes and represents complete displacement from OC_1. Displacement was not observed from any catenane species (all the radioactive signal remains in the CCC_L and OC_L species). Therefore, despite the proximity of the catenane rings, translocation on one ring cannot displace a triplex on the other. If HsdR could make long-range hops during translocation, then some of those events would have resulted in the HsdR binding the interlinked ring and displacing the triplex. Cleavage assays on DNA catenanes also indicate that hopping events do not occur (ref. 5).



Supplementary Figure 3 Generation of gapped DNA. (**A**) Schematic example of the method employed to generate gapped DNA, based on the protocol of Wang and Hays (ref. 7). Two closely-spaced (20 nt shown), directly-repeated BbvCI restriction sites are cloned downstream of an EcoR124I site. Using commercial mutant restriction enzymes (Nt- and Nb.BbvCI, New England Biolabs), the sites could be nicked specifically on either the top (5′-3′) or bottom (3′-5′) strand. In the next step, the short oligonucleotide is removed by heating in the presence of the complementary oligonucleotide. This produces a ssDNA gap in either the 5′-3′ or 3′-5′ strand depending on the orientation of the sites and the nicking endonuclease used. To generate gaps of different lengths, a synthetic oligonucleotide can then be annealed into the gap. Small circles represent the positions of 5′ phosphate groups – these are optional for the rescue oligonucleotides. See Supplementary Tables 1-3 for more details of rescue oligonucleotides used. (**B**) Plasmid substrates used for the tweezers assays (pRSgap) and the triplex assays (pMDS41.21/pMDS41.1). The triplex binding site (TBS) was the same as used previously (ref. 2). See Supplementary Tables 1-3 for further details.



Supplementary Figure 4 Comparisons of the translocated distances on intact and gapped DNA from tweezers experiments. The plots show histograms of the maximum translocated distance, i.e., the distance at which the enzyme stops translocating and dissociates from the DNA, of all recorded events. Note for the gapped substrates the peak at about 1200 bp in agreement with the distance between EcoR124I site and gap. This indicates that the gap is a barrier, which favours termination of translocation. Note also the absence of translocation events beyond the 20 bp 3´-5´gap at 1.6 pN. The broadening of the peak at 1200 bp for the traces recorded at 0.4 pN compared to the data recorded at 1.6 pN is due to the increased noise in the DNA end-to-end distance at lower force.

Modification 5´-3´	Modification 3´-5´	Substrate	BbvCI	Oligonucleotide	*D* (bp)
20 nt		pRSgap	Nb	n/a	1180
	20 nt	pRSgap	Nt	n/a	1183
5nt		pRSgap	Nb	GGCTGTGTCTCTGGC	1180
	5nt	pRSgap	Nt	CAGAGACACAGCCCC	1198
nick (-PO_4)		pRSgap	Nb	TGAGGGGCTGTGTCTCTGGC	1180
	nick (-PO_4)	pRSgap	Nt	TCAGCCAGAGACACAGCCCC	1203

Supplementary Table 1 Details of the gapped DNA used in the tweezers assays. *D* represents the distance from the base pair 3′ to the EcoR124I binding site (Supplementary Figure 3) to the base pair immediately 5′ adjacent (on the 5′-3′ strand) or 3′ adjacent (on the 3′-5′ strand) to the gap.

Modification 5′-3′	Modification 3′-5′	Substrate	BbvCI	Oligonucleotide	*D* (bp)
30 nt		pMDS41.21	Nb	n/a	589
	30 nt	pMDS41.1	Nb	n/a	596
30 nt		pMDS41.1	Nt	n/a	593
	30 nt	pMDS41.21	Nt	n/a	592
	15 nt	pMDS41.1	Nb	ATCATTCCATCTGGC	611
10 nt		pMDS41.21	Nb	ACGATATCATTCCATCTGGC	589
	10 nt	pMDS41.1	Nb	ACGATATCATTCCATCTGGC	616
	9 nt	pMDS41.1	Nb	AACGATATCATTCCATCTGGC	617
	8 nt	pMDS41.1	Nb	CAACGATATCATTCCATCTGGC	618
	7 nt	pMDS41.1	Nb	TCAACGATATCATTCCATCTGGC	619
	6 nt	pMDS41.1	Nb	CTCAACGATATCATTCCATCTGGC	620
5 nt		pMDS41.21	Nb	ACTCAACGATATCATTCCATCTGGC	589
	5 nt	pMDS41.1	Nb	ACTCAACGATATCATTCCATCTGGC	621
5 nt (+PO4)		pMDS41.21	Nb	P-ACTCAACGATATCATTCCATCTGGC	589
	5 nt (+PO4)	pMDS41.1	Nb	P-ACTCAACGATATCATTCCATCTGGC	621
	4 nt	pMDS41.1	Nb	GACTCAACGATATCATTCCATCTGGC	622
	3 nt	pMDS41.1	Nb	GGACTCAACGATATCATTCCATCTGGC	623
	2 nt	pMDS41.1	Nb	AGGACTCAACGAT- -ATCATTCCATCTGGC	624
1 nt		pMDS41.21	Nb	GAGGACTCAACGAT- -ATCATTCCATCTGGC	589
	1 nt	pMDS41.1	Nb	GAGGACTCAACGAT- -ATCATTCCATCTGGC	625
nick (-PO4)		pMDS41.21	Nb	TGAGGACTCAACGAT- -ATCATTCCATCTGGC	n/a
	nick (-PO4)	pMDS41.1	Nb	TGAGGACTCAACGAT- -ATCATTCCATCTGGC	n/a
nick (+PO4)		pMDS41.21	Nb	P-TGAGGACTCAACGAT- -ATCATTCCATCTGGC	589
	nick (+PO4)	pMDS41.1	Nb	P-TGAGGACTCAACGAT- -ATCATTCCATCTGGC	626

Supplementary Table 2 Details of the gapped DNA used in the triplex assays. P represents a 5′ phosphate group added during synthesis. *D* represents the distance from the base pair 3′ to the EcoR124I binding site (Supplementary Figure 3) to the base pair immediately 5′ adjacent (on the 5′-3′ strand) or 3′ adjacent (on the 3′-5′ strand) to the gap.

Modification 5′-3′	Modification 3′-5′	Substrate	BbvCI	Oligonucleotide	*D* (bp)
PEG 601		pMDS41.21	Nb	TGAGGACTCAA- -CC18C18ATTCCATCTGGC	601
	PEG 601	pMDS41.21	Nt	TCAGCCAGATGGA- -ATC18C18GTTGAGTCC	601
	PEG 605	pMDS41.1	Nb	TGAGGACTCAACG- -ATC18C18CCATCTGGC	605
	PEG 606	pMDS41.1	Nb	TGAGGACTCAACG- -AC18C18TCCATCTGGC	606
	PEG 607	pMDS41.1	Nb	TGAGGACTCAACG- -C18C18TTCCATCTGGC	607
PEG 608		pMDS41.1	Nt	TCAGCCAGATGGA- -ATC18C18GTTGAGTCC	608
	PEG 608	pMDS41.1	Nb	TGAGGACTCAA- -CC18C18ATTCCATCTGGC	608
	PEG 609	pMDS41.1	Nb	TGAGGACTCAA- -C18C18CATTCCATCTGGC	609
	PEG 610	pMDS41.1	Nb	TGAGGACTCA- -C18C18TCATTCCATCTGGC	610
	PEG 611	pMDS41.1	Nb	TGAGGACTC- -C18C18ATCATTCCATCTGGC	611
1 nt THF		pMDS41.21	Nb	TGAGGACTCAACG- -ATSTCATTCCATCTGGC	604
	1 nt THF	pMDS41.1	Nb	TGAGGACTCAACG- -ATSTCATTCCATCTGGC	610
5 nt THF		pMDS41.21	Nb	TGAGGACTCAACG- -SSSSSATTCCATCTGGC	602
	5 nt THF	pMDS41.1	Nb	TGAGGACTCAACG- -SSSSSATTCCATCTGGC	608
5 nt reversed		pMDS41.21	Nb	TGAGGACTCAACG- -[ATATC]ATTCCATCTGGC	602
	5 nt reversed	pMDS41.1	Nb	TGAGGACTCAACG- -[ATATC]ATTCCATCTGGC	608
5 nt PNA		pMDS41.21	Nb	TGAGGACTCAA- -C**gatat**CATTCCATCTGGC	601
	5 nt PNA	pMDS41.1	Nb	TGAGGACTCAA- -C**gatat**CATTCCATCTGGC	609

Supplementary Table 3 Details of the modified DNA used in the triplex assays. "C18" represents a C18 polyethylene glycol linker; "S" represents an abasic site (tetrahydrofuran, THF); "[]" represents a region of reversed sequence (3′-5′ as drawn); bold lower case characters represents a region of peptide nucleic acid. *D* represents the distance from the base pair 3′ to the EcoR124I binding site (Supplementary Figure 3) to the base pair immediately 5′ adjacent (on the 5′-3′ strand) or 3′ adjacent (on the 3′-5′ strand) to the modified region.