

Structure of the metal-independent restriction enzyme Bfil reveals fusion of a specific DNA-binding domain with a nonspecific nuclease

Saulius Grazulis^{*†}, Elena Manakova^{*}, Manfred Roessle[‡], Matthias Bochtler^{§¶}, Giedre Tamulaitiene^{*}, Robert Huber^{†||}, and Virginijus Siksnys^{*†}

^{*}Laboratory of Protein–DNA Interaction, Institute of Biotechnology, Graiciuno 8, LT-02241 Vilnius, Lithuania; [†]European Synchrotron Radiation Facility, B.P. 220, F-38043 Grenoble Cedex, France; [‡]Department of Structural Biology, International Institute of Molecular and Cell Biology, Księcia J. Trojdena Street 4, 02-109 Warsaw, Poland; [§]Max-Planck-Institut für Molekulare Zellbiologie und Genetik, Pfotenhauerstrasse 108, 01307 Dresden, Germany; and ^{||}Abteilung Strukturforschung, Max-Planck-Institut für Biochemie, Am Klopferspitz 18a, D-82152 Martinsried, Germany

Contributed by Robert Huber, September 20, 2005

Among all restriction endonucleases known to date, Bfil is unique in cleaving DNA in the absence of metal ions. Bfil represents a different evolutionary lineage of restriction enzymes, as shown by its crystal structure at 1.9-Å resolution. The protein consists of two structural domains. The N-terminal catalytic domain is similar to Nuc, an EDTA-resistant nuclease from the phospholipase D superfamily. The C-terminal DNA-binding domain of Bfil exhibits a β -barrel-like structure very similar to the effector DNA-binding domain of the Mg^{2+} -dependent restriction enzyme EcoRII and to the B3-like DNA-binding domain of plant transcription factors. Bfil presumably evolved through domain fusion of a DNA-recognition element to a nonspecific nuclease akin to Nuc and elaborated a mechanism to limit DNA cleavage to a single double-strand break near the specific recognition sequence. The crystal structure suggests that the interdomain linker may act as an autoinhibitor controlling Bfil catalytic activity in the absence of a specific DNA sequence. A PSI-BLAST search identified a Bfil homologue in a *Mesorhizobium* sp. BNC1 bacteria strain, a plant symbiont isolated from an EDTA-rich environment.

restriction endonuclease | x-ray crystallography

Restriction endonucleases (REases) protect bacteria by hydrolyzing invading viral or other foreign DNA. Type II REases perform their function by catalyzing the sequence-specific cleavage of double-stranded DNA molecules in the presence of Mg^{2+} ions within or close to their recognition sites (1). Surprisingly, orthodox REases and bacteriophage λ -exonuclease that binds a free end of double-stranded DNA and processively degrades one strand in the 5' to 3' direction share a similar catalytic mechanism and a common structural ancestor (2). A λ -exonuclease-like domain has been identified in many Mg^{2+} -dependent nucleases involved in DNA recombination and repair (3–5), suggesting that it has been remodeled during evolution to perform different functions. To constrain cleavage at specific sites, REases had to develop effective means to control nucleolytic activity of λ -exonuclease-like catalytic domain.

A variety of mechanisms have evolved to maintain REases in inactive configuration to avoid uncontrolled DNA cleavage and couple it to the recognition of specific nucleotide sequence. Orthodox REases, like EcoRI or BamHI, are homodimers that make largely symmetrical interactions with palindromic DNA sequences and contain two distinct sites each responsible for catalyzing cleavage in one DNA strand (1). In orthodox restriction enzymes, structural elements involved in sequence recognition are grafted on the conserved λ -exonuclease-like scaffold that makes a catalytic core (6). Structural analysis indicates that EcoRI amino acid residues involved in specific DNA binding are coupled to the catalytic residues through the intricate network of interactions (7). Therefore, any perturbation in the recognition site (for example, an incorrect base pair) would propagate via the

network to the cleavage site and thus prevent DNA cutting at the incorrect site.

Different mechanisms of the activity control are used by type IIS REases. Most type IIS REases are monomers and recognize asymmetric nucleotide sequences. The archetypal type IIS enzyme FokI exhibits a modular architecture and consists of two functional domains: the N-terminal DNA recognition domain and the C-terminal catalytic domain (8). FokI binds DNA as a monomer, with the N-terminal recognition domain making all of the base-specific contacts at the recognition site (8). The catalytic C-terminal domain of FokI exhibits a λ -exonuclease-like fold similar to the orthodox REases (8) and possesses a weak nonspecific nuclease activity (9). Control of nuclease activity in FokI is achieved by two different mechanisms. First, the catalytic domain is sequestered by the recognition domain in a “piggyback” fashion preventing contact with DNA at the cleavage site (8). Second, FokI has only one catalytic site per monomer; therefore, dimerization of two C-terminal domains is required to achieve a double-strand break in DNA (10–12). Association of the catalytic domains of FokI is much more effective when FokI monomers are bound to the separate DNA sites on the same DNA molecule (10). Thus, rearrangement of FokI triggered by specific DNA binding promotes DNA cleavage and limits it to the specific sites.

Yet, another type of mechanism of DNA cleavage control is used by the EcoRII enzyme that belongs to the type IIE subtype of REases. EcoRII requires two copies of the CCWGG recognition sequence for the optimal activity but cleaves only one copy before the first C; the second copy acts as an allosteric activator for the cleavage of the first DNA copy. Structural and biochemical studies demonstrate a modular architecture of EcoRII (13, 14). Proteinase treatment of EcoRII generates the N-terminal allosteric DNA-binding domain and C-terminal domain dimer that cleaves DNA in a site-specific manner (14). Apo-structure of EcoRII suggests (13) that N-terminal allosteric DNA-binding domain sterically blocks DNA access to the catalytic C-terminal domain. Binding of the first DNA copy at the allosteric site triggers enzyme rearrangement and promotes binding and cleavage of the second DNA copy at the C-terminal domain.

Bfil is a type IIS REase that acts at an asymmetric sequence, 5'-ACTGGG-3', and cuts top and bottom strands at fixed

Conflict of interest statement: No conflicts declared.

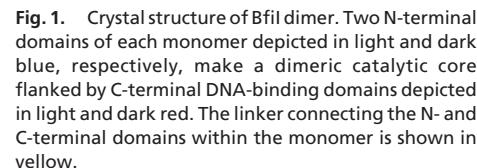
Freely available online through the PNAS open access option.

Abbreviations: PLD, phospholipase D; REase, restriction endonuclease.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2C1L).

[†]To whom correspondence may be addressed. E-mail: grazulis@ibt.lt, huber@biochem.mpg.de, or siksnys@ibt.lt.

© 2005 by The National Academy of Sciences of the USA



Grazulis et al.

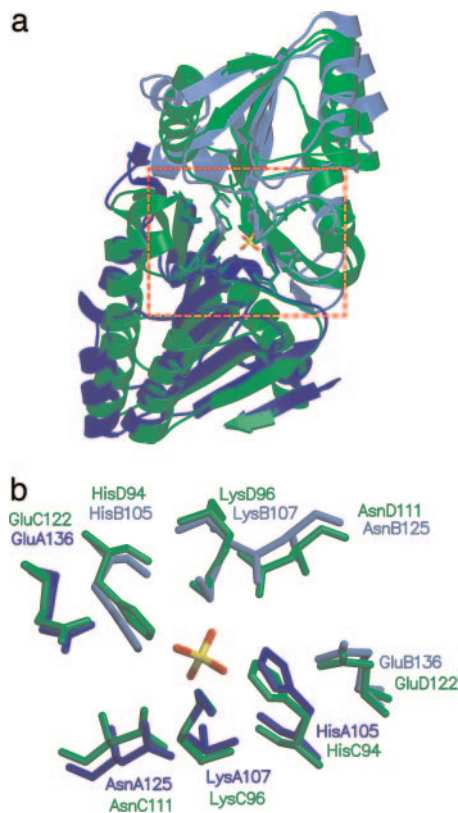


Fig. 2. Comparison of BfiI with EDTA-resistant nonspecific Nuc nuclease from *S. typhimurium*. (a) Structural superposition of the N-terminal domain dimer of BfiI (light and dark blue) and Nuc nuclease (green). The active-site region of Nuc nuclease is boxed in red. (b) Close-up view of the active-site residues of Nuc nuclease (green) and their structural equivalents in BfiI (light and dark blue) shown in the same orientation as in a. The tungstate molecule coordinated at the putative scissile DNA phosphate-binding site in Nuc is shown in stick representation colored by atom type. Amino acid residues from both subunits of Nuc and BfiI contribute to a single active site located in the intersubunit cleft.

stretch of amino acid residues (171–198) mostly in extended conformation, which is flanked by short 3_{10} -helices on both ends. The N-terminal domains form a dimer (Fig. 1), which presumably represents the structure of the functionally active BfiI dimer in solution (18). C-terminal domains within this dimer are situated on the opposite sides of the N-terminal domain core and make no contacts to each other. The overall shape of the BfiI molecule resembles a two-axial ellipsoid with a short axis of ≈ 55 Å and a long axis of 100 Å, respectively. All four domains are positioned along the long ellipsoid axis in the order C-NN-C.

The N-Terminal Domain of BfiI: Similarities to Nuc Nuclease. The N-terminal domain of BfiI extends from Met-1 to Met-170. It is arranged as an eight-stranded β -sheet flanked by three α -helices on one side and two α -helices on the other side (Figs. 1 and 6). The N-terminal domains of BfiI are folded into an α/β -globule through dimerization of two β -sheets of individual chains. A DALI (32) database search for similar three-dimensional structures revealed a high similarity ($Z = 15.4$) to the *Salmonella typhimurium* Nuc nuclease (17) [Protein Data Bank (PDB) entry 1BYR] that belongs to the PLD superfamily (Fig. 2a). Human tyrosyl-DNA phosphodiesterase Tdp1 (33) (PDB entry 1JY1) and PLD (34) (PDB entry 1F0I) from *Streptomyces*, both members of the PLD superfamily, also show a high degree of structural similarity to the N-terminal domain of BfiI (DALI Z scores of 8.9 and 9.6, respectively). Structural similarities con-

firm earlier biochemical and mutational studies indicating that BfiI is a member of PLD family (16, 18).

Upon structural superposition with TOP3D (35), amino acid residues His-105, Lys-107, Asn-125, and Glu-136 of BfiI that are important for the catalytic activity (18) closely superimpose with His-94, Lys-96, Asn-111, and Glu-122 residues at the active site of Nuc (17, 36) (Fig. 2b). These residues of BfiI and Nuc display the same side-chain rotamers in both structures and can be superimposed with an rms deviation of 1.4 Å. The conservation of the active-site residues between Nuc and the N-terminal domain of BfiI (Fig. 2b) provides important insights into the mechanism of catalysis by BfiI. In the Nuc dimer, the conserved His-94, Lys-96, and Glu-122 residues from both subunits contribute to the single active site located in the cleft between the two monomers (Fig. 2b). It has been suggested that Nuc catalysis proceeds with the formation of a phosphoenzyme intermediate that is subsequently hydrolyzed to form the final product (17, 36). The conserved active-site His-94 residue in one of the monomers has been directly implicated as a nucleophile in the reaction leading to the covalent intermediate, while the symmetry-related His from another subunit acts as a general acid protonating the leaving group (36). Conservation of the key catalytic residues between Nuc and BfiI implies similar chemistry of phosphodiester bond cleavage; however, the formation of the covalent intermediate for BfiI has yet to be demonstrated.

The N-terminal domain of BfiI, similarly to the human tyrosyl-DNA phosphodiesterase Tdp1, lacks the conserved aspartate residue of HKD motif characteristic to Nuc and *Streptomyces* PLD (37). According to the crystal structures, the latter residue is not a part of the active site of Nuc (17) and *Streptomyces* PLD (34), and it is thought to be important for the stabilization of the tertiary structure. Because of this distinction, human tyrosyl-DNA phosphodiesterase Tdp1 and its orthologs were assigned into a separate subclass within the PLD superfamily (37). BfiI is suggested to be a member of this subclass.

The C-Terminal Domain of BfiI: Similarities to Mg^{2+} -Dependent Rease and Plant Transcription Factors. The C-terminal domain of BfiI extends from Val-199 to Phe-358. It is comprised of two β -sheets (Figs. 1 and 3a) that are interconnected through β_{13} and β_{14} strands, shared by both sheets (Fig. 6). The two β -sheets are sandwiched together into a barrel-like structure, composed of eight β -strands (β_{10} , β_{11} , β_{13} – β_{17} , and β_{19}). The “barrel,” however, is not closed because strands β_{11} and β_{19} are not hydrogen-bonded. Two additional short β -strands (β_{12} and β_{18}) and three α -helices are located outside the pseudobarrel.

A DALI database search revealed that the C-terminal domain of BfiI is significantly similar ($Z = 8.8$) to the N-terminal domain of Mg^{2+} -dependent REase EcoRII that belongs to the type IIE subtype. Another significant similarity ($Z = 7.1$) is to the B3 DNA-binding domain of *Arabidopsis* cold-responsive transcription factor RAV1 (38) (PDB entry 1WID), specific for the CACCTG sequence (39). A structural superposition of the C-terminal domain of BfiI, the N-terminal domain of EcoRII REase, and the RAV1 transcription factor of *Arabidopsis* (Fig. 3a) clearly shows that all three domains represent the same fold.

The structures of BfiI, EcoRII (13), and the B3-domain of RAV1 transcription factor of *Arabidopsis* (38) were solved in a DNA-free form. Structural similarities between the C-terminal domain of BfiI, the N-terminal domain of EcoRII, and the B3-domain of RAV1 (Fig. 3a), however, provide some clues regarding the putative DNA-binding surface. Topologically similar concave clefts laid by the positively charged residues and large enough to accommodate a DNA molecule could be identified in all domains (see Fig. 7, which is published as supporting information on the PNAS web site). Cross-linking and mutational studies of EcoRII demonstrated the importance of His-36, Tyr-41, Lys-92, Arg-94, Glu-96, Lys-97, and Arg-98

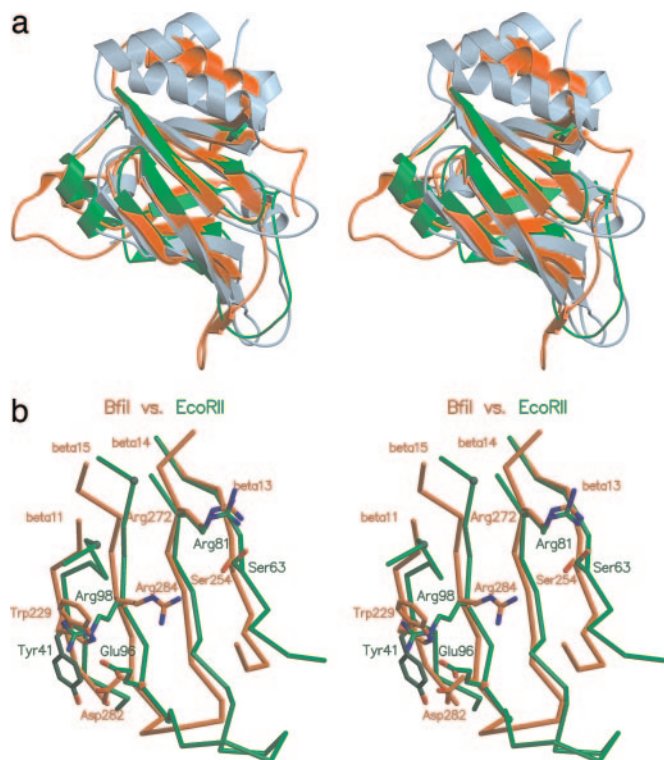


Fig. 3. Structural comparison of the C-terminal DNA-binding domain of the metal-independent REase BfiI, the N-terminal effector DNA-binding domain of type IIE restriction enzyme EcoRII, and the B3 DNA-binding domain of plant transcription factor RAV1. (a) Structural overlay of DNA-binding domains of BfiI (red), EcoRII (green), and RAV1-B3 (blue). (b) Stereoview of the backbone atom superposition in the putative DNA-binding clefts of BfiI (red) and EcoRII (green). The EcoRII residues important for DNA binding are shown as sticks colored by atom type, with carbon atoms in green. Structurally equivalent BfiI residues are shown as sticks with carbon atoms in red.

residues located in the cleft for the effector DNA binding (40, 41). NMR-titration experiments revealed a similar protein–DNA interface in RAV1-B3 (38). The equivalent cleft in BfiI is suggested to contribute to specific DNA binding. C_{α} atoms of BfiI residues Trp-229, Asp-282, and Arg-284 (Fig. 3b) spatially correspond to EcoRII residues Tyr-41, Glu-96, and Arg-98, which are important for DNA binding (41). Their side chain rotamers differ but may change in the presence of DNA. In addition, Arg-272 and Ser-254 residues of BfiI coincide with Arg-81 and Ser-63 residues of EcoRII. The functional importance of these residues so far has not yet been tested. The conservation of some residues in the DNA-binding clefts of BfiI and EcoRII is intriguing in view of the different recognition sequences (ACTGGG in the case of BfiI and CCWGG for EcoRII).

Interdomain Interface. N- and C-terminal domains in a BfiI monomer make an extensive set of hydrophobic contacts between the “inner” faces of the β -sheet. Approximately 1,700 Å² are buried at the interdomain interface in BfiI monomer. Contacts of the C-terminal domain to the N-terminal domain of the neighboring subunit (i.e., C terminus of chain A with the N terminus of the chain B) contribute an additional 500 Å² to the interdomain interface. Thus, the two domains have a relatively large interdomain interface of $\approx$ 2,200 Å², characteristic of specific protein–protein interactions (42). In comparison, the main dimer interface of the BfiI formed by the N-terminal domains buries 3,500 Å² (Fig. 1).

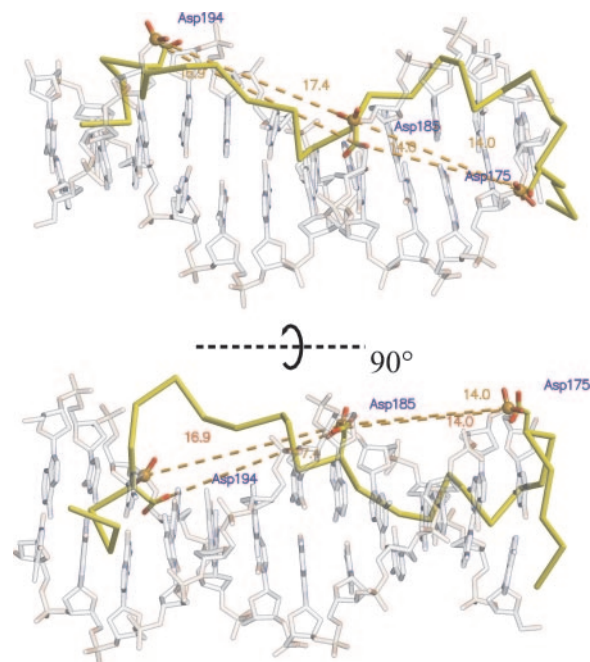


Fig. 4. DNA backbone mimic by the interdomain linker of BfiI. The linker region is shown as a C_{α} trace; Asp residues are labeled and shown in stick representation. Distances between C_{γ} atoms of Asp residues are shown as yellow dashed lines. The B-DNA molecule (Nucleic Acid Database ID bd0001) is shown in gray. DNA backbone phosphate atoms equivalent to Asp C_{γ} atoms are shown as yellow spheres. Superposition of BfiI linker peptide and DNA was obtained by superimposing C_{γ} atoms of linker aspartates (Asp-175, -185, and -194) onto the DNA's closest possible phosphate and C_{α} of the same residues onto DNA O_4' atoms of the corresponding nucleotides.

The C-terminal DNA-binding domain of BfiI and the N-terminal effector DNA-binding domain of EcoRII are similar to the B3-like DNA-binding domain of plant-specific transcription factors (Fig. 3a). Hence, the transcription factor-like DNA-binding domain has been used not only to confer DNA specificity to a PLD family nuclease, to give the metal-free BfiI enzyme as shown here, but also to create allosteric site(s) in an orthodox PD ... EXK family enzyme, to generate the metal-dependent type IIE restriction enzyme EcoRII. Although the C-terminal DNA-binding domain of BfiI is structurally similar to the N-terminal effector DNA-binding domain of EcoRII, they employ different interfaces to interact with the corresponding catalytic domains (see Fig. 8, which is published as supporting information on the PNAS web site). In BfiI, the DNA-binding domain makes contacts predominantly with the catalytic N-terminal domain of the same chain, making hydrogen bonds and hydrophobic contacts through residues coming from the edge of the first barrel-forming β -sheet (strands β 13– β 15 and extended region connecting strands β 11 and β 12). In contrast, in EcoRII the effector DNA-binding N-terminal domain contacts the catalytic domain through the residues located on α -helices and their flanking loop regions (13). Thus, DNA-binding domains of BfiI and EcoRII possess a similar DNA-binding cleft but are oriented differently with respect to their catalytic cores. Moreover, whereas the DNA-binding domains of BfiI are spatially separated, the N-terminal domains of EcoRII contact each other and contribute to the dimer interface.

Interdomain Peptide Linker: Possible Role in Activity Control. The linker connecting the N- and C-terminal domains of the BfiI contains a stretch of 28 amino acid residues (171–198) that predominantly has an extended conformation (Fig. 4). The

linker runs across the putative DNA-binding surface at the N-terminal domain of BfiI close to the entrance to the active site buried at the interface between the two domains (Fig. 1). It contains a number of negatively charged Asp and Glu residues. Interestingly, three Asp residues (Asp-175, -185, and -194) are spaced at distances very close to those between the phosphates of a B-DNA backbone (Fig. 4). One might speculate that the interdomain linker in BfiI mimics the DNA backbone similarly to the Ocr antirestriction protein (43) and acts as an intramolecular inhibitor that represses nucleolytic activity of the catalytic N-terminal domain by occluding the DNA-binding surface.

Discussion

The crystal structure of BfiI (Fig. 1) reveals a novel domain architecture consistent with previously reported limited proteolysis (44) data, where the N terminus is shown to be an unspecific nuclease and the C terminus is a DNA-binding domain. The crystal structure of BfiI demonstrates how nature can reconfigure protein domains for new functions. In BfiI, a DNA-recognition element is fused to a nonspecific nuclease core family to generate a REase. In the asymmetric unit, two N-terminal domains of BfiI are arranged into a dimer structurally similar to the nonspecific EDTA-resistant Nuc nuclease (Fig. 2) that belongs to the PLD superfamily. Structural similarities between BfiI, Nuc nuclease (17), tyrosyl-DNA phosphodiesterase Tdp1 (33), and PLD (34) provide a direct link between BfiI and the PLD superfamily, a diverse group of proteins that includes phospholipases, phospholipid synthases, bacterial toxins, pox virus envelope proteins, phosphodiesterases, and bacterial nucleases (45). PLD-family enzymes exhibit a conserved catalytic core and share a common catalytic mechanism; however, these enzymes encompass a very broad range of substrate specificities (45). Different enzymes in the PLD family may therefore have evolved through the fusion of a common catalytic core to separate domains for substrate recognition.

The C-terminal DNA-binding domain of BfiI (Fig. 3*a*) is structurally similar to the N-terminal DNA-binding domain of Mg²⁺-dependent REase EcoRII (13). This finding suggests that different subtypes and lineages of REases might have evolved through the recombination of conserved structural blocks that became fused to different catalytic cores.

Strikingly, DNA-binding domains of EcoRII/BfiI and B3 DNA-binding domain of *Arabidopsis* cold-responsive transcription factor RAV1 (38) share a similar structural fold (Fig. 3*a*). The B3-like DNA-binding domain was thought to be plant-specific and belongs to the family that currently includes ≈300 proteins (InterPro IPR 003340) involved in the auxin-regulated and abscisic acid-regulated transcription (46). Identification of structural and functional homologues of the “plant-specific” B3 DNA-binding domain in the bacteria kingdom provokes speculations on the possible evolutionary relationships.

A PSI-BLAST search of the protein sequence database by using BfiI sequence as a query reveals significant similarities to the hypothetical protein MBNC03000766 of *Mesorhizobium* sp. BNC1 strain (see Fig. 9, which is published as supporting information on the PNAS web site). The degree of sequence similarity (60% of identical and 76% of similar amino acids) indicates that the unknown protein MBNC03000766 of *Mesorhizobium* sp. BNC1 strain is likely to be a restriction enzyme similar to BfiI. Moreover, typically for type IIS restriction modification system gene organization two ORFs possessing similarities to the DNA-methyltransferases are found in the vicinity of the gene of putative *Mesorhizobium* sp. BNC1 endonuclease (data not shown). Interestingly, the *Mesorhizobium* sp. BNC1 strain subjected for the genome sequencing has been isolated from an enrichment of industrial sewage receiving EDTA-containing wastewater effluents added to surface soil. This strain is able to use EDTA as a sole source of carbon and

nitrogen, and thrives in an EDTA-rich environment (47). In such a metal-ion hostile environment, the presence of EDTA-resistant REase in bacteria might be an advantage.

Both the N- and C-terminal domains of BfiI (Fig. 1) represent independent structural and functional modules. The isolated DNA-binding domain binds specifically to the 5'-ACTGGG sequence characteristic for BfiI, whereas the N-terminal domain possesses a nonspecific nuclease activity (44). Mixing of the isolated N- and C-terminal domains, however, does not reconstitute the ability of BfiI to cleave phosphodiester bonds at specific sites (M. Zaremba and V.S., unpublished data). Thus, to function as a restriction enzyme, BfiI had to elaborate a nuclease activity control mechanism to limit double-strand breaks at the specific recognition sequence. The crystal structure of BfiI suggests how interdomain communications can lead to novel protein functions, not possessed by either domain alone.

BfiI, like many orthodox Mg²⁺-dependent REases, is a dimer in solution (18). In contrast to the dimeric enzymes that recognize a single copy of their DNA, the dimeric form of BfiI binds two copies of its recognition sequence (18). Yet, when bound to two copies of its recognition site, the BfiI dimer cleaves only one phosphodiester bond at a time (48). The crystal structure reveals that BfiI, like Nuc from *S. typhimurium*, has a single active site located at the interface between the N-terminal domains (Fig. 2*a* and *b*). Similarly to Nuc, an isolated N-terminal domain of BfiI degrades DNA containing or lacking BfiI recognition sequence by introducing random cuts (44). Structural superposition between BfiI and Nuc reveals that in the wild-type BfiI the linker region interconnecting N- and C-terminal domains crosses the putative DNA-binding surface of the catalytic domain (Fig. 1) and sterically interferes with DNA binding. Distribution of negatively charged residues in the interdomain linker suggests that it might mimic the DNA backbone (Fig. 4) and autoinhibit the DNA cleavage. The BfiI conformation seen in the crystal may therefore represent a latent form. The active site of BfiI, however, remains available for small artificial substrates like bis-*p*-nitrophenyl phosphate, which is hydrolyzed by BfiI (16). Thus, the negatively charged linker in BfiI does not disrupt the active site but rather creates a steric block for DNA binding.

To cleave DNA, the binding surface at the N-terminal domain must become available for DNA binding. Proteolytic cleavage of BfiI opens an access for DNA binding at the active site located at the catalytic domain, but phosphodiester bond hydrolysis then occurs at random rather than at specific sites (44). Thus, to focus cleavage at particular DNA sites, the opening of the catalytic domain in the wild-type BfiI has to be coupled to the specific DNA binding (see Fig. 10, which is published as supporting information on the PNAS web site). We propose that specific DNA binding to the C-terminal domain of BfiI may trigger a conformational rearrangement that leads to repositioning of the linker and unveils a surface for DNA binding at the catalytic N-terminal domain (Fig. 10). Such a cognate DNA-binding-induced conformational switch juxtaposes the catalytic domain in a proper configuration to enable BfiI to cleave a phosphodiester bond (48). The single active site in the dimeric catalytic domain then acts sequentially on the four target phosphodiester bonds. It has yet to be determined how the four target bonds, two from each recognition site, are placed in turn into the single active site of the enzyme.

We thank V. Stonyte for help with purification of the BfiI protein and S. Halford of the University of Bristol (Bristol, UK) and M. Zaremba, G. Sasnauskas, and C. Venclovas of the Institute of Biotechnology (Vilnius) for discussions and critical reading of the manuscript. Data sets were collected on the ID13 beamline of the European Synchrotron Radiation Facility (Grenoble), the European Molecular Biology Laboratory X13 beamline of the Deutsches Elektronen Synchrotron (Hamburg), and the International Institute of Molecular and Cell Biology (Warsaw). Measurements at the European Molecular Biology Labora-

tory outstation were supported by the European Community Access to Research Infrastructure Action of the Improving Human Potential Program. V.S. is a recipient of Lithuanian State Fellowship. The work at

the Institute of Biotechnology (Vilnius) was supported in part by the Howard Hughes Medical Institute, the Wellcome Trust, the Max Planck Society, and the Lithuania State Science and Studies Foundation.

1. Pingoud, A., Fuxreiter, M., Pingoud, V. & Wende, W. (2005) *Cell Mol. Life Sci.* **62**, 685–707.
2. Kovall, R.A. & Matthews, B.W. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 7893–7897.
3. Ban, C. & Yang, W. (1998) *EMBO J.* **17**, 1526–1534.
4. Lilley, D. M. & White, M. F. (2000) *Proc. Natl. Acad. Sci. USA* **97**, 9351–9353.
5. Aravind, L., Makarova, K. S. & Koonin, E. V. (2000) *Nucleic Acids Res.* **28**, 3417–3432.
6. Bujnicki, J. M. (2004) in *Restriction Endonucleases*, ed. Pingoud, A. (Springer, Berlin), pp. 63–93.
7. Kurpiewski, M.R., Engler, L. E., Wozniak, L. A., Kobylanska, A., Koziolkiewicz, M., Stec, W. J. & Jen-Jacobson, L. (2004) *Structure* **12**, 1775–1788.
8. Wah, D. A., Hirsch, J. A., Dorner, L. F., Schildkraut, I. & Aggarwal, A. K. (1997) *Nature* **388**, 97–100.
9. Li, L., Wu, L. P. & Chandrasegaran, S. (1992) *Proc. Natl. Acad. Sci. USA* **89**, 4275–4279.
10. Bath, A. J., Milsom, S. E., Gormley, N. A. & Halford, S. E. (2002) *J. Biol. Chem.* **277**, 4024–4033.
11. Wah, D. A., Bitinaite, J., Schildkraut, I. & Aggarwal, A. K. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 10564–10569.
12. Bitinaite, J., Wah, D. A., Aggarwal, A. K. & Schildkraut, I. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 10570–10575.
13. Zhou, E. X., Wang, Y., Reuter, M., Mücke, M., Krüger, D., Meehan, E. J. & Chen, L. (2004) *J. Mol. Biol.* **335**, 307–319.
14. Mücke, M., Grelle, G., Behlke, J., Kraft, R., Krüger, D. H. & Reuter, M. (2002) *EMBO J.* **21**, 5262–5268.
15. Vitkute, J., Maneliene, Z., Petrusyte, M. & Janulaitis, A. (1998) *Nucleic Acids Res.* **26**, 3348.
16. Sapranasauskas, R., Sasnauskas, G., Lagunavicius, A., Vilkaitis, G., Lubys, A. & Siksnys, V. (2000) *J. Biol. Chem.* **275**, 30878–30885.
17. Stuckey, J. A. & Dixon, J. E. (1999) *Nat. Struct. Biol.* **6**, 278–284.
18. Lagunavicius, A., Sasnauskas, G., Halford, S. E. & Siksnys, V. (2003) *J. Mol. Biol.* **326**, 1051–1064.
19. Cusack, S., Belrhali, H., Bram, A., Burghammer, M., Perrakis, A. & Riekel, C. (1998) *Nat. Struct. Biol.* **5**, Suppl., 634–637.
20. Leslie, A. G. W. (2003) *MOSFLM 6.2.3 User Guide* (MRC Lab. of Mol. Biol., Cambridge, U.K.).
21. Evans, P. R. (1997) *Joint CCP4 and ESF-EACBM Newsletter* **33**, 22–24.
22. French, G. S. & Wilson, K. S. (1978) *Acta Crystallogr. A* **34**, 517–525.
23. Collaborative Computational Project Number 4 (1994) *Acta Crystallogr. D* **50**, 760–763.
24. Cowtan, K. (1994) *Joint CCP4 and ESF-EACBM Newsletter on Protein Crystallography* **31**, 34–38.
25. Read, R. J. (1986) *Acta Crystallogr. A* **42**, 140–149.
26. Morris, R. J., Perrakis, A. & Lamzin, V. S. (2002) *Acta Crystallogr. D* **58**, 968–975.
27. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997) *Acta Crystallogr. D* **53**, 240–255.
28. Laskowski, R. A., MacArthur, M. W., Moss, D. S. & Thornton, J. M. (1993) *J. Appl. Crystallogr.* **26**, 283–291.
29. Kraulis, P. J. (1991) *J. Appl. Crystallogr.* **24**, 946–950.
30. Merritt, E. A. & Bacon, D. J. (1997) *Methods Enzymol.* **277**, 505–524.
31. Koradi, R., Billeter, M. & Wüthrich, K. (1996) *J. Mol. Graphics* **14**, 51–55.
32. Holm, L. & Sander, C. (1998) *Proteins* **33**, 88–96.
33. Davies, D. R., Interthal, H., Champoux, J. J. & Hol, W. G. J. (2002) *Structure* **10**, 237–248.
34. Leiros, I., Secundo, F., Zambonelli, C., Servi, S. & Hough, E. (2000) *Structure* **8**, 655–667.
35. Lu, G. (2000) *J. Appl. Crystallogr.* **33**, 176–183.
36. Gottlin, E. B., Rudolph, A. E., Zhao, Y., Matthews, H. R. & Dixon, J. E. (1998) *Proc. Natl. Acad. Sci. USA* **95**, 9202–9207.
37. Interthal, H., Pouliot, J. J. & Champoux, J. J. (2001) *Proc. Natl. Acad. Sci. USA* **98**, 12009–12014.
38. Yamasaki, K., Kigawa, T., Inoue, M., Tateno, M., Yamasaki, T., Yabuki, T., Aoki, M., Seki, E., Matsuda, T., Tomo, Y., et al. (2004) *Plant Cell* **16**, 3448–3459.
39. Kagaya, Y., Ohmiya, K. & Hattori, T. (1999) *Nucleic Acids Res.* **27**, 470–478.
40. Mücke, M., Pingoud, V., Grelle, G., Kraft, R., Krüger, D. H. & Reuter, M. (2002) *J. Biol. Chem.* **277**, 14288–14293.
41. Reuter, M., Schneider-Mergener, J., Kupper, D., Meisel, A., Mackeldanz, P., Krüger, D. H. & Schroeder, C. (1999) *J. Biol. Chem.* **274**, 5213–5221.
42. Janin, J. (1997) *Nat. Struct. Biol.* **4**, 973–974.
43. Walkinshaw, M. D., Taylor, P., Sturrock, S. S., Atanasiu, C., Berge, T., Henderson, R. M., Edwardson, J. M. & Dryden, D. T. F. (2002) *Mol. Cell* **9**, 187–194.
44. Zaremba, M., Urbanke, C., Halford, S. E. & Siksnys, V. (2004) *J. Mol. Biol.* **336**, 81–92.
45. Ponting, C. P. & Kerr, I. D. (1996) *Protein Sci.* **5**, 914–922.
46. Ulmasov, T., Hagen, G. & Guilfoyle, T. J. (1997) *Science* **276**, 1865–1868.
47. Bohuslavsek, J., Payne, J. W., Liu, Y., Bolton, H. J. & Xun, L. (2001) *Appl. Environ. Microbiol.* **67**, 688–695.
48. Sasnauskas, G., Halford, S. E. & Siksnys, V. (2003) *Proc. Natl. Acad. Sci. USA* **100**, 6410–6415.