1554–1564 *Nucleic Acids Research, 2011, Vol. 39, No. 4* *Published online 8 October 2010*
doi:10.1093/nar/gkq821

Hpy188I–DNA pre- and post-cleavage complexes—snapshots of the GIY-YIG nuclease mediated catalysis

Monika Sokolowska[1,2], Honorata Czapinska[1,2] and Matthias Bochtler[1,2,3,*]

[1]International Institute of Molecular and Cell Biology, Trojdena 4, 02-109 Warsaw, Poland, [2]Max Planck Institute of Molecular Cell Biology and Genetics, Pfotenhauerstr. 108, 01309 Dresden, Germany and [3]Schools of Chemistry and Biosciences, Main Building, Park Place, Cardiff University, Cardiff CF10 3AT, UK

Received August 11, 2010; Revised and Accepted September 1, 2010

ABSTRACT

The GIY-YIG nuclease domain is present in all kingdoms of life and has diverse functions. It is found in the eukaryotic flap endonuclease and Holliday junction resolvase Slx1–Slx4, the prokaryotic nucleotide excision repair proteins UvrC and Cho, and in proteins of 'selfish' genetic elements. Here we present the structures of the ternary pre- and post-cleavage complexes of the type II GIY-YIG restriction endonuclease Hpy188I with DNA and a surrogate or catalytic metal ion, respectively. Our structures suggest that GIY-YIG nucleases catalyze DNA hydrolysis by a single substitution reaction. They are consistent with a previous proposal that a tyrosine residue (which we expect to occur in its phenolate form) acts as a general base for the attacking water molecule. In contrast to the earlier proposal, our data identify the general base with the GIY and not the YIG tyrosine. A conserved glutamate residue (Glu149 provided in *trans* in Hpy188I) anchors a single metal cation in the active site. This metal ion contacts the phosphate proS oxygen atom and the leaving group 3′-oxygen atom, presumably to facilitate its departure. Taken together, our data reveal striking analogy in the absence of homology between GIY-YIG and ββα-Me nucleases.

INTRODUCTION

The GIY-YIG catalytic module (also called a URI domain) is found in all kingdoms of life (1). It assumes some house-keeping functions, but is often also associated with 'selfish' genetic elements. Biochemical and structural studies of GIY-YIG nucleases have elucidated the fold of these enzymes and provided a list of conserved residues that are thought to play a role in the catalysis (2–7). However, the details of substrate binding and the individual roles of catalytic residues have remained unclear because structures of GIY-YIG nuclease–DNA complexes have not been available. In the following, we briefly present the main groups of GIY-YIG enzymes, summarize the existing information about the fold and active site, and introduce the Hpy188I restriction endonuclease (8) that we have used for our work.

Prokaryotic GIY-YIG nucleases play house-keeping roles in nucleotide excision repair (9). The pathway deals with lesions that affect only a single DNA strand and are too bulky or otherwise unsuitable for the base excision or mismatch repair machinery (9). Nucleotide excision repair begins with the identification of the damage by UvrA (10), which recruits UvrB to the site (10). Subsequently, the DNA is nicked upstream and downstream of the lesion by the RNase H and GIY-YIG domains of UvrC, respectively (11,12). Alternatively, the downstream nick can be introduced by the GIY-YIG Cho nuclease (13). A structure of the GIY-YIG domain of UvrC has been reported. However, the complex with DNA has not yet been crystallized, perhaps because the enzyme requires other proteins to be guided to a DNA lesion (9).

Eukaryotic GIY-YIG nucleases with a well-known housekeeping function include Slx1 proteins, which form heterodimeric complexes with their Slx4 partners (14). The two enzymes were originally identified in yeast synthetic lethal screens designed to isolate proteins redundant with Sgs1 of the Sgs1/Top3 complex (15). Biochemical data suggested that the Slx1–Slx4 heterodimers are active on single-Y, 5′-flapped and replication fork structures, which implied a role of these nucleases in the processing of stalled replication forks (16). More recently, Holliday junction resolving activity was reported for the Slx1–Slx4

*To whom correspondence should be addressed. Tel: +48 22 5970732/+44 29 208 70625; Fax: +48 22 5970715; Email: mbochtler@iimcb.gov.pl; bochtlerm@cardiff.ac.uk

complex and cited to explain its role in meiosis (17). At present, no structural data are available for the components of the complex, but the alignment of the Slx1 sequence to the GIY-YIG consensus is convincing and supported by a loss of activity when predicted active site residues are altered (15).

The Penelope GIY-YIG nuclease domain provides an example of a 'selfish' enzyme in this class (18). Penelope is a *Drosophila* retrotransposon that spreads by host-mediated forward and Penelope mediated reverse transcription (19). The second step is primed by a nicked substrate, generated by the GIY-YIG nuclease domain. Its structure has not been reported, but site-directed mutagenesis data are consistent with the GIY-YIG classification (19).

GIY-YIG homing endonucleases further illustrate the involvement of this group of enzymes in 'selfish' genetic elements. Examples have been found in prokaryotes, viruses, and also in the bi-parentally inherited mitochondria of fungi (20). Homing endonucleases do not provide short term benefits to their hosts. They can spread because they bias gene conversion in their own favor by catalyzing double strand breaks in target DNA (21). Homing endonucleases come in several phylogenetically unrelated groups including LAGLIDADG, PD-(D/E)XK, ββα-Me (also known as HNH and His-Cys box) and GIY-YIG enzymes (22). Some genes encoding the latter are found as free standing open reading frames, but others are located in type I introns, presumably in order to mitigate their fitness cost (1). Bacteriophage T4 intron-resident I-TevI is the prototype for this group of the GIY-YIG homing endonucleases. It is composed of an N-terminal non-specific nuclease domain (2) and a C-terminal sequence specific DNA binding domain (23). The two parts of the enzyme have been crystallized separately, but a DNA complex has only been reported for the latter (2,23).

GIY-YIG nucleases have also been found among type II restriction-modification systems (3,24,25). These systems provide the host (and themselves) with immunity against invading DNA in return for their own replication. Their restriction endonuclease components were originally regarded as evolutionarily unrelated. They were then suspected to be monophyletic, but are now known to belong to several different catalytic classes (26). The largest group consists of the PD-(D/E)XK enzymes, and the next most abundant one of the ββα-Me enzymes (26). Both classes are metal dependent and catalyze a direct attack of a water molecule (or a hydroxide ion) on the scissile bond phosphorus atom (27,28). In contrast, PLD restrictases, represented by the extensively studied BfiI (29), are metal independent and catalyze DNA cleavage by two sequential substitution reactions (30). This leaves half-pipe enzymes, represented by PabI (31), and GIY-YIG restriction endonucleases as the least understood groups. The latter form now the only well-studied monophyletic group of type II restriction endonucleases without a structurally characterized representative in the Protein Data Bank.

Superpositions of the GIY-YIG nuclease structures reveal a conserved core that is present almost without additional 'decorations' in the UvrC and I-TevI. The 'GIY-YIG hairpin' anchors the GIY and YIG motifs, which are both subject to substantial variation. In the first motif (GIY), 'Y' is conserved, but 'G' and 'I' are present only in some enzymes. In the second motif (YIG), 'Y' is almost universal, but replaced by a lysine in Hpy188I, 'I' can be substituted with a valine, and 'G' is strongly conserved (1) The next core element is an α-helix which we name the 'arginine helix' because of an invariant arginine residue. This is followed by a 'linker strand', which extends the GIY-YIG hairpin to an anti-parallel β-sheet. The last conserved secondary structure element is another α-helix that we call the 'glutamate

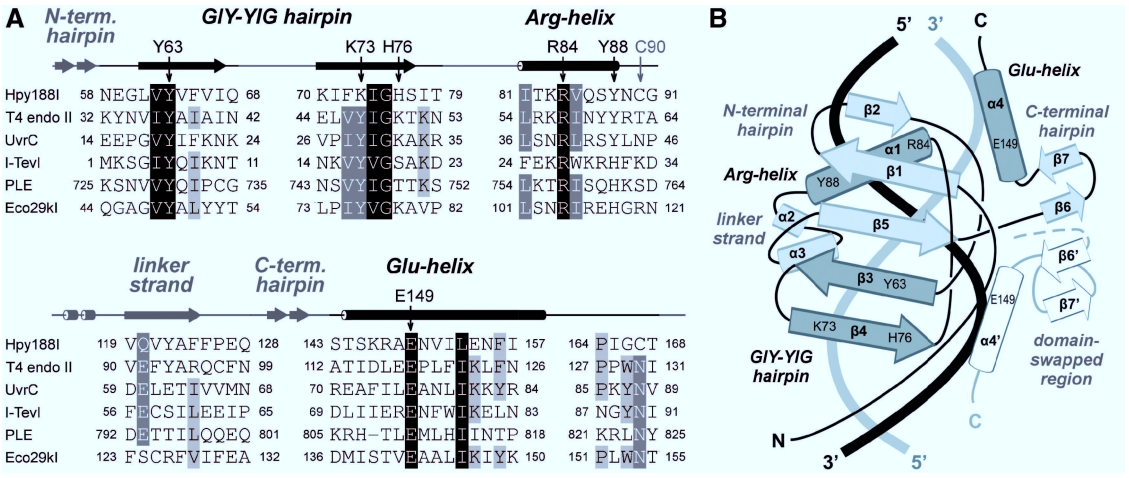

Figure 1. GIY-YIG nuclease alignment and Hpy188I fold: (**A**) Representative GIY-YIG nucleases were aligned by a combination of structure- and sequence-based methods. The alignment was corrected manually to take into account mutagenesis data that have identified active site residues and previously published alignments (19,24,26). The numbering of conserved residues and secondary structure information in the first and second lines refer to Hpy188I. (**B**) A single Hpy188I subunit and the domain swapped region of the other subunit are shown in symbolic representation, with cylinders for α-helices and arrows for β-strands. Short 3/10 helices are omitted for clarity. The DNA is represented by the smoothed backbones (proximal strand in gray and distal strand in black).

helix' because of a glutamate residue that stands out in the alignment (Figure 1A).

This general understanding has been the basis for excellent predictions of the core structures of uncharacterized GIY-YIG nucleases, including Hpy188I (3,24). However, in the absence of any experimental structures in complex with DNA, the bioinformatic analysis did not make it possible to discuss the mode of the substrate binding or to explain the roles of the most conserved active site residues. The lack of a DNA bound structure of any GIY-YIG nuclease has also prevented a detailed understanding of the mechanism of these enzymes. Most authors assume that the reaction proceeds by a direct attack of a water molecule (or a hydroxide) on the scissile bond phosphorus atom (32). However, a two-step mechanism with a phosphotyrosine intermediate as for some topoisomerases (33,34) and recombinases (35) remains a formal possibility. We have crystallized Hpy188I and obtained specific substrate-like and product ternary complexes. Together, the two structures make it possible to propose a catalytic mechanism for Hpy188I and by implication for GIY-YIG nucleases in general.

MATERIALS AND METHODS

Cloning

Codon optimized Hpy188I REase (*hpy188IR*) and Hpy188I MTase (*hpy188IM*) synthetic genes in pBluescript II SK (+) vectors were purchased from Epoch Biolabs, Inc. (TX, USA). The *hpy188IM* gene was PCR-amplified using primers that were designed to introduce a Shine-Dalgarno sequence 8 nt upstream of the ATG start codon and HindIII and BamHI cloning sites upstream and downstream of the gene. Using these sites, the fragment was then placed into the tetracycline region of the pACYC184 (Cm^r) vector. For the expression of the REase, we used pET15bmod (Ap^r) (28), a derivative of pET15b (+) (Ap^r). The *hpy188IR* gene was cloned into this vector using the EcoRI and XhoI. The resulting construct coded for the full-length protein with the N-terminal tag MGHHHHHHEF.

REase expression

Expression experiments were done in *Escherichia coli* ER2566 strain (F^- λ^- *fhuA2 [Ion] ompT lacZ::T7geneI gal sulA11 Δ(mcrC-mrr)114::IS10 R(mcr-73::*miniTn*10)2 R(zgb-210::*Tn*10)1* (Tet^s) *endA1 [dcm]*). The strain was transformed sequentially with plasmids pACYC184 (Cm^r) bearing the *hpy188IM* gene and pET15bmod (Ap^r) containing the *hpy188IR* gene. Cells were grown in LB medium with appropriate antibiotics at 37°C to OD_{600} of 0.6 and induced with 0.2 mM IPTG. After 4 h of induction at 22°C the cells were harvested by centrifugation and the pellet was stored at −20°C. Expression of the selenomethionine variant of Hpy188I was done in the minimal M9 medium (36) in the presence of 0.05 mg/ml D,L-selenomethionine (Sigma) following the published procedure (37), optimized to suppress methionine biosynthesis.

REase purification

Frozen cells expressing Hpy188I REase were thawed and resuspended in buffer A (20 mM Tris/HCl pH 7.6, 500 mM NaCl and 1 mM PMSF). Cells in suspension were opened by sonication and the cell debris was removed by centrifugation at 145 000*g* for 40 min. Hpy188I was purified by affinity chromatography on Nickel Nitrilotriacetic Acid (Ni-NTA) agarose column (Qiagen). The protein was eluted using an imidazole gradient in buffer B (20 mM Tris/HCl pH 7.6, 200 mM NaCl, 5 mM 2-mercaptoethanol). Fractions containing Hpy188I were combined and concentrated using Vivaspin concentrators (10 kDa MWCO). The protein was purified further by size-exclusion chromatography on HiLoad 16/60 Superdex 75 column (GE Healthcare), equilibrated with buffer C (20 mM Tris/HCl pH 7.6, 200 mM NaCl, 1 mM EDTA and 10 mM DTT). Fractions containing Hpy188I REase activity were pooled and concentrated to 24–36 mg/ml. From a 2 l culture, ~12 mg of protein was obtained that appeared homogeneous on Coomassie-stained SDS–PAGE. The selenomethionine variant of Hpy188I was purified following the protocol for the wild-type enzyme. The overall yield of selenomethionine derivative of Hpy188I was 7 mg from 1 l of culture.

Crystallization

Oligonucleotides 5′-GA*TCTGA*AC-3′ and 5′-GT*TCAGA* TC-3′ were purchased from Metabion. They were dissolved in 10 mM Tris/HCl pH 7.9 and annealed by heating to 95°C followed by slow cooling to 4°C to yield a duplex with blunt ends. Hpy188I in buffer C was concentrated to 24–36 mg/ml and mixed with the oligoduplex at a 1:1.1 molar ratio (protein dimer:duplex DNA). The protein–DNA complex was allowed to stand on ice for at least 1 h before crystallization. Hpy188I–DNA co-crystals were grown by the vapor diffusion technique. Initial high-throughput screens were set up at the 200-nl scale using a Cartesian pipetting robot and 96-well Greiner sitting drop plates. Crystallization trials with larger drop volumes were pipetted in CRYSCHEM plates (Hampton Research). Crystals of Hpy188I–DNA complex were obtained with the reservoir containing 0.1 M MES/NaOH pH 6.2 and 30% MPD. Drops were formed by mixing 2 µl of the protein–DNA solution with 2 µl of the reservoir buffer and equilibrated over the reservoir at 18°C. Crystals of Hpy188I–DNA complex grew within 5 days. Crystals of the selenomethionine variant of Hpy188I with DNA were grown in the same conditions within 3 days. Ca^{2+} ions were soaked into the crystals for 18 h by adding 100 mM $CaCl_2$. All Hpy188I–DNA crystals could be flash-cryocooled directly from the crystallization drop without additional cryoprotection.

Structure determination

Structure determination was carried out exclusively for the much better diffracting crystals of Hpy188I with selenomethionine substitutions. Both reported crystals

Table 1. Data collection and refinement statistics

	Hpy188I–DNA pre-cleavage complex	Hpy188I–DNA post-cleavage complex
Data collection statistics		
Space group	P4(1)2(1)2	P4(1)2(1)2
a (Å)	65.236	65.395
b (Å)	65.236	65.395
c (Å)	220.630	221.783
Resolution range (Å)	20–1.75	20–1.95
Total reflections	268 209	494 630
Unique reflections	48 330	35 348
Completeness (%) (last shell)	98.0 (97.1)	97.5 (95.8)
I/σ (last shell)	24.6 (3.3)	37.8 (6.5)
R(sym) (%) (last shell)	6.1 (44.0)	6.8 (40.6)
B(iso) from Wilson (Å^2)	23.2	25.2
Refinement statistics		
Protein atoms excluding H	2894 (3304)[a]	2844 (3245)[a]
DNA atoms excluding H	363 (726)[a]	365 (730)[a]
Solvent molecules	362	318
R-factor (%)	16.5	16.8
R-free (%)	19.2	20.1
Rmsd bond lengths (Å)	1.2	1.3
Rmsd angles (°)	0.010	0.012
Ramachandran core region (%)	94.4	92.2
Ramachandran allowed region (%)	5.6	7.5
Ramachandran additionally allowed region (%)	0	0.3
Ramachandran disallowed region (%)	0	0

[a]Double conformations are either counted only once (no brackets) or twice (in brackets).

(with Na^+ and Ca^{2+} ions in the active sites) belonged to the tetragonal space group P4(1)2(1)2 and contained a complete Hpy188I dimer with target oligoduplex in the asymmetric unit. Three-wavelength (peak, inflection point and low energy remote) multiple anomalous diffraction (MAD) was used for phase determination. Although the correlation of the anomalous signal for peak and inflection point was relatively weak (<60% throughout and down to ~30% at 3.0 Å resolution), the SHELXD program (38) identified a convincing selenium substructure (CC All/Weak 40.47/23.72). It included two very pronounced and two weaker sites, followed by a continuum of 'noise', in agreement with the presence of the initiator and an additional methionine residue in each of the Hpy188I protomers. Density modification using the SHELXE program (39) showed a clear preference for the original (pseudo-free CC 60.51%, contrast after 15 cycles 0.66) over the alternative (pseudo-free CC 50.32%, contrast after 15 cycles 0.47) choice of hand. As we could not build the Hpy188I–DNA structure with the SHELXE calculated phases, the experimental phases were recalculated with the MLPHARE program (40). At this stage, the figure of merit for combined centric and acentric reflections was 0.28 for the range from 10 to 2.5 Å. For efficient density modification with the program DM, an averaging operator was required. The straightforward choice would have been the Hpy188I dimer axis, located based on the self-rotation function and the heavy atom structure. However, we noticed that the Patterson function for the Hpy188I–DNA complex contained a strong pseudoorigin peak at (1/2, 1/2, 1/2) in fractional coordinates and used this pseudotranslation as the initial averaging operator. The refined translation differs slightly from (1/2, 1/2, 1/2), but due to the symmetry of Patterson space this difference is obliterated in the Patterson map. Using the MLPHARE phases up to 2.5 Å resolution, the DM program (41) immediately calculated a correlation between the NCS related densities >85%, which did not change much with phase extension. After DM averaging and solvent flattening (in resolution extension mode) the overall figure of merit for the full resolution range was 0.55. The resulting protein density was readily interpretable by the ARP/wARP program (42). The DNA model was built starting with an idealized B-DNA model of the correct sequence generated with the program 3DNA (43). As the oligoduplex has lower symmetry than the surrounding Hpy188I dimer, two binding modes were seen that had to be modeled as alternative conformations of the DNA. Refinement was carried out in the usual way, using the programs COOT (44) and XTALVIEW (45) for model building and the program REFMAC (46) for refinement. The data collection and refinement statistics for the final models are summarized in Table 1.

RESULTS

Structures

Hpy188I (5′-TCN|GA-3′, where '|' marks the cleavage site) was expressed and purified as described in 'Materials and Methods' section, and shown to be a dimer by gel filtration chromatography (data not shown). The obtained protein was mixed with pre-annealed blunt-ended 9-mer dsDNA that had an A:T pair at the center of the recognition sequence and crystallized. Crystals contained one Hpy188I dimer with bound oligoduplex in the asymmetric unit and diffracted to ~3.0 Å resolution on a second generation synchrotron beamline. As we had difficulty to interpret these data by molecular replacement (using various GIY-YIG nucleases as search models), we prepared the selenomethionine variant of Hpy188I and confirmed its activity. Surprisingly, selenomethionine Hpy188I–DNA crystals diffracted much better (to 1.75 Å resolution) than the native ones, although they belonged to the same form and were grown under identical conditions.

Crystals that were obtained in the presence of the divalent metal cation chelator EDTA contained one Na^+ ion per active site and the uncleaved DNA duplex. They could be soaked with Ca^{2+} ions to replace the Na^+ ions and drive the DNA cleavage reaction to completion. Fortunately, the soaking procedure had only a minor influence on the diffraction quality of the crystals, so that substrate and product complexes could be compared at good resolution. The crystallographically independent Hpy188I protomers were very similar in the substrate (main chain rmsd 0.2 Å) and product complexes (main chain rmsd 0.3 Å), despite differences in their local environment. DNA cleavage introduced local changes, but

left the overall structures unperturbed (main chain rmsd between protein dimers in the two structures 0.6 Å). In both complexes, the DNA did not distinguish between protomers and adopted two possible binding modes (with the central adenine or thymine proximal to the A protomer). With the exception of the middle base pair, the specifically recognized DNA followed the 2-fold symmetry of the Hpy188I dimer. Therefore the presence of the two binding modes did not complicate the interpretation of the X-ray data.

Overall fold

Hpy188I has the expected core structure built of the GIY-YIG hairpin (with LVY and KIG instead of the GIY and YIG motifs), arginine helix, linker strand and glutamate helix. The conserved antiparallel β-strand is expanded by an extra 'N-terminal hairpin', which corresponds to an irregular loop and a single additional β-strand of T4 endonuclease II (T4 endo II) but lacks counterparts in UvrC and I-TevI. In Hpy188I, the arginine helix and linker strand are connected by two helices. A connection of these secondary structure elements by helices is typical, but their number varies in GIY-YIG nucleases. The linker strand and glutamate helix of Hpy188I are separated by a 'C-terminal hairpin' which contributes to dimer contacts. Uncharacteristically for GIY-YIG nucleases, the C-terminal hairpin and the glutamate helix are domain swapped.

Our co-crystal structures make it possible to define the orientation of the central β-sheet with respect to the DNA. The conserved GIY-YIG hairpin is closer to the 5′-end of the proximal DNA strand, and the N-terminal hairpin nearer to its 3′-end (Figure 1B). The two Hpy188I subunits interact extensively across a 2500 $Å^2$ interface. The biggest contribution to this large area is from the domain-swapped regions, which alone account for 2200 $Å^2$. In the specific substrate and product complexes, the Hpy188I dimer embraces the bound DNA completely (Figure 2). As the enzyme can cleave circular DNA, conformational changes must accompany the transition from the scanning complex to the specific complex. This is not surprising and has been reported before for restriction endonucleases in the PD-(D/E)XK (47) and ββα-Me families (28).

Intercalation of cysteine residues into the DNA

The most prominent feature of the Hpy188I–DNA binding mode is the 'shallow' insertion of the Cys90 side chains into the major groove between the outermost specifically recognized and the flanking base pairs (Figure 3A). The lack of deep penetration of the base stack is typical for protein–DNA intercalation complexes (48). However, in most cases the insertion takes place from the minor and not the major groove side. Moreover, the intercalating proteins typically introduce a substantial kink or bend in the DNA (49), which is not the case in the Hpy188I–DNA complex. As the two cysteine insertions are separated by 5 bp or approximately half of a DNA turn, their long-distance effects mostly cancel out. Nevertheless, they do induce significant local DNA distortions (Figure 3B). All base pairs of the TCNGA target sequence are severely inclined with respect to the long DNA axis.

Sequence recognition

Unlike some other pseudopalindrome cleaving restriction enzymes, Hpy188I does not distinguish W (A:T) from S (G:C) pairs at the center of its recognition sequence. The crystal structures of the Hpy188I substrate-like and product complexes are consistent with this lack of specificity. The hydrogen bonding requirements of the A:T pair at the center of the target DNA are entirely satisfied by water molecules. Hpy188I interacts with the inner G:C pairs of its recognition sequence only on the major groove side. The side chain OH group of Ser102 donates

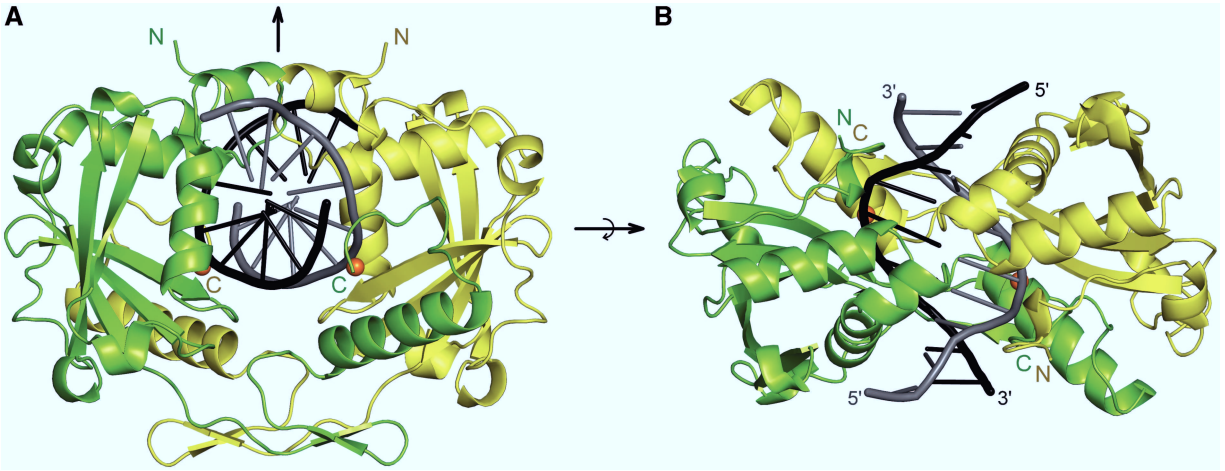

Figure 2. Overview of the Hpy188I–DNA complex: The two subunits of Hpy188I are shown in ribbon representation in green and yellow color. The DNA is depicted by its smoothed backbones with sticks for the bases. (**A**) View along the DNA. In the region of the central base pair, the major groove points upwards. In the vicinity, the DNA backbone is oriented toward the reader on the left and away from the reader on the right side (in the 5′–3′ direction). (**B**) View toward the major groove of the central DNA base pair. The two orientations differ by a 90° rotation about the horizontal axis.

a hydrogen bond to the N7 atom of the guanine. Moreover, the main chain carbonyl oxygen atom of Thr100 accepts a hydrogen bond from the cytosine. Hpy188I contacts the outer A:T pair on the major groove side via Ser87. This residue accepts a hydrogen bond from the adenine N6 and donates another one to its N7. In the substrate complex, side chain carboxamide group of Gln169 donates a hydrogen bond to the adenine N3 in the outer minor groove, but this interaction is lost in the product complex. The thymine in this base pair is only in contact with solvent (Figure 4).

The interactions of Hpy188I with the outer A:T base pair explain its methylation sensitivity. The enzyme is paired with an N6 adenine methyltransferase, which protects genomic DNA against its nucleolytic activity. As the lone pair of the N6 nitrogen of adenine is conjugated to the electrons of the purine ring, the methyl group should be located in the plane of the base. On the side of the Watson–Crick edge, it would clash with the O4 atom of thymine, on the other side with Ser87. Therefore productive binding of the TCNGm6A target sequence should be prevented, which agrees with the genetic and biochemical data.

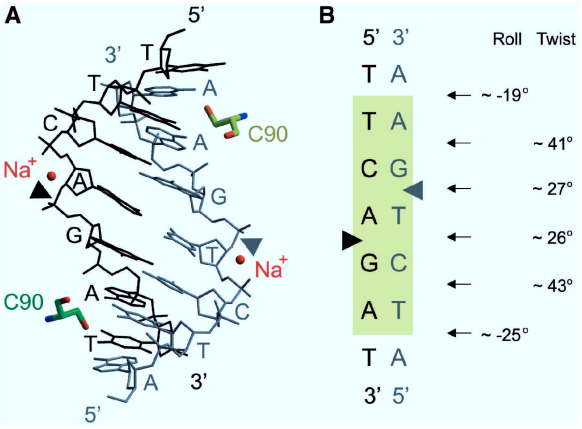

Figure 3. Distorted DNA in the Hpy188I–DNA complex. (**A**) All-atom representation of the DNA target sequence and the most proximal flanking base pairs in the substrate complex. The intercalating cysteine residues are shown as sticks. The Na^+ ions in the active sites are represented by spheres, and the scissile bonds are indicated by triangles. (**B**) Schematic representation of the DNA. The most significant deviations from standard B-DNA geometry were calculated with the 3DNA program (50).

Active site of the substrate complex

The crystal that was grown in the absence of divalent metal cations contains a single Na^+ ion per active site. It shows robust electron density for the scissile phosphoester bond, indicating that the DNA was not cleaved as expected. In the active site of this complex, we noticed electron density that was interpreted as a (badly ordered) water molecule. It is positioned 'at the back' of the scissile bond phosphorus atom, almost ideally for in-line displacement of the 3′-oxygen atom. This water molecule accepts a hydrogen bond from the Nε atom of Arg84 (from the arginine helix) and donates a hydrogen bond to the main chain carbonyl oxygen atom of His76. In addition, it is in contact with the phenolic oxygen atom of the GIY tyrosine Tyr63, which is in turn interacting with the OH group of Tyr88. The latter is also hydrogen bonded with the ε-amino group of Lys73, replacing the YIG tyrosine. The proR oxygen atom of the scissile bond phosphate is in contact with this lysine side chain and the guanidino Nη atom of Arg84. The proS oxygen atom accepts a hydrogen bond from the main chain of His76. Robust electron density in its vicinity was interpreted as a Na^+ ion, octahedrally coordinated by three solvent molecules, the side chain carboxyl oxygen of Glu149, and the proS and leaving group 3′-oxygen atoms of the phosphate. Comparison of the pre- and post-cleavage complexes shows that the Na^+ ion takes the place of the Ca^{2+} ion in the product complex. Therefore we assume that the catalytically unproductive substrate complex represents a fairly good approximation of the true Michaelis complex (Figure 5, top row).

Active site of the product complex

A crystal structure with clearly cleaved DNA and electron density for three oxygen atoms of the 5′-phosphate was obtained by soaking the original crystals with Ca^{2+} ions. The first oxygen atom of the 5′-phosphate is located within

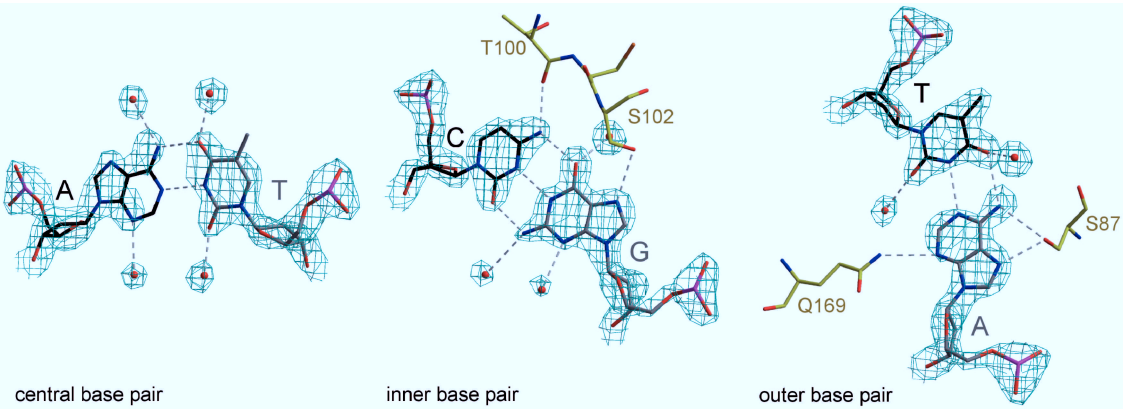

Figure 4. Target sequence recognition by Hpy188I: Base pairs, interacting amino acids and water molecules are shown in all-atom representation. Hydrogen-bonding interactions are indicated by dashed lines. The density is the original ARP/wARP map contoured at 1.5σ.

hydrogen bonding distance of Arg84 Nε, His76 main chain O and N atoms, and the phenolic oxygen atom of Tyr63. These contacts are very reminiscent of the nucleophilic water molecule interactions in the substrate complex. We therefore conclude that this oxygen atom is most likely derived from the attacking water molecule. The second oxygen of the 5′-phosphate has hydrogen bonding interactions with the Nη atom of Arg84 and the ε-amino group of the Lys73. Thus we take this atom to match the substrate proR oxygen. This leaves the third 5′-phosphate oxygen to be identified with the proS oxygen of the substrate. This atom could accept a hydrogen bond from the main chain of His76 like in the pre-cleavage complex, but the distance indicates that this hydrogen bond is weak at best, despite the otherwise favorable geometry. The same applies to the interaction of this oxygen atom with the metal ion, which appears weaker than in the substrate complex because the distance is slightly larger. The strongest hydrogen bond of the former proS oxygen atom links it to the newly formed 3′-OH. The hydroxyl is only ~2.7 Å away from this oxygen atom, but ~3.5 Å away from the phosphorus atom to which it was covalently connected. There are also other changes in the vicinity of the metal ion: instead of three close water ligands, we see only two at marginally relevant distance (~3.4 Å). The coordination sphere is completed by additional interaction with the Nδ atom of His76. It is too far away for efficient ligandation in the substrate complex (~3.8 Å), but comes sufficiently close in the product complex to lie in the first coordination sphere (~2.6 Å; Figure 5B, bottom row).

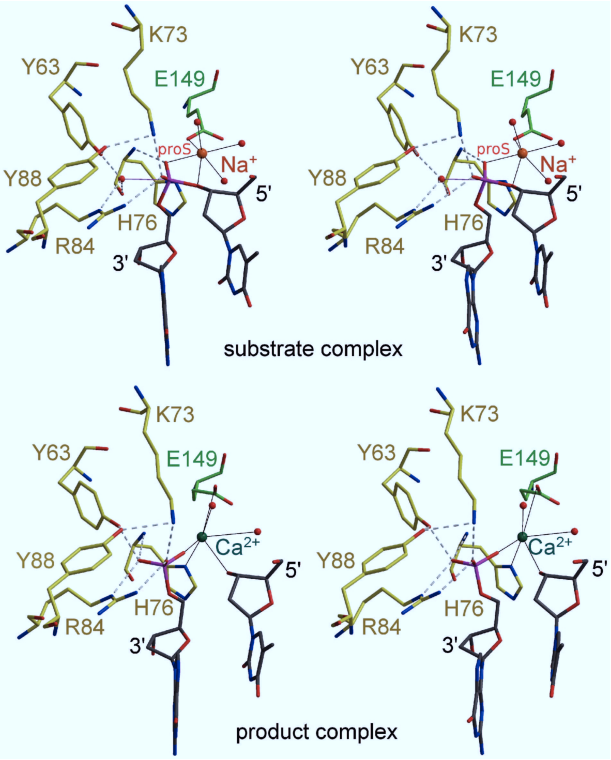

Figure 5. Active site of the pre- and post-cleavage complexes: Selected residues, water molecules and the metal ions are shown in all-atom representation. Hydrogen bonds are indicated by dashed lines, the coordination of the catalytic metal ion or its surrogate is shown by continuous black lines. The link between the nucleophilic water molecule and the phosphorus atom in the pre-cleavage state is shown in magenta. In the product complex, the density for the metal ion is slightly smeared in both subunits. It is therefore possible that solvent ligands are missed or not accurately placed due to disorder.

DISCUSSION

DNA deformation and cleavage stagger

Hpy188I restriction endonuclease cleaves DNA to products with single nucleotide 3′-overhangs. In contrast, the related Eco29kI enzyme cuts its substrate DNA with two nucleotide stagger. Therefore it is tempting to speculate that the inclination of the bases compensates for an 'absent' base pair in the Hpy188I recognition sequence (relative to the Eco29kI target). However, the comparison of the Hpy188I-bound DNA with regular B-DNA shows that this hypothesis can at best be qualitatively correct. Further analysis will require crystal structures of GIY-YIG restriction endonucleases that cleave DNA with two base pair stagger.

One or two substitution reactions

Previous biochemical and crystallographic data related to the GIY-YIG nucleases do not distinguish clearly between mechanisms with either a single or two sequential substitutions. To our knowledge, mechanistically informative hydrolysis of diastereomerically pure phosphorothioate oligonucleotides in oxygen isotope labeled water (50) has not yet been carried out with any GIY-YIG nuclease. In the absence of 'hard' stereochemical data, analogies are the only source of information. On the one hand, conserved tyrosine and arginine residues are found not only in GIY-YIG nucleases, but also in topoisomerases and recombinases. The two groups of enzymes catalyze transient DNA cleavage in a two-step reaction via the phosphotyrosine intermediates (33,35). On the other hand, analogies between GIY-YIG and ββα-Me nucleases have been previously noted (25,32). The latter enzymes are widely believed to carry out single step substitution reactions (28,51). However, in some structures (28,52), the spatial arrangement of the general base and of the attacking water molecule does not appear ideal and double substitution has been discussed as a possibility (52). The location of the 'attacking' water molecule in the crystal structure of the pre-cleavage Hpy188I–DNA complex speaks strongly for a single substitution mechanism. Moreover, similar interactions make it possible to identify individual phosphate oxygen atoms in the substrate and product complexes. Their arrangement in the two structures provides an independent confirmation for a single step mechanism. Unlike sulfur substitution, the crystallographic 'labeling' is certain not to affect the cleavage, but requires the assumption that interactions are not scrambled during the reaction. Discounting 'label' swaps as unlikely we conclude that GIY-YIG nucleases catalyze the direct attack of a water molecule on the scissile bond phosphorus atom in a single step

nucleophilic substitution. In the absence of detailed quantum chemical/molecular dynamics simulations, we cannot distinguish between S_N1 (dissociative), S_N2 (associative) and intermediate (concerted) mechanisms.

Activation of the water molecule

In the Hpy188I GIY-YIG nuclease active site the nucleophilic water molecule is positioned by a conserved arginine (Arg84), main chain carbonyl oxygen atom (His76) and GIY tyrosine (Tyr63). The X-ray data are not nearly of adequate resolution to judge the protonation state of the tyrosine. However, we suspect that its pK_a (typically ~10 in solution) is sufficiently decreased for the side chain to be present in the phenolate form and act as a proton acceptor. Even in this form, the GIY tyrosine is not basic enough to abstract a proton from a water molecule (pK_a ~15 in solution). Therefore we suggest that the proton transfer occurs concomitantly with the nucleophilic attack of the water molecule on the phosphorus atom of the scissile bond. We have noticed that another tyrosine [Tyr88 in Hpy188I, a tyrosine or histidine in other GIY-YIG nucleases (1)] is in close contact with the GIY tyrosine (2.5 Å distance between the Oη atoms), possibly to help shuttle protons. It is unclear whether the oxygen from the attacking water molecule looses the second proton during the reaction, because its hydrogen bonding interactions in the product complex are somewhat ambiguous. Irrespective of whether one or two protons need to be lost from the nucleophile, the structure suggests that the two tyrosine residues (Tyr63 and Tyr88) could serve as an efficient shuttle. Eventually a proton has to end up on the leaving group 3′-oxygen atom, but there is no clear path and transfer via bulk solvent could be involved.

On the basis of the structure of UvrC in the absence of DNA, Truglio and colleagues have previously suggested that a tyrosine in the active site acts as the general base that accepts a proton from the attacking water molecule [Figure 7 of (12)] and loses its OH proton in the process. While this remains possible, we favor conversion of the tyrosine to the phenolate form prior to the reaction, because the neutral side chain is itself somewhat acidic (pK_a ~10 in water) and therefore not very suitable as a proton acceptor. A more important difference between our proposal for the reaction mechanism and the one by Truglio and co-workers concerns the identity of the general base. On the basis of circumstantial evidence, Truglio and colleagues assign this role to the YIG tyrosine (12). In contrast, we propose based on the crystal structures that the GIY tyrosine is the general base. Mutagenesis experiments show that both tyrosines are required for folding and/or activity of GIY-YIG nucleases (6,12). The conservation of the GIY tyrosine and simultaneous 'natural' mutation of the YIG tyrosine to lysine in Hpy188I is consistent with our proposal. However, in the absence of the structural data, the sequence information alone would not be conclusive, because a lysine (with a pK_a for the protonated form similar to the pK_a of tyrosine) could in principle also act as a general base.

Facilitation of the departure of the leaving group

As for other single-step substitution reactions, the departure of the leaving group can be facilitated by hydrogen bonds or ionic interactions that reduce its nucleophilicity. This seems to be true for Hpy188I, and by implication for other GIY-YIG nucleases, with a direct contact of the 3′-oxygen atom to the active site metal. The distance between the two partners is almost 1 Å longer in the product complex, anticipating the eventual release of the hydrolysis products. The contact of the proS oxygen atom to the metal ion is also longer in the post- than in the pre-cleavage structure, but the difference is less pronounced. Unfortunately, the alterations to the metal ion coordination spheres in the two complexes cannot be straightforwardly treated as changes that occur during the reaction. As Na^+ ions do not support catalysis, it remains possible that some detailed properties of the pre-cleavage complex are only 'substrate-like' and do not represent the physiological situation.

Explaining residue conservation in the GIY-YIG family

The proposed catalytic mechanism explains the conservation of the GIY-YIG nuclease active site residues. The GIY tyrosine serves as a general base with respect to the attacking water molecule. It is in contact with another spatially invariant tyrosine (or histidine in some enzymes) which sets up the proton shuttle. The YIG tyrosine (replaced by lysine in Hpy188I) is needed as a hydrogen bond donor to the proR oxygen atom. The YIG glycine is conserved, because a side chain of this residue would clash with the GIY tyrosine and the 5′-phosphate in the product complex (provided the spatial arrangement of the active site is preserved). The increased backbone flexibility of glycine is not needed, because its main chain torsion angles have typical β-strand values. The conservation of the α-helix arginine residue is explained by its interactions with the attacking water molecule and proR oxygen atom in the substrate, which are retained in the product complex. The α-helix glutamate is needed to anchor the active site metal ion.

Striking similarities of the GIY-YIG and ββα-Me active sites

Analogies between GIY-YIG and ββα-Me nucleases at the biochemical level have been noted earlier by other authors (25,32). Both groups bind only a single metal ion per active site. Moreover, both can accept a wide array of metal cations that support DNA cleavage. A structural comparison of the Hpy188I and Hpy99I (28) restriction endonucleases as representatives of the GIY-YIG and ββα-Me enzymes shows that the mechanistic similarities go yet further. (i) In both cases, the place of the divalent metal ion can be occupied by a Na^+ ion from the buffer if no suitable divalent cation is available. (ii) The metal ion is anchored by an acidic residue (Glu149 in Hpy188I and Asp148 in Hpy99I), which however need not be its only amino acid ligand. (iii) The metal ion contacts the proS oxygen atom of the scissile bond phosphate, and the leaving group 3′-oxygen atom. (iv) In both cases, a

water molecule attacks the phosphorus atom of the scissile bond from the back, most likely in a single substitution reaction. (v) The water molecule is activated by a general base in spatially conserved position (Tyr63 in Hpy188I and His149 in Hpy99I). (vi) The secondary structure elements that anchor key catalytic residues are analogous. The general base for activating the water molecule is located in a β-hairpin, and the metal ligand is found in a spatially equivalent α-helix. Do these similarities imply that GIY-YIG and ββα-Me nucleases are monophyletic? While it is difficult to definitely exclude this possibility, there are several arguments against it. First, there is no significant sequence similarity between the two nuclease groups to suggest divergent evolution from a common ancestor protein. Second, the β-hairpins in GIY-YIG and ββα-Me nucleases are traversed in opposite directions. Third, many ββα-Me, but not GIY-YIG active site motifs are stabilized by structural Zn^{2+} ions. Finally, strong mechanistic requirements for substrate binding and catalysis place multiple constraints on active sites. Therefore we think that the GIY-YIG and ββα-Me nucleases represent an example of convergent rather than divergent evolution (Figure 6).

Implications for other GIY-YIG nucleases

GIY-YIG nucleases with different cellular functions have been crystallized previously [T4 endo II (5), UvrC (12) and

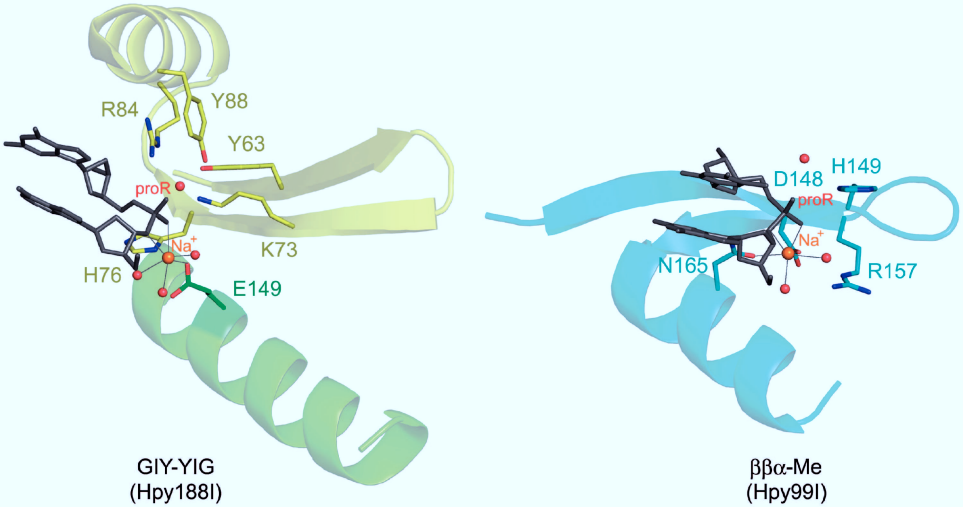

Figure 6. Comparison of the GIY-YIG and ββα-Me nuclease active sites: The dinucleotide around the scissile bond, key active site residues and water molecules are shown in all-atom representation. Secondary structure elements anchoring the catalytic residues are indicated in faint colors.

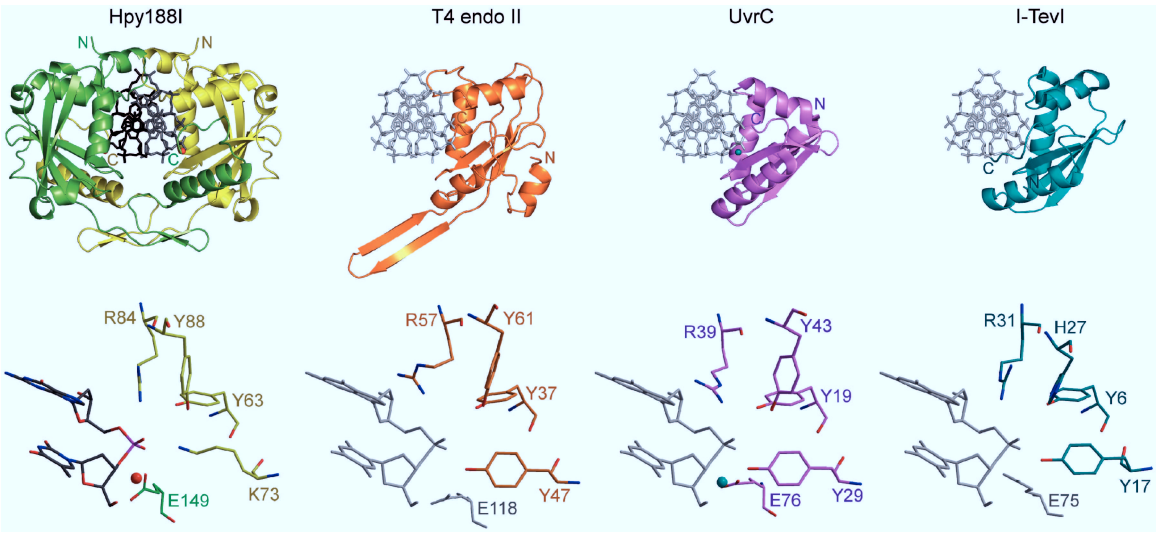

Figure 7. Observed and predicted GIY-YIG nuclease-DNA complexes: Known structures of GIY-YIG nucleases that were crystallized in the absence of DNA were globally superimposed on a single Hpy188I subunit. The top row panels show the composite overall models, the bottom row panels details upon zooming into the active sites. In some cases, an inactivating mutation in the crystallized protein was substituted with its wild-type version (a rotamer choice was suggested by other structures). No further structural adjustments were made, leaving even unlikely conformations (like the extended hairpin in T4 endo II) unaltered. The same applies to the zoom panels of the active sites, which must also require slight adjustments to assign analogous roles to equivalent residues.

I-TevI (2)]. However, because the nuclease domains of these enzymes are not sequence specific, the structures of their DNA complexes were not obtained. We have used the TOP3D program (40) to superimpose these structures on Hpy188I in complex with substrate DNA. In this way the 'composite' models were obtained with the original coordinates for the enzymes (after fixing point mutations that were introduced for crystallization purposes) and the Hpy188I-bound DNA. In T4 endo II the lone hairpin is likely to move, but otherwise all models appear grossly plausible. We then zoomed into the active sites of the composite structures. The catalytic residues in the DNA-free structures are found in correct or nearly correct conformations. Apart from suggesting a fairly rigid active site, this result supports our belief that the catalytic mechanism that we have described for Hpy188I is general for GIY-YIG nucleases (Figure 7).

NOTE ADDED IN PROOF

While this manuscript was under review, similar work by Prof Barry Stoddard and co-workers on the structure of the GIY-YIG restrictase Eco29kI has appeared online [Structure (2010), doi:10.1016/j.str.2010.07.006].

ACCESSION NUMBERS

Hpy188I substrate complex (rcsb061444, PDB ID code 3OQG) Hpy188I product complex (rcsb061467, PDB ID code 3OR3).

ACKNOWLEDGEMENTS

The authors are grateful for access to the DESY (Hamburg, Germany), BESSY (Berlin, Germany) and DIAMOND (Didcot, UK) protein crystallography beamlines. In particular, the authors thank Hans Bartunik (BW6, DORIS, DESY) for very generous allocation of beamtime on short notice.

FUNDING

Foundation for Polish Science START Scholarship (to M.S.); EC FP7 grant 'Proteins in Health and Disease' (HEALTH-PROT, GA No 229676); EMBO/HHMI Young Investigator Award; Polish Ministry of Science and Higher Education grant (0295/B/PO1/2008/34); Cardiff University funding (to M.B.). Funding for open access charge: Polish Ministry of Science and Higher Education grant (0295/B/PO1/2008/34).

Conflict of interest statement. None declared.

REFERENCES

1. Dunin-Horkawicz,S., Feder,M. and Bujnicki,J.M. (2006) Phylogenomic analysis of the GIY-YIG nuclease superfamily. *BMC Genomics*, **7**, 98.
2. Van Roey,P., Meehan,L., Kowalski,J.C., Belfort,M. and Derbyshire,V. (2002) Catalytic domain structure and hypothesis for function of GIY-YIG intron endonuclease I-TevI. *Nat. Struct. Biol.*, **9**, 806–811.
3. Ibryashkina,E.M., Zakharova,M.V., Baskunov,V.B., Bogdanova,E.S., Nagornykh,M.O., Den'mukhamedov,M.M., Melnik,B.S., Kolinski,A., Gront,D., Feder,M. *et al.* (2007) Type II restriction endonuclease R.Eco29kI is a member of the GIY-YIG nuclease superfamily. *BMC Struct. Biol.*, **7**, 48.
4. Lagerback,P. and Carlson,K. (2008) Amino acid residues in the GIY-YIG endonuclease II of phage T4 affecting sequence recognition and binding as well as catalysis. *J. Bacteriol.*, **190**, 5533–5544.
5. Andersson,C.E., Lagerback,P. and Carlson,K. (2010) Structure of bacteriophage T4 endonuclease II mutant E118A, a tetrameric GIY-YIG enzyme. *J. Mol. Biol.*, **397**, 1003–1016.
6. Kowalski,J.C., Belfort,M., Stapleton,M.A., Holpert,M., Dansereau,J.T., Pietrokovski,S., Baxter,S.M. and Derbyshire,V. (1999) Configuration of the catalytic GIY-YIG domain of intron endonuclease I-TevI: coincidence of computational and molecular findings. *Nucleic Acids Res.*, **27**, 2115–2125.
7. Derbyshire,V., Kowalski,J.C., Dansereau,J.T., Hauer,C.R. and Belfort,M. (1997) Two-domain structure of the td intron-encoded endonuclease I-TevI correlates with the two-domain configuration of the homing site. *J. Mol. Biol.*, **265**, 494–506.
8. Xu,Q., Stickel,S., Roberts,R.J., Blaser,M.J. and Morgan,R.D. (2000) Purification of the novel endonuclease, Hpy188I, and cloning of its restriction-modification genes reveal evidence of its horizontal transfer to the *Helicobacter pylori* genome. *J. Biol. Chem.*, **275**, 17086–17093.
9. Truglio,J.J., Croteau,D.L., Van Houten,B. and Kisker,C. (2006) Prokaryotic nucleotide excision repair: the UvrABC system. *Chem. Rev.*, **106**, 233–252.
10. Croteau,D.L., DellaVecchia,M.J., Perera,L. and Van Houten,B. (2008) Cooperative damage recognition by UvrA and UvrB: identification of UvrA residues that mediate DNA binding. *DNA Repair*, **7**, 392–404.
11. Karakas,E., Truglio,J.J., Croteau,D., Rhau,B., Wang,L., Van Houten,B. and Kisker,C. (2007) Structure of the C-terminal half of UvrC reveals an RNase H endonuclease domain with an Argonaute-like catalytic triad. *EMBO J.*, **26**, 613–622.
12. Truglio,J.J., Rhau,B., Croteau,D.L., Wang,L., Skorvaga,M., Karakas,E., DellaVecchia,M.J., Wang,H., Van Houten,B. and Kisker,C. (2005) Structural insights into the first incision reaction during nucleotide excision repair. *EMBO J.*, **24**, 885–894.
13. Moolenaar,G.F., van Rossum-Fikkert,S., van Kesteren,M. and Goosen,N. (2002) Cho, a second endonuclease involved in Escherichia coli nucleotide excision repair. *Proc. Natl Acad. Sci. USA*, **99**, 1467–1472.
14. Coulon,S., Gaillard,P.H., Chahwan,C., McDonald,W.H., Yates,J.R. III and Russell,P. (2004) Slx1-Slx4 are subunits of a structure-specific endonuclease that maintains ribosomal DNA in fission yeast. *Mol. Biol. Cell*, **15**, 71–80.
15. Fricke,W.M. and Brill,S.J. (2003) Slx1-Slx4 is a second structure-specific endonuclease functionally redundant with Sgs1-Top3. *Genes Dev.*, **17**, 1768–1778.
16. Rouse,J. (2009) Control of genome stability by SLX protein complexes. *Biochem. Soc. Trans.*, **37**, 495–510.
17. Fekairi,S., Scaglione,S., Chahwan,C., Taylor,E.R., Tissier,A., Coulon,S., Dong,M.Q., Ruse,C., Yates,J.R. III, Russell,P. *et al.* (2009) Human SLX4 is a Holliday junction resolvase subunit that binds multiple DNA repair/recombination endonucleases. *Cell*, **138**, 78–89.
18. Evgen'ev,M.B., Zelentsova,H., Shostak,N., Kozitsina,M., Barskyi,V., Lankenau,D.H. and Corces,V.G. (1997) Penelope, a new family of transposable elements and its possible role in hybrid dysgenesis in *Drosophila virilis*. *Proc. Natl Acad. Sci. USA*, **94**, 196–201.
19. Pyatkov,K.I., Arkhipova,I.R., Malkova,N.V., Finnegan,D.J. and Evgen'ev,M.B. (2004) Reverse transcriptase and endonuclease activities encoded by Penelope-like retroelements. *Proc. Natl Acad. Sci. USA*, **101**, 14719–14724.
20. Saguez,C., Lecellier,G. and Koll,F. (2000) Intronic GIY-YIG endonuclease gene in the mitochondrial genome of *Podospora*

curvicolla: evidence for mobility. *Nucleic Acids Res.*, **28**, 1299–1306.
21. Burt,A. and Trivers,R. (2006) *Genes in Conflict: The Biology of Selfish Genetic Elements*. Belknap Press of Harvard University Press, Cambridge, MA.
22. Stoddard,B.L. (2005) Homing endonuclease structure and function. *Q. Rev. Biophys.*, **38**, 49–95.
23. Van Roey,P., Waddling,C.A., Fox,K.M., Belfort,M. and Derbyshire,V. (2001) Intertwined structure of the DNA-binding domain of intron endonuclease I-TevI with its substrate. *EMBO J.*, **20**, 3631–3637.
24. Kaminska,K.H., Kawai,M., Boniecki,M., Kobayashi,I. and Bujnicki,J.M. (2008) Type II restriction endonuclease R.Hpy188I belongs to the GIY-YIG nuclease superfamily, but exhibits an unusual active site. *BMC Struct. Biol.*, **8**, 48.
25. Gasiunas,G., Sasnauskas,G., Tamulaitis,G., Urbanke,C., Razaniene,D. and Siksnys,V. (2008) Tetrameric restriction enzymes: expansion to the GIY-YIG nuclease family. *Nucleic Acids Res.*, **36**, 938–949.
26. Orlowski,J. and Bujnicki,J.M. (2008) Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses. *Nucleic Acids Res.*, **36**, 3552–3569.
27. Pingoud,A. and Jeltsch,A. (2001) Structure and function of type II restriction endonucleases. *Nucleic Acids Res.*, **29**, 3705–3727.
28. Sokolowska,M., Czapinska,H. and Bochtler,M. (2009) Crystal structure of the beta beta alpha-Me type II restriction endonuclease Hpy99I with target DNA. *Nucleic Acids Res*, **37**, 3799–3810.
29. Grazulis,S., Manakova,E., Roessle,M., Bochtler,M., Tamulaitiene,G., Huber,R. and Siksnys,V. (2005) Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc. Natl Acad. Sci. USA*, **102**, 15797–15802.
30. Sasnauskas,G., Connolly,B.A., Halford,S.E. and Siksnys,V. (2007) Site-specific DNA transesterification catalyzed by a restriction enzyme. *Proc. Natl Acad. Sci. USA*, **104**, 2115–2120.
31. Miyazono,K., Watanabe,M., Kosinski,J., Ishikawa,K., Kamo,M., Sawasaki,T., Nagata,K., Bujnicki,J.M., Endo,Y., Tanokura,M. *et al.* (2007) Novel protein fold discovered in the PabI family of restriction enzymes. *Nucleic Acids Res.*, **35**, 1908–1918.
32. Belfort,M., Stoddard,B.L., Wood,D.W. and Derbyshire,V. (2005) *Homing Endonucleases and Inteins*. Springer, Berlin.
33. Wang,J.C. (1996) DNA topoisomerases. *Annu. Rev. Biochem.*, **65**, 635–692.
34. Gellert,M. (1981) DNA topoisomerases. *Annu. Rev. Biochem.*, **50**, 879–910.
35. Guo,F., Gopaul,D.N. and van Duyne,G.D. (1997) Structure of Cre recombinase complexed with DNA in a site-specific recombination synapse. *Nature*, **389**, 40–46.
36. Power,D.A. and Zimbro,M.J. (2003) *Difco & BBL Manual: Manual of Microbiological Culture Media*. BD Diagnostic Systems, Sparks, MD.
37. Van Duyne,G.D., Standaert,R.F., Karplus,P.A., Schreiber,S.L. and Clardy,J. (1993) Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.*, **229**, 105–124.
38. Schneider,T.R. and Sheldrick,G.M. (2002) Substructure solution with SHELXD. *Acta. Crystallogr. D Biol. Crystallogr.*, **58**, 1772–1779.
39. Sheldrick,G.M. (2002) Macromolecular phasing with SHELXE. *Zeitschrift für. Kristallographie*, **217**, 644–650.
40. Collaborative Computational Project Number 4. (1994) The CCP4 suite: programs for protein crystallography. *Acta. Crystallogr. D Biol. Crystallogr.*, **50**, 760–763.
41. Cowtan,K. (1994) DM: an automated procedure for phase improvement by density modification. *Joint CCP4 ESF-EACBM Newslett. Protein Crystallogr.*, **31**, 34–38.
42. Morris,R.J., Perrakis,A. and Lamzin,V.S. (2003) ARP/wARP and automatic interpretation of protein electron density maps. *Methods Enzymol.*, **374**, 229–244.
43. Lu,X.J. and Olson,W.K. (2003) 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.*, **31**, 5108–5121.
44. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta. Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
45. McRee,D.E. (1999) XtalView/Xfit–A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.*, **125**, 156–165.
46. Murshudov,G.N., Vagin,A.A. and Dodson,E.J. (1997) Refinement of macromolecular structures by the maximum-likelihood method. *Acta. Crystallogr. D Biol. Crystallogr.*, **53**, 240–255.
47. Pingoud,A. and Wende,W. (2007) A sliding restriction enzyme pauses. *Structure*, **15**, 391–393.
48. Horton,N.C., Dorner,L.F. and Perona,J.J. (2002) Sequence selectivity and degeneracy of a restriction endonuclease mediated by DNA intercalation. *Nat. Struct. Biol.*, **9**, 42–47.
49. Werner,M.H., Gronenborn,A.M. and Clore,G.M. (1996) Intercalation, DNA kinking, and the control of transcription. *Science*, **271**, 778–784.
50. Mizuuchi,K., Nobbs,T.J., Halford,S.E., Adzuma,K. and Qin,J. (1999) A new method for determining the stereochemistry of DNA cleavage reactions: application to the SfiI and HpaII restriction endonucleases and to the MuA transposase. *Biochemistry*, **38**, 4640–4648.
51. Galburt,E.A., Chevalier,B., Tang,W., Jurica,M.S., Flick,K.E., Monnat,R.J. Jr and Stoddard,B.L. (1999) A novel endonuclease mechanism directly visualized for I-PpoI. *Nat. Struct. Biol.*, **6**, 1096–1099.
52. Shen,B.W., Heiter,D.F., Chan,S.H., Wang,H., Xu,S.Y., Morgan,R.D., Wilson,G.G. and Stoddard,B.L. (2010) Unusual target site disruption by the rare-cutting HNH restriction endonuclease PacI. *Structure*, **18**, 734–743.