Published online 29 May 2008 *Nucleic Acids Research, 2008, Vol. 36, No. 12* ***3939–3949***
doi:10.1093/nar/gkn333

A RecB-family nuclease motif in the Type I restriction endonuclease EcoR124I

Eva Šišáková[1], Louise K. Stanley[2], Marie Weiserová[1] and Mark D. Szczelkun[2,*]

[1]Institute of Microbiology v.v.i., Academy of Sciences of the Czech Republic, Prague, Czech Republic and [2]DNA-Protein Interactions Unit, Department of Biochemistry, University of Bristol, Bristol, BS8 1TD, UK

Received April 10, 2008; Revised April 30, 2008; Accepted May 8, 2008

ABSTRACT

The Type I restriction-modification enzyme EcoR124I is an ATP-dependent endonuclease that uses dsDNA translocation to locate and cleave distant non-specific DNA sites. Bioinformatic analysis of the HsdR subunits of EcoR124I and related Type I enzymes showed that in addition to the principal PD-(E/D)xK Motifs, I, II and III, a QxxxY motif is also present that is characteristic of RecB-family nucleases. The QxxxY motif resides immediately C-terminal to Motif III within a region of predicted α-helix. Using mutagenesis, we examined the role of the Q and Y residues in DNA binding, translocation and cleavage. Roles for the QxxxY motif in coordinating the catalytic residues or in stabilizing the nuclease domain on the DNA are discussed.

INTRODUCTION

EcoR124I is a Type I restriction-modification (RM) system that uses a combination of methyltransferase, ATPase and endonuclease activities to protect bacteria from infection by phage DNA (1). Following recognition of a specific DNA sequence GAA(n)$_6$RTCG, DNA cleavage occurs at a non-specific location which can be many thousands of base pairs distant from the recognition site (2). The communication between specific binding and non-specific cleavage is catalysed by the ATP-dependent dsDNA translocase activity of the helicase motor domain of the HsdR subunit (3–5; Figure 1A). The nuclease activity is found in 'Region X', a domain N-terminal to the motor (6–9; Figure 1A; E.S., unpublished data). Region X can be classified as part of the PD-(E/D)xK superfamily (10,11), which includes not only Type I, II and III RM enzymes, but also a great many repair and recombination nucleases. The conserved structural unit is an αααββα core that serves as a scaffold for more-weakly conserved catalytic residues (11), principally the charged amino acids clustered in Motifs I, II and III. Variations in the precise arrangement of the residues can occur (12), but the overall mechanistic outcome is the same: phosphodiester bond hydrolysis is catalysed in a metal ion-assisted mechanism, which is characterized by inversion of configuration at the scissile phosphorous (13).

The weak sequence conservation of the PD-(E/D)xK motifs has historically made identification of family members difficult. Previous studies, undertaken when a more limited number of HsdR primary sequences were known, identified Motifs II and III, and subsequent mutagenesis studies confirmed the fundamental importance of the catalytic residues from these motifs (6–9), consistent with the well-established roles in the Type II RM enzymes (13). More recently, Kneale and co-workers were able to identify Motif I and produced a 3D model of the nuclease fold based on the Type II restriction endonuclease NgoMIV (14). We have extended the analysis of Region X and have been able to identify a fourth motif, located in a region of predicted α-helix immediately C-terminal to Motif III and with the conserved sequence QxxxY (Figure 1; see Results section). This motif is characteristic of the RecB-family of nucleases (10). Whilst the QxxxY sequence was not highlighted in the recent study by Kneale and co-workers (14), Dryden and co-workers have previously presented alignments of HsdR sequences that indicated that the tyrosine residue at least is conserved across HsdR subunits (8). A subset of these sequences, including EcoR124I HsdR, also retained the glutamine residue.

What is the role of the QxxxY motif in the RecB nucleases? *Escherichia coli* RecB, the archetypical member of the family, is a helicase-nuclease fusion that in combination with RecC and RecD forms a DNA-end processing machine that plays a central role in homologous recombination (15). The RecB nuclease domain cleaves both unwound ssDNA strands that are produced by the dual action of the antipolar RecB and RecD helicases (Figure 2A). The RecBCD crystal structure shows the RecB nuclease domain sitting over an exit

*To whom correspondence should be addressed. Tel: +44 117 331 2158; Fax: +44 117 331 2168; Email: mark.szczelkun@bristol.ac.uk
Present address:
Louise K. Stanley, Institute of Human Genetics, Molecular Genetics, International Centre for Life, Central Parkway, Newcastle Upon Tyne, NE1 3BZ, UK

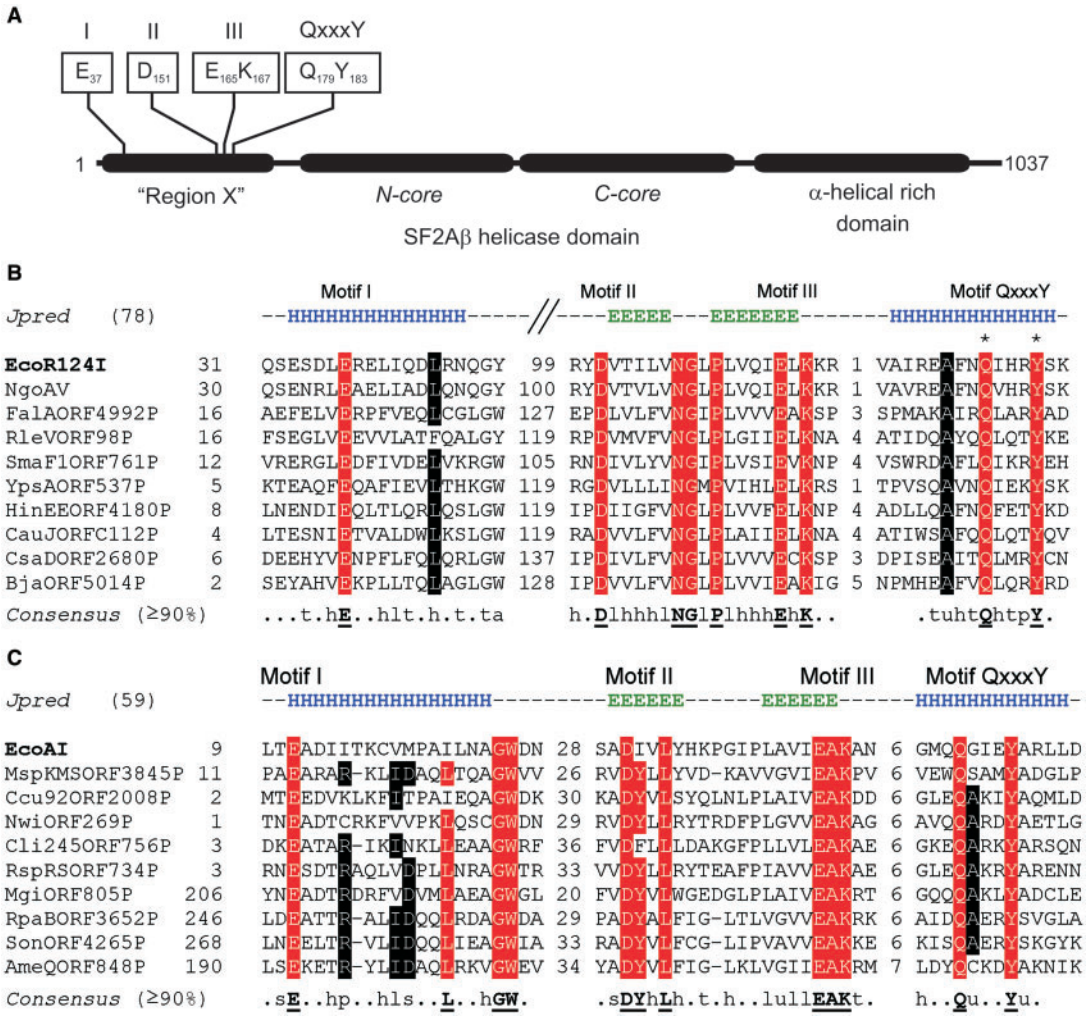

Figure 1. Nuclease motifs in the HsdR subunit of Type I enzymes. **(A)** Cartoon of the EcoR124I HsdR subunit with domains indicated (37). The locations of PD-(ExK) nuclease superfamily I motifs I, II, III and QxxxY are shown with residues highlighted that are conserved across all HsdRs examined. The N- and C-core of the helicase domain relate to the tandem RecA folds of the motor (37,38). SF2Aβ indicates a superfamily 2, dsDNA motor with 3′–5′ polarity (5,38). The α-helical domain is believed to be involved in protein–protein interactions with the MTase core (14). **(B)** Multiple alignment of representative HsdR sequences of the EcoR124I-family (10/78 sequences shown). **(C)** Multiple alignment of representative HsdR sequences of the EcoAI-family (10/59 sequences shown). Amino acid sequence alignments were carried out using ClustalW (21). Target sequences have been filtered to remove redundant sequences and truncated HsdRs. Secondary structure predictions were made by JPred using the full alignment (24). Consensus sequences were also calculated using the full alignments. For further details, please refer to Materials and methods section. Each protein is identified by the gene name as given in REBASE (22). Numbers to the left indicate the starting residues. Poorly conserved regions are replaced by numbers indicating the number of residues omitted. Red highlighting indicates sequences with >95% identity, black highlighting indicates sequences with >85% identity. The predicted secondary structure is shown above the alignments with H for an α-helix and E for a β-strand. The ≥90% consensus is shown below each alignment with the following key: *t*, turnlike (ACDEGHKNQRST); *h*, hydrophobic (ACFGHIKLMRTVWY); *l*, aliphatic (ILV); *a*, aromatic (FHWY); *u*, tiny (AGS); *p*, polar (CDEHKNQRST) and, *s*, small (ACDGNPSTV). Residues in bold and underlined are 100% conserved. Residues mutated in this study are marked with an asterisk.

channel in RecC (Figure 2A and B). The hypothetical route of the 3′–5′ strand passes through this channel and directly over the RecB nuclease (15; Figure 2A). Motifs I, II and III are arranged closely in 3D space, with Motifs II and III coordinating a Ca^{2+} ion in the crystal structure (Figure 2C). The Q1110 and Y1114 residues of the QxxxY motif point towards this cluster of residues.

A role for the conserved glutamine has not been proposed and the effect of mutations at this residue have yet to be reported. Aravind et al (10) first proposed that the conserved tyrosine might play a central catalytic role by forming a covalent intermediate with the cleaved DNA strand. However, this now appears less likely for a number of reasons: Firstly, Julin and co-workers have suggested that mutations in Y1114 do not have a significant detrimental effect on DNA catalysis (17). A similar conclusion was drawn using equivalent mutations in *E. coli* RecE (an ATP-independent 5′–3′ exonuclease), although substitution with asparagine reduced activity to 1% of wild-type (19); Secondly, numerous studies of Type II RM enzymes have shown that the three canonical PD-(E/D)xK motifs are sufficient for endonuclease activity. The arrangement of Motifs I, II and III in RecB is analogous, consistent with a similar mechanism of DNA cleavage, i.e. *without* the formation of a covalent intermediate (13,17). Mutations in Motifs II and III of RecB and related

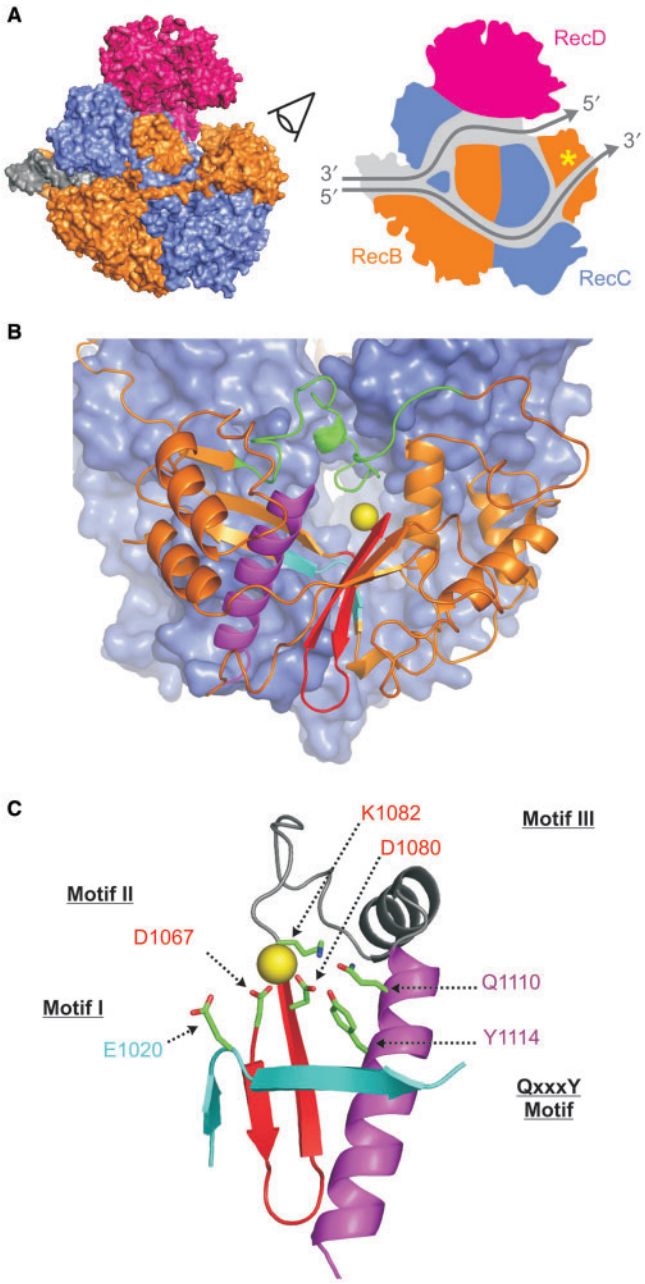

Figure 2. The RecB nuclease domain and location of the QxxxY motif. (**A**) Structure of RecBCD. The left panel shows a space filling model of RecBCD taken from the crystal structure (PDB 1W36; 15) with RecB in orange, RecC in blue, RecD in magenta and the DNA in grey. The eye symbol shows the orientation of the RecB nuclease view in (**B**). The right panel shows a cartoon representation of a cross-section through RecBCD showing the putative path of the separated DNA strands, coloured as in the left panel (15). The yellow asterisk indicates the location of the RecB nuclease domain. Note that the nuclease domain can also introduce strand breaks into the ssDNA exiting from the RecD channel. (B) View of the RecB nuclease domain of RecBCD viewed towards the RecC channel (equivalent to 'Figure 4b' in ref. 15), with RecB shown in cartoon form and the RecC domain in blue space-filling form. Structural elements that contain the Motifs are coloured: I (cyan), II/III (red) and QxxxY (purple). The bound Ca^{2+} ion is shown as a yellow sphere. Residues 909-930 that block the tunnel from RecC are coloured green (15). The region of RecB linking Motifs II/III and QxxxY (1083-1106) is omitted for clarity. (**C**) The coloured structural elements and conserved residues of the *E. coli* RecB nuclease motifs from (B) are shown, rotated by 180° (i.e. viewed from the RecC tunnel).

recombinases have confirmed their central role in catalysis (16–20). These observations suggest that the conserved QxxxY motif plays a secondary, yet important, role in the activity of RecB-family nucleases.

Here we extend the analysis of the QxxxY motif by examining its role in the Type I RM enzyme EcoR124I. We find that mutation of the glutamine and/or tyrosine residues alters DNA-binding affinity and reduces the rate and efficiency of DNA cleavage. Two distinct roles for the QxxxY motif are proposed and their consequences for DNA cleavage discussed.

MATERIALS AND METHODS

HsdR sequence analysis

Amino acid sequence alignments were carried out using ClustalW (21; http://www.ebi.ac.uk/Tools/clustalw/index.html). HsdR sequences were obtained from REBASE (22). Sequences that were shorter than the EcoAI HsdR minimal template were rejected. Two hundred and forty representatives were identified as potentially carrying a QxxxY motif. Preliminary alignments identified two main families, characterized by either EcoR124I or EcoAI; further analysis treated these groups separately. Manual realignments were made using Jalview (23). The final alignments were filtered to remove redundancy and were submitted to JPred (24) to predict average amino acid secondary structure. Consensus sequences were calculated at http://coot.embl.de/Alignment//consensus.html.

HsdR mutagenesis and protein purification

Wild-type HsdR was expressed and purified from pACR124 as described in ref. (1). QxxxY mutants were generated from pACR124 with QuikChange site-directed mutagenesis (Strategene, La Jolla, CA, USA) using the following oligodeoxynucleotide primers (where the mutated residues are in bold and the mutated codon underlined): Oli#259F (5′-TCGTG AGGCTTTCAAC**GC**GATACATCGTTACAG-3′) and Oli#259R (5′-CTGTAACGATGTATCGCGTTGAAAG CCTCACGA-3′) were used to produce the Q179A mutation; Oli#260F (5′-TCGTGAGGCTTTCAAC**A**AGATA CATCGTTACAG-3′) and Oli#260R (5′-CTGTAACGAT GTATCTTGTTGAAAGCCTCACGA-3′) were used to produce the Q179K mutation; Oli#261F (5′-CCAGATA CATCGT**GC**CAGTAAAGAGAGTTTTAACAGCG-3′) and Oli#261R (5′-CGCTGTTAAAACTCTCTTTACTG GCACGATGTATCTGG-3′) were used to produce the Y183A mutation; and, Oli#262F (5′-CCAGATACAT CGTT**T**CAGTAAAGAGAGTTTTAACAGCG-3′) and Oli#262R (5′-CGCTGTTAAAACTCTCTTTACTGAAA CGATGTATCTGG-3′) were used to produce the Y183F mutation. Primers Oli#263F (5′-TCGTGAGGCTTTCAA C**GC**GATACATCGT**GC**CAG-3′) and Oli#263R (5′-CTG GCACGATGTATC**GC**GTTGAAAGCCTCACGA-3′) were used to mutate pACR124(Y183A) to produce the Q179A,Y183A double mutant. All mutants were fully sequenced, and then expressed and purified as for the wild-type enzyme. EcoR124I MTase was expressed and purified from pJS4M as described previously (1).

DNA-binding assays

The extent of DNA binding was assayed using an electrophoretic mobility shift assay (EMSA). A 39 bp ^{32}P-labelled dsDNA substrate containing the EcoR124I site (25) was made by annealing oligodeoxynucleotide R124-bind1 (5′-^{32}P-CTACGGTACCGAAACGCGTGTCGGGCCCGCGAAGCTTGC-3′) with a 1.5-fold molar excess of R124-bind2 (5′-GCAAGCTTCGCGGGCCCGACACGCGTTTCGGTACCGTAG-3′) (EcoR124I recognition sequence underlined). A Hoefer SE 600 Ruby vertical slab gel system (GE Healthcare, Amersham, Bucks, UK) was used for native gel electrophoresis. This system includes a heat exchanger which, when connected to a circulating water bath, allows the temperature in the lower buffer chamber, and therefore the gels, to be maintained at a constant temperature. $1.5 \times 18 \times 16$ cm^3 gels [5% (w/v) acrylamide (37.5:1), 134 mM Tris-borate, 2.55 mM EDTA] were pre-run at 30 W (constant power) and at 5°C for >45 min. A total of 20 µl samples containing 5 nM dsDNA, and MTase and HsdR as indicated in the text, were incubated in buffer R (50 mM Tris–Cl, 10 mM $MgCl_2$, 1 mM DTT, pH 8.0) for >5 min at 25°C. (Longer incubation times did not affect the extent of DNA binding—data not shown). In turn, each sample was mixed with 4 µl of loading buffer (50 mM Tris–Cl, 10 mM $MgCl_2$, 1 mM DTT, 30% (w/v) sucrose, pH 8.0) and immediately loaded onto the gels at 40 V (constant voltage). Once all samples were loaded, electrophoresis was continued at 40 W (constant power) and 5°C, until the dye band had travelled ~12 cm. Gels were fixed in 40% (v/v) methanol, 5% (v/v) acetic acid for >30 min before being dried onto DE81 chromatography paper (Whatman) in a Biorad Model 583 system set on the 'normal cycle' and 80°C.

Dried gels were visualized using a Molecular Dynamics Typhoon PhosphorImager (GE Healthcare). The 16-bit images were analysed without editing and using a linear intensity scale in ImageQuant TL v2005 (GE Healthcare). To quantify the data, we first considered a simplified four-state model:

$$[D] + [M] \underset{K_{DM}}{\longleftrightarrow} [D \cdot M] + [R] \underset{K_1}{\longleftrightarrow} [D \cdot M \cdot R_1] + [R] \underset{K_2}{\longleftrightarrow} [D \cdot M \cdot R_2] \qquad \textbf{Scheme1}$$

where D is the free DNA, M is the free M_2S_1 (MTase) complex, R is the free HsdR, $D{\cdot}M$ is the DNA-MTase complex, $D{\cdot}M{\cdot}R_1$ is the DNA–$R_1M_2S_1$ complex, $D{\cdot}M{\cdot}R_2$ is the DNA–$R_2M_2S_1$ complex, K_{DM} is the dissociation constant for the DNA–MTase complex, K_1 is the dissociation constant for the DNA–$R_1M_2S_1$ complex and K_2 is the dissociation constant for the DNA–$R_2M_2S_1$ complex. However, we were unable to resolve a clear MTase–DNA band in our EMSA assays. Therefore, the occupancy of the R_1 and R_2 species were calculated in ImageQuant TL as a fraction of the total input DNA concentration from the 'volume' of the R_1 and R_2 bands relative to the total volume of that lane. Apparent values of K_1 and K_2 were then estimated by global fitting of the R_1 and R_2 species only by non-linear least squares regression in Scientist (v2.01 MicroMath Scientific Software, Salt Lake City, UT, USA). The Scientist software requires the programming of a model file describing the system (26). Based on Scheme 1, but taking into account that we could not resolve the DNA–MTase complex, the following model was used:

$$\text{IndVars}: E_{tot}$$

$$\text{DepVars}: E_f, D_f, DE, D2E$$

$$\text{Params}: K_1, K_2, D_{tot}$$

$$DE = D_f \times E_f / K_1$$

$$D2E = DE \times E_f / K_2$$

$$E_{tot} = E_f + DE + (2 \times D2E)$$

$$D_{tot} = D_f + DE + D2E0 < E_f < E_{tot}$$

$$0 < D_f < D_{tot} \qquad \textbf{Scheme 2}$$

Where D_f is the concentration of free DNA and DNA–MTase complexes, D_{tot} is the total DNA concentration, DE is the DNA–$R_1M_2S_1$ complex concentration, $D2E$ is the DNA–$R_2M_2S_1$ complex concentration, E_f is the free HsdR concentration and E_{tot} is the total HsdR concentration. D_{tot} was taken as 5 nM and fixed during the fitting. Initial estimates for $K_{1,app}$ and $K_{2,app}$ were obtained using a simplex fit within Scientist. The fitted profiles are shown in the Supplementary Figure 1, the $K_{1,app}$ and $K_{2,app}$ values in Table 1.

Table 1. DNA translocation, binding and cleavage by the wild type and QxxxY mutant HsdRs

	DNA binding (−ATP)		Translocation (+ATP)		DNA cleavage (+ATP)		
	$K_{1,app}$ (nM)	$K_{2,app}$ (nM)	k_{step} (bp/s) (±SE)	$k_{ini,app}$/s (±SE)	Relative 1st strand rate	Lag offset (S) (±SE)	Amplitude (%) (±SE)
wt	15	900	599 ± 5	2.5 ± 0.1	1.00	10.3 ± 1.5	94.8 ± 0.3
Q179A	12	1500	601 ± 8	2.6 ± 0.2	0.51	25.0 ± 1.6	81.2 ± 0.4
Q179K	10	1100	611 ± 19	2.6 ± 0.5	NC		
Y183F	14	2100	621 ± 11	2.2 ± 0.2	0.87	17.2 ± 3.1	92.5 ± 0.8
Y183A	15	2700	596 ± 20	2.1 ± 0.3	0.42	35.7 ± 1.3	66.8 ± 0.6
Q179A Y183A	20	3100	684 ± 38	1.8 ± 0.4	NC		

DNA binding data from EMSA assays (Figure 3, Supplementary Figure 1, Materials and methods section). Translocation data from triplex displacement assays (Figure 4, Materials and methods section). DNA cleavage data from the agarose gel assays (Figure 5, Supplementary Figure 2, Materials and methods section). The relative 1st strand cleavage rates (CCC disappearance) were determined by estimating the initial rates of DNA cleavage using Equation 1. and then normalizing these rates relative to the wild type rate. NC, no cleavage; SE, standard error from least squares regression analysis.

DNA translocation assays

Triplex displacement measurements were carried out in an SF61-DX2 stopped-flow fluorimeter (TgK Scientific, Bradford on Avon, UK) as described previously (4,27).

Temperature was maintained at 25.0 ± 0.1°C by a water-bath connected to the chamber housing the syringes and flow cell. Linear DNA substrates were produced by digestion of pLKS5 with ApaI (28). Final reaction conditions were 1 nM linear DNA (0.5 nM tetramethylrhodamine triplex), 30 nM MTase, 120 nM HsdR and 4 mM ATP in Buffer R. Lag times were estimated in GraFit 5.0.8 (Erithacus Software, UK) by fitting the profiles to a triple-exponential relationship:

$$y = A_1 \cdot (1 - e^{-k_1(t-T_{\text{lag}})}) + A_2 \cdot (1 - e^{-k_2(t-T_{\text{lag}})}) + A_3 \cdot (1 - e^{-k_3(t-T_{\text{lag}})}) \quad \mathbf{1}$$

where k_n and A_n are the rate and amplitude, respectively, of the n-th phase and T_{lag} is the sum of all time constants for the initiation and translocation steps (4). k_{step} (the translocation speed) and $k_{\text{ini,app}}$ (the initiation speed) were then determined from the linear relationship between T_{lag} and distance (d bp):

$$T_{lag} = \left[\left(\frac{1}{k_{\text{step}}}\right) \cdot d\right] + \left(\frac{1}{k_{\text{ini}}}\right) \quad \mathbf{2}$$

DNA-cleavage assays

Cleavage reactions contained 5 nM pLKS5 (28), 40 nM MTase and 100 nM HsdR in Buffer R at 25°C. Reactions were started by addition of ATP to a final concentration of 4 mM. Reaction aliquots were quenched after the required incubation times by addition of 0.5 volumes of STEB [0.1 M Tris–Cl, pH 7.5, 0.2 M EDTA, 40% (w/v) sucrose, 0.4 mg/ml bromophenol blue]. The covalently closed circular DNA substrate (CCC), open circle/nicked intermediate (OC) and full length linear product (FLL) were separated by agarose gel electrophoresis and the percentage of DNA in each band evaluated by scintillation counting (29).

DNA cleavage by Type I enzymes cannot be described by simple exponential functions as there are multiple steps leading to cleavage; initiation, translocation and collision of two motors are required for each dsDNA break (Frank Peske, Sarah McClelland and M.D.S., unpublished data). In order to compare the cleavage rates of the wild-type and mutant enzymes, we analysed the apparent rate of CCC cleavage using:

$$y = A \cdot \left(1 - e^{-k_{cut,\,app} \cdot (t - offset)}\right) \quad \mathbf{3}$$

Where y represents the appearance of the first strand break in CCC, A is the percentage of starting CCC cut, $k_{\text{cut,app}}$ is the apparent cleavage rate and *offset* is a time lag that represents the continuum of initiation/translocation/collision states prior to cleavage. Experimental y-values were calculated by subtracting the CCC concentration at each time point from the starting concentration and normalizing the result to a 100% scale. The cleavage of the CCC DNA does not go to 100% because of a background inhibition activity that competes with the translocation/cleavage process (F.P. and M.D.S., unpublished data). The data were fitted to Equation 3 with A, $k_{\text{cut,app}}$ and *offset* allowed to float. The resulting profiles are shown in the Supplementary Figure 2. Cleavage rates are reported in Table 1 as relative values normalized to the wild-type rate, *offset* and A values are reported directly.

RESULTS

Identification of a RecB-family nuclease motif in Type I HsdR subunits

We aligned the primary amino acid sequences around Region X of characterized and putative HsdR subunits from REBASE (22; Materials and methods section). Of more than 200 sequences examined, >80% contained a putative QxxxY motif. From these, we identified two distinct sets of sequence alignments, one exemplified by EcoR124I (Figure 1B), the other exemplified by EcoAI (Figure 1C). As well as differences in the primary amino acid sequences, the two classifications show differences in the locations of Motif I relative to Motifs II and III. Within the EcoR124I-related sequences, Motif I resides in a region of predicted α-helix at the extreme N-terminus of Region X, separated from Motif II by an ~117 amino acid (aa) linker. Within the EcoAI-related sequences, Motif I resides in a region of predicted α-helix immediately adjacent to the predicted ββα secondary structure core containing Motifs II, III and QxxxY. A subset of these sequences, for example EcoAI itself (Figure 1C), show extremely compact HsdR structures. The other subset has large (>190 aa) extensions that are N-terminal to the nuclease motifs; the role of these additional domains is unknown. It is worth noting that EcoAI displays all the activities of a typical Type I enzyme (4,30) and thus represents the minimal, indispensable HsdR structure—additional domains seen in other sequences are likely to have secondary roles. Given the extensive characterization of EcoR124I (1–5), we chose this HsdR for further analysis.

Mutagenesis of the QxxxY motif of EcoR124I and analysis of DNA-binding activity

The Q179 and Y183 residues of EcoR124I HsdR were mutated in the expression plasmid pACR124 using the 'QuikChange' mutagenesis protocol (Materials and methods section). A double mutant (Q179A,Y183A) was produced by two consecutive rounds of mutagenesis. The wild-type and mutant HsdRs were purified as described (Materials and methods section); all the mutants behaved as wild-type during expression and purification (not shown). The complete RM enzymes were reconstituted by mixing the separately purified HsdR and MTase; reconstitution of Type I RM systems from separate protein preparations produces enzyme activity equivalent to that of the holoenzyme (27,31).

We first compared the ability of the mutant HsdRs to assemble onto a DNA-bound MTase. This was carried out by incubating a fixed concentration of ^{32}P-labelled dsDNA substrate and MTase with different concentrations of HsdR in the absence of ATP, and then separating the unbound DNA from the different protein–DNA species by an EMSA. EcoR124I can form three distinct and separable protein–DNA complexes: MTase–DNA;

R_1–MTase–DNA and R_2–MTase–DNA. In the absence of ATP, the binding of the first HsdR is relatively efficient ($K_{1,app} << 1$ nM) whereas the binding of the second HsdR is considerably weaker ($K_{2,app} = \sim 240$ nM). The DNA substrate and conditions used (Figure 3A; Materials and methods section) were duplicated from the earlier study by Janscak *et al.* (25). This allowed us to compare our binding data with an independent study. Currently, the assembly of the different species cannot be directly measured using EMSA under the same conditions as DNA translocation or cleavage (i.e in the presence of ATP). It should therefore be noted that whilst the binding affinities measured here are comparative with respect to the wild-type and mutant proteins, they may not reflect the binding affinities during translocation and cleavage (27).

An example EMSA for Y183F HsdR is shown in Figure 3B. Although we could resolve the free DNA and R_1 and R_2 complexes, we were unable to resolve clearly the MTase–DNA complex. In our hands it appears that DNA binding by the EcoR124I MTase in the absence of HsdR cannot be efficiently caged in the EMSA assay. In contrast, binding by HsdR appears to stabilize the MTase–DNA complex. Quantification of the relative levels of free DNA, and R_1 and R_2 complexes for the wild-type and all mutant proteins are shown in Figure 3C.

A simple model (Scheme 2, Materials and methods section) was used to fit the data to obtain apparent dissociation constants for the R_1 and R_2 species (Supplementary Figure 1, Table 1). The binding of the first HsdR ($K_{1,app}$) is largely similar for all HsdRs. More significant, systematic differences were observed upon binding of the second HsdR ($K_{2,app}$). Mutations at Q179 produced a 1.2- to 1.7-fold reduction in affinity, whilst mutations at Y183 produced a 2.3- to 3-fold reduction in affinity. The double mutant (Q179A,Y183A) showed the biggest effect (>3.4-fold reduction in affinity). As pointed out in the Supplementary section and Materials and methods section, Scheme 2 was unable to completely describe our observed binding data and therefore the values we obtained can only be used as relative/comparative values and not absolute values. Nonetheless, a clear trend is apparent—mutation of the QxxxY motif destabilizes the assembly of the R_2 complex, with the biggest contribution coming from the Y183 residue. Since formation of the R_2 complex is required for DNA cleavage (1), then it is likely that these mutations will also affect DNA cleavage efficiency.

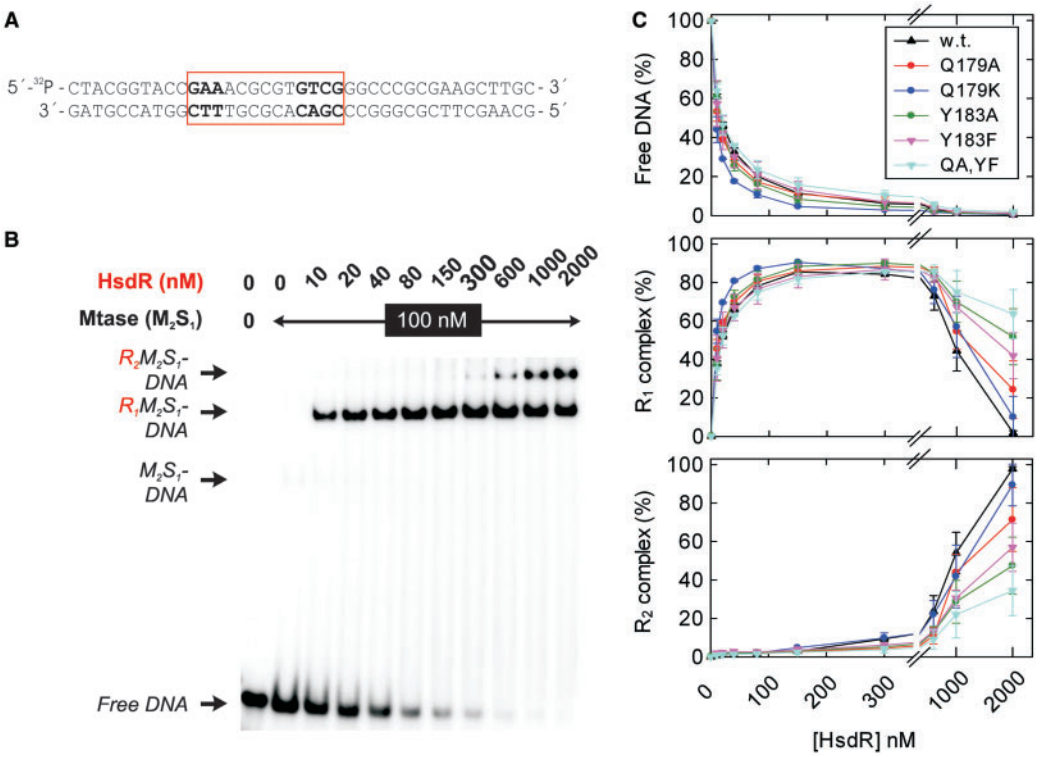

Figure 3. Assembly of EcoR124I on DNA in the absence of ATP. (**A**) Sequence of the dsDNA used in the EMSA assays (25). The bipartite EcoR124I recognition sequence is boxed. (**B**) Example EMSA gel for Y183F HsdR. 5 nM ^{32}P-labelled dsDNA was incubated with MTase (0 or 100 nM) and HsdR (0–2000 nM) as indicated for >5 min at 25°C. Samples were then separated on a 5% (w/v) native acrylamide gel at 5°C (to prevent dissociation of the bound species during electrophoresis). The contrast of the image in the figure has been increased to enhance clarity. Densitometric analysis was carried out on uncorrected images (16-bit linear scale). Full details of the assay and analysis of the gels is given in Materials and methods section. (**C**) The percentage of free DNA (*top graph*), R_1 complex (*middle graph*) and R_2 complex (*bottom graph*) was calculated for each of the HsdRs indicated. Points are the average of 2–3 repeat experiments, error bars are the standard deviations of the averages. Lines show the linear connections between the points. Fits of the data to a simple binding model (Scheme 2) are shown in the Supplementary Figure 1. Estimates derived for the apparent R_1 and R_2 dissociation constants based on Scheme 2 are given in Table 1.

Translocation assays with QxxxY mutants

DNA cleavage by Type I enzymes requires DNA translocation by the helicase motor. Therefore, any mutation that affects the translocase activity could alter DNA cleavage activity without necessarily changing the capacity for DNA hydrolysis (i.e. the delivery of the nuclease would be affected rather than the nuclease activity *per se*). To determine if the QxxxY mutants had altered DNA translocation properties, we assayed the motor activity using the triplex displacement assay (4; Figure 4A). Linear DNA substrates were bound by one of four fluorescent-labelled triplexes (28) and pre-incubated with an excess of EcoR124I MTase and HsdR, sufficient to saturate the DNA recognition sites with R_2 complex (Materials and methods section). Translocation reactions were initiated by addition of ATP to 4 mM, and the fluorescence monitored. A typical profile is shown (for Q179K, Figure 4B). For each triplex substrate, a clear lag in the kinetics was observed before a characteristic increase in fluorescence that corresponds to the displacement of the triplex (4). The lag times (T_{lag}) were derived by fitting each profile to Equation 1 (Materials and methods section). The rate of initiation ($k_{ini,app}$/s) and of translocation (k_{step}/s) could then be derived from the linear relationship between the triplex distance and T_{lag} (e.g. Figure 4B, *inset*). The k_{step} and $k_{ini,app}$ values derived in this way for the wild-type and each of the HsdR mutants are given in Table 1.

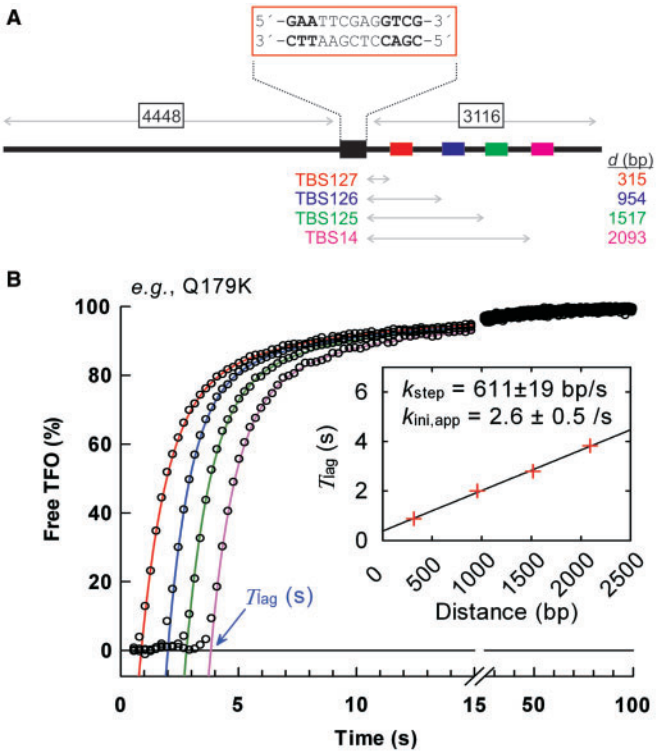

Figure 4. DNA translocation by a QxxxY mutant. (**A**) DNA substrate used for the translocation assays. pLKS5 (28) was cut at one site using ApaI to generate the linear DNA shown. By using different TFOs in each translocation assay (TFOs 127, 126, 125 and 14 bound to TBSs 127, 126, 125 and 14, respectively; 28), spacing between the EcoR124I site and the triplex could be varied as shown. (**B**) Example stopped flow traces using Q179K HsdR. 1 nM DNA (0.5 nM triplex) was pre-incubated with 30 nM MTase and 120 nM Q179K HsdR and the reaction initiated by rapid mixing with ATP (see Materials and methods section). Four traces are shown corresponding, from left to right, to spacings of 315 bp (TFO127, red), 954 bp (TFO126, blue), 1517 bp (TFO125, green) and 2093 bp (TFO14, magenta). Each profile was fitted by least squares regression to Equation 1 to give T_{lag}. (*inset*) Graph showing relationship between T_{lag} and triplex spacings. The line shows a fit by least squares regression to Equation 2 to give the translocation rate and apparent initiation rate. The standard errors in the T_{lag} are smaller that the data points and smaller than the scatter between the points. Rates for the wild-type and all mutant proteins are given in Table 1.

With the exception of the Q179A,Y183A double mutant, all the HsdR mutants showed translocation kinetics within experimental error of the wild-type HsdR. The triplex displacement profiles were all consistent with wild-type-like translocation profiles. The data suggest that mutation of the QxxxY motif does not alter the translocation properties of EcoR124I significantly. Because of the larger experimental error in the Q179A,Y183A data, we cannot explicitly state whether or not this mutant was a slightly faster motor. The overall conclusion from these data is that all the proteins are still capable of translocating and that gross changes in translocation mechanism of the R_2 complex have not occurred.

DNA cleavage by QxxxY mutants is slowed

To test the endonuclease activity of the QxxxY mutants, we followed DNA cleavage activity on a tritium-labelled, single-site plasmid using conditions that were saturating with respect to the HsdR concentration (29; Materials and methods section). Reactions were pre-incubated with DNA, MTase and HsdR, and then initiated by the addition of ATP to 4 mM. Aliquots were removed from the reaction at the time points indicated, quenched immediately, and the quenched reactions separated by agarose gel electrophoresis. This allowed us to resolve the substrate (supercoiled plasmid), intermediates (nicked plasmid) and products (linear DNA) of the cleavage reaction. The percentage of each species was then calculated by scintillation counting of the individual bands (29). Time courses are shown in Figure 5.

DNA cleavage by wild-type EcoR124I shows a characteristic profile (29; Figure 5A). The first cleavage event (measured as disappearance of the supercoiled substrate) occurs within ~2 min (under conditions used here). A peak of ~38% nicked DNA is observed, which could be an on- or off-pathway intermediate, followed rapidly by the linear DNA cut in both strands. DNA cleavage does not go to completion; ~22% DNA remains cut in only one strand, ~4% remains intact. Very little change in the relative levels of the species is observed beyond ~5 min. This inhibition of cleavage appears to be dependent upon changes to the DNA that make it resistant and is not dependent upon inhibition of the protein (the exact nature of the inhibition of DNA cleavage is under investigation; F.P., unpublished data).

What are the reaction steps that lead to this kinetic profile? Cleavage of a one-site circular DNA is believed to result from the collision of HsdRs originating from the

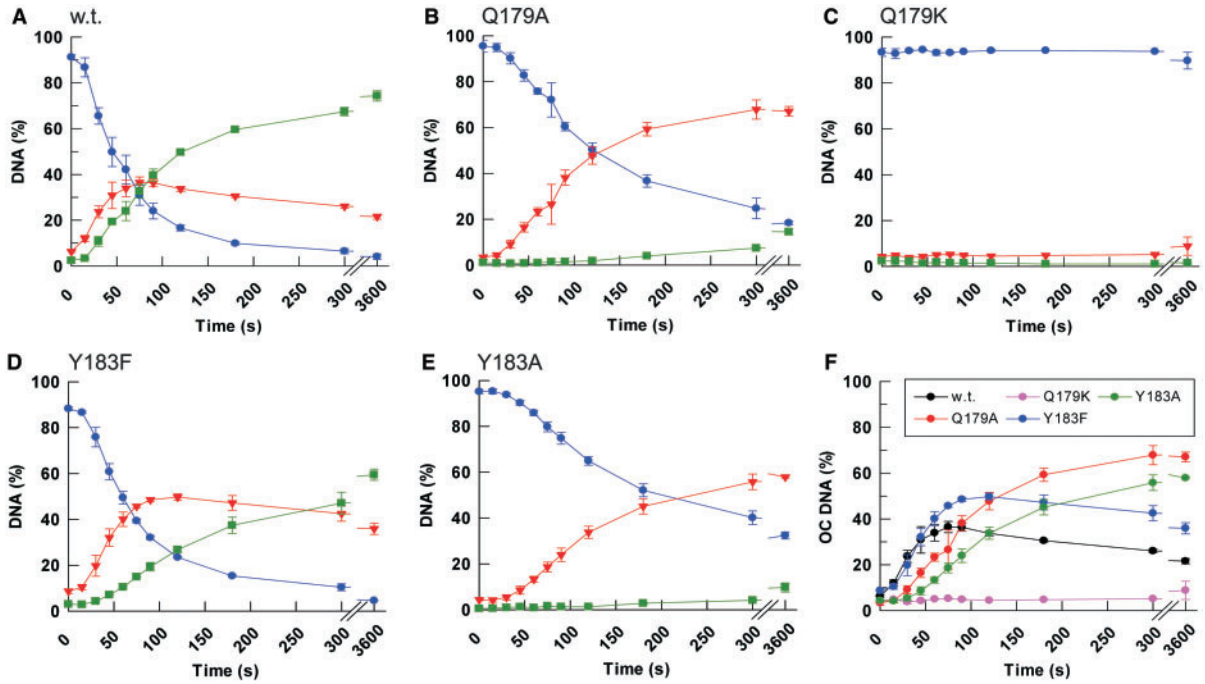

Figure 5. Effect of QxxxY HsdR mutants on the rate of DNA cleavage by EcoR124I. Cleavage of pLKS5 (28) was monitored over a 1 h time course (Materials and methods section). The reaction profiles are shown for (**A**) wild-type HsdR, (**B**) Q179A HsdR, (**C**) Q179K HsdR, (**D**) Y183F HsdR and (**E**) Y183A HsdR. Supercoiled substrate plasmid DNA (blue circles), nicked (open circle) intermediate DNA (red triangles) and linear product DNA (green squares) are shown. Error bars represent the standard deviations from three repeat experiments. (**F**) Comparison of the production of open circle DNA for the wild-type and mutant proteins as indicated. Error bars as above.

same EcoR124I complex, that bidirectionally translocate the entire length of the plasmid DNA (29). Consequently, the cleavage kinetics are complex, with the DNA hydrolysis steps preceded by multiple initiation, translocation, dissociation, re-initiation and collision events (27,32). Because of the continuum of these states (and in particular, the multiple dissociation/re-initiation cycles of translocating HsdRs), cleavage occurs after a lag phase and the observed cleavage rates do not reflect the absolute strand hydrolysis rates. We are currently developing a mathematical model to fully describe these kinetics (F.P., unpublished data). A simple way to compare the wild-type cleavage data with that of the mutants is to fit the rate of the first strand cleavage to an offset exponential function (Equation 3, Materials and methods section). This requires three parameters: The apparent cleavage rate (note, not the true, microscopic strand-hydrolysis rate); the lag offset (the average time taken to reach a cleavable state); and the amplitude (a measure of the extent of inhibition of cleavage). Fits to the data are shown in Supplementary Figure 2 with the fitted parameters reported in Table 1. The data for each mutant will be described in turn.

Q179A. Mutation of Q179 to an alanine reduces the apparent rate for cleavage of the first strand by ~2-fold (Figure 5B, Table 1). There is also an ~2.5-fold increase in the lag offset, which suggests that it takes longer to achieve the first successful cleavage, despite the wild-type-like translocation properties. The apparent rate of appearance of linear DNA is markedly reduced compared to the nicking rate. The amount of DNA cleaved is also reduced; when the reaction stops after ~5 min, the DNA is predominantly nicked (~67%), with very little linear DNA produced (~14%). It may be that the repression of EcoR124I activity is competitive with respect to cleavage, such that a reduction in the cleavage rate allows a greater proportion of the DNA to become irreversibly inhibited.

Q179K. Despite the normal DNA translocation properties of the Q179K mutant, DNA cleavage was not observed under any conditions (Figure 5C, Table 1). This suggests that either strand hydrolysis is no longer possible or that cleavage is very slow (very little measurable cleavage occurring before the reaction stops at 5 min).

Y183F. The Y183F mutation showed a moderately reduced rate for the first strand cleavage but an almost 2-fold increase in the lag offset (Figure 5D, Table 1). As observed with the Q179A mutant, the longer lag offset corresponds to a less concerted cleavage of the DNA; more nicked DNA is generated. The apparent rate of linear DNA production was markedly slower, with the result that only ~60% linear DNA was generated before the reaction was irreversibly inhibited.

Y183A. The Y183A mutation showed a >2-fold decrease in the apparent rate for the first strand cleavage and a >3.5-fold increase in the lag offset (Figure 5E, Table 1). Consequently, very little second strand cleavage was observed; when the reaction stopped at 5 min, only 65%

of the DNA had been cut, with the majority (58%) nicked in only one strand.

Q179A,Y183A. The additive effects of the Q179A,Y183A mutant were such that no DNA cleavage was observed, similar to the Q179K mutant (data not shown).

CONCLUSIONS

All the QxxxY motif mutants that we examined were affected in their assembly of an R_2-complex in the absence of ATP and in their rate and efficiency of DNA cleavage. A comparison in the peak amounts of nicked DNA for each HsdR is shown in Figure 5F. This illustrates that mutation of the QxxxY motif produced not only a decrease in the observed rate of the first strand hydrolysis, but that the subsequent second strand cleavage was affected to an apparently greater extent such that nicked DNA became a more significant product of the reaction. Similar results were observed with QxxxY mutants of EcoAI (Mark Sanderson and M.D.S., unpublished data). These results are in contrast to mutations in Motifs II and III, where DNA cleavage activity is completely lost (or at least reduced to a level significantly slower than the inhibitory rate) (E.S., M.W., M.D.S. and Ralf Seidel, unpublished data). Whilst not being critical components of the nuclease motif, Q179 and Y183 clearly play an important role, consistent with their conservation amongst the majority of HsdR subunits (Figure 1).

We can rationalize our observations using two general models: In the first model, the QxxxY motif directly influences the microscopic DNA hydrolysis rates; In the second model, the QxxxY motif does not play a direct role in DNA cleavage, but instead stabilizes the catalytic residues on the DNA. We will consider each model in turn.

The most straightforward explanation for our observations is that the rate of hydrolysis of the first and second strands, v_1 and v_2, respectively, are changed by mutations in the QxxxY motif. Irrespective of the rates for any preceding sequential steps (translocation, etc.), if $v_1 = v_2$ then the maximal amount of nicked DNA intermediate observed cannot exceed 37% of the total DNA (33; modelling not shown). To get the maximal levels of nicked DNA observed with the QxxxY mutants (all of which exceed 49%; Figure 5F), then v_2 must be slower than v_1. If $v_{1,\mathrm{mutant}} < v_{1,\mathrm{wt}}$ and $v_{2,\mathrm{mutant}} < v_{1,\mathrm{mutant}}$, then nicked DNA would accumulate more slowly and to a greater percentage than wild-type, as observed. Because the cleavage reaction does not proceed beyond 5 min, if $k_{2,\mathrm{mutant}}$ were very slow, then the principal end product of the reaction would be nicked DNA (i.e. as seen for both the Q179A and Y183A mutants; Figure 5B and E).

The QxxxY motif could play a number of roles which would produce changes in v_1 and v_2. There may be direct contacts to the magnesium ion(s) required for DNA cleavage (13)—mutations would therefore disrupt metal ion binding and reduce the cleavage rate. Alternatively, the motif may have an indirect effect on magnesium binding by coordinating a network of interactions with the catalytic motifs I, II and III.

Although we cannot completely exclude the above schemes given our data, we believe that it is less likely that v_1 and v_2 would be affected differently by the mutations examined here (e.g. changes in magnesium binding or in the arrangement of the active site would be more likely to have equal effects on v_1 and v_2). Moreover, given that the core Motifs I, II and III are sufficient for DNA cleavage in a great many PD-(E/D)xK superfamily enzymes (10,11,13), it is also less likely that the QxxxY motif directly affects the absolute cleavage rates. Independently assessing changes in v_1 or v_2 is difficult as the apparent rate of DNA cleavage is affected by the continuum of slower states during the lag phase (F.P., unpublished data) and currently we do not have a method to measure the absolute rates in the absence of the preceding motor activity.

The alternative explanation for our observations is that the QxxxY motif plays an upstream role, perhaps in the collision process that leads to DNA cleavage. We must therefore consider what happens during these upstream steps. DNA cleavage by EcoR124I requires the interaction of two independent HsdR motors that stochastically move over many thousands of base pairs (29,32). If each collision event were to result in cleavage of only one strand, then many collision/cleavage events would need to occur before two strand scissions were ideally located to produce a dsDNA break (the probability of two successive collision events occurring at the same location being very small). There would therefore be a marked difference in the cleavage kinetics; the nicked DNA would accumulate because multiple collision/nicking events would be required to generate a dsDNA break. Alternatively, if each collision event were to result in the concerted cleavage of both strands, the rate of accumulation of the linear DNA would be markedly faster and nicked DNA would not accumulate. The cleavage profile observed for the wild-type HsdR is more consistent with the second scenario in which there is a concerted cleavage of both strands during one collision event. The probability of concerted cleavage will be dictated by the lifetime of the collision complex versus the time required to cut both strands. If a mutation in HsdR were to reduce this lifetime, then the probability of cleavage before dissociation would also be reduced even if the cleavage rate were unaffected. This would have several effects: The lag before appearance of the first strand break would increase; The apparent rate of the first strand cleavage would decrease; The apparent rate of cleavage of the second strand would decrease as a function of the number of random collision/nicking events; and, because of the inhibition of the EcoR124I reaction with time, the amount of uncut DNA and DNA cut in only one strand would increase. A number of observations made with the mutants are consistent with just such a role for the QxxxY motif in stabilizing the catalytic site on the DNA during collision:

(1) For every QxxxY mutant examined which could still cleave DNA, the lag before the accumulation of the first cleaved species was measurably longer than wild-type (Table 1). Since the mutants showed conventional DNA translocation properties

(Figure 4, Table 1), this suggests that collision events were less successful.
(2) The QxxxY mutants all showed a greater accumulation of nicked DNA intermediate compared to wild-type (Figure 5).
(3) The accumulation of linear DNA by the QxxxY mutants was even more inefficient than the accumulation of nicked DNA.
(4) The stability of the R_2-complex was, in each case, moderately but systematically reduced (Figure 3, Table 1).

What role could the glutamine and tyrosine residues play that would give the above effects? Given that the mutations at Y183 had the greatest effect on both DNA binding and cleavage (Figures 3 and 5, Table 1), we suggest that the tyrosine residue of the QxxxY motif could be involved in stabilizing the catalytic domain on the DNA. The hydroxyl residue is clearly important since the Y183F mutant still showed a reduced DNA binding affinity and a reduced DNA cleavage rate. A DNA-binding role may be equivalent to the role of residues K148, N149 and W150 of the Type II RM enzyme EcoRI, that make direct contacts to the non-bridging oxygens of the 'primary clamp phosphate' (34,35; Supplementary Figure 3). These residues are located on the α-helix immediately C-terminal to Motif III, equivalent to the location of the QxxxY motif, and face towards the phosphodiester backbone. It is proposed that the primary clamp phosphates are intimately coupled with the entire recognition/cleavage network of EcoRI and assist in distorting the DNA so that cleavage can occur (34,35). The exact location of the DNA relative to the PD-(E/D)xK motifs of RecB-family nucleases is not known, but the hypothetical route of the DNA from the RecC channel in RecBCD would allow the DNA to approach the nuclease catalytic centre of RecB in the same manner as the EcoRI example (17; Supplementary Figure 3). The tyrosine residue would then be in an ideal location to play a role in binding and stabilizing the DNA during cleavage, perhaps via a clamp phosphate adjacent to the scissile phosphodiester. A similar role for Y183 in EcoR124I HsdR would explain the results we observed; a reduced affinity of the nuclease motif would mean a reduced lifetime of the nuclease motif on the DNA and so fewer successful cleavages during a collision event.

The glutamine residue of the QxxxY motif could play a related role, either in directly contacting the DNA, or in coordinating the tyrosine residue. The combined loss of both Q179 and Y183 residues in EcoR124I disrupted the interaction sufficiently to completely prevent DNA cleavage. However, the glutamine residue appears dispensable in some cases. There are examples of repair enzymes that show close homology to RecB nucleases but which lack this residue [e.g. The Slr0569 protein from Synechocystis species strain PCC 6803, has a TxxxY version (10)]. We also found that ~20% of those HsdR examined did not appear to have a clear QxxxY motif. However, many of these HsdRs, including EcoKI (8), retain the tyrosine residue (data not shown). It is possible that the role of the glutamine can be fulfilled by other residues (as in the alternative RecB nucleases). It is also possible that the glutamine is donated from a different part of the HsdR that is distant in the primary sequence but proximal in the 3D structure. Alternative spatial arrangements of motifs are also seen in Type II restriction endonucleases (12).

If the QxxxY motif plays a role in stabilizing the nuclease domain on the DNA, why would the RecB-family nucleases and Type I enzymes in particular have retained this motif? What features of the DNA cleavage mechanism are peculiar to these enzymes? In simple terms, RecB is a ssDNA-specific endonuclease. Although there is no evidence that DNA unwinding plays a part in the cleavage mechanism of Type I RM enzymes, we cannot rule out that DNA distortion in the collision complex leads to strand separation [the DNA between the enzymes would be expected to be overwound (29), and a large increase in twist at high stretching force can induce a transition in which the DNA bases are believed to be exposed (36)]. A role for the QxxxY motif in stabilizing protein–protein contacts between Type I enzymes during collision and DNA cleavage is unlikely to be important since RecB does not require nuclease domain dimerization. Instead, the importance of this motif may be because many of these nucleases move along and cleave non-specific DNA, either as passive exonucleases (e.g. RecE), or as nucleases that process DNA that is translocated via a helicase motor (e.g. RecBCD, EcoR124I). The nuclease domain of the Type I enzymes is believed to be ahead of the dsDNA motor in the direction of motion (37). The probability of the nuclease domain binding the DNA at a given site will therefore be affected by the rate of DNA translocation. The QxxxY motif may act as a molecular anchor, opposing DNA motion long enough for the nuclease motifs to engage, and potentially cleave, the polynucleotide track.

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Mark Dillingham, Frank Peske and Ralf Seidel for comments on the article, and Mark Sanderson for preliminary experiments on EcoAI. This work was supported by a Wellcome Trust Senior Research Fellowship in Basic Biomedical Sciences (067439 to M.D.S.); The Grant Agency of the Czech Republic (204/07/0325 to M.W.); and, The European Commission through the BioNano-Switch project (043288 to M.W. and E.S.). Funding to pay the Open Access publication charges for this article was provided by Wellcome Trust Value-in-People Award 078595.

Conflict of interest statement. None declared.

REFERENCES

1. Janscak,P., Abadjieva,A. and Firman,K. (1996) The type I restriction endonuclease R.EcoR124I: over-production and biochemical properties. *J. Mol. Biol.*, **257**, 977–991.

2. Szczelkun,M.D., Janscak,P., Firman,K. and Halford,S.E. (1997) Selection of non–specific DNA cleavage sites by the type IC restriction endonuclease EcoR124I. *J. Mol. Biol.*, **271**, 112–123.
3. Seidel,R., van Noort,J., van der Scheer,C., Bloom,J.G., Dekker,N.H., Dutta,C.F., Blundell,A., Robinson,T., Firman,K. and Dekker,C. (2004) Real-time observation of DNA translocation by the type I restriction modification enzyme EcoR124I. *Nat. Struct. Mol. Biol.*, **11**, 838–843.
4. 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.
5. Stanley,L.K., Seidel,R., van der Scheer,C., Dekker,N.H., Szczelkun,M.D. and Dekker,C. (2006) When a helicase is not a helicase: dsDNA tracking by the motor protein EcoR124I. *EMBO J.*, **25**, 2230–2239.
6. Titheradge,A.J.B., Ternent,D. and Murray,N.E. (1996) A third family of allelic hsd genes in Salmonella enterica: sequence comparisons with related proteins identify conserved regions implicated in restriction of DNA. *Mol. Microbiol.*, **22**, 437–447.
7. Janscak,P., Sandmeier,U. and Bickle,T.A. (1999) Single amino acid substitutions in the HsdR subunit of the type IB restriction enzyme EcoAI uncouple the DNA translocation and DNA cleavage activities of the enzyme. *Nucleic Acids Res.*, **27**, 2638–2643.
8. Davies,G.P., Martin,I., Sturrock,S.S., Cronshaw,A., Murray,N.E. and Dryden,D.T.F. (1999) On the structure and operation of type I DNA restriction enzymes. *J. Mol. Biol.*, **290**, 565–579.
9. Davies,G.P., Kemp,P., Molineux,I.J. and Murray,N.E. (1999) The DNA translocation and ATPase activities of restriction–deficient mutants of Eco KI. *J. Mol. Biol.*, **292**, 787–796.
10. Aravind,L., Makarova,K.S. and Koonin,E.V. (2000) SURVEY AND SUMMARY: Holliday junction resolvases and related nucleases: identification of new families, phyletic distribution and evolutionary trajectories. *Nucleic Acids Res.*, **28**, 3417–3432.
11. Bujnicki,J.M. (2004) Molecular phylogenetics of restriction endonucleases. In Pingound,A. (Ed.), *Restriction Endonucleases, Nucleic Acids and Molecular Biology*, Vol. 14, Springer, Germany, pp. 63–93.
12. Skirgaila,R., Grazulis,S., Bozic,D., Huber,R. and Siksnys,V. (1998) Structure–based redesign of the catalytic/metal binding site of Cfr10I restriction endonuclease reveals importance of spatial rather than sequence conservation of active centre residues. *J. Mol. Biol.*, **279**, 473–481.
13. Pingoud,A., Fuxreiter,M., Pingoud,V. and Wende,W. (2005) Type II restriction endonucleases: structure and mechanism. *Cell Mol. Life Sci.*, **62**, 685–707.
14. Obarska–Kosinska,A., Taylor,J.E., Callow,P., Orlowski,J., Bujnicki,J.M. and Kneale,G.G. (2008) HsdR subunit of the type I restriction–modification enzyme EcoR124I: biophysical characterisation and structural modelling. *J. Mol. Biol.*, **376**, 438–452.
15. Singleton,M.R., Dillingham,M.S., Gaudier,M., Kowalczykowski,S.C. and Wigley,D.B. (2004) Crystal structure of RecBCD enzyme reveals a machine for processing DNA breaks. *Nature*, **432**, 187–193.
16. Yu,M., Souaya,J. and Julin,D.A. (1998) Identification of the nuclease active site in the multifunctional RecBCD enzyme by creation of a chimeric enzyme,. *J. Mol. Biol.*, **283**, 797–808.
17. Wang,J., Chen,R. and Julin,D.A. (2000) A single nuclease active site of the Escherichia coli RecBCD enzyme catalyzes single–stranded DNA degradation in both directions. *J. Biol. Chem.*, **275**, 507–513.
18. Quiberoni,A., Biswas,I., El Karoui,M., Rezaïki,L., Tailliez,P. and Gruss,A. (2001) *In vivo* evidence for two active nuclease motifs in the double–strand break repair enzyme RexAB of Lactococcus lactis. *J. Bacteriol.*, **183**, 4071–4078.
19. Chang,H.W. and Julin,D.A. (2001) Structure and function of the Escherichia coli RecE protein, a member of the RecB nuclease domain family. *J. Biol. Chem.*, **276**, 46004–46010.
20. Yeeles,J.T. and Dillingham,M.S. (2007) A dual-nuclease mechanism for DNA break processing by AddAB-type helicase-nucleases. *J. Mol. Biol.*, **371**, 66–78.
21. Thompson,J.D., Higgins,D.G. and Gibson,T.J. (1994) CLUSTAL W: improving the sensitivity of progressivemultiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res.*, **22**, 4673–4680.
22. Roberts,R.J., Vincze,T., Posfai,J. and Macelis,D. (2007) REBASE-enzymes and genes for DNA restriction and modification. *Nucleic Acids Res.*, **35**, D269–D270.
23. Clamp,M., Cuff,J., Searle,S.M. and Barton,G.J. (2004) The Jalview Java alignment editor. *Bioinformatics*, **20**, 426–427.
24. Cuff,J.A., Clamp,M.E., Siddiqui,A.S., Finlay,M. and Barton,G.J. (1998) JPred: a consensus secondary structure prediction server. *Bioinformatics*, **14**, 892–893.
25. Janscak,P., Dryden,D.T.F. and Firman,K. (1998) Analysis of the subunit assembly of the type IC restriction-modification enzyme EcoR124I. *Nucleic Acids Res.*, **26**, 4439–4445.
26. Reid,S.L., Parry,D., Liu,H.H. and Connolly,B.A. (2001) Binding and recognition of GATATC target sequences by the EcoRV restriction endonuclease: a study using fluorescent oligonucleotides and fluorescence polarization. *Biochemistry*, **40**, 2484–2494.
27. 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.
28. Stanley,L.K. and Szczelkun,M.D. (2006) Direct and random routing of a molecular motor protein at a DNA junction. *Nucleic Acids Res.*, **34**, 4387–4394.
29. Szczelkun,M.D., Dillingham,M.S., Janscak,P., Firman,K. and Halford,S.E. (1996) Repercussions of DNA tracking by the type IC restriction endonuclease EcoR124I on linear, circular and catenated substrates. *EMBO J.*, **15**, 6335–6347.
30. Suri,B. and Bickle,T.A. (1985) EcoA: the first member of a new family of type I restriction modification systems. Gene organization and enzymatic activities. *J. Mol. Biol.*, **186**, 77–85.
31. Dryden,D.T., Cooper,L.P., Thorpe,P.H. and Byron,O. (1997) The in vitro assembly of the EcoKI type I DNA restriction/modification enzyme and its in vivo implications. *Biochemistry*, **36**, 1065–1076.
32. Szczelkun,M.D. (2002) Kinetic models of translocation, head-on collision, and DNA cleavage by type I restriction endonucleases. *Biochemistry*, **41**, 2067–2074.
33. Bilcock,D.T., Daniels,L.E., Bath,A.J. and Halford,S.E. (1999) Reactions of type II restriction endonucleases with 8-base pair recognition sites. *J. Biol. Chem.*, **274**, 36379–36386.
34. Kurpiewski,M.R., Koziolkiewicz,M., Wilk,A., Stec,W.J. and Jen-Jacobson,L. (1996) Chiral phosphorothioates as probes of protein interactions with individual DNA phosphoryl oxygens: essential interactions of EcoRI endonuclease with the phosphate at pGAATTC. *Biochemistry*, **35**, 8846–8854.
35. Grigorescu,A., Horvath,M., Wilkosz,K., Chandrasekhar,K. and Rosenberg,J.M. (2004) The integration of recognition and cleavage: X-ray structures of pre-transition state complex, post-reactive complex and the DNA-free endonuclease. In Pingound,A. (Ed.), *Restriction Endonucleases, Nucleic Acids and Molecular Biology*, Vol. 14, Springer, Germany, pp. 137–177.
36. Allemand,J.F., Bensimon,D., Lavery,R. and Croquette,V. (1998) Stretched and overwound DNA forms a Pauling-like structure with exposed bases. *Proc. Natl Acad. Sci. USA*, **95**, 14152–14157.
37. McClelland,S.E. and Szczelkun,M.D. (2004) Molecular motors that process DNA in restriction enzymes. In Pingound,A. (Ed.), *Restriction Endonucleases, Nucleic Acids and Molecular Biology*, Vol. 14, Springer, Germany, pp. 111–135.
38. Singleton,M.R., Dillingham,M.S. and Wigley,D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, **76**, 23–50.