The Type ISP Restriction–Modification enzymes LlaBIII and LlaGI use a translocation–collision mechanism to cleave non-specific DNA distant from their recognition sites

Eva Šišáková, Kara van Aelst, Fiona M. Diffin and Mark D. Szczelkun*

DNA-Protein Interactions Unit, School of Biochemistry, Medical Sciences Building, University of Bristol, Bristol BS8 1TD, UK

ABSTRACT

The Type ISP Restriction–Modification (RM) enzyme LlaBIII is encoded on plasmid pJW566 and can protect *Lactococcus lactis* strains against bacteriophage infections in milk fermentations. It is a single polypeptide RM enzyme comprising Mrr endonuclease, DNA helicase, adenine methyltransferase and target-recognition domains. LlaBIII shares >95% amino acid sequence homology across its first three protein domains with the Type ISP enzyme LlaGI. Here, we determine the recognition sequence of LlaBIII (5′-TnAGCC-3′, where the adenine complementary to the underlined base is methylated), and characterize its enzyme activities. LlaBIII shares key enzymatic features with LlaGI; namely, adenosine triphosphate-dependent DNA translocation ($\sim$309 bp/s at 25°C) and a requirement for DNA cleavage of two recognition sites in an inverted head-to-head repeat. However, LlaBIII requires K^+ ions to prevent non-specific DNA cleavage, conditions which affect the translocation and cleavage properties of LlaGI. By identifying the locations of the non-specific dsDNA breaks introduced by LlaGI or LlaBIII under different buffer conditions, we validate that the Type ISP RM enzymes use a common translocation–collision mechanism to trigger endonuclease activity. In their favoured *in vitro* buffer, both LlaGI and LlaBIII produce a normal distribution of random cleavage loci centred midway between the sites. In contrast, LlaGI in K^+ ions produces a far more distributive cleavage profile.

INTRODUCTION

The Restriction–Modification (RM) enzymes play an important role in bacteria and archaea by preventing infection by bacteriophages and other foreign DNA elements (1,2). Analysis of these enzymes has, in addition to throwing light on their biological role, provided more general information on nuclease, methyltransferase and helicase mechanisms and on mechanisms of molecular recognition of DNA sequences. We recently characterized LlaGI, a prototype of a subclass of adenosine triphosphate (ATP)-dependent RM enzymes collectively called type ISP enzymes (for single polypeptide) (3–5). LlaGI comprises a single polypeptide with, from N- to C-terminus, Mrr endonuclease, superfamily 2 (SF2) helicase and γ-type adenine methyltransferase (MTase) domains (Figure 1A). LlaGI recognizes the sequence 5′-CTnGAYG-3′. The site is methylated on only one of the two conserved adenine residues (i.e. it is hemimethylated) (3), which leads to a potential problem after replication where one daughter molecule contains an unmethylated site. LlaGI overcomes this by requiring the presence of two sites in an inverted, head-to-head (HtH) repeat on the same DNA molecule to activate dsDNA cleavage. On the host genome, one of the two sites will always be hemimethylated after replication, thus protecting both sites (7). As these sites can be thousands of base pairs apart, LlaGI must use long-range interactions to communicate the dual methylation status. On the basis of its enzyme activities, and the long-range communication mechanism for the multi-subunit Type I RM enzymes (8), we proposed that LlaGI uses its helicase domain to couple ATP hydrolysis to DNA translocation in search of a second enzyme complex (5). We have interpreted the transient trapping of topological domains during translocation as evidence of loop translocation, where the

---

*To whom correspondence should be addressed. Tel: +44 117 331 2158; Fax: +44 117 331 2168; 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.
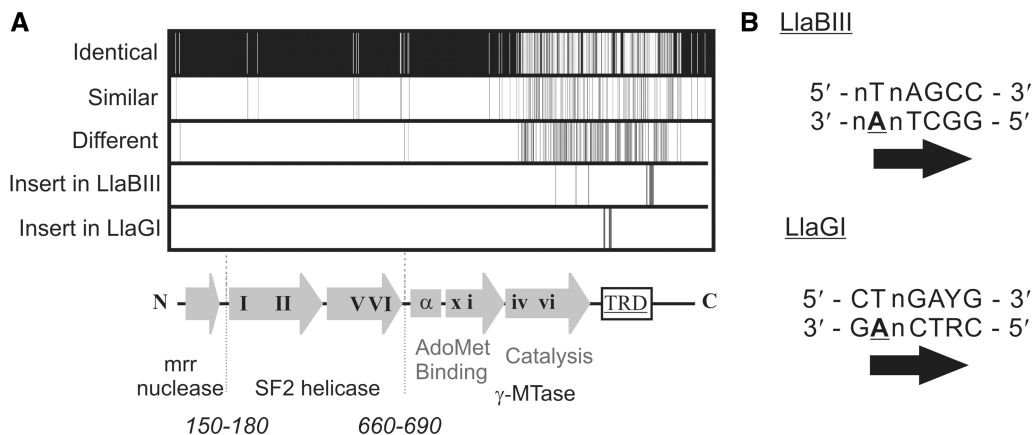
Figure 1. LlaGI and LlaBIII are related RM enzymes with similar protein domain structures. (**A**) Pairwise amino acid sequence alignment of LlaGI and LlaBIII by Basic Local Alignment Search Tool was used to identify identical, similar and different amino acids and regions with inserts/deletions. These are shown as vertical lines within horizontal rows, which merge into solid areas at high density. Below is shown a cartoon representation of the putative domain structure for LlaGI/LlaBIII. Vertical dotted lines show domain boundaries within loop regions identified from secondary structure predictions using PSIPRED (6) (data not shown). Numbering is for LlaGI. Some principal conserved amino acid motifs are labelled within the helicase and methyltransferase domains. (**B**) Recognition sites for LlaBIII (identified here) and LlaGI (3). The adenine residue modified by the MTase is underlined in bold. The arrows indicate the direction of translocation. Also see Figure 6 for further discussion of the site format.

enzyme remains bound at the site during motion. However, direct evidence for this has yet to be collected. When two such enzyme complexes converge and collide, DNA cleavage is activated at the collision site. These sites will, therefore, be distant from the recognition site, distributed halfway between the sites on average, with a normal distribution dependent on the relative translocation and cleavage kinetics (9). Although our previous studies were consistent with this scheme, the locations of the non-specific cleavage sites were not mapped, and the generality of the mechanism was untested.

LlaBIII is a putative Type ISP RM enzyme encoded on the plasmid pJW566 isolated from *Lactococcus lactis* subspecies *cremoris* W56 (10,11). It has been shown to function *in vivo* as a bacteriophage defence mechanism, efficiently protecting *L. lactis* in milk fermentations. The amino acid sequences of LlaBIII and LlaGI were aligned using Basic Local Alignment Search Tool, and the pairwise comparison is shown pictorially in Figure 1A (10). For the first two-thirds of the sequences, LlaBIII and LlaGI are >95% identical. These regions correspond to domains associated with the nuclease, helicase and *S*-adenosyl methionine-(AdoMet)-binding activities. In the final third of the alignment, there are more significant number of differences, both in sequence and in register of the amino acids, with <50% identity. This region corresponds to the domain(s) likely to be associated with the catalysis of methyl transfer and to target recognition (the target recognition domain or TRD). The final ~100 amino acids are more closely related. On the basis of these sequence relationships, LlaBIII would be expected to have similar structural and catalytic properties to LlaGI but to recognize a different specific DNA sequence (10).

A key question for the Type ISP class of RM enzymes is, how are the helicase and nuclease domains interacting to generate DNA cleavage? Furthermore, how does this differ from the classical Type I enzymes where the nuclease–helicase, methyltransferase and target-recognition activities are on separate subunits (HsdR, HsdM and HsdS, respectively)? These questions were addressed further here. We determined the recognition sequence of LlaBIII, determined the identity of the methylated adenine and measured the accompanying endonuclease/helicase activities. We compared the distribution of the non-specific cleavage loci of both LlaGI and LlaBIII to address whether the results are consistent with the previously proposed ATP-dependent translocation–collision mechanism for activation of nuclease activity (5). The similarity between the nuclease and helicase domains of LlaGI and LlaBIII (10) also offered the possibility that the two enzymes could cooperate to cleave DNA, that is, on DNA with sites for LlaGI and LlaBIII in HtH orientation, addition of both enzymes would lead to dsDNA breaks. In the accompanying article (12), we confirm this property, and use it to explore the relative roles of the endonucleases in the collision complex. The characterization of LlaBIII in these two studies allows us to refine our understanding of how the Type ISP class of RM enzyme are activated to cut DNA. In combination, our data have wider implications on the understanding of how nucleases are controlled to prevent unwanted and potentially toxic DNA cleavage events.

MATERIALS AND METHODS

DNA

The *llabiiiRM* gene was amplified from pJW566 (11), (a kind gift of Dr Jytte Josephsen) by polymerase chain reaction (PCR) using oligonucleotides 5′-GCGTAAGTC<u>CCATGG</u>TGGCATTTTGGAAGG-3′ (NcoI site underlined) and 5′-GCGTAAGTC<u>GGATCC</u>ATTATAGTCCCTGTACTACTCCTTG-3′ (BamHI site underlined). The PCR product was digested with NcoI and BamHI,

and it was inserted into pET28a (Novagen, Germany) to produce pET28aLlaBIII, and the *llabiiiRM* gene was sequenced fully. We note that the 4741 nucleotide sequence differed from the published sequence between nucleotides 440–550 because of errors in the original sequencing.

DNA for biochemical assays was prepared, and ^{3}H-labelled where needed as described previously (13). The LlaBIII DNA substrates were constructed using QuikChange mutagenesis (Stratagene) based on pUC19 (14), which contains three LlaBIII sites at 144, 1137 and 1734, and one LlaGI site at 1866: pInvR contains mutations at C141A and A144T, knocking out LlaBIII site 144; pDirR contains a mutation at G1140T, knocking out LlaBIII site 1137; pOne contains mutations at C141A, A144T, C1731A and A1734T, knocking out LlaBIII sites 144 and 1734; pZero contains mutations at C141A, A144T, G1140T, C1731A and A1734T, knocking out all three LlaBIII sites. For the methylation competition assays, the 'top' and 'bottom' substrates were made by directional cloning of oligonucleotide duplexes containing overlapping LlaBIII/EcoAI sites into pOne digested with SphI and BamHI. For the triplex displacement assays, PCRs were carried out using pRMA03 [a plasmid with four triplex binding sites originally cloned for use with LlaGI, (5)] and the following primer pairs: reaction 1, 5'-GCGTAAGTGGATCCGATCCCAGGTTGTTGTA CAGAAC-3' and 5'-GCGTAAGTAAGCTTGGTTATG CTAGTTATTGCTCAG-3'; reaction 2, 5'-GCGTAAGT GACGTCCGATGCGTCCGGCGTAGAG-3' and 5'-GC GTAAGTGACGTCGCCAACAATAGCGTCCCATAA TG-3'. The product of PCR reaction 1 was digested with SalI and HindIII and was ligated into pOne digested with SalI and HindIII. The resulting DNA product and PCR reaction 2 were digested with AatII, ligated and the correct orientation of the triplex insert selected by sequencing to give pOne-triplex.

Proteins

Wild-type LlaGI and LlaBIII were purified according to the protocol in Smith *et al.* (3) using either *Escherichia coli* BL21 (DE3) cells (Novagen, Germany) transformed with pRSFLlaGI (for LlaGI) or from *E. coli* Rosetta 2 (DE3) cells (Novagen, Germany) transformed with pET28aLlaBIII (for LlaBIII). The concentration of LlaBIII was determined from the absorbance at 280 nm using a theoretical molar extinction coefficient determined from the amino acid composition. M.EcoAI was supplied by Michelle Simons (University of Bristol).

Determining the recognition site of LlaBIII

An *in vitro* approach was used similar to that developed for LlaGI (3). First, a library of arbitrary DNA plasmids was screened for susceptibility to cleavage by LlaBIII in the presence of 4 mM ATP. PCR substrates were generated from these DNA using primer walking, in which one end of the DNA was iteratively shortened, and these DNA substrates were tested for cleavage. Loss of DNA cleavage activity was taken as evidence that an LlaBIII recognition site resided in the upstream deleted region. By repeating this with a number of DNA, common sequences were identified. This process assumed that LlaBIII would require a pair of sites in inverted HtH repeat, as subsequently proven.

Methylation competition assays using EcoAI

The circular 'top' and 'bottom' substrates (5 nM) were methylated for 4 h with 100 μM of AdoMet and either 200 nM LlaBIII in TMDK buffer (50 mM Tris–Cl pH 8.0, 10 mM MgCl₂, 150 mM KCl, 1 mM DTT) or 40 nM M.EcoAI in TMDN buffer (50 mM Tris–Cl pH 8.0, 10 mM MgCl₂, 50 mM NaCl, 1 mM DTT). Each reaction was supplemented with 100 μM AdoMet each hour. The DNA was purified by phenol–choloform extraction followed by isopropanol precipitation. The methylated substrates were then treated with 200 nM LlaBIII in TMDK supplemented with 4 mM ATP for 10 min at 25°C. Reactions were stopped with 0.5 volumes of 3× STEB [0.1 M Tris pH 7.5, 0.2 M ethylenediaminetetraacetic acid, 40% (w/v) sucrose, 0.4 mg/ml bromophenol blue]. The DNA substrates and products were separated by agarose gel electrophoresis.

DNA cleavage assays

Cleavage assays contained 5 nM DNA (supercoiled or linear), 4 mM ATP and 200 nM LlaGI or LlaBIII in either TMDK buffer or TMD buffer (50 mM Tris–Cl pH 8.0, 10 mM MgCl₂, 1 mM DTT). Reactions were started by adding ATP and incubated at 25°C for the times indicated. Other Type II RM endonucleases were added at concentrations recommended by the suppliers. Samples were stopped using STEB and were analysed by agarose gel electrophoresis, and, where required, the percentage of ^{3}H-labelled DNA in each band per lane was ascertained by scintillation counting (13).

DNA translocation assays

To produce linear substrates, pOne-triplex was digested using either PciI (for LlaBIII reactions) or pRMA03R (5) was digested using BlpI (for LlaGI reactions). All substrates were purified by phenol–chloroform extraction and isopropanol precipitation. Triplexes were formed overnight as described previously using tetramethylrhodamine-labelled triplex forming oligonucleotides (15). Reactions were carried out using 1 nM DNA (0.5 nM triplex), 100 nM enzyme, 4 mM ATP in TMDK or TMD at 25°C. Fluorescence intensity measurements were performed using an SF61-DX2 stopped-flow fluorimeter (TgK Scientific, Bradford-upon-Avon). DNA and enzymes were pre-mixed, and the reactions were initiated by mixing with ATP. The exponential phases of the triplex displacement profiles were fitted to a single or double exponential increase with time offset (5):

$$y = A_1 \cdot \left(1 - \exp^{-k_1\left(t - T_{lag}\right)}\right) + A_2 \cdot \left(1 - \exp^{-k_2\left(t - T_{lag}\right)}\right) \quad (1)$$

where y is the triplex displaced (%), A_1 and A_2 are the amplitudes of the phases, k_1 and k_2 are the corresponding rates of those phases, t is the time after mixing with ATP

and T_{lag} is the lag time. The variation of T_{lag} (s) with distance (d, bp) was fitted to:

$$T_{lag} = \left(\frac{1}{k_{step}} \cdot d\right) + T_i \tag{2}$$

where k_{step} is the rate of translocation, and T_i is a macroscopic reaction time that is governed by either the initiation phase of the reaction or by steps at triplex displacement. T_{lag} changes linearly with d, but T_i remains constant. The shape of the triplex displacement profile is identical if the values for the initiation rate (i.e. the rate limiting step before initiation of translocation) and the triplex displacement rate (i.e. the rate limiting step after stalling of the enzyme at the triplex) are swapped (16). Therefore, we cannot distinguish what the T_I value refers to without further analysis. However, with the triplex substrates used here, the d values were relatively large and subsequently T_i could not be reliably fitted. Consequently, these values are not reported. This does not affect the reliability of the k_{step} values.

DNA cleavage mapping assays

To produce linear DNA substrates, pInvR was digested using NdeI (for LlaBIII reactions) or pHH-3 (3) was digested using Eco109I (for LlaGI reactions). The linear DNA was dephosphorylated, purified by phenol–chloroform extraction and isopropanol precipitation and the 5′ overhangs labelled with $[\gamma^{32}P]$-ATP (Perkin Elmer) and T4 polynucleotide kinase (New England Biolabs). To distinguish the 5′–3′ strand products from the 3′–5′ strand products, the linear DNA was further digested (Figure 5): for the LlaBIII substrates, the DNA was cut with AatII (to leave a labelled 'top' strand) or with SphI (to leave a labelled 'bottom' strand); for the LlaGI substrate, the DNA was cut with XmnI (to leave a labelled 'top' strand) or NdeI (to leave a labelled 'bottom' strand). A ladder with known DNA sizes was produced by a mixing of the substrate DNA, the full-length linear DNA (FLL) produced previously, and cleavage products from digestion of the substrate DNA with XmnI, AhdI or AlwNI.

Type ISP cleavage assays contained 2 nM DNA, 4 mM ATP and 200 nM LlaGI or LlaBIII in either TMDK buffer or TMD buffer. Reactions were started by adding ATP and were incubated at 25°C for the times indicated. Samples were analysed by alkaline denaturing agarose gel electrophoresis. After electrophoresis, the gels were neutralized, compressed for 1 h, dried under vacuum for 1 h and imaged using a Typhoon phosphoimager (GE). The results were analysed using the 1D gel analysis software of ImageQuant (GE).

A 16-bit densitometric scan of each lane was recorded. To convert pixel positions down the lane scans to DNA lengths in base pairs, the known sizes of the DNA bands in the ladder were plotted against the pixel positions of the peak maxima and fitted to a one phase exponential decay:

$$y = S * \exp^{-k \cdot x} + P \tag{3}$$

where y is the DNA length in base pairs, x is the pixel position of the peak maximum, k is the decay constant, $(S + P)$ is the theoretical pixel position as $y \to 0$ and P is the theoretical pixel position as $y \to \infty$. The values returned from the fit were used to convert pixel positions in the reaction lanes to DNA lengths in bp. Given our gel electrophoresis conditions, accurate conversion could be made across the range 2.7–1.0 kb. For easier gel-to-gel comparison, the intensities of the pixels are reported as percentages normalized in relation to the peak maximum of the full-length linear band at zero time.

RESULTS

The DNA recognition site of LlaBIII, and the effect of KCl on endonuclease activity

The LlaBIII recognition sequence (5′-TnAGCC-3′) was determined using a brute force approach similar to that used previously for LlaGI (see 'Materials and Methods' section) (Figure 1B) (3). Although this sequence is one base pair shorter than that for LlaGI, it has a similar format (see 'Discussion' section). To explore the catalytic requirements of LlaBIII in more detail, we first tested DNA cleavage activity on a range of plasmid substrates (Figure 2A). Under the same reaction conditions as used previously for LlaGI, namely, TMD buffer (50 mM Tris–Cl pH 8.0, 10 mM MgCl₂, 1 mM DTT) supplemented with 4 mM ATP, we unexpectedly observed non-specific DNA nicking activity using a zero-site plasmid substrate that, at elevated LlaBIII concentrations, even led to production of linear DNA (Figure 2B, upper gel). We tested a number of alternative buffers and found that the addition of KCl removed this non-specific activity (Figure 2B, lower gel for KCl data, other data not shown). The reaction buffer for LlaBIII in the subsequent experiments, therefore, contained, in addition, 150 mM of KCl and is called TMDK. The effect of the TMDK buffer on LlaGI activity is considered below. Enzyme concentration was maintained in excess over DNA, conditions which favour DNA cleavage. In our previous analysis of LlaGI, we found that although translocation could occur at 10 nM enzyme, cleavage required >100 nM enzyme. The relationship of protein complex stoichiometry to cleavage activity is still being investigated and is not considered further here.

The efficiency of cleavage by LlaBIII in TMDK buffer on plasmid substrates with zero, one or two inverted repeat LlaBIII sites at a single time point (10 min) is shown in Figure 2C. dsDNA cleavage was only observed with two sites on the same DNA in an inverted repeat, whereas an inefficient DNA nicking activity was observed on the one-site plasmid. The interaction between two sites was further tested using plasmid and linear DNA substrates in which the pair of asymmetric sites were oriented HtH, tail-to-tail (TtT) or head-to-tail (HtT) (Figure 2D); dsDNA cleavage was only observed with the HtH arrangement as anticipated from LlaGI (3). The LlaBIII cleavage products separate as a characteristic smear on the agarose gel, representing a mixture of linear DNA of different lengths that result from cleavage in the DNA region between the HtH
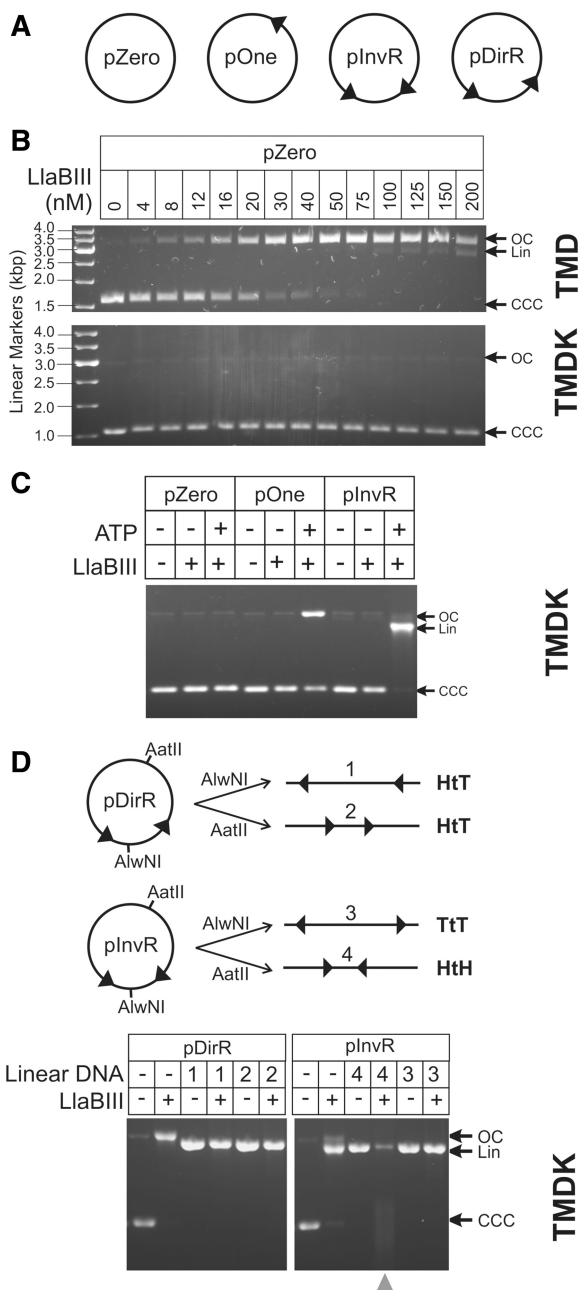
Figure 2. DNA substrate requirements for LlaBIII endonuclease activity. TMD/TMDK refers to the buffers used in the reactions (see main text). (**A**) Plasmid DNA substrates. The arrowheads represent the orientation of the LlaBIII site as in Figure 1B. (**B**) Non-specific DNA cleavage by LlaBIII in the absence of K$^+$ ions. Five nanomolars of pZero was incubated with the concentrations of LlaBIII shown for 10 min in either TMD or TMDK with 4 mM ATP and the reactions separated by agarose gel electrophoresis. CCC (covalently closed circular), open circle (OC) and Lin (linear) bands are labelled. (**C**) dsDNA cleavage requires two sites, but a one-site DNA is nicked. pZero, pOne or pInvR (5 nM) were incubated with 200 nM LlaBIII for 10 min. The reactions separated by agarose gel electrophoresis. (**D**) dsDNA cleavage requires two sites in inverted HtH repeat. Plasmid DNA, or DNA pre-linearized with AlwNI or AatII as shown (DNA 1–4) (5 nM), was incubated with 200 nM LlaBIII for 10 min with 4 mM ATP. The reactions were separated by agarose gel electrophoresis. Linear DNA contained pairs of sites in head-to-head (HtH) (DNA 4), tail-to-tail (TtT) (DNAs 1 and 2) or tail-to-tail (TtT) (DNA 3) repeat. On linear DNA, only the HtH arrangement produces cleavage (highlighted by the grey arrowhead).
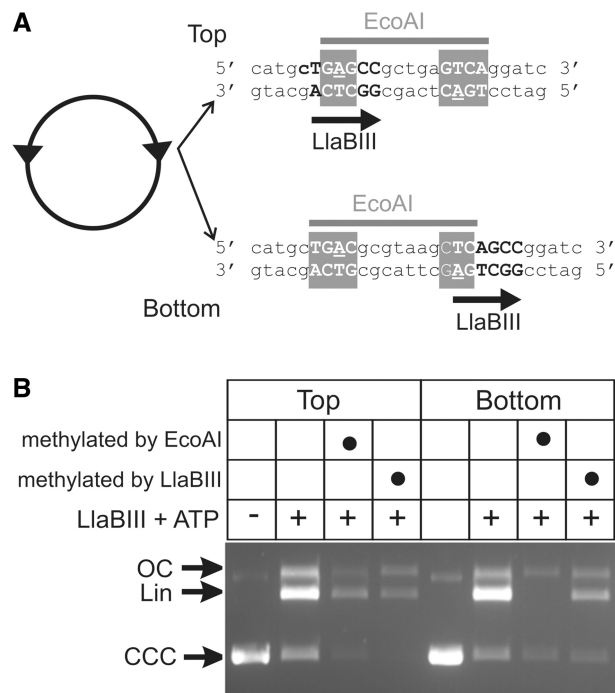
Figure 3. Determining the methylation sensitivity of LlaBIII. (**A**) Plasmid DNA substrates 'top' and 'bottom' used in the assays in which the sequence of one of the two HtH LlaBIII sites (in bold) overlapped with an EcoAI site (in grey). Adenines methylated by M.EcoAI are underlined. (**B**) Pre-incubation with EcoAI blocks LlaGI cleavage of the 'bottom' DNA but not the 'top' DNA. DNA was either used directly or pre-incubated with M.EcoAI or LlaGI in the absence of ATP to allow methylation (indicated by dots above the lanes). The DNA was then incubated with LlaBIII and 4 mM of ATP for 10 min, and the reactions were separated by agarose gel electrophoresis. See main text for full details.

sites (8,9). The distribution of cleavage sites is considered in more detail later in the text. Note that the plasmid with two sites in direct repeat (pDirR) was nicked more efficiently than the one-site plasmid; this most likely reflects the increased probability of inter- or intramolecular DNA interaction when two sites are available.

The LlaBIII site has two possible adenines that could be modified by the methyltransferase domain; one on the top strand and one on the bottom strand (Figure 1B). To test which was the methylation target, we cloned a pair of two-site HtH LlaBIII substrates, in which one of the two sites overlaps with a recognition site for the multi-subunit Type I RM enzyme EcoAI (Figure 3A). Treatment of these DNA with the methyltransferase M.EcoAI would lead to methylation at two locations within the EcoAI site (site boxed in grey, methylated adenines underlined) (17), one of which would correspond to a putative target site in the LlaBIII sequence. Treatment of the unmodified DNA with LlaBIII in the absence of ATP but presence of AdoMet should lead to methylation of one adenine (or both) within the LlaBIII sites, which, depending on the DNA substrate, would block EcoAI (hemimethylation of Type I sites being sufficient to block nuclease activity).

After incubation with M.EcoAI or LlaBIII, the DNA was purified and re-treated with LlaBIII plus ATP.

The 'top' substrate was cleaved by LlaBIII to produce linear DNA products, indicating that methylation of the adenine on the top strand (5′-Tn<u>A</u>GCC-3′) does not block LlaBIII. In contrast, the 'bottom' substrate pre-treated with M.EcoAI was only nicked by LlaBIII, similar to what would be observed on a one-site DNA. This indicates that methylation of the adenine on the bottom strand (5′-GGCTn<u>A</u>-3′) blocks LlaBIII cleavage at that site, leaving the other site unmodified and, hence, a target for nicking activity. Surprisingly, pre-treatment with LlaBIII did not block subsequent LlaBIII cleavage, suggesting that the methylation reaction is inefficient under these conditions. Similarly, LlaGI is inefficient at *in vitro* DNA methylation (3). This also meant that we could not confirm the results by repeating the experiment using EcoAI cleavage as the readout after LlaBIII methylation. Nonetheless, we are confident in our interpretation, as the location of the target adenine matches that seen for

LlaGI (Figure 1B), and it also matches a pattern seen in the Type II MmeI-family of RM enzymes (see 'Discussion' section) (17,18).

Comparison of DNA endonuclease and translocase activities of LlaBIII and LlaGI

The endonuclease activity of LlaBIII was analysed in more detail by measuring the rate of DNA cleavage using a plasmid substrate with two sites in HtH repeat and a fixed 40-fold molar excess of LlaBIII over DNA (Figure 4A, upper graph). The linear product was rapidly produced, with a peak of ∼40% nicked DNA observed in 7 s. The overall cleavage velocity of LlaBIII in TMDK buffer is similar to LlaGI in TMD buffer (3). However, LlaGI produced a slightly larger peak of nicked DNA (∼50%). Given that the reaction mechanism of the Type I enzymes is in essence a single turnover reaction
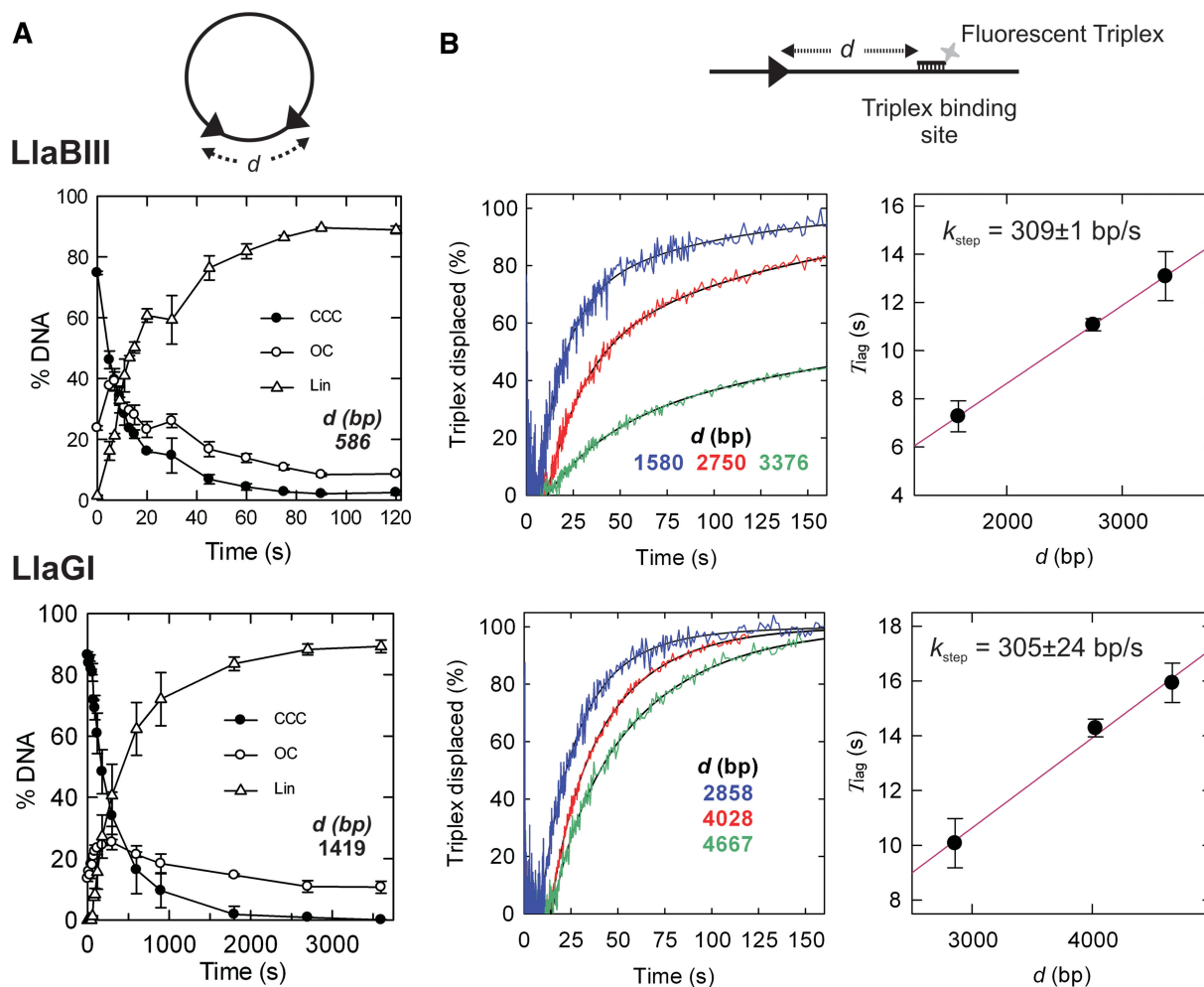
Figure 4. Cleavage and translocation kinetics for LlaGI and LlaBIII in TMDK buffer at 25°C. (**A**) Quantification of cleavage reactions using 5 nM HtH DNA, 200 nM LlaBIII or LlaGI and 4 mM ATP. The reactions were separated by agarose gel electrophoresis, and the CCC (covalently closed circular), open circle (OC) and Lin (linear) bands were quantified. Mean values and standard deviations are shown for at least two repeat experiments. (**B**) Triplex displacement reactions on linear DNA. Linear triplex DNA was pre-incubated with LlaGI or LlaBIII, and the reaction initiated by mixing with ATP, to give final concentrations of 1 nM DNA, 100 nM LlaBIII or LlaGI and 4 mM ATP. Triplex displacement profiles (left-hand graphs) were fitted to Equation 1 (solid lines) to obtain the estimated lag time (T_{lag}). The linear relationship between T_{lag} and d, (the distance between the binding site and triplex—see cartoon) was used to estimate the translocation rate, k_{step} (right-hand graphs). Mean values and standard deviations at each spacing are from at least three repeat experiments.

(the DNA substrate is saturated with a molar excess of enzyme, and multiple turnovers are not required), we suggest that the nicked intermediate is not released from the enzyme complex, but represents a bound intermediate on pathway to the dsDNA-cleaved product.

We repeated our analysis using LlaGI and TMDK buffer using a cognate HtH substrate (Figure 4A, lower graph—note the difference in the *x*-axis scale with the upper graph). Surprisingly, LlaGI showed a marked reduction in the DNA cleavage rate in the presence of K^+ ions, being ~30-fold slower than LlaGI in TMD buffer (3) or LlaBIII in the same buffer. The percentage of nicked DNA intermediate observed was also lower (~20%). This may reflect changes in the relative cleavage rates or the slowing of rates before or after cleavage that, in turn, mask the appearance of the intermediate.

To test whether LlaBIII has a directional DNA translocase activity similar to LlaGI (5), we measured the motor activity using the triplex displacement assay in TMDK buffer (Figure 4B). The translocation activity of LlaBIII (upper graph) can be seen by the characteristic distance-dependent lag before an exponential triplex displacement phase (15). A translocation rate (k_{step}) of ~309 bp/s could be estimated from the linear relationship in the lag times (Figure 4B, upper graph). In the TMDK buffer, LlaGI had a nearly identical stepping rate of ~305 bp/s (Figure 4B, lower graph). Therefore, the slower cleavage of plasmid DNA by LlaGI in the TMDK buffer cannot be accounted for by a slower rate of translocation.

We note that the exponential phases of the triplex displacement profiles are slower and more biphasic than seen previously with LlaGI in TMD buffer. The triplex substrates used in these assays had relatively long spacings between the recognition sequence and the triplex. This can affect the spatial distribution of the translocating population, which, in turn, can result in slower displacement profiles. Even so, the effect is marked for LlaBIII. This can also be explained if one assumes a relatively high off rate during translocation. As *d* is increased and the cumulative probability of dissociation also increases, then the exponential phases can become more extended as observed (15). Dissociation during translocation is also consistent with the cleavage distribution profiles, discussed later in the text.

Our previous analysis of LlaGI in TMD was undertaken at 20°C, and it gave a translocation rate of 171 bp/s (5). We repeated the analysis at 25°C, to give translocation a stepping rate of 250 ± 15 bp/s (data not shown). It appears, therefore, that both LlaBIII and LlaGI can translocate moderately more quickly in the presence of K^+ ions, yet the enzymes show striking differences in their cleavage rates. This is all the more surprising given the structural similarity in the nuclease domains (Figure 1A).

Locations of DNA cleavage are consistent with a translocation–collision mechanism for both LlaGI and LlaBIII

We previously suggested that LlaGI uses a DNA translocation/collision mechanism to cut the DNA (5). As first suggested for the multi-subunit Type I enzymes (7), such a scheme leads to collision events, and thus cleavage, at loci distant from the sites that will be located on average halfway between a pair of sites (8). However, this was not unambiguously demonstrated for LlaGI. To test this, we mixed enzyme with linear DNA substrates labelled on one or other strand at the 5′-end with phosphorus-32 and followed the time courses of DNA cleavage. Aliquots at each time point were separated by alkaline denaturing agarose gel electrophoresis alongside single-stranded DNA length markers (Figure 5A).

For LlaBIII in TMDK, a 'smear' of DNA fragments was observed centred approximately midway between the recognition sites, consistent with random cleavage dictated by stochastic translocation and collision of two enzyme complexes (Figure 5A) (9). However, the distribution was broader than expected, with cleavage seen along the full length of the intervening DNA. As a comparison, we repeated the mapping experiment using LlaGI in both TMD and TMDK buffers (Figure 5B and C). In its 'ideal' TMD buffer, LlaGI produced a more tightly distributed smear of fragments between the sites (note the larger inter-site distance for the LlaGI DNA compared with the LlaBIII DNA). In contrast, in the TMDK buffer, the LlaGI reaction resulted in a far more distributive range of cleavage products, indicating cleavage occurring at almost any location between the two sites. The linear cleavage rate was slower than in the TMD buffer, but not as slow as on the plasmid DNA in TMD (Figure 4A, lower panel), suggesting that DNA topology additionally inhibits LlaGI in the TMDK buffer.

To provide a more objective analysis of the gel data, the cleavage profiles were quantified by densitometry, with data for each strand shown in Figures 5D and E, respectively. For LlaBIII in TMDK, a broad distribution was observed covering the entire DNA region between the sites, which showed a maximum at the approximate midpoint. The peak intensity increased to a maximum at 45 s, but it decreased thereafter. For LlaGIII in TMD, the peaked profile was more tightly distributed at the midpoint, with a 95% confidence interval (CI) of $\sim \pm 300$ bp. Again the maximum peak intensity was reached early in the reaction (15–30 s) with a subsequent decrease in intensity. For LlaGI in TMDK, the distribution was markedly broader, with cleavage showing an asymmetric bimodal distribution. The decrease in intensity observed with all conditions represents further DNA processing, which is discussed in the accompanying article (12).

For the collision between two identical processive enzymes (i.e. where dissociation during translocation does not occur), one would expect a normal distribution of collision/cleavage sites with a CI of $\pm d^{1/2}$, where *d* is the distance between the sites (9). For the LlaGI DNA used here, where $d = 1412$ bp, the distribution would be expected to have 95% CI of $\sim \pm 38$ bp, whereas for the LlaBIII DNA used here, where $d = 724$ bp, the distribution would be expected to have a 95% CI of $\sim \pm 27$ bp. The empirical distributions are clearly larger than this. One simple reason is that the agarose gel method results in band broadening (19). To give an estimate of the effect on band broadening, we can estimate the 95% CI for a
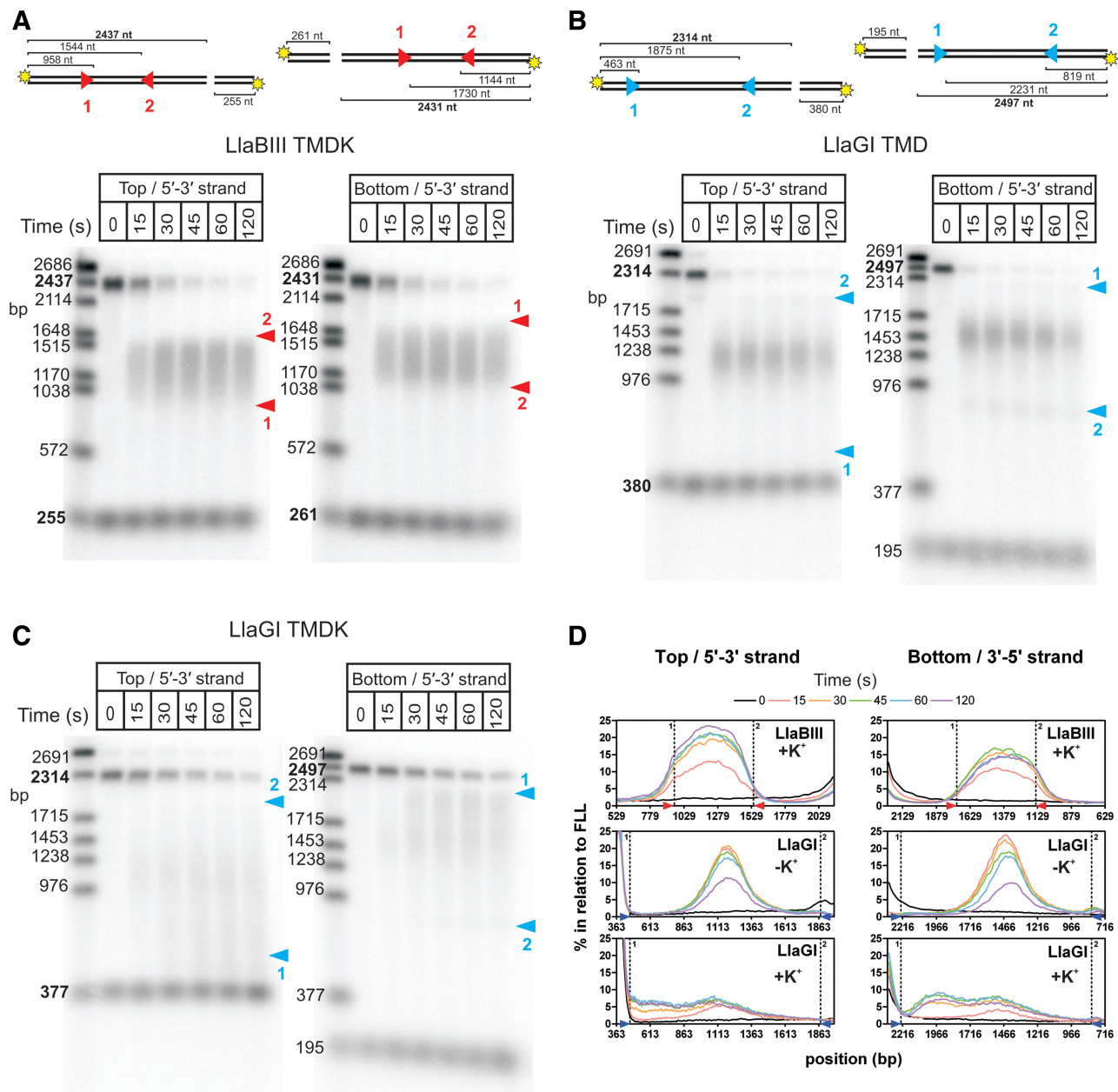
Figure 5. Mapping the positions of linear DNA cleavage by LlaBIII and LlaGI. (**A–C**) Maps of the linear DNA are shown, with the yellow sun symbols indicating the 5′-ends labelled with 32-phopshorous. The smaller 255/261-bp fragments were side products of substrate production. Two nanomolars of DNA was incubated with 200 nM of enzyme in either TMD or TMDK buffer, as indicated, with 4 mM of ATP. Aliquots were removed and quenched at the times indicated and then separated by alkaline denaturing agarose gel electrophoresis alongside labelled linear DNA markers. (**D**) Gel images from A–C were quantified using ImageQuant TL. The DNA marker lanes were used to calibrate and correct the pixel positions to DNA lengths. The y-axis intensity values were calculated as a percentage relative to the uncut FLL bands. The 'linear range' of the gel, that is, the range in which fragment sizes can be confidently estimated, is 2.7—1.0 kb (see 'Materials and Methods' section).

species of a single size using the marker lanes. Across the DNA sizes relevant to the cleavage loci, this gives approximately a 95% CI of ±132 bp for a DNA species of a single defined size. Based on this expected error range, LlaGI and LlaBIII both show broader distributions than expected for simple processive enzymes.

For non-processive enzymes (i.e. those that have a finite chance of dissociation during translocation or on collision), the distributions can be further broadened and can even become completely uniform (9). The triplex

displacement profiles for LlaBIII and LlaGI in TMDK show evidence that dissociation during translocation may occur. In addition, the distributions could be affected by static disorder in the enzyme populations, as if two enzymes converge at different speeds, the sites of collision will become more broadly distributed.

For LlaGI in TMDK, the cleavage showed an asymmetric bimodal distribution with one peak centred midway and one at a location approximately halfway between the midpoint and site 1 (Figure 5D and E). This

suggests that the two enzymes on the DNA are not acting identically, and there is some induced asymmetry. This could be because of local sequence effects during initiation, during translocation or during cleavage. Our reason for testing LlaGI in the TMDK buffer was that in the accompanying article (12), we wanted to mix both enzymes in the same reaction buffer. However, we showed that where LlaGI and LlaBIII cooperate, the cleavage kinetics and distribution matches more closely the data seen with LlaBIII alone. This suggests that the reduced rates and changes in cleavage site distribution for LlaGI alone in TMDK are because of events after collision.

DISCUSSION

The analysis of LlaBIII presented here confirms a number of enzymatic features that were also observed with LlaGI (3,5): recognition of a short asymmetric DNA sequence with methylation of only one strand; unidirectional DNA translocation dictated by the site orientation; a requirement for two sites in HtH repeat for dsDNA cleavage; an inefficient DNA nicking activity on circular DNA in the absence of a HtH interaction. We suggest that these activities will be common features of the wider type ISP family of enzymes.

We also examined the locations of DNA cleavage for both LlaGI and LlaBIII in more detail than our previous studies. This analysis shows evidence of a preference for cleavage at loci midway between the pair of HtH sites (at least under ideal *in vitro* buffer conditions). This is consistent with convergent DNA translocation, and collision being the trigger to activate nuclease activity (9). The distributions for both enzymes, and in particular for LlaBIII, were broader than expected for the collision of completely processive enzymes, even taking into account the resolution limits of the agarose gel method. There are a number of possibilities (dissociation during translocation/cleavage, static/dynamic protein disorder and so forth) that will need to be tested in a future study. The triplex dissociation profiles in the TMDK buffer are at least consistent with dissociation during translocation being a principal cause.

An unexpected observation given the similarity between the LlaGI and LlaBIII sequences (Figure 1A) was the requirement for different buffer conditions for 'ideal' *in vitro* enzyme activity. Moreover, the effects of K^+ ions were different for the two enzymes. LlaGI in TMDK had a more distributive cleavage profile, yet had translocation properties that were more similar to those observed in TMD. This could be explained if there are rearrangements of the collision complex in TMDK that move the enzymes from their original collision location, that is, the cleavage loci do not report accurately on the location of the initial translocation collision loci. In the accompanying article (12), we examine the DNA cleavage mechanism of LlaGI and LlaBIII in more detail, showing that each type ISP enzyme in a collision complex is targeted to a specific DNA strand. When LlaGI and LlaBIII cooperate to cleave DNA in TMDK buffer, we observed cleavage kinetics and distributions more akin to LlaBIII than to LlaGI. This suggests that LlaBIII may

stabilize the LlaGI protein–protein and/or protein–DNA interactions in the TMDK buffer. It is hard to say exactly what the cause of these different effects is. We and others have also noted striking effects of anions and cations on the cleavage activity of the Type III RM enzymes (20,21). Interpretation of such effects really requires detailed information on protein structure.

Although we only have the sequences of two recognition sites for the Type ISP enzymes, we note a relationship to another family of RM enzymes characterized by MmeI (17,18,22). These enzymes are ATP-independent and are classified as Type IIL (for lone strand DNA modification). They comprise a single polypeptide fusion of a PD-ExK family endonuclease and a γ-family N6-methyl adenine DNA methyltransferase. The Type IIL DNA recognition sites have common characteristics of being asymmetric, having a single target adenine for methyltaion (i.e. the sites are hemimethylated) and containing variable or unspecified nucleotide positions (17,18). dsDNA cleavage occurs outside of the recognition site. The adopted format for writing the Type IIL sequences is to show the strand where the cleavage occurs downstream of the site (Figure 6). This is also the strand containing the single methylated adenine. The Type ISP sequences can now be aligned with the Type IIL sequences using their methylated strands. Note that this differs from how we usually display Type ISP sequences (Figure 1B), where the convention is to show the opposite strand. This was chosen arbitrarily by us (3), to highlight the unidirectional translocation as occurring in a western-style
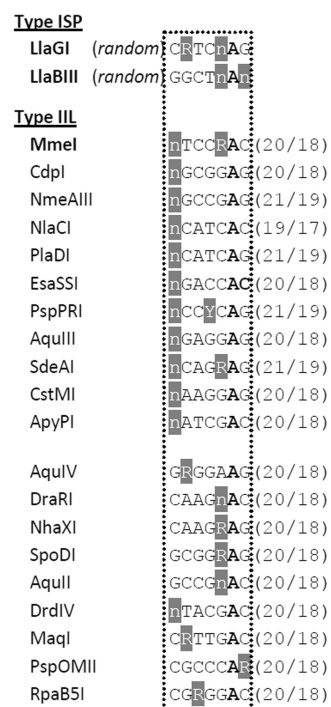
Figure 6. Similar recognition sequence formats in Type ISP and Type IIL RM enzymes. Sequences (3,17,18) are aligned according to the single methylated adenine residue (bold). Cleavage loci are indicated by the brackets (top/bottom strand positions indicated). Variable residues highlighted in grey. The Type IIL enzymes are split into two main phylogenetic branches based on amino acid sequence (18).

left-to-right direction. However, we suggest that when viewed as in Figure 6, the Type ISP and Type IIL enzymes have recognition sites with similar sequence formats and nucleotide preferences.

What is notable from the comparison in Figure 6 is that despite the similarities, the Type ISP and Type IIL sites differ in the relative locations of DNA cleavage: for the MmeI-related enzymes, the cleavage loci are downstream of the site as drawn in Figure 6, 19–21 bp on the top strand, 18–19 bp on the bottom strand; in contrast, the cleavage loci of the Type ISP enzymes are on the opposite side as drawn and in addition occur at loci, which can be many thousands of base pairs distant. Notwithstanding that the Type ISP enzymes use a translocation mechanism to communicate between the recognition and cleavage sites, one might have expected the cleavage activities to be targeted to the same side of the sites. To help understand this, a useful future analysis to understand this would be to compare at an amino acid sequence level all the Type I SP and IIL enzymes along with a wider cohort of related RM enzymes, some of which also have associated helicase subunits or domains, and which utilize different endonucleases families (e.g. GIY-YIG). Although classified as different RM Type, the domain structure of these enzymes may represent an example of highly adaptable evolution in which new RM systems can be produced by gene shuffling, facilitated by horizontal gene transfer, which is a common feature of bacterial host defence mechanisms. Further interpretation of this and of the enzymatic mechanisms will be greatly assisted by structural information on both Type ISP and Type IIL families of RM enzyme (23).

REFERENCES

1. Wilson,G.G. and Murray,N.E. (1991) Restriction and modification systems. *Annu. Rev. Genet.*, **25**, 585–627.
2. Labrie,S.J., Samson,J.E. and Moineau,S. (2010) Bacteriophage resistance mechanisms. *Nat. Rev. Microbiol.*, **8**, 317–327.
3. Smith,R.M., Diffin,F.M., Savery,N.J., Josephsen,J. and Szczelkun,M.D. (2009) DNA cleavage and methylation specificity of the single polypeptide restriction-modification enzyme LlaGI. *Nucleic Acids Res.*, **37**, 7206–7218.
4. Smith,R.M., Josephsen,J. and Szczelkun,M.D. (2009) An Mrr-family nuclease motif in the single polypeptide restriction-modification enzyme LlaGI. *Nucleic Acids Res.*, **37**, 7231–7238.
5. Smith,R.M., Josephsen,J. and Szczelkun,M.D. (2009) The single polypeptide restriction-modification enzyme LlaGI is a self-contained molecular motor that translocates DNA loops. *Nucleic Acids Res.*, **37**, 7219–7230.
6. Buchan,D.W., Ward,S.M., Lobley,A.E., Nugent,T.C., Bryson,K. and Jones,D.T. (2010) Protein annotation and modelling servers at University College London. *Nucleic Acids Res.*, **38**, W563–W568.
7. Meisel,A., Bickle,T.A., Kruger,D.H. and Schroeder,C. (1992) Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. *Nature*, **355**, 467–469.
8. Studier,F.W. and Bandyopadhyay,P.K. (1988) Model for how Type I restriction enzymes select cleavage sites in DNA. *Proc. Natl Acad. Sci. USA*, **85**, 4677–4681.
9. Szczelkun,M.D. (2002) Kinetic models of translocation, head-on collision, and DNA cleavage by type I restriction endonucleases. *Biochemistry*, **41**, 2067–2074.
10. Madsen,A. and Josephsen,J. (2001) The LlaGI restriction and modification system of *Lactococcus lactis* W10 consists of only one single polypeptide. *FEMS Microbiol. Lett.*, **200**, 91–96.
11. Kong,J. and Josephsen,J. (2002) The ability of the plasmid-encoded restriction and modification system LlaBIII to protect *Lactococcus lactis* against bacteriophages. *Lett. Appl. Microbiol.*, **34**, 249–253.
12. van Aelst,K., Sisakov,E. and Szczelkun,M.D. (2012) DNA cleavage by Type ISP Restriction-Modification enzymes is initially targeted to the 3′-5′ strand. *Nucleic Acids Res.*
13. Vipond,I.B., Baldwin,G.S., Oram,M., Erskine,S.G., Wentzell,L.M., Szczelkun,M.D., Nobbs,T.J. and Halford,S.E. (1995) A general assay for restriction endonucleases and other DNA-modifying enzymes with plasmid substrates. *Mol. Biotechnol.*, **4**, 259–268.
14. Yanisch-Perron,C., Vieira,J. and Messing,J. (1985) Improved M13 phage cloning vectors and host strains: nucleotide sequences of the M13mp18 and pUC19 vectors. *Gene*, **33**, 103–119.
15. McClelland,S.E., Dryden,D.T. and Szczelkun,M.D. (2005) Continuous assays for DNA translocation using fluorescent triplex dissociation: application to type I restriction endonucleases. *J. Mol. Biol.*, **348**, 895–915.
16. Seidel,R., Bloom,J.G., van Noort,J., Dutta,C.F., Dekker,N.H., Firman,K., Szczelkun,M.D. and Dekker,C. (2005) Dynamics of initiation, termination and reinitiation of DNA translocation by the motor protein EcoR124I. *EMBO J.*, **24**, 4188–4197.
17. Roberts,R.J., Vincze,T., Posfai,J. and Macelis,D. (2010) REBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.*, **38**, D234–D236.
18. Morgan,R.D., Dwinell,E.A., Bhatia,T.K., Lang,E.M. and Luyten,Y.A. (2009) The MmeI family: type II restriction-modification enzymes that employ single-strand modification for host protection. *Nucleic Acids Res.*, **37**, 5208–5221.
19. Shadle,S.E., Allen,D.F., Guo,H., Pogozelski,W.K., Bashkin,J.S. and Tullius,T.D. (1997) Quantitative analysis of electrophoresis data: novel curve fitting methodology and its application to the determination of a protein-DNA binding constant. *Nucleic Acids Res.*, **25**, 850–860.
20. Peakman,L.J., Antognozzi,M., Bickle,T.A., Janscak,P. and Szczelkun,M.D. (2003) S-adenosyl methionine prevents promiscuous DNA cleavage by the EcoP1I type III restriction enzyme. *J. Mol. Biol.*, **333**, 321–335.
21. Moncke-Buchner,E., Rothenberg,M., Reich,S., Wagenfuhr,K., Matsumura,H., Terauchi,R., Kruger,D.H. and Reuter,M. (2009) Functional characterization and modulation of the DNA cleavage efficiency of type III restriction endonuclease EcoP15I in its interaction with two sites in the DNA target. *J. Mol. Biol.*, **387**, 1309–1319.
22. Morgan,R.D., Bhatia,T.K., Lovasco,L. and Davis,T.B. (2008) MmeI: a minimal Type II restriction-modification system that only modifies one DNA strand for host protection. *Nucleic Acids Res.*, **36**, 6558–6570.
23. Callahan,S.J., Morgan,R.D., Jain,R., Townson,S.A., Wilson,G.G., Roberts,R.J. and Aggarwal,A.K. (2011) Crystallization and preliminary crystallographic analysis of the type IIL restriction enzyme MmeI in complex with DNA. *Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun.*, **67**, 1262–1265.