DNA cleavage site selection by Type III restriction enzymes provides evidence for head-on protein collisions following 1D bidirectional motion

Friedrich W. Schwarz¹, Kara van Aelst², Júlia Tóth², Ralf Seidel¹,* and Mark D. Szczelkun²,*

¹Biotechnology Center, Dresden University of Technology, 01062 Dresden, Germany and ²DNA-Protein Interactions Unit, School of Biochemistry, Medical Sciences Building, University of Bristol, Bristol, BS8 1TD, UK

Received April 8, 2011; Revised May 27, 2011; Accepted June 2, 2011

ABSTRACT

DNA cleavage by the Type III Restriction–Modification enzymes requires communication in 1D between two distant indirectly-repeated recognitions sites, yet results in non-specific dsDNA cleavage close to only one of the two sites. To test a recently proposed ATP-triggered DNA sliding model, we addressed why one site is selected over another during cleavage. We examined the relative cleavage of a pair of identical sites on DNA substrates with different distances to a free or protein blocked end, and on a DNA substrate using different relative concentrations of protein. Under these conditions a bias can be induced in the cleavage of one site over the other. Monte-Carlo simulations based on the sliding model reproduce the experimentally observed behaviour. This suggests that cleavage site selection simply reflects the dynamics of the preceding stochastic enzyme events that are consistent with bidirectional motion in 1D and DNA cleavage following head-on protein collision.

Type III restriction endonucleases (REs) will cut DNA molecules in an ATP-dependent manner when there are two copies of the cognate asymmetric recognition site in an inverted repeat arrangement (1–5). These sites can be many thousands of base pairs apart, but must be on the same DNA chain (6). Whilst sites in either head-to-head (HtH) or tail-to-tail (TtT) orientation can efficiently induce the nuclease activity, DNA with single sites or with pairs of sites in direct (head-to-tail, HtT) repeat do not, except under conditions which promote non-specific association with the DNA (e.g. K⁺ ions and/or elevated enzyme concentrations) (4,7). The Type III REs can therefore be said to exhibit ‘site orientation selectivity’ (1,5). This selectivity is important since these enzymes only hemimethylate their recognition sites but pairs of sites in inverted repeat will always remain protected following semi-conservative replication (2,8–10).

To account for how ATP hydrolysis allows long-range communication between sites in an oriented manner, a number of models have been suggested based on unidirectional 1D DNA translocation accompanied in part by accessory DNA loops (3,11–13). These models are principally based on similarity to the domain organization of the Type I REs, for which motor driven unidirectional DNA translocation has been unambiguously shown [(14) and references therein, (15–17)]. However, Type III REs utilize at least 1000-fold less ATP than their Type I counterparts (3,10,18,19), which has significantly challenged the 1D translocation hypothesis. To resolve this contradiction, additional 3D DNA looping steps preceding 1D translocation have also been suggested to shorten the necessary translocation distance (12).

Using single-molecule and bulk solution DNA cleavage assays we recently provided evidence that the communication occurs in 1D along the DNA contour (5,19). Additionally the communication was found to be bidirectional, since the cleavage efficiency on linear DNA was significantly increased if a bulky protein roadblock was incorporated at the DNA ends. Based on these observations, we proposed an alternative scheme in which ATP hydrolysis by an enzyme at a target site initiates a diffusive bidirectional motion on DNA (Figure 1) (1,19). In this model diffusion does not require ATP hydrolysis and maintains the orientation of the enzyme on the DNA as set by the direction of the target site. Cleavage occurs once a sliding enzyme collides ‘head-on’ (i.e. in a HtH manner), with a second enzyme that is bound to its target site.

*To whom correspondence should be addressed. Tel: +49 351 463 40333; Fax: +49 351 463 40342; Email: ralf.seidel@biotec.tu-dresden.de

Correspondence may also be addressed to Mark D. Szczelkun. Tel: +44 117 928 7439; Fax: +44 117 928 8274; Email: mark.szczelkun@bristol.ac.uk

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

© The Author(s) 2011. Published by Oxford University Press.
This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (http://creativecommons.org/licenses/by-nc/3.0), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.
motion on DNA with cleavage occurring upon head-on collision between a sliding and a static enzyme.

**MATERIALS AND METHODS**

**Proteins**

EcoPI and EcoP15I were purified as described previously (6). Protein concentrations were determined from the absorption at 280 nm using extinction coefficients derived from the aromatic amino acid composition in the predicted amino acid sequences. Concentrations are reported in terms of a Res2Mod2 heterotetramer. All other enzymes were obtained from New England Biolabs (MA, USA) and used as recommended by the manufacturer.

**DNA**

Synthetic oligonucleotides were purchased from Eurofins-MWG Operon (Ebersberg, Germany). pK9 (~R1, +N) was generated as follows. pMDS33 (4), was partially digested by SfiI, ligated with annealed oligonucleotides 5'-ACCTCAGCATGCGGCCGCTAACTA and 5'-TTGGCTTCAGCAGAGCG-3', and clones selected for insertion at the SfiI site located at 1611 bp to generate pMDS33 (MCS). An EcoPI site in the kanamycin resistance cassette of pUC4K (22), was mutated by QuikChange Mutagenesis (Agilent Technologies, CA, USA) using the primers 5'-GAGGTACGAATCGTGGTGTT-3' and 5'-ATCCGGGCTGCTTGCAAAC-3'. The PCR product and pMDS33 (MCS) were cleaved with KpnI, the DNA re-ligated and recombinants selected by the absence of cleavage with NdeI. Finally, with NdeI and BsaAI, the DNA re-ligated and clones selected for loss of the kanamycin resistance gene to generate pK9. A 727-bp region of pK9 containing a single Tn21-resolvase res site was then removed by digestion with NdeI and BsaAI, the DNA re-ligated and clones selected by the absence of cleavage with NdeI. Finally, an NdeI site was reintroduced by PCR using the primers 5'-ATGTCGCAGTGTTTCCGTGATGAC-3' and 5'-ATGGGCAGTGCTGCTTGCAAAC-3'.
ACGAAAGGGCTCGTAC-3' to produce pKA9 (−R1, +N).

To prepare DNA for biochemical assays, *Escherichia coli* Top10 (Invitrogen, CA, USA) or XL10-Gold (Agilent Technologies) were transformed with the required plasmid, grown in LB medium and the DNA extracted using either commercial protocols (Qiagen, Hilden, Germany) or by density gradient centrifugation in CsCl-ethidium bromide (23). We have not noted any difference in Type III activity between the different preparation methods. For experiments requiring 3H-labelled DNA, transformants were grown in M9 minimal medium supplemented with 37 MBq/L [3H-methyl] thymidine (PerkinElmer, MA, USA).

Linear DNA substrates were generated by PCR amplification or by digestion by Type II restriction enzymes. For substrates with two pairs of EcoPI and EcoP15I sites (Table 1, Figures 2 and 3), regions of either pJT1.1 (HtH) or pJT1.3 (TtT) (5) were amplified using Pfu polymerase (Promega) or Phusion polymerase (Finnzymes, New England Biolabs), following the manufacturer’s standard protocols. To generate DNA with biotin labels at either both ends (‘blocked end’) or at just one end (‘open end’), the PCR reactions contained the 5’-biotinylated primer 5’-Biotin-CAGTTTATCAGCTAC GGTATTGGCGACA-3’ and either 5’-biotinylated or unmodified versions of 5’-CTTTGATCCTCCGCCTG TATGGTG-3’ for the ~100-bp spacers, 5’-ATGCAGC TGCAACACAGGTTCG-3’ for the ~200-bp spacers, 5’-GGGTGTTGCCACCTGACTTGG-3’ for the ~500-bp spacers and 5’-CGCGTAATCTGGTGGCTT CAAAACAAAAAC-3’ for the ~1000-bp spacers (Figure 2A and Table 1). The DNA was then purified using a PCR purification kit (Qiagen) and eluted in H2O. For mixed substrates with a pair of EcoPI and EcoP15I sites in HtH repeat (Table 1 and Figure 4), pKA9 (−R1, +N) was digested with NdeI and the linear DNA purified by phenol/chloroform extraction followed by ethanol precipitation. End-labelling of the DNA with biotin was carried out using Klenow (3’-5’ exo-) polymerase and biotin-dUTP (5). The labelled DNA was further purified by phenol/chloroform extraction followed by ethanol precipitation.

DNA concentrations were determined from absorbance at 260 nm, assuming an extinction coefficient of 0.02 ml µg⁻¹ cm⁻¹ and a DNA molecular weight of 6.6 × 10⁵ Da kbp⁻¹.

### Cleavage reactions

DNA substrates (2 or 10 nM, as indicated) were preincubated for at least 1 min at 25°C in Buffer R⁺ [50 mM Tris–HCl, pH 8.0, 50 mM KCl, 10 mM MgCl₂, 1 mM DTT, 0.01% (w/v) BSA] supplemented with 4 mM ATP and streptavidin at a 50-fold molar excess over DNA. Reactions were initiated by adding the enzyme(s) at the concentration(s) indicated, incubated at 25°C for either 5 min (EcoP15I on HtH DNA), 10 min (EcoPI and EcoP15I on the mixed DNA) or 15 min (all other reactions). We verified that the reactions were run to completion in each case (data not shown). Reactions were subsequently quenched by addition of 1/2 volume of STEB buffer [0.1 M Tris–HCl, pH 7.5, 0.2 M EDTA, 40% (w/v) sucrose, 0.4 mg/ml bromophenol blue]. Then 0.1 volume of biotin [0.2 mg/ml in 50 mM Tris–HCl (pH 8.0)] was added to each aliquot and the samples heated for at least 5 min at 80°C to prevent anomalous electrophoretic mobility due to bound streptavidin and restriction enzymes. The DNA substrates and products were separated by agarose gel electrophoresis. Due to the low DNA concentrations used, some small DNA fragments were not detected. To calculate cleavage site distributions, the DNA bands for the longest cleavage product at either site (Figures 2B, 4B and Supplementary Figure S1) were quantified either from digital images of ethidium bromide-stained gels captured on a UVP GelDoc-It Imaging System using a linear intensity scale or by scintillation counting (23). In each case quantified values were corrected for differences in DNA size assuming a linear relationship between dsDNA length and signal.

We estimated the error of the measured cleavage site distribution from all measurements made on symmetric substrates (i.e. with both ends being closed and with a = ~1000 bp), which should nominally be cleaved with a 50:50 distribution (Figure 3). From 13 individual experiments (two different enzymes, two different site orientations and three to four repeats for each condition) we obtained a mean of 48.7% and a standard deviation of 4.0% for cleavage at Site 1—the latter value equating the standard error of an individual measurement. In order to exclude accidental errors and potential bias from a particular DNA or enzyme preparation, all cleavage experiments were repeated between two to four times. The repetitions always reproduced the initial measurements.

---

**Table 1. Distances for the DNA substrates**

<table>
<thead>
<tr>
<th>Enzyme(s)</th>
<th>Site orientation DNA #</th>
<th>dsDNA distances (bp)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>a</td>
</tr>
<tr>
<td>EcoPI</td>
<td>HtH</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>204</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>505</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1004</td>
</tr>
<tr>
<td></td>
<td>TtT</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>219</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>518</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1019</td>
</tr>
<tr>
<td>EcoP15I</td>
<td>HtH</td>
<td>1</td>
</tr>
<tr>
<td></td>
<td>2</td>
<td>200</td>
</tr>
<tr>
<td></td>
<td>3</td>
<td>499</td>
</tr>
<tr>
<td></td>
<td>4</td>
<td>1000</td>
</tr>
<tr>
<td></td>
<td>TtT</td>
<td>5</td>
</tr>
<tr>
<td></td>
<td>6</td>
<td>222</td>
</tr>
<tr>
<td></td>
<td>7</td>
<td>521</td>
</tr>
<tr>
<td></td>
<td>8</td>
<td>1022</td>
</tr>
</tbody>
</table>

Distances are defined as follows (Figures 2-4): a is the distance from the leftward dsDNA end to the base pair 5’ to Site 1; b is the distance from the base pair 3’ to Site 1 to the base pair 5’ to Site 2; and, c is the distance from the base pair 3’ to Site 2 to the rightward dsDNA end. To minimize DNA handling, DNA #1 to #8 had a pair of HH or TT sites for both EcoPI and EcoP15I; one DNA could therefore be used as a substrate for either EcoPI or EcoP15I.
within small errors (see standard deviations in Supplementary Figure S2).

MC simulations
MC simulations as described in the main text were programmed in LabVIEW (National Instruments Inc.) and run on a workstation with a dual quad core 2.66 GHz processor (Dell, T5400-Intel Xeon E5430). The statistical error of the simulations for the normalized cleavage probability \( p_x \) at a given site \( X \) \((X = 1, 2)\) was calculated as twice the standard error (95% confidence interval) of a binomial distribution normalized by the total number of simulations \( N_{tot} \) that ended in cleavage: 
\[
\text{Err}(p_x) = 2\sqrt{p_x(1-p_x)/N_{tot}} \quad \text{with} \quad p_x = \frac{N_X}{N_{tot}} \quad \text{and} \quad N_X \quad \text{being the number of cleavage events at site} \quad X.
\]

RESULTS AND DISCUSSION
Experimental rationale
Our previous analysis of the Type III REs has concentrated on the related proteins EcoPI (which recognizes 5′-AGACC-3′) and EcoP15I (which recognizes 5′-CGACG-3′). To test the sliding model we investigated DNA cleavage rates as an indirect measure of the site-to-site communication rate. By changing the length of the DNA between a pair of identical sites or by varying symmetrically the distance between the sites and the DNA ends we sought to observe changes in the cleavage rate that could be related to possible communication models (5,19). Because the sites were identical in these cases, site selection was 50:50 regardless of orientation [as in (4)]. However, we were unable to observe any systematic change in the cleavage rate that could be fitted to either 1D diffusion or translocation models. An explanation for this may be that the communication process is significantly faster than the binding or cleavage steps and is thus kinetically masked.

As an alternative approach to test the validity of the sliding model and its mechanistic assumptions we addressed here how this mode of communication affects the relative cleavage site distribution on asymmetric two-site linear DNA substrates. To introduce sufficient asymmetry into the reaction pathway we applied two different strategies.

Introducing bias due to end-blocking and end-spacing effects. The sliding model indicates that on linear DNA a Type III RE cannot exit at a DNA end that is ‘blocked’ (e.g. by biotin–streptavidin) but it can exit from an ‘open’ end (1,19). Consequently, RE activity is efficient when both ends are blocked but inefficient when both ends are open. This difference could be exploited by using linear DNA molecules in which one DNA end differed from the other end in terms of distance to a site and by being either blocked (by biotin–streptavidin) or open. A single open end will asymmetrically deplete the sliding enzyme population producing characteristic changes in the cleavage site distribution (discussed below).

Introducing bias due to increased initiation at one site. As a comparative approach, the cleavage at one site over another could be biased by increasing the total amount of sliding enzyme that originated from only one of the two sites. For this a mixed HtH linear DNA substrate could be used with one recognition site for EcoPI and one recognition site for EcoP15I (Table 1 and Figure 4A). EcoPI and EcoP15I are very similar enzymes (amino acid identity >90% in the Res subunit and >60% in the Mod subunit) and can cooperate to cleave a mixed DNA substrate (24,25). By varying the concentration of one enzyme over the other, the corresponding increase in communicating species originating from one site will also influence the relative cleavage distributions in a characteristic manner (discussed below).

Distribution of cleavage on head-to-head substrates with end bias
To apply the first strategy described above, we generated a family of linear HtH and TtT DNA substrates with biotin–streptavidin blocks at either one end (‘open end’ substrates) or both ends (‘blocked end’ substrates). We first tested the HtH DNA illustrated in Figure 2A. As drawn, we arbitrarily define the site closest to the left end as ‘Site 1’ and the other site as ‘Site 2’. The distance \( a \) from Site 1 to its neighbouring open or blocked end was varied whilst the inter-site distance \( b \) of ~100 bp and the distance \( c \) from Site 2 to its neighbouring blocking end of ~1000 bp were kept constant (see ‘Materials and Methods’ section and Table 1). To avoid any differences resulting from a preference for cleavage at particular sequences flanking the target site, we used pairs of Type III sites with identical sequences up- and down-stream of the sites. Thus, any observed asymmetry in the cleavage preference would be due to variation in distance \( a \) and/or to the difference in end-capping.

Based on an intuitive consideration of the sliding model (Figure 1), an ‘open end’ HtH substrate would asymmetrically deplete the sliding enzyme population as follows. Enzymes that initiate sliding from Site 1 would be more likely to dissociate from the neighbouring open end than enzymes that initiate sliding from Site 2. This would reduce the chance of enzymes initiating from Site 1 reaching Site 2. Correspondingly, cleavage would occur more often at Site 1. If distance \( a \) were then shortened, we would expect this bias in cleavage at Site 1 to increase. In contrast, on the ‘blocked end’ HtH substrates, enzymes initiating from either Site 1 or Site 2 will have an equal chance of dissociation and the cleavage will distribute 50:50, regardless of any change in distance \( a \). The bias in cleavage site distribution on the open end substrates would thus be expected to vary proportional to the difference in distance \( a \).
Figure 2. Using end-spacing effects to introduce bias into the distribution of cleavage between two sites. (A) DNA substrates (see ‘Materials and Methods’ section). Triangles denote the orientation of the EcoPI (5'–AGACC–3') or EcoP15I (5'–CAGCAG–3') recognition sites. Sites 1 and 2 are defined as the sites being closest to the leftward or rightward end, respectively. The Site 1 triangles are either 'solid' (to denote a blocked end) or 'open' (to denote an open end). Green ovals represent streptavidin molecules attached to biotin-labelled DNA ends (blue circles). Distances a, b and c for each substrate are listed in Table 1. Distance a was varied as indicated by the dashed line. Each DNA had a pair of sites for both EcoPI and for EcoP15I (only one pair of sites for one enzyme is shown for clarity) (Table 1). (B) Example agarose gel for the open end HH substrates with distance a = ~100, ~200 and ~500 bp. See Supplementary Figure S1 for the same gel including a DNA size marker and the lanes for a = ~1000 bp. Reactions contained 10 nM DNA and 75 nM EcoPI or EcoP15I and were incubated for 15 min with EcoPI or 5 min with EcoP15I. DNA samples were separated by agarose gel electrophoresis. Resulting cleavage products are illustrated by cartoons. Except for a = ~500 bp, the two smaller cleavage products that include the short end a were not visible on the gel. Cleavage at Site 1 was quantified from the DNA product containing domains b and c: cleavage at Site 2 was quantified from the DNA product containing domain c alone (see ‘Materials and Methods’ section). (C) Percentages of Site 1 cleavage (red open triangles), of Site 2 cleavage (solid black triangles) and total cleavage (solid grey circles) for the gel shown in B as function of distances a. Points and error bars are the average and standard deviation from two to four repeat experiments. See Supplementary Figure S2 for all data, including the standard deviations, which illustrate the high reproducibility of the measurements.
of 1. For the initial modelling of the HtH and TtT substrates, \(k_{\text{bind}}\) and \(k_{\text{ini}}\) were adapted to fit the experimental data. For subsequent simulations we modelled 5 bp instead of 1 bp random walk steps in order to save computation time. This provided identical results compared to the 1-bp steps (data not shown) but required on average 25-fold less simulation steps. For each condition (DNA end length, end capping), at least 500 cleavage events were simulated to obtain acceptable statistical accuracy.

Both MC model data and empirical data showed good qualitative agreement, independent of the particular value of \(k_{\text{ini}}\) (Figure 3A). Best overall agreement (also including TtT and mixed substrates, see below) was obtained by choosing \(k_{\text{bind}} = 4 \times 10^{-8}\) for both enzymes, and \(k_{\text{ini}} = 4 \times 10^{-6}\) for EcoPI and \(k_{\text{ini}} = 4 \times 10^{-8}\) for EcoP15I. Note that other sets of \(k_{\text{bind}}\) and \(k_{\text{ini}}\) can also describe the experimental data in Figure 3 as long as the ratio between the rates is preserved (Supplementary Figure S3).

Considering published diffusion constants for enzymes sliding along DNA in the order of \(10^6 \text{ bp}^2 \text{ s}^{-1}\) (26,27), the chosen value for \(k_{\text{bind}} = 4 \times 10^{-8}\) in these simulations would correspond to a ‘real’ enzyme binding rate in the order of \(0.1 \text{ s}^{-1}\). A faster initiation rate for EcoPI, and thus shorter-lived binding state, compared to EcoP15I was also suggested previously to explain the difference between the enzymes in relative cleavage rates of TtT substrates compared to HtH substrates (5). The agreement between both the simulated and empirical data sets in Figure 3A validates our predictions and provides evidence that the sliding model can account for the relative cleavage of a site in a HtH oriented substrate dependent on its proximity to a DNA end.

### Distribution of cleavage on tail-to-tail substrates

In addition to sites in HtH arrangements we also tested similar substrates with the sites in TtT arrangement (Figure 3B, Table 1, and Supplementary Figure S2). On these substrates, cleavage that would be triggered by head-on collision requires that the sliding enzyme actually bypasses a site before the second enzyme binds (1,5). Consequently, the efficiency of cleavage of TtT DNA can be affected by the binding lifetime at the site: the more often the site is occupied, the less chance that a sliding enzyme can bypass the site. The reactions were carried out and analysed as above. Quantified and normalized data from repeat experiments using either EcoPI or EcoP15I are shown in Figure 3B. In contrast to the data on HtH DNA, we observed quite different outcomes on the TtT DNA that depended on both the enzyme used and the characteristics of the DNA substrate:

**EcoPI on blocked end TtT DNA.** Where distance \(a\) was 518 or 1019 bp, the cleavage distribution between Sites 1 and 2 was equal, within experimental variation. However, at the shorter \(a\) distances (120 or 219 bp), cleavage at Site 2 increased moderately.
EcoPI on open end TiT DNA. Where distance \( a \) was 518 or 1019 bp, a small but significant preference was observed for cleavage at Site 1. However, as \( a \) was shortened, the relative preference switched until Site 2 was preferred at \( a = 120 \) bp.

EcoP15I on blocked end TiT DNA. Where \( a = 1022 \) bp, the cleavage distribution between Sites 1 and 2 was equal, within experimental variation. However, as \( a \) was shortened, cleavage at Site 2 increased significantly. When \( a = 123 \) bp, cleavage at Site 2 accounted for \( >80\% \) of the total cleavage. A reduction in the total cleavage efficiency was also observed as \( a \) was reduced (Supplementary Figure S2).

EcoP15I on open end TiT DNA. As for EcoPI above, Site 1 was preferred at longer distances of \( a \), whilst Site 2 was preferred at shorter distances of \( a \). However, the contrast was more marked. There was also a more marked reduction in overall cleavage as \( a \) reduced which was not observed to the same extent with EcoPI (Supplementary Figure S2). Strikingly, the overall pattern in this data looks similar to that observed using EcoP15I and the blocked end TiT DNA.

The feature common to all the TiT data is that cleavage of Site 2 is preferred to Site 1 at short distances of \( a \), the opposite of what was observed with HtH DNA. Moreover, this preference was still observed when both ends were blocked, again the opposite of what was observed with HtH DNA.

We tested whether the observed profiles could still be explained within the framework of the sliding model by simulating the cleavage of TiT DNA using the MC simulations, as above (Movie S2 in Supplementary Data). Both simulated and empirical data show very clear agreement using the same rates chosen for the HtH simulations (Figure 3B). This demonstrates that the measured cleavage site selection is exactly what one would expect from a bidirectional sliding model with head-on collision causing cleavage at a site.

The increase in cleavage at Site 2 relative to Site 1 at small distances may seem counterintuitive, particularly when compared to the data obtained using HtH site arrangements. However, our observations can also be rationalized using the sliding model: On both HtH and TiT DNA, the population of enzymes that initiates sliding from Site 1 will be reduced if there is a proximal open end. Therefore, less of these enzymes reach Site 1 and cleavage at Site 1 increases. However, for TiT DNA, cleavage at Site 1 additionally requires that the sliding enzyme coming from Site 2 enters the DNA domain between the DNA end and Site 1 (i.e. defined as domain \( a \), Figure 2A) and subsequently becomes trapped there by a second enzyme binding to Site 1. As domain \( a \) shortens, the chance of this trapping event reduces. Thus, less of the sliding enzymes that started at Site 2 reach Site 1 in a head-on orientation and relative cleavage at Site 2 increases. It appears that the reduced ‘trapping efficiency’ that is specific to TiT DNA is the dominant effect for short distances of \( a \) while the depletion of enzymes that start near an open end becomes more dominant for longer distances of \( a \). The clear switch in preference on the open end substrates can therefore be explained.

An additional limitation for TiT DNA is that binding of an enzyme to Site 1 blocks sliding enzymes from entering domain \( a \). For EcoP15I this effect is more dominant than the trapping effect above since TiT DNA is cleaved less efficiently than HtH DNA (5). A tight binding of EcoP15I most likely causes the reduction in overall cleavage efficiency seen here as domain \( a \) shortens (Supplementary Figure S2).

Overall, the observation of differences between HtH and TiT DNA are consistent with a sliding model where the cleavage state requires a head-on protein-protein contact.

Distribution of cleavage as a function of enzyme loading

As an alternative way to bias cleavage at one location over another, we generated a mixed DNA substrate with recognition sites for EcoPI and EcoP15I in HtH orientation (Figure 4A and Table 1). The substrate was capped at both ends with streptavidin to ensure efficient cleavage and to prevent bias due to end effects (i.e. as in Figure 3). DNA cleavage only occurs when both enzymes are added (data not shown). By varying, for example, the concentration of EcoPI whilst keeping EcoP15I constant, the sliding model predicts that the population of EcoPI molecules sliding on DNA will increase relative to EcoP15I. In turn this would result in a concentration-dependent increase in cleavage at the EcoP15I site relative to the EcoPI site.

To test this prediction the mixed DNA substrate was incubated with a fixed concentration of EcoP15I and with a varying concentration of EcoPI for a fixed time, the reactions stopped, the substrates and products separated by agarose gel electrophoresis, and the DNA quantified by scintillation counting (see ‘Materials and Methods’ section). An example gel is shown in Figure 4B, whilst the average quantified data is shown in Figure 4C, left panel. With an equal concentration of EcoPI and EcoP15I, a small but significant preference for cleavage at the EcoPI site was observed. As the concentration of EcoPI was increased relative to EcoP15I, the total cleavage reduced from \( \sim96\% \) to \( \sim82\% \), reflecting either an increase in inhibition by co-methylation or a general decrease in activity due to non-specific inhibition by the excess EcoPI. In addition, the distribution of cleavage loci changed—cleavage at the EcoPI site decreased to a minimum of \( \sim25\% \) while the total cleavage at the EcoP15I site increased to a maximum of \( \sim75\% \) (as percentages of the total cleavage). Overall this pattern matches what one would expect based on a simple consideration of the sliding model.

The data was normalized by the total cleavage and compared to MC simulations using the same \( k_{\text{bind}} \) and \( k_{\text{ini}} \) values as used above, i.e. \( k_{\text{bind}} = 4 \times 10^{-8} \) for both enzymes, and \( k_{\text{ini}} = 4 \times 10^{-8} \) for EcoPI and \( k_{\text{ini}} = 4 \times 10^{-8} \) for EcoP15I (Figure 4C, right panel). The empirical and simulated data show very good correspondence, providing further evidence of the validity of the simple sliding scheme in Figure 1. Note that in contrast to the single
enzyme substrates, faster values for $k_{\text{bind}}$ and $k_{\text{ini}}$ fail to describe the data obtained with the mixed substrates (Supplementary Figure S3).

**CONCLUSION**

It is well known that non-specific DNA sequences adjacent to a RE target site can affect protein binding due to indirect interactions. Additionally for the Type III REs, cleavage is targeted to the non-specific DNA 25–28 bp downstream of the site. Therefore the nuclease domain may show a catalytic preference for certain sequences. Using EcoP15I plasmid substrates, a systematic study showed that variations in sequence around the CAGCAG sequence at one site can cause alterations in the relative cleavage pattern independent of DNA ends (21). Whilst this data indicates that Type III REs can have a preference for cleaving certain sequences, it does not address the underlying mechanism of communication. In the experiments presented here we instead observed long-range effects on the cleavage preference at a pair of identical sites due to variations in site location (in particular the distance to a DNA end and whether it was blocked or free) or due to variations in the concentration of a Type III enzyme initiating from one of the two sites. In addition, these variations in preference were dependent on whether the sites were in HtH or TiT orientation. In each case the data could be explained using a simple DNA sliding model, backed-up by MC simulations, that is based on a 1D bidirectional motion of the enzymes on DNA and cleavage that is activated upon head-on collision between one sliding enzyme and one static enzyme (Figure 1, Movies S1 and S2 in Supplementary Data).

In combination with previous bulk solution and single molecule results, the empirical and simulated data presented here provide strong evidence in favour of the proposed sliding mode for long-range communication of Type III enzymes. However, there is still a considerable debate in the field as to whether or not DNA loops play any role in the communication scheme (1,20). A series of AFM studies argued for extensive looping being induced by passive 3D diffusive loop capture and by active 1D translocation driven loop expansion (12,13). In contrast single-molecule magnetic tweezers assays at fN forces that hardly perturb the random coil configuration of the DNA did not obtain evidence for loops that persisted longer than 1 s (19).

So far the discussion about whether ‘to loop or not to loop’ has not been very clear in distinguishing whether looping itself is responsible for the inter-site communication or whether looping just plays an accessory role. The data we present here cannot provide evidence to refute the existence of short-lived loop states. However, it would be very difficult to explain the different cleavage preferences we observed in terms of any bias due to communication by 3D looping. It is a long-held view in the field of DNA–protein transactions that passive 3D DNA looping cannot provide: (i) site orientation selectivity, except under special circumstances of DNA topology or short inter-site spacings and (ii) a bias for interactions between sites on the same DNA as opposed to separate but topologically linked DNA.

The first point is illustrated by a study of the Type II RE BspMI, where site orientation selectivity was observed as a function of inter-site spacing and DNA topology (28). A bias for the geometry of a DNA loop [preferentially trapping a negative or positive node (29)] was only observed for very short inter-site spacings (<170 bp). The Type III REs still show orientation selectivity over many thousands of bp. A bias for the interaction geometry over such long distances is well established for many
site-specific recombinases, which however requires an absolute dependence on negatively supercoiled DNA (30–32). In contrast, cleavage rates for Type III REs on linear and negatively supercoiled DNA are more or less identical (4,19). Inhibition on positively supercoiled DNA (25), most likely reflected the extreme levels of DNA twist introduced into those DNA preparations (33).

The second point is illustrated by a wide range of studies of 3D DNA looping by REs using concatenated DNA [e.g. (34–38)], where cleavage could occur when sites were both on the same DNA and also on separate but topologically interlinked DNA rings. This simply reflects the fact that the local site concentration can be elevated by either 1D or 3D connectivity. Again, and in contrast to all these examples, the Type III REs could not cut sites on separate rings of a catenane (6), arguing against 3D looping as a critical step in selecting directionality.

Within any communication model however, DNA looping could play an accessory role rather than being a stringent requirement. This is consistent with the fact that inhibition of looping by DNA stretching has no significant influence on the communication and cleavage kinetics of Type III REs (19), in agreement with data using restriction enzymes that interact only with a single site (39), but in stark contrast to restriction enzymes that obligatorily use 3D looping (40,41). Also the sliding model, and thus the data presented here, is still compatible with the formation of transient loops during bidirectional 1D motion that are shorter lived than ~1 s (19). This may in fact be an attractive option to resolve the controversy between AFM measurements and single-molecule magnetic tweezers measurements (20). However, we also note that there are still severe discrepancies between the two data sets, since in the AFM experiments the equilibrium is shifted towards a compacted and heavily looped configuration, while for the tweezers experiments the equilibrium configuration is predominantly non-looped, and only allows for rare transient looping. Thus, further work is still required to resolve the contribution of DNA-looped states in inter-site communication by Type III restriction enzymes.

We suggest that the positional effects observed here (Figures 2 and 3), might also be observed in other processes, such as DNA mismatch repair, where communication by DNA sliding has also been suggested to occur (1). Similar experiments could therefore provide corroboratory evidence for other diffusing enzyme systems. In addition, and beyond the academic interest in understanding the cleavage site selection, our observations may be useful for biotechnological reasons. Since these enzymes cut at ‘long’ distances from their target sites they are used in mate-pair library preparations for next generation sequencing to study chromosome structural rearrangements (42). Our observations could be exploited to help increase the efficiency of library preparations by the Type III REs.

SUPPLEMENTARY DATA
Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS
The authors thank Ross (Guanshen) Cui for protein preparations, Fiona Diffin and Alice Sears for assistance with DNA cloning, Dominik Kauert for advice on programming and Andreas Dahl for helpful discussions.

FUNDING
The Wellcome Trust (084086 to M.D.S.), the EU Marie Curie Research Training Network ‘DNA Enzymes’ (MRTN-CT-2005-019566 to M.D.S.), the Deutsche Forschungsgemeinschaft (DFG, SE 1646/1-1 and SE 1646/2-1 to R.S.) and the Dresden International Graduate School for Biomedicine and Bioengineering, funded by the DFG (to F.W.S.). Funding for open access charge: Wellcome Trust Value in People Award.

Conflict of interest statement. None declared.

REFERENCES


