Two crystal forms of the restriction enzyme MspI–DNA complex show the same novel structure

QIAN STEVEN XU,^{1,3} RICHARD J. ROBERTS,² AND HWAI-CHEN GUO¹

¹Department of Physiology and Biophysics, Boston University School of Medicine, Boston, Massachusetts 02118, USA

²New England Biolabs, Beverly, Massachusetts 01915, USA

(RECEIVED May 2, 2005; FINAL REVISION July 5, 2005; ACCEPTED July 5, 2005)

Abstract

The crystal structure of the Type IIP restriction endonuclease MspI bound to DNA containing its cognate recognition sequence has been determined in both monoclinic and orthorhombic space groups. Significantly, these two independent crystal forms present an identical structure of a novel monomer–DNA complex, suggesting a functional role for this novel enzyme–DNA complex. In both crystals, MspI interacts with the CCGG DNA recognition sequence as a monomer, using an asymmetric mode of recognition by two different structural motifs in a single polypeptide. In the crystallographic asymmetric unit, the two DNA molecules in the two MspI–DNA complexes appear to stack with each other forming an end-to-end pseudo-continuous 19-mer duplex. They are primarily B-form and no major bends or kinks are observed. For DNA recognition, most of the specific contacts between the enzyme and the DNA are preserved in the orthorhombic structure compared with the monoclinic structure. A cation is observed near the catalytic center in the monoclinic structure at a position homologous to the 74/45 metal site of EcoRV, and the orthorhombic structure also shows signs of this same cation. However, the coordination ligands of the metal are somewhat different from those of the 74/45 metal site of EcoRV. Combined with structural information from other solved structures of Type II restriction enzymes, the possible relationship between the structures of the enzymes and their cleavage behaviors is discussed.

Keywords: restriction enzyme; MspI; protein–DNA complex; crystal structure

Type II restriction enzymes comprise one of the major families of endonucleases. Of more than 3600 restriction endonucleases currently described in REBASE (<http://rebase.neb.com>), over 98% are of Type II (Roberts et al. 2005). They usually recognize short DNA sequences of 4–8 bp in length, requiring only Mg^{2+} as a cofactor for DNA cleavage (Roberts and Halford 1993). Compari-

son of the protein sequences reveals little or no similarity, suggesting that diverse strategies are used for DNA recognition. Recognition is also highly specific, since a single base pair change in the cognate DNA sequence can reduce catalytic efficiency of the enzyme by as much as a million fold (Roberts and Halford 1993). All of these remarkable features make this family of enzymes an ideal system for structural studies of protein–DNA interactions.

To date, at least 17 crystal structures of Type II endonucleases have been reported, including EcoRI (Kim et al. 1990), EcoRV (Winkler et al. 1993), BamHI (Newman et al. 1994, 1995), PvuII (Athanasiadis et al. 1994; Cheng et al. 1994), Cfr10I (Bozic et al. 1996), BglI (Newman et al. 1998), FokI (Wah et al. 1997, 1998), MunI

³Present address: Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA.

Reprint requests to: Qian Steven Xu, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA; e-mail: QSXu@lbl.gov; fax: (510) 486-6059.

Article and publication are at <http://www.proteinscience.org/cgi/doi/10.1110/ps.051565105>.

(Deibert et al. 1999), BglII (Lukacs et al. 2000), NgoMIV (Deibert et al. 2000), NaeI (Huai et al. 2000, 2001), BsoBI (van der Woerd et al. 2001), HincII (Horton et al. 2002), Bse634I (Grazulis et al. 2002), EcoRII (Zhou et al. 2004), BstYI (Townson et al. 2004), and HinPII (Yang et al. 2005). Most of them belong to a subtype (some of them, e.g., NaeI, belong to more than one subtype), called Type IIP, which recognize palindromic DNA sequences of from 4 to 8 bp and cleave both strands of the DNA at fixed symmetrical locations either within the sequence or immediately adjacent to it (Roberts et al. 2003). Consistent with their symmetric DNA recognition and cleavage patterns, they are usually assembled as symmetric dimers or tetramers. The subunit fold of these solved structures revealed that despite their lack of sequence similarity, they are all $\alpha\beta$ proteins with a similar central core consisting of a mixed β -sheet flanked by α -helices on both sides (Aggarwal 1995; Pingoud and Jeltsch 1997, 2001). The conserved core harbors the catalytic center in all enzymes. Nevertheless, beyond the common core these structures show great diversity (Pingoud and Jeltsch 1997, 2001) with the strongest additional structural similarity being exhibited by endonucleases that share a similar cleavage pattern. For instance, EcoRV (Winkler et al. 1993) and PvuII (Athanasiadis et al. 1994), both blunt-end cutters, show remarkable structural similarity. They form similar homodimers and bind their 6-bp recognition sequences from the minor groove side. In contrast, BamHI (Newman et al. 1994, 1995) and EcoRI (Kim et al. 1990), both 5' four-base sticky end cutters, are similar to each other but are topologically distinct from EcoRV and PvuII. They have very different dimerization modes as well and bind to their 6-bp recognition sequences from the major groove side. Hence, it has been proposed that Type II endonucleases could be classified into subgroups according to their cleavage patterns instead of recognition sequences (Aggarwal 1995; Guo 2003).

The crystal structures of some specific enzyme–DNA complexes revealed certain striking conformational changes of the DNA induced by the protein. A typical example is a specific EcoRV–DNA complex (Winkler et al. 1993), in which the DNA has a sharp bend of $\sim 50^\circ$. It is also unwound in the middle, and the two central base pairs are unstacked. The bend leads to a compressed major groove and a wide and shallow minor groove. A similar central kink has also been observed in some other specific enzyme–DNA complexes, such as EcoRI (Kim et al. 1990) and MunI (Deibert et al. 1999), accompanied by unwinding of the DNA in the middle. However, not all specific enzyme–DNA complexes display this type of major kink or distortion of the DNA. Some have only localized unwinding and overall bending but without a central kink, e.g., BglI–DNA (Newman et al. 1998) and

BglII–DNA (Lukacs et al. 2000) complexes, others do not present significant DNA deformation, e.g., BamHI–DNA (Newman et al. 1995) and PvuII–DNA complexes (Cheng et al. 1994). Therefore, whether and to what extent DNA distortion would occur may still depend on individual endonucleases and their interactions with the DNA substrates.

Mg^{2+} is an essential cofactor for almost all Type II restriction endonucleases (Roberts and Halford 1993). Although several models have been proposed to account for the catalytic mechanism of restriction enzymes (Pingoud and Jeltsch 1997, 2001), each individual enzyme structure revealed great diversity in the details of the catalytic process, one of which is the number of cations involved in catalysis (Kovall and Matthews 1999). For instance, three metal ion binding sites have been identified for EcoRV, although not all are occupied simultaneously. Based on structural information and complementary biochemical studies, several mechanisms, involving one (Jeltsch et al. 1992, 1993), two (Kostrewa and Winkler 1995; Vipond et al. 1995), or even three metals (Horton et al. 1998), have been proposed for EcoRV, yet the precise function of each metal ion in catalysis is still a matter of debate (Pingoud and Jeltsch 1997, 2001).

MspI is a 262-amino-acid Type IIP endonuclease, originally isolated from *Moraxella species* (Nwankwo and Wilson 1988; Lin et al. 1989), that recognizes and cleaves the DNA sequence 5'-CCGG between the two cytosines to produce two-base 5' overhangs. To understand further the structural basis of DNA recognition, the catalytic mechanism, and how the structural organization is related to the recognition sequences and cleavage patterns for Type II restriction enzymes, we have solved the crystal structures of a specific MspI–DNA complex in a monoclinic space group and an orthorhombic space group, at 1.95 Å and 2.7 Å resolution, respectively. An unusual feature of the protein–DNA interaction, based on the high-resolution structure in the monoclinic crystal form, has been reported elsewhere (Xu et al. 2004). Here we present all other structural details for both crystal forms, as well as the details of the crystallographic methods for both space groups. In addition, structural implications inspired by the MspI structure are discussed.

Results and Discussion

The crystal structure of MspI in complex with its cognate DNA in the orthorhombic space group $P2_12_12_1$ has now been determined at 2.7 Å resolution. Like the structure of the same complex in a monoclinic space group (Xu et al. 2004), an MspI monomer and not a dimer binds to the palindromic DNA sequence, showing asym-

metric DNA recognition by the MspI subunit. Indeed, these two structures in the different space groups are nearly identical, suggesting that this novel monomer–DNA complex is not merely a crystallographic artifact. Superimposition of all main chain atoms of the two enzyme subunits in the asymmetric unit between the two structures gives an RMSD of 0.44 Å. If all atoms of the equivalent DNA base pairs in the two structures are taken into account, the difference would be 0.56 Å.

In the asymmetric unit, there are two enzyme–DNA complexes that are related by noncrystallographic symmetry (NCS) (red and green in Fig. 1). The two NCS-related enzyme subunits make very limited contacts. Calculation of the accessible surface area indicates that a surface of only 220 Å² from each monomer is buried at the interface between the two MspI monomers. This value is much smaller than the lower limit of the range (670 Å²) reported for specific protein–protein contacts of typical oligomeric proteins (Janin et al. 1988), indicating that these two monomers most likely form just a crystallographic dimer. Intriguingly, the two DNA molecules in the two complexes appear to stack with each other forming an end-to-end pseudo-continuous 19-mer duplex (Figs. 1, 2). In the current model, there are 9 bp on strand W1 and C1 of Complex I and 10 bp on strand W2 and C2 of Complex II. The missing base pair of Complex I was opened and disordered at the joint region. Strand W1 and strand C1 in Complex I are NCS related to strand W2 and strand C2 in Complex II, respectively. Due to the DNA helix, however, strand W1 is pseudo-connected to strand C2, and strand W2 to strand C1 in an appropriate orientation (5'→3'). As a consequence, the end base pair Gua10(W2):Cyt11(C2) in Complex II stacks with the base pair Gua9(W1):Cyt12(C1) in Complex I (guanine with

cytosine) to form the pseudo-continuous 19-mer duplex and the helical axes of the two DNA molecules appear to join up smoothly (Fig. 2). Nonetheless, some helical parameters at the junction deviate significantly from the normal values to permit this intermolecular stacking interaction. The twist angles between Gua9(W1):Cyt12(C1) and Gua10(W2):Cyt11(C2) base pairs, and Gua10(W2):Cyt11(C2) and Gua9(W2):Cyt12(C2) base pairs are 49° and 42°, respectively, as calculated using the 3DNA algorithm (Lu et al. 2000), in contrast to the mean value of 36° for B-DNA. The values of rise for these two base pair steps are both 3.7 Å.

Apart from the joint region, the two DNA duplexes are essentially identical to each other with an RMSD of 0.10 Å for all atoms of equivalent fragments. They are primarily B-form and no major bends or kinks of the type present in EcoRI- (Kim et al. 1990) or EcoRV-cognate DNA complexes (Winkler et al. 1993) are observed. Taking the DNA duplex in Complex I as an example, the average helical and conformational parameters are 33° for twist and 3.5 Å for rise as determined by CURVES (Lavery and Sklenar 1988, 1989) (Table 1), with most of the sugar puckers in the C2'-endo conformation preferred by B-DNA (Table 2). Notably, the helical axis of the DNA shows a sigmoid curvature, in which a point between the base pairs Gua6:Cyt15 and Gua7:Cyt14 seems to be the joint point between the two halves of the curvature (Fig. 2). This finding is consistent with the change in the inclination angles of the base pairs along the global axis of the DNA (Table 1). The duplex between base pairs Cyt1:Gua20 and Gua6:Cyt15 bends away from the protein by 21°, whereas the remaining part bends toward the protein by 9°. It is noteworthy that most of the larger deviations from ideal B-form DNA occur in the

Figure 1. Two NCS-related MspI–DNA complexes I and II in the asymmetric unit. The enzyme subunits are represented as red and green ribbons, and the DNA molecules are shown as blue and brown bond models in Complex I and II, respectively. The view in *B* is related to that in *A* by a rotation of 90° along the vertical axis.

Figure 2. Two NCS-related DNA molecules in the asymmetric unit form a pseudo-continuous 19-mer duplex. (A) Two NCS-related DNA molecules are shown in blue and brown. The corresponding DNA strands in the two monomers are labeled W1, C1 and W2, C2. The lines colored in red and green mark the sigmoid curvature of the helical axes of two DNA molecules. The MspI subunits are shown as ribbons in gray and the upper MspI unit is made partly transparent to reveal the DNA molecule behind it. The view is in the same orientation as that in Figure 1A. (B) Electron density around the junction region of the two DNA duplexes. The DNA molecules are shown in bond representation with carbon in gray, nitrogen in blue, oxygen in red, and phosphorus in pink. The nucleotide bases are labeled and the strand names are in parentheses.

part of the DNA duplex that bends toward the protein. Within this region, the twist can be as small as 24° and the roll as large as 10° . The largest positive propeller angle (15°) occurs in this part as well (Table 1). For the sugar-phosphate backbone, some torsion angles at base pair Gua6:Cyt15 (α and γ for both Gua6 and Cyt15 bases, and χ for Gua6 base) lie outside of the usual range of B-

form DNA (Schneider et al. 1997) (Table 2). Moreover, the sugar pucker modes of Cyt5 base and Gua6 base are C1'-exo and C4'-exo, respectively, while all other bases show a C2'-endo conformation.

DNA recognition in the orthorhombic structure is nearly identical to that in the monoclinic structure. The enzyme approaches and recognizes the DNA mainly from the

Table 1. Selected helical parameters of the cognate DNA fragment as determined by CURVES (Lavery and Sklenar 1988, 1989)

Base pair	Global axis	Intrabase pair			Interbase pair			
	Inclination	Buckle ($^\circ$)	Propel ($^\circ$)	Open ($^\circ$)	Rise ($\text{\AA}$)	Tilt ($^\circ$)	Roll ($^\circ$)	Twist ($^\circ$)
C1:G20	5.41	-2.80	-8.36	-2.07	3.42	-5.63	1.26	38.17
C2:G19	2.40	-8.08	-3.86	-1.09	3.24	-3.67	2.22	32.38
C3:G18	0.74	-5.51	2.69	-0.85	3.10	-3.78	-2.20	33.93
C4:G17	-1.95	4.80	-0.54	-1.92	3.65	-7.05	7.54	37.46
C5:G16	-9.63	-7.33	9.59	0.49	4.21	5.80	7.70	29.34
G6:C15	-11.57	-18.90	-0.78	1.35	3.49	-4.76	10.36	23.92
G7:C14	-12.52	1.95	15.13	1.16	3.70	3.79	-0.16	35.48
G8:C13	-6.49	7.40	-8.08	-1.60	3.06	1.63	4.74	34.22
G9:C12	-4.73	18.80	-6.52	-6.22				
Average	-4.26	-1.08	-0.08	-1.20	3.48	-1.71	3.93	33.11
B-DNA ^a	1.5	0.0	-13.3	0.0	3.38	0.0	0.0	36.0
A-DNA ^a	20.7	0.0	-7.5	0.0	2.56	0.0	0.0	32.7

^a Average values from Hartmann and Lavery (1996).

Table 2. Selected sugar-phosphate conformational parameters of the cognate DNA fragment as determined by CURVES (Lavery and Sklenar 1988, 1989)

	Base	Alpha (°)	Gamma (°)	Delta (°)	Chi (°)	Sugar pucker mode
Strand W1	C1	—	60.86	146.77	−113.37	C2'-endo
	C2	−70.13	53.19	136.02	−113.15	C2'-endo
	C3	−56.22	48.34	137.30	−110.79	C2'-endo
	C4	−51.80	41.34	135.22	−106.31	C2'-endo
	C5	−48.51	30.98	119.66	−133.27	C1'-exo
	G6	112.24	−147.66	109.63	−169.73	C4'-exo
	G7	−38.37	41.69	139.96	−133.61	C2'-endo
	G8	−39.76	44.98	146.90	−109.25	C2'-endo
	G9	−62.60	48.33	142.14	−97.32	C2'-endo
Strand C1	G20	−51.87	42.89	149.26	−109.01	C2'-endo
	G19	−42.78	45.78	139.57	−110.03	C2'-endo
	G18	−60.72	44.33	141.83	−101.28	C2'-endo
	G17	−53.26	44.21	145.03	−115.11	C2'-endo
	G16	−60.54	46.55	146.20	−116.72	C2'-endo
	C15	45.92	−78.33	139.50	−113.14	C2'-endo
	C14	−55.36	44.22	141.83	−107.55	C2'-endo
	C13	−66.23	51.09	135.25	−109.77	C2'-endo
	C12	—	70.70	127.96	−136.02	C2'-endo
B-DNA ^a		−30 – −90	20–80	70–180	−60 – −160	—
A-DNA ^a		−170 – −180 or −40 – −100	30–80 or 140–160	60–100	−110 – −200	—

^a Populated ranges of torsion angles from Schneider et al. (1997).

major groove through the small three-stranded recognition β -sheet, and the two symmetric half-sites of the DNA are recognized asymmetrically by the MspI monomer, although the numbers of specific and nonspecific interactions (especially the water-mediated ones) between the enzyme and the DNA are smaller as compared with the monoclinic structure. However, this may be due to the low resolution of the structure (Fig. 3). Within the 5'-CCGG palindromic recognition sequence, five of six direct contacts between the enzyme and the DNA are retained with the same hydrogen-bond distance criteria of < 3.2 Å as in the monoclinic structure. One difference is the hydrogen bond between the N4 atom of the Cyt15 base and the main chain carbonyl oxygen of Tyr249. The distance between these two atoms is 3.24 Å, compared with 3.10 Å in the monoclinic structure. For the water-mediated interactions, the one between Lys261 and the Gua7 base and the one between Ser251 and the Gua17 base are lost, while three others to the Cyt4, Cyt5, and Gua16 bases are preserved. Together, there are still five direct and three water-mediated hydrogen bonds from the MspI monomer to the 4-bp recognition sequence. For the minor groove side, the water-mediated contact to the Cyt5 base still exists, while the one to the Gua17 base seems too long. The only direct hydrogen bond between the enzyme and the DNA outside the recognition sequence in the monoclinic structure is between Gly252 and the Cyt3 base and is also outside the 3.2 Å cutoff in the orthorhombic structure.

The MspI–DNA complex structure in both crystal forms shows that instead of the PDX_{10–30}(D/E)XK signature sequence usually found in the active center of other restriction enzymes (Anderson 1993), the structural equivalent of this motif in MspI is T⁹⁸DX₁₇NXK, with Asn117 in the place of an Asp or Glu found in other enzymes. Interestingly, a sequence alignment between MspI and its isoschizomer BsuFI shows that the equivalent residue at this position in BsuFI is neither Asp nor Glu, but rather is a Ser (Ser254) (Fig. 4). Otherwise, the catalytic sequence motif of MspI is conserved in BsuFI. Of notable interest is that almost all of the residues of MspI that are involved in DNA recognition from the major groove are also strictly conserved in BsuFI with only one exception, Ser127 is replaced by Thr264 in BsuFI (Fig. 4). The same alignment was also given by the threading of the BsuFI sequence through the MspI structure using the GenThreader (Jones 1999) and SAM-T99 servers (Krogh et al. 1994). These results suggest that the DNA recognition mode of MspI might be conserved in BsuFI, which recognizes the same DNA sequence as MspI.

In the monoclinic structure, a single sodium ion is identified near the catalytic site at a position homologous to the 74/45 site of EcoRV (Fig. 5A) (Kostrewa and Winkler 1995). Similarly, in the orthorhombic structure, a significant electron density peak can also be observed at the same position, suggesting the existence of a binding cation at the same site as well. The sodium ion is coordinated as a

Figure 3. Schematic diagram of the hydrogen bonding between MspI and DNA in Complex I in the orthorhombic crystal form. The DNA recognition sequence is shaded in gray, with the scissile phosphate (C4-C5) indicated with an arrow. One DNA base pair (G10:C11) is omitted in the final model. Amino acids contacting DNA bases are displayed between the two strands of the phosphate backbone. They all bind to DNA bases in the major groove, with the exception of Ser27 that contacts a DNA base in the minor groove. Amino acids that bind to the phosphate backbone of the DNA are displayed on either side of the DNA double strands. Solid lines represent direct hydrogen bonds as follows: Ser127-G6 (OG-N7), Thr248-C14 (OG1-N4), Ser251-G16 (OG-N7), Gln259-G6 (NE2-O6), and Lys261-G7 (NZ-O6) in the recognition sequence. Dotted lines represent water-mediated hydrogen bonds as follows: Glu130-C4 (OE1-N4) through a water that also contacts Gly252 (N), Glu130-C5 (OE1-N4) through another water that also contacts Gln259 (NE2), and Ser251-G16 (N-O6) in the major groove; and Ser27-C5 (OG-O2) in the minor groove.

pentagonal pyramid (Fig. 5B), involving six ligands: the carbonyl oxygen of Asp99 (2.7 Å), the two side chain oxygen atoms of Glu35 (2.4 Å and 2.8 Å, respectively), one

carboxylate oxygen of Asp84 (2.4 Å), and two water molecules (2.3 Å and 2.4 Å, respectively). Comparison of the active sites shows that Glu35 of MspI is structurally equivalent to Glu45 of EcoRV (Fig. 5A), both of which are situated on a strictly conserved α -helix (α 2 in MspI and α B in EcoRV). Glu45 of EcoRV provides side chain carboxylates to the coordination of the divalent cation (Kostreva and Winkler 1995). It has been proposed that Glu45 could either be a catalytic residue in a two-metal-ion mechanism (Kostreva and Winkler 1995; Vipond et al. 1995), or it could play a strictly structural role in a three-metal-ion mechanism (Horton et al. 1998) to correctly orient the crucial carboxylate of Asp74 that is equivalent to Asp99 of MspI. In either case, the carboxylate group of Asp74 in EcoRV is bound to the metal (Fig. 5C). In the MspI-DNA complex, however, no coordination bonds are formed between the sodium ion and the side chain carboxylate group of Asp99, although its main chain carbonyl is ligated to the metal and one of its side chain oxygen atoms makes a hydrogen bond to one of the coordinating water molecules of the cation (Fig. 5B). In fact, the cation is positioned at distances of 4.4 Å and 5.7 Å to the two carboxylate oxygen atoms of Asp99, which is too far to form coordination bonds. Therefore, Glu35 does not interact with the carboxylate of Asp99 via the metal like their counterparts in EcoRV. Furthermore, the distances between the sodium ion and the scissile phosphorus and the leaving O3' oxygen atom are 8.7 Å and 8.9 Å, respectively, which are much too far away for any divalent metal ion binding at this site to stabilize the pentacovalent transition state and/or the leaving anion. However, the possibility cannot be ruled out at this point that big conformational changes occur on the DNA and/or the protein upon the formation of a cleavage-capable MspI-DNA complex (possibly though dimerization). This could result in a

MspI	94	GSPKTDATIRFTFHNQSSRLVP-LNKHSSKKKVSIAEYDVETICTGVGISD-----	144
BsuFI	232	GKPKTDVSV--TIKTNTKELIRNISIKNTREKTVTIHEGSVSDLISRLKLSETDPLSQAL	289
MspI	145	-----GELKELIRKHONDQSAKLETPVQKQRLTELLEPYRERFIRWCVTLRAEKSEGN	198
BsuFI	290	IHFKEVGSKKKLLAEHPNSDKI-----LEENLKLYNRELTEF---LHSPLLNDKI	336
MspI	199	LHPDLLI---RFQVIDREYVDVTIKNIDDYVSDRIAEGSKARKPGFGTGLNWTYASGSKA	255
BsuFI	337	QMVDLIIFTNKEFAVWNR-----DDYIKHYIEEYSG--KGQFGTPFKWTYPSKKRG	384
MspI	256	KMQFKG	262
BsuFI	385	QKIQIKG	391

Figure 4. Sequence alignment between MspI and its isoschizomer BsuFI. Only part of the sequence for each enzyme is shown because there is no significant sequence homology present for the remaining part of the enzymes. Identical residues are shaded in gray. The active-site residues and the residues of MspI involved in DNA recognition from the major groove are indicated by open and closed arrowheads, respectively.

Figure 5. (A) A stereo view of the superimposition of the active sites of MspI and EcoRV. Active site residues and the phosphates at the cleavage sites are shown as ball-and-stick representations. Other parts of the DNA molecules are shown in bond representations. The DNA bases have been omitted for clarity. All atoms in EcoRV are gray and atoms in MspI are colored according to atom type as follows: gold for carbons, blue for nitrogens, red for oxygens, purple for phosphorus, and cyan for sodium. The coordinates of the crystal structure of the T93A EcoRV–DNA–Ca²⁺ complex were used to draw the relevant part of EcoRV (PDB code 1BSS). (B,C) The coordination geometries of the Na⁺ ion near the active site of MspI (B) and the Ca²⁺ ion bound to the 74/45 site of EcoRV (C) (Horton et al. 1998). Both views are in a similar orientation to A. The Na⁺ and Ca²⁺ ions are shown as cyan and green spheres, respectively. Other atoms are colored according to atom type as follows: gray for carbons, blue for nitrogens, and red for oxygens. Coordination bonds to the metal ions and the hydrogen bond are shown in gold and black dashed lines, respectively. The coordinates of the crystal structure of the T93A EcoRV–DNA–Ca²⁺ complex were used to draw the figure for EcoRV (PDB code 1BSS).

change in the coordination geometry of the metal ion and bring it closer to the DNA allowing it to play a structural or a functional role.

It has been shown recently that HinP1I, a Type IIP restriction enzyme, displays striking structural similarity to MspI despite the lack of significant sequence homology (Yang et al. 2005). Similar to MspI, HinP1I recognizes a

palindromic DNA sequence 5′-GCGC and cleaves between the first and second nucleotides, producing two-base 5′ overhangs. Structurally, the two enzymes are highly similar and share almost all of the same structural elements, including both the DNA recognition and catalytic elements (Yang et al. 2005). Although no structure of a specific HinP1I–DNA complex is available so far, the remarkable

structure similarity between HinP1I and MspI suggests that HinP1I is likely to be the second example of asymmetric recognition of a palindromic DNA sequence (5'-GCGC) by two different structural motifs in one polypeptide. Besides HinP1I, MspI shows some structural similarities with several other restriction enzymes, including BglI (Newman et al. 1998), EcoRV (Winkler et al. 1993), PvuII (Athanasiadis et al. 1994), NaeI (N-terminal domain) (Huai et al. 2000), and HincII (Horton et al. 2002), as well as one endonuclease involved in DNA mismatch repair, MutH (Ban and Yang 1998). The superimposition of MspI with these structures indicates that one α -helix ($\alpha 2$), five β -strands (four in MutH) in the central β -sheet, and three β -strands (two in PvuII) in the small antiparallel β -sheet are common to all of them, which form the catalytic core and DNA recognition elements. This common core reflects the common DNA cleavage activity for all of these enzymes. In addition, structural equivalents of helix $\alpha 1$ of MspI can also be found in NaeI and MutH. Except for HinP1I, whose functional form is still unclear, and MutH, a monomeric endonuclease, all other structural homologs of MspI form homodimers and belong to subtype P of the Type II restriction endonucleases (NaeI is also a Type IIE enzyme) that recognize specific palindromic DNA sequences and symmetrically cleave the DNA either within or immediately adjacent to that sequence (Roberts et al. 2003). It is interesting to note that the distance between the DNA recognition half-sites and the corresponding cleavage points of these enzymes is either 0 bp (EcoRV, PvuII, NaeI, and HincII) or 1 bp (BglI, MspI, and HinP1I). In contrast, for other structures of Type IIP enzymes that don't show significant structure similarity with MspI, the cleavage distances are 2 bp (EcoRI, BamHI, MunI, BglII, Cfr10I, NgoMIV, BsoBI, Bse634I, and BstYI, all 5' four-base sticky ends cutters) or 2.5 bp (EcoRII, a 5' five-base sticky end cutter). These enzymes share a different common core structure and use predominantly an α -helix and a loop for DNA recognition, different from MspI and its structural homologs that interact with the cognate DNA sequence via a small β -sheet as described above. More structures of restriction enzymes are necessary to confirm the relationship between the cleavage distance and the core structure organization of restriction enzymes.

In spite of the common core structure and the similar DNA recognition elements among MspI and its structural homologs, however, these enzymes produce distinct staggered ends upon cleavage of DNA. EcoRV (Winkler et al. 1993), PvuII (Athanasiadis et al. 1994; Cheng et al. 1994), NaeI (Huai et al. 2000, 2001) and HincII (Horton et al. 2002) bind a 6-bp sequence (degenerate in the case of HincII) and cleave the DNA to give blunt ends. They also share a similar dimerization scheme. BglI recognizes an interrupted sequence and cleaves the DNA to generate 3' sticky ends (Newman et al. 1998). Correspondingly, its di-

merization mode is different from those of enzymes producing blunt ends. MutH binds a 4-bp palindromic sequence, yet it cleaves only the unmethylated strand of the hemimethylated recognition sequence to form a nick (Ban and Yang 1998). This is consistent with the fact that MutH uses a monomer to bind the hemimethylated palindromic sequence and contains only one catalytic center. In contrast to these enzymes, MspI and HinP1I recognize similar 4-bp DNA sequences and cleave to produce two-base 5' overhanging ends. It remains unclear how MspI can manage to cut both strands of DNA symmetrically to achieve such a pattern since there is only one catalytic site present in the crystal structure of the MspI–DNA complex. Notably, there have been biochemical studies suggesting that some Type IIP enzymes can act as monomers, such as BspRI (Koncz et al. 1978), BsuRI (Bron and Horz 1980), and Sau96I (Szilak et al. 1990). The crystal structure of the monomeric MspI–DNA complex and its novel asymmetric DNA recognition mode raise the possibility that MspI may also function as a monomer despite the lack of biochemical evidence. On the other hand, it cannot be ruled out at this point that, like other Type IIP enzymes, MspI forms a homodimer to recognize its specific DNA sequence, which then achieves double-strand DNA cleavage. A few possible mechanisms for DNA cleavage have been proposed, such as flipping and dimerization, based on the monomer and dimer model, respectively (Xu et al. 2004). Regardless of what mechanism is used, however, the identical structure of the MspI–DNA complex in two different crystal forms suggests a potential functional role of this monomeric enzyme–DNA complex. Altogether, the different patterns of DNA cleavage for all of these enzymes are presumably related to their structural differences including dimerization schemes (if applicable). As more structures of Type II enzymes become available, a better understanding of the relationship between the structural organization and the corresponding cleavage behavior can be anticipated.

Materials and methods

Crystallization

Crystals of MspI in complex with a DNA decamer (5'-CCCCGGGG-3') containing the MspI recognition sequence (underlined) were grown as previously described (O'Loughlin et al. 2000). In the crystals, DNA cleavage was prevented by removing the required metal ion (Mg^{2+}) with EDTA. Crystals of the MspI–DNA complexes grew in two different space groups in different batches even under the same crystallization conditions. One belongs to the monoclinic space group $P2_1$, with unit cell constants $a = 50.2$ Å, $b = 131.6$ Å, $c = 59.3$ Å, and $\beta = 109.7^\circ$, while the other one falls into the orthorhombic space group $P2_12_12_1$, with unit cell constants $a = 50.3$ Å, $b = 111.3$ Å, and $c = 130.7$ Å. Matthews coefficient (V_M) calculations (Matthews 1968) suggest a solvent content of about 53% for both crystals (V_M about 2.6 Å³/Da).

Table 3. Data collection and refinement statistics

	Natives		Derivatives						
	Nat 1	Nat 2	Hg(OAc) ₂ #1		Hg(OAc) ₂ #2		CH ₃ HgCl #1	CH ₃ HgCl #2	Sm(OAc) ₃
Space group	P2 ₁	P2 ₁ 2 ₁ 2 ₁	P2 ₁		P2 ₁		P2 ₁	P2 ₁	P2 ₁
Resolution (Å)	42.6–1.95	46.9–2.7	25–3.2		25–2.9		25–3.4	25–2.0	25–2.7
Wavelength(Å)	0.9207	0.9040	edge	peak	remote	peak	remote	0.9287	0.9207
Total reflections	153,014	213,354	1.0077	1.0060	0.9763	1.0065	1.5418		
Unique reflections	52,241	21,426	70,526	68,630	70,497	104,638	31,615	168,498	65,782
R _{sym} (%) ^{a,b}	4.3 (27.5)	7.2 (28.2)	12,066	12,018	12,068	16,067	11,251	44,579	21,624
< I/σ > ^a	24.5 (3.6)	28.7 (7.0)	8.9 (22.0)	9.5 (23.2)	9.0 (22.5)	7.1 (25.0)	14.2 (23.8)	5.7 (18.4)	5.3 (20.4)
Completeness (%) ^a	97.4 (99.9)	99.9 (100.0)	17.2 (5.8)	16.1 (5.3)	16.5 (5.7)	21.9 (5.1)	7.4 (4.2)	18.1 (6.2)	20.2 (5.4)
Phasing			97.8 (99.9)	97.8 (99.9)	97.6 (98.9)	97.4 (98.5)	96.7 (97.9)	89.3 (74.7)	94.4 (96.2)
No. of sites			MAD phasing		MAD phasing		MIRAS phasing		
Phasing power			6		4		4		
centric			—	0.21	0.26	—	1.33	1.08	0.54
acentric			—	0.36	0.43	—	1.90	1.55	0.74
Figure of merit (f.o.m.)			0.35		0.25		0.45		
Phase combination									
(overall f.o.m.)					0.54				
Density modification					0.79				
(overall f.o.m.)									
Refinement									
Reflections									
(working/test)	46,088/5172	18,789/2088							
R _{cyst} / R _{free} ^c	0.223/0.252	0.226/0.274							
Protein residues/DNA									
base pairs	524/18	524/19							
Nonhydrogen atoms built in the asymmetric unit									
Protein	4188	4206							
DNA	726	767							
Water molecules	331	53							
Ions	2	0							
RMS deviation									
Bonds (Å)	0.005	0.007							
Angles (°)	1.16	1.21							
Average B value (Å ²)									
(protein/DNA/water)	27.9/27.2/29.2	28.0/26.5/23.0							
Ramachandran analysis									
Most favored (%)	90.5	88.6							
Allowed (%)	9.5	11.4							

^a Numbers in parentheses are for the highest resolution shell.^b $R_{\text{sym}} = (\sum_h \sum_i |I_{hi} - \langle I_h \rangle|) / (\sum_h \sum_i I_{hi})$, where $\langle I_h \rangle$ is the average intensity of n independent observations I_{hi} of a given reflection h .^c $R_{\text{cryst}} = \sum_h |F_{\text{obs}} - F_{\text{calc}}| / \sum_h |F_{\text{obs}}|$, where F_{obs} and F_{calc} are the observed and calculated structure factor amplitudes, respectively. The R_{free} was calculated using randomly selected 10% reflections that were omitted from all stages of the refinement.

Data collection and processing

Two native data sets have been collected to 1.95 Å for a monoclinic crystal and 2.7 Å for an orthorhombic crystal, respectively, at 100 K at the National Synchrotron Light Source (NSLS) beamline X12C, Brookhaven National Laboratory. The MspI–DNA complex crystals have been soaked in a cryoprotecting buffer containing 20% (v/v) glycerol (other components are the same as in the crystallization buffer) typically for 5 min, and flash-frozen into liquid nitrogen for storage and transfer to the synchrotron. Heavy atom derivatives were obtained by soaking the native monoclinic crystals of the MspI–DNA complex in various heavy atom solutions for 5 to 6 h. All soakings were carried out in the dark due to the vigorous photochemistry of many heavy atom compounds. One three-wavelength MAD (multiwavelength anomalous diffraction) data set and one two-wavelength MAD data set were collected at the NSLS beamline X12C for two Hg(OAc)₂ derivatives, respectively. For two CH₃HgCl derivatives, diffraction data were measured using the NSLS beamline X4A (with anomalous signal) and on a Rigaku RU-300 rotating anode generator equipped with a RAXIS-IIc image plate detector, respectively. Furthermore, a Sm(OAc)₃ derivative data set has been collected using the NSLS beamline X12C. All native and derivative data were processed using the HKL suite (Otwinowski and Minor 1997) and converted to structure factors using TRUNCATE from the CCP4 suite (CCP4 1994). The data collection statistics are shown in Table 3.

Structure determination and refinement

For the monoclinic crystals, the heavy atom sites of two non-identical (independent) Hg(OAc)₂ derivative MAD data sets were determined and refined, and the MAD phases were calculated using the program SOLVE (Terwilliger and Berendzen 1999). The heavy atom positions in two other nonidentical (independent) CH₃HgCl and one Sm(OAc)₃ derivative data sets were identified by inspection of isomorphous and anomalous difference Patterson maps calculated in XtalView (McRee 1999). Cross-phased difference Fourier syntheses were used to verify and place all of the heavy atom sites on the same origin, then the positions were refined and the MIRAS (multiple isomorphous replacement with anomalous scattering) phases were calculated using MLPHARE from the CCP4 suite (CCP4 1994). The combination of the MAD phases and the MIRAS phases were carried out using CNS (Brunger et al. 1998) by a simple addition of the phase probability distribution coefficients, also known as Hendrickson–Lattman coefficients (Hendrickson and Lattman 1970) giving an overall figure of merit of 0.54 at 2.7 Å resolution. The electron density maps after density modification were significantly improved with an overall figure of merit of 0.79 and allowed sequence assignment and initial model building with the program O (Jones et al. 1991). Refinement was carried out in CNS (Brunger et al. 1998) against the monoclinic native data set and the model was completed after several alternating cycles of manual building, stepwise resolution extension, and refinement to 1.95 Å resolution. Program SIGMAA (CCP4 1994) was used in the early cycles of refinement and manual building to combine model phases with experimental phases. Strict noncrystallographic symmetry (NCS) restraints or tight NCS restraints were initially applied, and in later stages of refinement, NCS restraints were loosened as guided by the behavior of R_{free} . Finally, ions and water molecules have been added. The refinement statistics of the structure are shown in Table 3.

For the orthorhombic crystal, the structure was determined by molecular replacement (MR) with CNS (Brunger et al. 1998) using the structure of the monoclinic crystal (without DNA) as the initial probe model. Subsequent refinements were carried out in CNS (Brunger et al. 1998) to 2.7 Å resolution and reached convergence after several alternating cycles of manual building and refinement. NCS restraints were applied throughout the refinement and modified as guided by the behavior of R_{free} . The refinement statistics of the structure are shown in Table 3. Atomic coordinates and structure-factor amplitudes have been deposited into the Protein Data Bank (PDB) with accession code 1SA3 for the monoclinic structure and 1YFI for the orthorhombic structure.

Acknowledgments

We thank Dr. Ira Schildkraut for collaboration at the early stage of this research. We are grateful to Rebecca Kucera for her generous support during these studies. We also thank Dr. Debanu Das for helpful discussions and comments on the manuscript; and C. Ogata, A. Saxena, Y. Wang, and X. Qian for assistance with data collection at the National Synchrotron Light Source, Brookhaven National Laboratory.

References

- Aggarwal, A.K. 1995. Structure and function of restriction endonucleases. *Curr. Opin. Struct. Biol.* **5**: 11–19.
- Anderson, J.E. 1993. Restriction endonucleases and modification methylases. *Curr. Opin. Struct. Biol.* **3**: 24–30.
- Athanasiadis, A., Vlassi, M., Kotsifaki, D., Tucker, P.A., Wilson, K.S., and Kokkinidis, M. 1994. Crystal structure of PvuII endonuclease reveals extensive structural homologies to EcoRV. *Nat. Struct. Biol.* **1**: 469–475.
- Ban, C. and Yang, W. 1998. Structural basis for MutH activation in *E. coli* mismatch repair and relationship of MutH to restriction endonucleases. *EMBO J.* **17**: 1526–1534.
- Bozic, D., Grazulis, S., Siksnys, V., and Huber, R. 1996. Crystal structure of *Citrobacter freundii* restriction endonuclease Cfr10I at 2.15 Å resolution. *J. Mol. Biol.* **255**: 176–186.
- Bron, S. and Horz, W. 1980. Purification and properties of the Bsu endonuclease. *Methods Enzymol.* **65**: 112–132.
- Brunger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography & NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D. Biol. Crystallogr.* **54**: 905–921.
- CCP4. 1994. The CCP4 suite: Programs for protein crystallography. *Acta Crystallogr. D. Biol. Crystallogr.* **50**: 760–763.
- Cheng, X., Balendiran, K., Schildkraut, I., and Anderson, J.E. 1994. Structure of PvuII endonuclease with cognate DNA. *EMBO J.* **13**: 3927–3935.
- Deibert, M., Grazulis, S., Janulaitis, A., Siksnys, V., and Huber, R. 1999. Crystal structure of MunI restriction endonuclease in complex with cognate DNA at 1.7 Å resolution. *EMBO J.* **18**: 5805–5816.
- Deibert, M., Grazulis, S., Sasnauskas, G., Siksnys, V., and Huber, R. 2000. Structure of the tetrameric restriction endonuclease NgoMIV in complex with cleaved DNA. *Nat. Struct. Biol.* **7**: 792–799.
- Grazulis, S., Deibert, M., Rimselienė, R., Skirgaila, R., Sasnauskas, G., Lagunavicius, A., Repin, V., Urbanke, C., Huber, R., and Siksnys, V. 2002. Crystal structure of the Bse634I restriction endonuclease: Comparison of two enzymes recognizing the same DNA sequence. *Nucleic Acids Res.* **30**: 876–885.
- Guo, H.-C. 2003. Type II restriction endonucleases: Structures and applications. In *Recent research developments in macromolecules*. (ed. S.G. Pandalai), pp. 225–245. Research Signpost, Trivandrum, India.
- Hartmann, B. and Lavery, R. 1996. DNA structural forms. *Q. Rev. Biophys.* **29**: 309–368.

- Hendrickson, W.A. and Lattman, E.E. 1970. Representation of phase probability distributions for simplified combination of independent phase information. *Acta Crystallogr. B* **26**: 136–143.
- Horton, N.C., Newberry, K.J., and Perona, J.J. 1998. Metal ion-mediated substrate-assisted catalysis in type II restriction endonucleases. *Proc. Natl. Acad. Sci.* **95**: 13489–13494.
- 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.
- Huai, Q., Colandene, J.D., Chen, Y., Luo, F., Zhao, Y., Topal, M.D., and Ke, H. 2000. Crystal structure of NaeI—an evolutionary bridge between DNA endonuclease and topoisomerase. *EMBO J.* **19**: 3110–3118.
- Huai, Q., Colandene, J.D., Topal, M.D., and Ke, H. 2001. Structure of NaeI–DNA complex reveals dual-mode DNA recognition and complete dimer rearrangement. *Nat. Struct. Biol.* **8**: 665–669.
- Janin, J., Miller, S., and Chothia, C. 1988. Surface, subunit interfaces and interior of oligomeric proteins. *J. Mol. Biol.* **204**: 155–164.
- Jeltsch, A., Alves, J., Maass, G., and Pingoud, A. 1992. On the catalytic mechanism of EcoRI and EcoRV. A detailed proposal based on biochemical results, structural data and molecular modelling. *FEBS Lett.* **304**: 4–8.
- Jeltsch, A., Alves, J., Wolfes, H., Maass, G., and Pingoud, A. 1993. Substrate-assisted catalysis in the cleavage of DNA by the EcoRI and EcoRV restriction enzymes. *Proc. Natl. Acad. Sci.* **90**: 8499–8503.
- Jones, D.T. 1999. GenTHREADER: An efficient and reliable protein fold recognition method for genomic sequences. *J. Mol. Biol.* **287**: 797–815.
- Jones, T.A., Zou, J.Y., Cowan, S.W., and Kjeldgaard, 1991. Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A* **47**: 110–119.
- Kim, Y.C., Grable, J.C., Love, R., Greene, P.J., and Rosenberg, J.M. 1990. Refinement of EcoRI endonuclease crystal structure: A revised protein chain tracing. *Science* **249**: 1307–1309.
- Koncz, C., Kiss, A., and Venetianer, P. 1978. Biochemical characterization of the restriction-modification system of *Bacillus sphaericus*. *Eur. J. Biochem.* **89**: 523–529.
- Kostrewa, D. and Winkler, F.K. 1995. Mg²⁺ binding to the active site of EcoRV endonuclease: A crystallographic study of complexes with substrate and product DNA at 2 Å resolution. *Biochemistry* **34**: 683–696.
- Kovall, R.A. and Matthews, B.W. 1999. Type II restriction endonucleases: Structural, functional and evolutionary relationships. *Curr. Opin. Struct. Biol.* **3**: 578–583.
- Krogh, A., Brown, M., Mian, I.S., Sjolander, K., and Haussler, D. 1994. Hidden Markov models in computational biology. Applications to protein modeling. *J. Mol. Biol.* **235**: 1501–1531.
- Lavery, R. and Sklenar, H. 1988. The definition of generalized helicoidal parameters and of axis curvature for irregular nucleic acids. *J. Biomol. Struct. Dyn.* **6**: 63–91.
- . 1989. Defining the structure of irregular nucleic acids: Conventions and principles. *J. Biomol. Struct. Dyn.* **6**: 655–667.
- Lin, P.M., Lee, C.H., and Roberts, R.J. 1989. Cloning and characterization of the genes encoding the MspI restriction modification system. *Nucleic Acids Res.* **17**: 3001–3011.
- Lu, X.J., Shakked, Z., and Olson, W.K. 2000. A-form conformational motifs in ligand-bound DNA structures. *J. Mol. Biol.* **300**: 819–840.
- Lukacs, C.M., Kucera, R., Schildkraut, I., and Aggarwal, A.K. 2000. Understanding the immutability of restriction enzymes: Crystal structure of BglII and its DNA substrate at 1.5 Å resolution. *Nat. Struct. Biol.* **7**: 134–140.
- Matthews, B.W. 1968. Solvent content of protein crystals. *J. Mol. Biol.* **33**: 491–497.
- McRee, D.E. 1999. XtalView/Xfit—A versatile program for manipulating atomic coordinates and electron density. *J. Struct. Biol.* **125**: 156–165.
- Newman, M., Strzelecka, T., Dorner, L.F., Schildkraut, I., and Aggarwal, A.K. 1994. Structure of restriction endonuclease BamHI phased at 1.95 Å resolution by MAD analysis. *Structure* **2**: 439–452.
- . 1995. Structure of BamHI endonuclease bound to DNA: Partial folding and unfolding on DNA binding. *Science* **269**: 656–663.
- Newman, M., Lunnen, K., Wilson, G., Greci, J., Schildkraut, I., and Phillips, S.E. 1998. Crystal structure of restriction endonuclease BglI bound to its interrupted DNA recognition sequence. *EMBO J.* **17**: 5466–5476.
- Nwankwo, D.O. and Wilson, G.G. 1988. Cloning and expression of the MspI restriction and modification genes. *Gene* **64**: 1–8.
- O’Loughlin, T.J., Xu, Q., Kucera, R.B., Dorner, L.F., Sweeney, S., Schildkraut, I., and Guo, H.C. 2000. Crystallization and preliminary X-ray diffraction analysis of MspI restriction endonuclease in complex with its cognate DNA. *Acta Crystallogr. D Biol. Crystallogr.* **56**: 1652–1655.
- Otwinowski, Z. and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**: 307–326.
- Pingoud, A. and Jeltsch, A. 1997. Recognition and cleavage of DNA by type-II restriction endonucleases. *Eur. J. Biochem.* **246**: 1–22.
- . 2001. Structure and function of type II restriction endonucleases. *Nucleic Acids Res.* **29**: 3705–3727.
- Roberts, R.J. and Halford, S.E. 1993. Type II restriction endonucleases. In *Nucleases* (eds. S.M. Linn et al.), pp. 35–88. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- Roberts, R.J., Belfort, M., Bestor, T., Bhagwat, A.S., Bickle, T.A., Bitinaite, J., Blumenthal, R.M., Degtyarev, S., Dryden, D.T., Dybvig, K., et al. 2003. A nomenclature for restriction enzymes, DNA methyltransferases, homing endonucleases and their genes. *Nucleic Acids Res.* **31**: 1805–1812.
- Roberts, R.J., Vincze, T., Posfai, J., and Macelis, D. 2005. REBASE—Restriction enzymes and DNA methyltransferases. *Nucleic Acids Res.* **33**: D230–D232.
- Schneider, B., Neidle, S., and Berman, H.M. 1997. Conformations of the sugar-phosphate backbone in helical DNA crystal structures. *Biopolymers* **42**: 113–124.
- Szilak, L., Venetianer, P., and Kiss, A. 1990. Cloning and nucleotide sequence of the genes coding for the Sau96I restriction and modification enzymes. *Nucleic Acids Res.* **18**: 4659–4664.
- Terwilliger, T.C. and Berendzen, J. 1999. Automated MAD and MIR structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **55**: 849–861.
- Townson, S.A., Samuelson, J.C., Vanamee, E.S., Edwards, T.A., Escalante, C.R., Xu, S.Y., and Aggarwal, A.K. 2004. Crystal structure of BstYI at 1.85 Å resolution: A thermophilic restriction endonuclease with overlapping specificities to BamHI and BglII. *J. Mol. Biol.* **338**: 725–733.
- van der Woerd, M.J., Pelletier, J.J., Xu, S., and Friedman, A.M. 2001. Restriction enzyme BsoBI–DNA complex: A tunnel for recognition of degenerate DNA sequences and potential histidine catalysis. *Structure* **9**: 133–144.
- Vipond, I.B., Baldwin, G.S., and Halford, S.E. 1995. Divalent metal ions at the active sites of the EcoRV and EcoRI restriction endonucleases. *Biochemistry* **34**: 697–704.
- Wah, D.A., Hirsch, J.A., Dorner, L.F., Schildkraut, I., and Aggarwal, A.K. 1997. Structure of the multimodular endonuclease FokI bound to DNA. *Nature* **388**: 97–100.
- Wah, D.A., Bitinaite, J., Schildkraut, I., and Aggarwal, A.K. 1998. Structure of FokI has implications for DNA cleavage. *Proc. Natl. Acad. Sci.* **95**: 10564–10569.
- Winkler, F.K., Banner, D.W., Oefner, C., Tsernoglou, D., Brown, R.S., Heathman, S.P., Bryan, R.K., Martin, P.D., Petratos, K., and Wilson, K.S. 1993. The crystal structure of EcoRV endonuclease and of its complexes with cognate and non-cognate DNA fragments. *EMBO J.* **12**: 1781–1795.
- Xu, Q.S., Kucera, R.B., Roberts, R.J., and Guo, H.C. 2004. An asymmetric complex of restriction endonuclease MspI on its palindromic DNA recognition site. *Structure* **12**: 1741–1747.
- Yang, Z., Horton, J.R., Maunus, R., Wilson, G.G., Roberts, R.J., and Cheng, X. 2005. Structure of HinPII endonuclease reveals a striking similarity to the monomeric restriction enzyme MspI. *Nucleic Acids Res.* **33**: 1892–1901.
- Zhou, X.E., Wang, Y., Reuter, M., Mucke, M., Kruger, D.H., Meehan, E.J., and Chen, L. 2004. Crystal structure of type IIE restriction endonuclease EcoRII reveals an autoinhibition mechanism by a novel effector-binding fold. *J. Mol. Biol.* **335**: 307–319.