doi:10.1016/j.jmb.2007.03.018

J. Mol. Biol. (2007) **369**, 722–734

JMB

Available online at www.sciencedirect.com

ScienceDirect

ELSEVIER

Monomeric Restriction Endonuclease BcnI in the Apo Form and in an Asymmetric Complex with Target DNA

**Monika Sokolowska[1,2], Magdalena Kaus-Drobek[1,2]
Honorata Czapinska[1,2], Gintautas Tamulaitis[3]
Roman H. Szczepanowski[1,2], Claus Urbanke[4]
Virginijus Siksnys[3]* and Matthias Bochtler[1,2]***

[1]*International Institute of Molecular and Cell Biology ul. Trojdena 4, 02-109 Warsaw, Poland*

[2]*Max-Planck-Institute for Molecular Cell Biology and Genetics, Pfotenhauerstr. 108 01309 Dresden, Germany*

[3]*Institute of Biotechnology Graiciuno 8, LT-02241 Vilnius, Lithuania*

[4]*Medizinische Hochschule Abteilung Strukturanalyse OE 8830, Carl Neuberg Str. 1 30625 Hannover, Germany*

**Corresponding authors*

Restriction endonuclease BcnI cleaves duplex DNA containing the sequence CC/SGG (S stands for C or G, / designates a cleavage position) to generate staggered products with single nucleotide 5′-overhangs. Here, we show that BcnI functions as a monomer that interacts with its target DNA in 1:1 molar ratio and report crystal structures of BcnI in the absence and in the presence of DNA. In the complex with DNA, BcnI makes specific contacts with all five bases of the target sequence and not just with a half-site, as the protomer of a typical dimeric restriction endonuclease. Our data are inconsistent with BcnI dimerization and suggest that the enzyme introduces double-strand breaks by sequentially nicking individual DNA strands, although this remains to be confirmed by kinetic experiments. BcnI is remotely similar to the DNA repair protein MutH and shares approximately 20% sequence identity with the restriction endonuclease MvaI, which is specific for the related sequence CC/WGG (W stands for A or T). As expected, BcnI is structurally similar to MvaI and recognizes conserved bases in the target sequence similarly but not identically. BcnI has a unique machinery for the recognition of the central base-pair.

Keywords: restriction endonuclease; BcnI, monomer; asymmetric recognition; 1 nt 5′-overhang; crystal structure

Introduction

Most type IIP restriction endonucleases (REases) are dimers that match the exact or approximate 2-fold symmetry of their palindromic or pseudopalindromic target sequences. However, the recently determined structures of MspI (CCGG),[1] HinP1I (GCGC),[2,3] and MvaI (CCWGG, W stands for A or T) show that these enzymes interact with their recognition sites as monomers. The latter enzymes are unusual also in other respects: MspI (C/CGG) and HinP1I (G/CGC) generate DNA products with 2 nt 5′-overhangs as indicated by the / while MvaI (CC/WGG) is the first structurally characterized REase that produces 1 nt 5′-overhangs upon DNA cleavage.

Restriction endonuclease BcnI attracted our interest because the enzyme shares protein sequence similarity with MvaI, recognizes a related target sequence CC/SGG (S stands for G or C) and cleaves it downstream of the inner cytosine residue as indicated by the /.[4] Moreover, it has been reported that BcnI is a monomer in solution in the apo form.[4] The BcnI REase is a part of restriction-modification system identified in *Bacillus centrosporus* strain RFL1,[5] where it occurs in one operon with the small accessory protein BcnIC and with the cytosine N4-methyltransferases BcnIMA and BcnIMB, which both methylate the N4 atom of the inner cytosine of the recognition sequence.[6,7] BcnIMB suffices to protect host cells against cleavage by BcnI REase, while BcnIMA is dispensable for this purpose.[7]

Here, we report a biochemical characterization of BcnI and crystallographic structures of the enzyme in the apo form and in complex with the cognate oligoduplex. We find that BcnI, like MvaI, is a monomer that recognizes its target asymmetrically. The crystal structure provides detailed information

Abbreviations used: REase, restriction endonuclease; MAD, multiple anomalous diffraction.

E-mail addresses of the corresponding authors: MBochtler@iimcb.gov.pl, siksnys@ibt.lt

on the protein–DNA interactions and explains the degenerate sequence recognition of the central basepair. A detailed comparison of BcnI and MvaI with the related DNA repair protein MutH will be presented elsewhere.

Results and Discussion

BcnI is a monomer and binds DNA in 1:1 stoichiometric ratio

The similarity between BcnI and MvaI and other available data suggested that BcnI might be another monomeric REase that binds its target DNA in stoichiometric ratio and prompted us to test this hypothesis by analytical ultracentrifugation, analytical gel-filtration and native gel electrophoresis.[4] Sedimentation velocity runs in the analytical ultracentrifuge with BcnI gave a sedimentation constant of $s_{20°C,W}=2.52$ S. Using a mass for the monomeric protein of 27.29 kg/mol, this sedimentation constant corresponds to a frictional ratio of 1.29. For spherical hydrated proteins, a frictional ratio of 1.1–1.2 is expected,[8] and thus BcnI can be viewed as a mostly globular, monomeric particle. Sedimentation equilibrium gave a molar mass of 28 kg/mol, which is in excellent agreement with the calculated mass of 27 292 Da for the BcnI monomer.

The oligomeric state of BcnI, both in the apo form and in complex with a blunt-ended, 9-mer cognate oligoduplex (see Materials and Methods for the details), was analyzed independently by analytical gel-filtration on a Superdex HR column. BcnI eluted from the column as a monomer, both in the absence and in the presence of cognate oligoduplex. DNA up to a molar ratio of one oligoduplex per BcnI monomer migrated with BcnI. Any oligoduplex in excess of this ratio eluted as free DNA (Figure 1(a)). Essentially the same result was obtained by native gel electrophoresis (Figure 1(b)). In the conditions of the experiment (see Materials and Methods for the details), free DNA did not enter the gel and BcnI migrated in the gel faster than the BcnI-DNA complex. Two bands, corresponding to free and DNA-bound BcnI were obtained when BcnI monomers were present in excess over DNA oligoduplexes. Only the DNA-bound species of BcnI was observed if BcnI was incubated with one molar equivalent or with an excess of DNA (Figure 1(b)).

BcnI is monomeric in the crystals without and with DNA

In the absence of divalent metal ions and DNA, BcnI formed tetragonal crystals in space group $P4_12_12$, which diffracted to 1.6 Å resolution. The BcnI structure was solved by the multiple anomalous diffraction (MAD) method, using the selenomethionine variant of the protein and a bromide soak. The crystals contained one molecule in the asymmetric unit. The interface between molecules that were related by a 2-fold crystallographic axis buried approximately 1230 Å^2. The small interface area, the lack of sequence conservation in the interface region and the odd shape of the resulting dimer suggested that this contact was not physiologically relevant, which is consistent with the biochemical data.

In order to obtain cocrystals with DNA, 11-mer oligoduplexes with the BcnI target sequence at the center and sticky 2 nt overhangs were designed (Figure 2(a)). The overhangs were chosen to allow polymerization of the DNA in one orientation only, because such oligonucleotides might form rods in the crystal, ideally with a crystallographic repeat coinciding with the repeat of the DNA. Using the sticky end 11-mer duplex, cocrystals of BcnI with DNA were indeed obtained. Molecular replacement with the BcnI apo structure as the search model was unsuccessful. Therefore, the BcnI–DNA complex structure was solved independently by the MAD method, again using the selenomethionine variant of the protein.

The crystal structure showed that the design of DNA rods in the crystal was successful, although not quite in the expected manner. The first surprise was that instead of one oligoduplex, the asymmetric unit contained two duplexes that were related by a

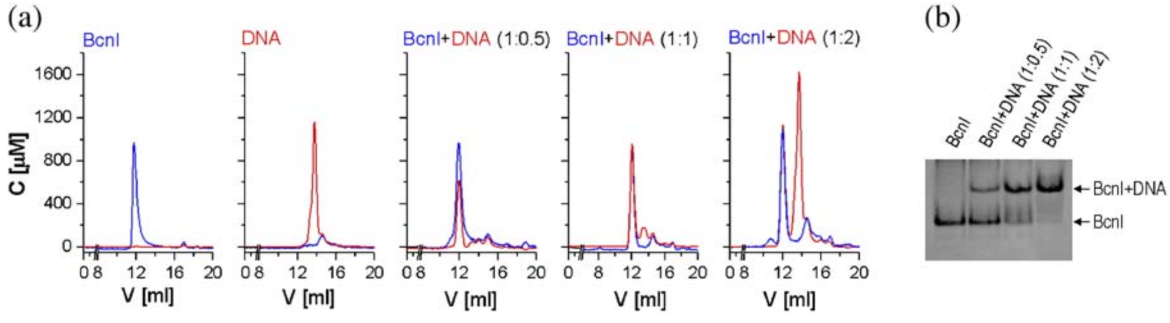


Figure 1. BcnI is a monomer in solution and binds DNA in 1:1 stoichiometric ratio. (a) Mixtures of BcnI and DNA in the indicated molar ratios were injected into a Superdex HR column. The elution profiles were recorded simultaneously at 260 nm and 280 nm, and deconvoluted to obtain separate curves for the BcnI (blue) and DNA (red) concentrations. (b) Mixtures of BcnI and DNA in the indicated molar ratios were subjected to native gel electrophoresis runs. Gels were stained by Coomassie brilliant blue.

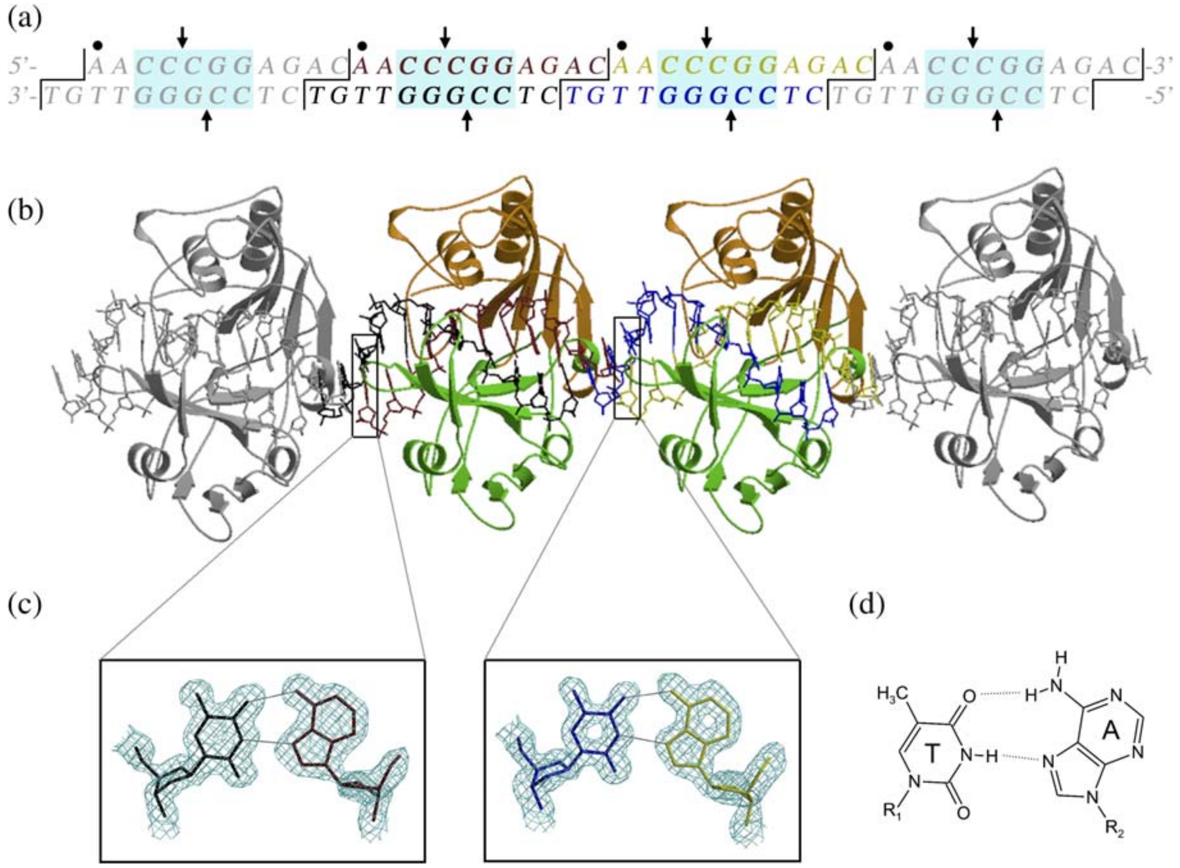


Figure 2. Arrangement of BcnI and DNA molecules in the cocrystals. (a) A representation of the staggered arrangement of oligonucleotide duplexes in the crystal. The brown, black, yellow and dark blue sequences represent the oligonucleotides in the asymmetric unit of the crystals. Symmetry mates are shown in grey. The BcnI recognition sequence is highlighted by blue boxes. Black dots mark adenine bases that interact with thymine bases on the other strand *via* their Hoogsteen rather than their Watson–Crick edge. (b) All-atom representation of the DNA and ribbon representation of BcnI. In the Figure, the contents of the asymmetric unit are colored, and symmetry-related molecules are presented in grey. The color code for DNA is the same as in (a) and BcnI catalytic and recognition lobes are in orange and green, respectively. (c) Enlarged view of the two non-Watson–Crick base-pairs in the asymmetric unit. The color code is the same as in (a) and (b), but the orientation has been changed to show the hydrogen bonding arrangement. The $2F_o$–F_c density from the final model has been contoured at 1.0 σ. (d) A diagram of the hydrogen bonds in the two non-Watson–Crick base-pairs.

simple non-crystallographic translation with only a minor rotational component (Figure 2(b)). As one molecule of BcnI was bound to each oligoduplex, this arrangement resulted in a packing with two molecules of BcnI in the asymmetric unit that were essentially related by a pseudotranslation, which was also readily detectable in the Patterson map. The second surprise came when the hydrogen bonding interactions of the bases were examined in detail: although the pairing of bases was as expected, the hydrogen bonding pattern was not. Two adenine bases (marked by dots in Figure 2(a)) hydrogen bonded with the thymine bases on the other strand *via* their Hoogsteen edge, not their Watson–Crick edge (Figure 2(c) and (d)). It is likely that this unusual feature of the crystal structure is a crystallization artifact, because the unusual base-pair is outside the recognition sequence and not in contact with protein. Moreover, there is no equivalent non-Watson–Crick base-pair in the complex of MvaI with DNA.[9] Therefore, we presume that the DNA "rods" in the BcnI–DNA crystals are probably under stress, which might be relieved, in part, by an exception to standard Watson–Crick hydrogen bonding. If this interpretation is correct, the detailed parameters for the DNA conformation in the BcnI–DNA complex (Supplementary Data Table 1) might be influenced by crystal packing forces.

BcnI has the expected two-lobe structure and a mobile hinge region

The BcnI structure shows the expected two-lobe architecture. According to their function, and by analogy with MvaI REase,[9] we suggest the terms catalytic lobe and recognition lobe, because the former contains all active site residues and the latter plays a key role in sequence recognition (see below). The recognition lobe is built from residues 68–153 and 186–230, which can be regarded as two large insertions in the catalytic lobe, which comprises the rest of the protein (Figure 3(a) and (b)).

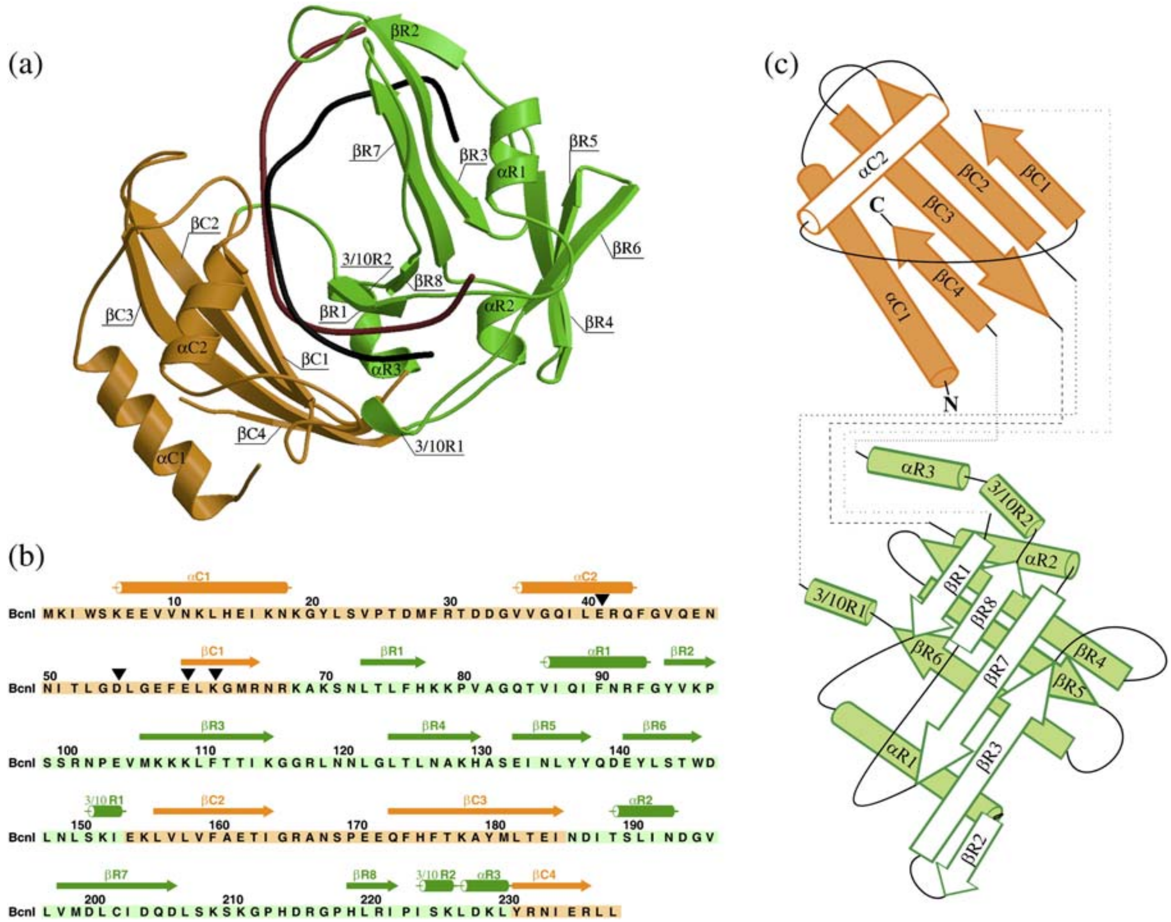


Figure 3. Gross BcnI structure. (a) A ribbon representation of BcnI in the cocrystals with DNA. The BcnI catalytic and recognition lobes are shown in orange and green, respectively. The DNA phosphodiester backbones have been smoothed and are presented in brown and black, respectively. The brown strand is the strand with C at the center of the recognition sequence. (b) BcnI amino acid sequence and secondary structure assignment. The filled triangles mark the active site residues. (c) A diagram of the BcnI fold. Coloring and labels in (b) and (c) are consistent with (a). Helix 3/10R1, helix αR2, strand βR8 and the mixed helix 3/10R2-αR3 in BcnI correspond to helix αR2, helix αR3, strand βR9 and helix 3/10R in MvaI, respectively. All other secondary structures in BcnI have direct counterparts with identical names in MvaI.

The catalytic lobe is built of two α-helices at the N terminus and of a four-stranded β-sheet with topology +1*x*, +1, +1, +1 according to the Richardson nomenclature (Figure 3(a) and (c)).[10] The recognition lobe consists of two β-sheets and several helices, including a mixed 3/10-α helix. The β-strands in the two sheets are oriented approximately perpendicular to each other, so that the structure can also be described as a very irregular β-barrel. Six of the eight strands of the barrel are contributed by the more N-terminal part of the recognition lobe, the remaining two strands are from the C-terminal part (Figure 3).

The hinge angles in the structures of BcnI in the apo form and in complex with cognate DNA are similar, but sufficiently different to prevent success of automatic molecular replacement procedures (Figure 4(a)). There are also other indications of hinge mobility: although the crystals of BcnI in the apo and DNA-bound forms diffract to almost exactly the same resolution, the crystallographic temperature factors for atoms of the recognition lobe are much higher in the apo form than in DNA complex (compare Figure 4(b) and (c)).

The active site is formed in the presence of divalent metal ions and DNA

The catalytic and recognition lobes of BcnI in the free and DNA-bound forms generally superimpose well, with one major exception (Figure 5). The loop region from residue 47 to residue 58 undergoes a major conformational change upon DNA binding. Interestingly, this flexible loop of BcnI and flanking regions anchor the putative active site residues Glu41, Asp55, Glu60 and Lys62 predicted from the alignment of BcnI and MvaI sequences (data not shown). The crystal structure of BcnI in complex with oligoduplex confirms the prediction, because all predicted active site residues are spatially close to each other and to the scissile phosphoester bond. As the cocrystals of BcnI and DNA were grown in the absence of Mg^{2+}, but in the presence of Ca^{2+}, we expected to find two calcium ions bound to each

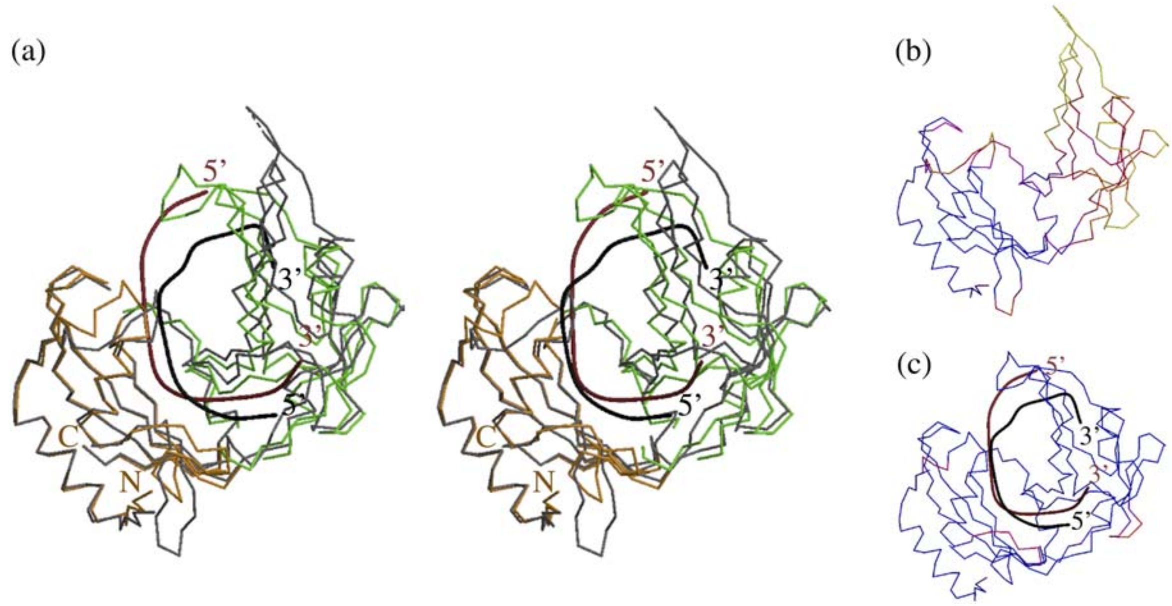


Figure 4. Apo-BcnI *versus* BcnI in complex with DNA. (a) A stereo diagram of the superimposed structures. Apo-BcnI is shown in grey, and the BcnI from the cocrystals with DNA is colored as in Figure 3. (b) The C^{α} trace of apo-BcnI color-ramped according to *B*-factor. (c) The C^{α} trace of BcnI from the cocrystals with DNA color-ramped according to *B*-factor. In the protein region, the correspondence between *B*-factor and color is identical in (b) and (c). Residues with the lowest *B*-factors are blue, and residues with the highest *B*-factors are yellow. DNA is presented in a smoothed representation and colored as in Figure 3.

active site. Unexpectedly, the final electron density map and an anomalous difference Fourier map for in-house, 1.54 Å wavelength diffraction data provide strong evidence only for one well-ordered calcium ion per active site. This calcium ion is hexacoordinated and has its ligands arranged in a nearly perfect octahedral coordination sphere. The ligands are an O^{δ} oxygen atom of Asp55, an O^{ε} oxygen atom of Glu60, a main-chain carbonyl oxygen atom of Leu61, an oxygen atom of the scissile phosphate and two water molecules (Figure 6). One of the water molecules is ideally positioned for in-line attack on the phosphorus atom of the scissile phosphoester bond. However, the reaction does not occur in the crystals, as judged both from the robust electron density for the cleavable phosphoester bond and from the distance of the water molecule to the phosphorus atom, which approximately equals the sum of the van der Waals radii of the two atoms. We presume that in a productive complex with Mg^{2+} instead of the calcium ions, the shorter metal ligand distance (2.1 Å for Mg^{2+}–O *versus* 2.4 Å for Ca^{2+}–O)[11] would place the water molecule closer to the phosphorus atom, promoting in-line attack to yield

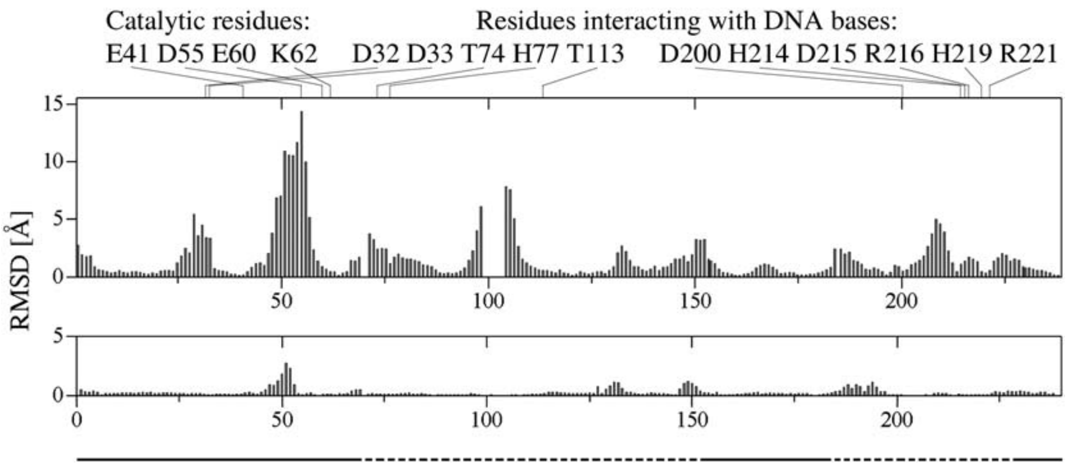


Figure 5. RMSD values between the BcnI structures. The catalytic lobe (continuous line) and the recognition lobe (broken line) were superimposed separately. The top panel shows the comparison of the structures with and without DNA, the bottom panel is a comparison of the two monomers of BcnI in the asymmetric unit of the crystals with DNA.

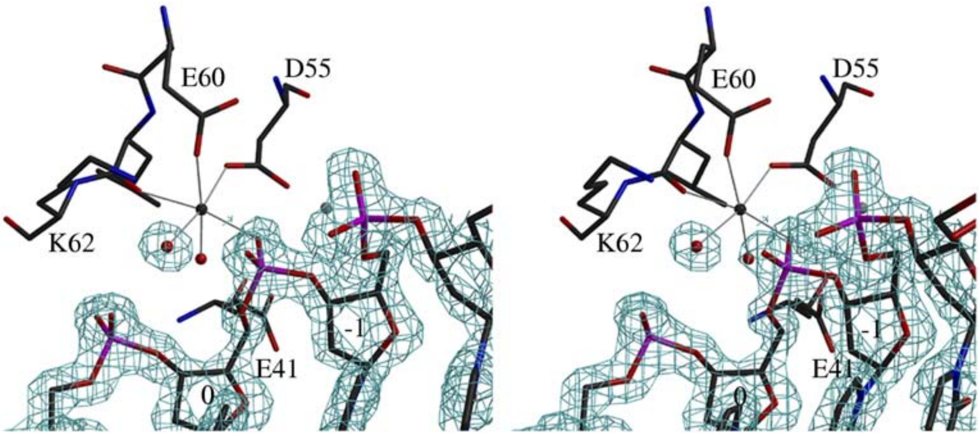


Figure 6. Stereo diagram of the BcnI active site in complex with DNA. BcnI active site residues are shown in all-atom representation, metal ions are presented as black and grey spheres, and some key water molecules are shown as red spheres. The 5′-end of the DNA is on the right and the 3′-end is on the left. The $2F_o$–F_c density was contoured at 1.5 σ and is shown only around the DNA and the nucleophilic water molecule. The assignment of the metal ion that is shown without detailed coordination is very uncertain.

cleavage products with a 5′-phosphate and 3′-OH, as expected for most REases (Figure 6). However, this explanation might be unduly simplistic, because modeling studies for other restriction endonucleases attribute the lack of activity of the Ca^{2+}-loaded enzymes to kinetic factors and not to the properties of the pre-reactive states.[12,13]

Interactions of BcnI with DNA

An overview of the BcnI–DNA interactions is presented in Figure 7(a). The phosphodiester backbone of the oligoduplex is involved in many charge–charge interactions with the enzyme. Interestingly, the basic residues of BcnI that neutralize the negative charge of the backbone phosphate groups are anchored almost exclusively in the recognition lobe. This feature is shared with the MvaI–DNA complex, despite poor conservation of the basic residues that are involved in the interactions with the DNA backbones (alignment not shown).

BcnI binds cognate DNA in such orientation that the DNA minor groove interacts with the catalytic lobe of the enzyme and the major groove faces the recognition lobe (Figures 3(a) and 7). Confident hydrogen bonding interactions between BcnI and the DNA are limited to the specifically recognized bases of the pseudopalindromic target sequence (Figure 7(b)). As the BcnI and MvaI target sequences differ only in the central base-pair, one might have expected that both enzymes would interact with conserved bases of the recognition sequence analogously. For most interactions, this turns out to be true, but there are some notable exceptions (Figure 7(c)). First, Asp33 in BcnI, which interacts with G+1, is substituted conservatively to Asn28 in MvaI. Second, there is no equivalent in BcnI for the interaction of MvaI Arg209 with G+1 and T0. Third, the interaction of BcnI Arg216 with C–1 and G+1 has its counterpart in the interaction of MvaI His225 with the same two bases. It is not clear whether these changes represent neutral drift, or whether they affect the recognition of the central base-pair indirectly (Figure 7(c)).

Interactions of BcnI with the central base-pair

Before the structure determination, we expected to find an equal or weighted mixture of the two possible binding modes of the DNA to the enzyme. Surprisingly, the binding mode that brings the C strand (the strand with C at the center of the recognition sequence) close to the active site is very dominant (Figure 7(d)). In part, this might be attributable to the assembly of the DNA molecules into rods (Figure 2), which interact with translationally related BcnI molecules, but then it remains to be explained why all DNA rods in the crystals are oriented in the same direction. Irrespective of the detailed explanation of the preference, it allows us to study the binding mode that would result in the cleavage of the C strand in a productive complex. On the minor groove site, the N2 atom of guanine donates a hydrogen bond to Asp32. In addition, the N3 atom of guanine and the O2 atom of cytosine accept indirect, water-mediated hydrogen bonds from BcnI. On the major groove side, the N4 atom of cytosine donates a hydrogen bond to the N^{ε} atom of His77 and the O6 atom of guanine accepts a hydrogen bond from the N^{ε} atom of His219. As the N^{δ} atoms of both His77 and His219 are in contact with hydrogen bond acceptors, the assignment implies that His77 is neutral and His219 is present in the charged form (Figure 7(e)).

Modeling the alternative binding mode

BcnI binding in the alternative orientation requires no major rearrangement on the minor groove side because the set of hydrogen bond donors and acceptors is nearly identical for G-C and C-G pairs.[14] However, on the major groove side, the alternative binding mode brings a hydrogen bond acceptor roughly in the place of the current donor and *vice versa*. As His77 and His219 are spatially close, they might accommodate DNA in the alternative binding mode by transferring the proton from His219 to His77 (Figure 7(e)). However, such a "proton toggle" would require a series of water molecules that could mediate the proton transfer. As such water molecules are not present or at least not crystallographically well defined in the cocrystal structure, the

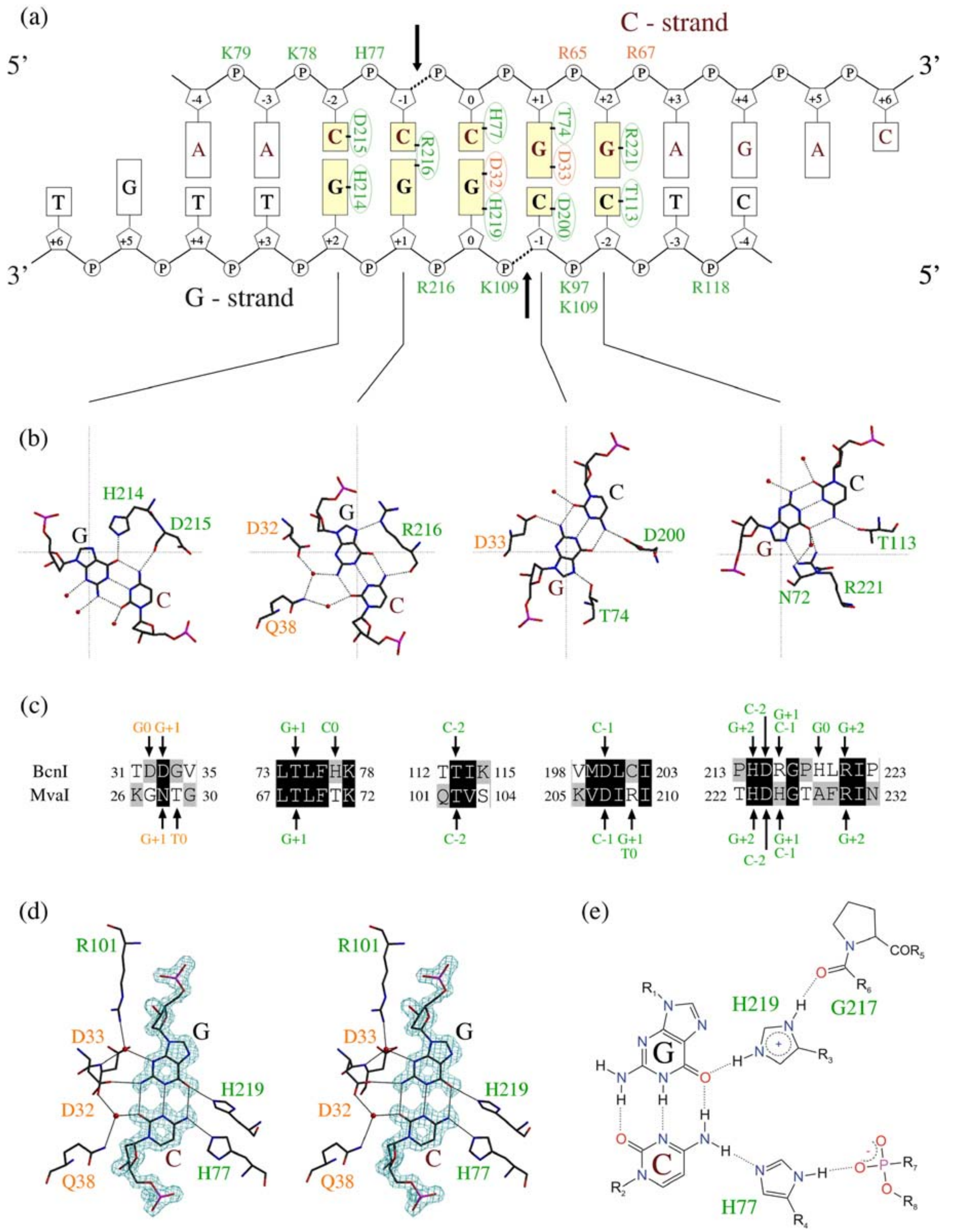


Figure 7 (*legend on next page*)

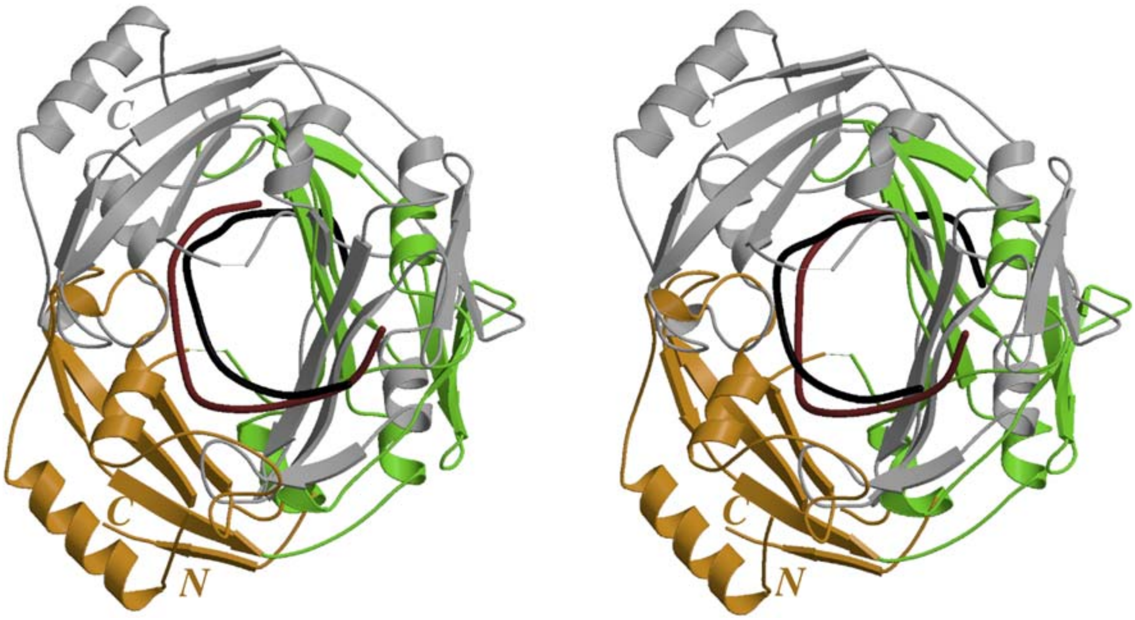


Figure 8. A stereo diagram of the *in silico* BcnI dimer. The protein is shown in ribbon representation and the DNA is represented by its smoothed backbone. The catalytic and recognition lobes of a BcnI molecule in the asymmetric unit are shown in orange and green, respectively. The DNA strand that comes close to the active site is shown in brown and the complementary strand in black. The grey BcnI molecule was generated *in silico* by application of the pseudo 2-fold symmetry of the DNA to the BcnI molecule in the crystal.

elucidation of BcnI interactions with DNA for the alternative binding mode will have to await a crystal structure.

No space for two BcnI monomers on a target site

In standard dimeric REases, each subunit interacts mainly with a half-site of the palindromic or pseudopalindromic recognition sequence. In contrast, the BcnI monomer interacts with the complete pseudopalindromic target site in an asymmetric manner (Figure 7). This unusual DNA binding mode implies that a second BcnI monomer cannot bind to the same site without competing for base interactions with the first monomer. Moreover, dimerization of BcnI *in silico* by application of the pseudo 2-fold symmetry of the target DNA leads to many severe clashes, including backbone atoms that could be relieved only by the major conformational change (Figure 8). Therefore, the crystallographic results strongly support the biochemical conclusion that BcnI is a monomer in the apo form, and remains monomeric in the complex with cognate DNA. Therefore, BcnI can introduce double-strand breaks in DNA only by sequential cleavage of the two DNA strands. This interpretation leaves open the question of whether one monomer of BcnI nicks both DNA strands sequentially, or whether a BcnI monomer recruits another monomer from the solution, which binds in the alternative orientation and displaces the originally bound monomer.

Similar structures in the Protein Data Bank

From the sequence information alone, we expected that the BcnI structure would be most similar to the recently determined MvaI structure and more distantly related to the structure of the DNA repair protein MutH. To obtain more quantitative results, we used the DALI server to search the Protein Data Bank (PDB) for structures with significant similarity to either full-length BcnI or to the catalytic lobe of BcnI.[15] With the exception of MutH, all hits with a DALI Z-score above 3.1 were restriction endonucleases.[15]

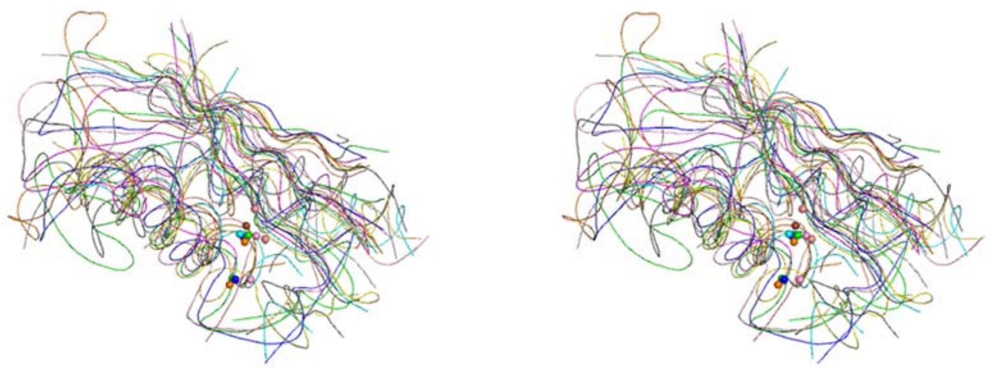
Figure 7. BcnI–DNA interactions. (a) A schematic representation. Arginine, lysine and histidine residues that have functional groups within 4.0 Å of phosphodiester backbone oxygen atoms are listed close to the P symbol that represents the phosphate. Hydrogen bond interactions with the bases are included only if they are direct and not water-mediated. (b) Detailed diagram of the hydrogen bonding interactions between BcnI and DNA (except for the interactions with the central base-pair). Thin lines indicate the correspondence between (a) and (b). (c) Structure-based alignment of BcnI and MvaI. Only the regions that interact with DNA are shown. Interactions of BcnI and MvaI with the bases are indicated above and below the alignment, respectively. Bases are numbered as in (a). Bases at the center of the recognition sequence are assigned base number 0. Numbering is continued in the 5′-3′ direction on each strand. (d) Recognition of the central base-pair. Hydrogen bonds are marked by thin lines. The density for the bases is from the original ARP/WARP run and was contoured at 1.5 σ. As ARP/WARP does not build DNA, this density is not biased by the interpretation that one binding mode for the DNA is very predominant. (e) A diagram of the major groove interactions of the central base-pair with His77 and His219.

As the DALI server uses only a non-redundant set of PDB structures for comparison and applies cut-off criteria, we expected the list of DALI hits to be incomplete. Therefore, we manually collected restriction endonuclease structures from the PDB and ran pairwise DALI comparisons with BcnI. The two lists of BcnI-related structures with *Z*-scores above 3 were combined and pruned to keep only the highest scoring representative of the DNA-bound form of each protein (Figure 9(a)). As expected, superposition of the BcnI catalytic lobe with its counterparts in the other restriction endonucleases and in MutH reveals a conserved fold (Figure 9(b)). After minor adjustment of the superposition operators, the active site residues overlap almost perfectly (Figure 9(c)). Together, these results show that BcnI is no exception to the rule that many or most restriction enzymes are similar in their active site regions,

(a)

REase	PDB ID	Z-score	RMSD	Aligned residues	Catalytic lobe Z-score	Catalytic lobe RMSD	Catalytic lobe Aligned residues	Active site residues
BcnI	2ODI							E41 D55 E60 K62
MvaI	2OAA	22.9	2.4	215	12.0	1.9	95	E36 D50 E55 K57
MutH	2AOR	9.0	3.5	151	8.3	3.2	93	E56 D70 E77 K79
HincII	2GIE	6.1	3.8	126	5.2	3.4	86	E38 D114 D127 K129
SfiI	2EZV	5.4	3.5	138	3.4	3.1	86	E55 D79 D100 K102
NaeI	1IAW	5.0	3.7	111	4.5	3.3	81	E70 D86 D95 K97
EcoRV	1RV5	5.0	3.5	113	3.9	3.9	83	E45 D74 D90 K92
HinP1I	2FKC	4.7	3.6	131	1.6	3.2	61	E18 D62 Q81 K83
BglI	1DMU	4.5	4.7	133	3.3	3.3	88	E87 D116 D142 K144
PvuII	1F0O	4.4	4.0	103	3.4	3.5	70	- D58 E68 K70
MspI	1YFI	3.8	4.2	124	2.7	3.3	64	E35 D99 N117 K119

(b)



(c)

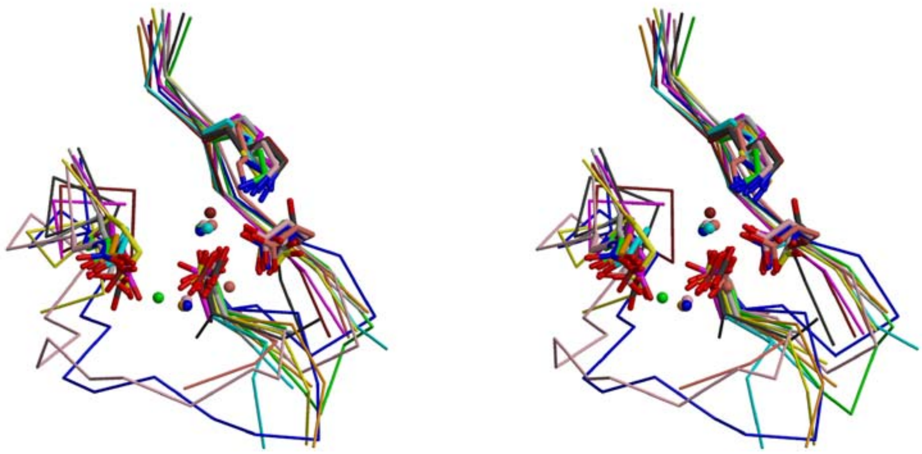


Figure 9. Structures with the highest similarity to BcnI. (a) A DALI quantitative structure comparison identifies significant similarities with MvaI,[9] MutH,[32] HincII,[33] SfiI,[35] NaeI,[34] EcoRV,[36] HinP1I,[3] BglI,[37] PvuII,[38] and MspI[1]. (b) Superposition of the BcnI catalytic lobe with its counterparts in the other restriction endonucleases and in MutH. Only smoothed C^{α} traces are shown. (c) Superposition of active site residues. Only side-chains, selected fragments of the C^{α} trace and divalent metal cations are shown. The color code is consistent in (a)–(c).

despite low or insignificant similarity at the sequence level. Interestingly, BcnI and MvaI, so far the only crystallographically characterized restriction enzymes that generate single nucleotide 5′-overhangs, are more similar to MutH, a nickase and component of the DNA repair machinery,[16] than to any other structurally characterized restriction endonuclease.

Materials and Methods

Strains

BcnI cloning and over-expression experiments were done in *Escherichia coli* strain ER2267 (F′ *proA*$^+$*B*$^+$ *lacI*^{q} *Δ(lacZ)M15 zzf*: :*mini-Tn*10 (Kan^r)/*Δ(argF-lacZ)U169 glnV44* e14$^-$(McrA$^-$) *rfbD1? recA1 relA1? endA1 spoT1? thi-1 Δ(mcrC-mrr)114*: :*IS10*). The selenomethionine variant of BcnI was expressed in strain BL21 (DE3) (*F– ompT gal [dcm] [lon] hsdSB (*r_B*–* m_B*–; an E. coli B strain)*), carrying the pVH1 plasmid (with lacI^q).[17]

BcnI cloning and expression

The BcnI REase (*bcnIR*) and BcnI MTase (*bcnIMB*) genes were amplified from the cloning vector pBR322_RM.BcnI (Ap^r) bearing *bcnIR, bcnIMAB* and *bcnIC* genes kindly provided by Fermentas (Vilnius, Lithuania).

The *bcnIR* and *bcnIMB* genes were re-cloned into the pBAD24 (Ap^r) and pACYC184 (Cm^r) vectors, respectively. The expression vector for the BcnI REase encoded for the full-length protein without the N or C-terminal tags and matched the *bcnIR* sequence in the GeneBank™ database exactly.

To express BcnI REase, the strain containing the pACYC184_MB.BcnI was transformed with pBAD24_R.BcnI. Cells were grown in LB medium with appropriate antibiotics at 37 °C to A_{600} 0.7 and protein expression induced by addition of arabinose to the final concentration of 0.2%. After 4 h of induction, the cells were harvested by centrifugation and the pellet was stored at −20 °C. Expression of the selenomethionine variant of BcnI was done in minimal M9 medium[18] in the presence of 0.05 mg/ml D,L-selenomethionine (Sigma) following the published procedure,[19] which is optimized to suppress methionine biosynthesis.

BcnI purification

Frozen cells expressing BcnI REase were thawed and resuspended in buffer A (10 mM potassium phosphate (pH 7.0) at 25 °C, 0.1 M NaCl, 1 mM EDTA, 7 mM 2-mercaptoethanol). The cell suspension was disrupted by sonication and the cell debris was removed by centrifugation at 35 000*g* for 1 h. BcnI was purified by subsequent chromatography on phosphocellulose P11 column (Whatman), Blue Sepharose and Heparin Sepharose (Amersham Biosciences) columns using linear NaCl gradients for protein elution. All purification steps were monitored by λ DNA cleavage and SDS-PAGE. Fractions containing BcnI REase activity were pooled and dialyzed against storage buffer (10 mM Tris–HCl (pH 7.4 at 25 °C), 0.2 M NaCl, 1 mM EDTA, 1 mM DTT, 50% (v/v) glycerol) and stored at −20 °C. The overall yield was 28 mg of BcnI from 5 l of *E. coli* culture. The selenomethionine variant of BcnI was purified following the protocol for the wild-type enzyme and stored in the same buffer. The overall yield of selenomethionyl derivative of BcnI was 23 mg from 2 l of culture.

Analytical ultracentrifugation

Analytical ultracentrifugation was done in an An-60 rotor with a Coulter-Beckman model XL-A analytical ultracentrifuge equipped with UV absorption detection using charcoal-filled epon centerpieces. Sedimentation velocity was measured at 4 °C and 44 000 rpm in a buffer containing 10 mM Tris–HCl (pH 7.4), 0.2 M KCl, 0.1 mM DTT and 0.1 mM EDTA at a loading concentration of 10 μM. The sedimentation coefficient distribution was analyzed with the program SEDFIT.[20] Viscosity and density corrections to calculate $s_{20°C,W}$ were done using the data supplied by the program package SEDNTERP†. Sedimentation equilibrium was measured in six-channel centerpieces at 18,000 rpm and 20 °C. Samples were spun until no change in absorbance profiles could be observed for at least 12 h at which time equilibrium was assumed to have been reached. Molar masses were evaluated from the concentration gradients observed in these last 12 h as described.[21]

Analytical gel-filtration

BcnI protein (50 μg; 1.8 nmol) was mixed with blunt-ended 9-mer cognate oligoduplex (the recognition sequence is shown in bold):

5′-CG**CCCGG**AC-3′
3′-GC**GGGCC**TG-5′

in buffer B (10 mM Tris–HCl (pH 7.4), 0.2 M NaCl, 5 mM $CaCl_2$, 1 mM DTT) at different stoichiometric ratios and samples were loaded on a Superdex 75 HR 10/300 (Amersham Biosciences) column, which was equilibrated with the same buffer. The column was calibrated by measuring the elution volumes of a series of standard proteins of known molecular mass (Bio-Rad). For the interpolation of unknown molecular mass, a linear dependence of the logarithm of the molecular mass on the elution time was assumed.

Elution profiles were monitored by an Ettan two-wavelength detector at 260 nm and 280 nm. The A_{260}/A_{280} ratios necessary for profile deconvolution were deduced from the ratios of A_{260} and A_{280} peak heights after injection of only protein or only DNA. For our system, we determined $A_{260}/A_{280}\approx 1.9$ for DNA and $A_{260}/A_{280}\approx 0.6$ for BcnI without DNA. Absolute absorbance values were calculated as follows: for the double-stranded DNA, an $A_{260}=1$ cm^{-1} was taken to correspond to 0.15 mM nucleotides or 8.33 μM 9-mer oligoduplex. For the BcnI apo form, an extinction coefficient of 21 430 M^{-1}cm^{-1} was calculated by ProtParam tool‡.[22] After some elementary algebraic manipulations, it follows that:

$$c_{\mathrm{protein}} = -35.8\,A_{260} + 68.6\,A_{280}$$

$$c_{\mathrm{DNA}} = 12.2\,A_{260} - 7.5\,A_{280}$$

if absorbances and concentrations are measured in units of cm^{-1} and μM, respectively.

† http://www.rasmb.bbri.org
‡ http://www.expasy.ch/

Native gel electrophoresis

Native gel electrophoresis was run in acidic conditions as described§. Electrophoresis was performed at 4 °C at 25 mA. Care was taken to reverse the polarity relative to the usual arrangement, because the proteins migrate as positively charged species at pH 4.3.

BcnI (5 μg; 0.18 nmol) was mixed in 10 μl of buffer B with the blunt-ended 9-mer duplex that was used for gel-filtration experiments in various stoichiometric ratios. Mixtures were kept on ice overnight, supplemented with 2 μl of loading dye (1.45 ml of 50% (v/v) glycerol, 0.5 ml of 0.25 M potassium acetate, a trace of methylene green) and samples were immediately loaded on the gel.

Crystallization

Crystallization was achieved by the sitting-drop, vapor-diffusion technique at room temperature. Initial high-throughput screens were set up at the 200 nl scale using a Cartesian pipetting robot and 96-well Greiner sitting-drop plates. CRYSCHEM plates (Hampton Research) were used for larger drop volumes.

Crystals without DNA

BcnI in storage buffer was dialyzed against buffer C (10 mM Tris–HCl (pH 7.4), 0.2 M NaCl, 1 mM EDTA, 1 mM DTT) and concentrated to 10–12 mg/ml by ultrafiltration (Amicon). Large crystals were grown in CRYSCHEM plates (Hampton Research) by mixing 2 μl of the protein solution with 2 μl of buffer containing 0.17 M ammonium acetate, 85 mM trisodium citrate (pH 5.6), 25.5% (w/v) PEG 4000 and 15% (v/v) glycerol, which served as the reservoir buffer. Crystals of the selenomethionine variant were grown in the same conditions. A bromide derivative was obtained by soaking a native crystal in crystallization buffer supplemented with 0.5 M NaBr for 30 s before flash cryo-cooling. All BcnI crystals could be flash cryo-cooled directly from the crystallization drop without additional cryoprotection.

Crystals with DNA

BcnI in storage buffer was dialyzed against buffer B and concentrated. Final protein concentrations were 8–10 mg/ml. The oligonucleotides used to make an 11-mer oligoduplex:

5′-AA**CCCGG**AGAC-3′
3′-TGTT**GGGCC**TC-5′

were purchased from Metabion (Germany), dissolved in 10 mM Tris–HCl (pH 7.4) and annealed by heating to 95 °C followed by slow cooling to 4 °C. BcnI and 11 bp oligoduplex (2 nt overhangs) solutions were mixed at a 1:1.1 molar ratio. Crystals were obtained using the sitting-drop, vapor-diffusion technique. The crystallization buffer was 20 mM calcium chloride, 0.1 M sodium acetate (pH 4.6) and 30% (v/v) 2-methyl-2,4-pentanediol.

Crystals of the SeMet variant of the BcnI with DNA were grown under identical conditions. All BcnI–DNA crystals appeared after ~48 h and were flash cryo-cooled without extra cryoprotection.

§ http://wolfson.huji.ac.il/purification/Protocols/PAGE_Acidic.html

Table 1. Data collection and refinement statistics

	BcnI without DNA (30.0 Å–1.60 Å)	BcnI with DNA (30.0 Å–1.45 Å)
Space group	$P4_12_12$	$P2_12_12$
Cell dimensions		
a (Å)	68.2	81.6
b (Å)	68.2	87.9
c (Å)	127.2	79.4
Resolution range (Å)	30.0–1.6	30.0–1.45
Wavelength (Å)	1.05	0.97955
Total reflections	231,587	366,066
Unique reflections	40,038	99,034
Completeness (%)	99.2 (99.4)	97.8 (96.3)
I/σ	4.0 (3.1)	8.5 (2.9)
R_{sym} (%)	7.2 (24.6)	5.1 (24.7)
B(iso) from Wilson plot (Å^2)	26.3	12.6
Protein atoms (excluding H)	1937	4038
DNA atoms (excluding H)	–	890
Solvent molecules	216	655
R-factor (%)	20.3	17.4
R_{free} (%)	21.5	19.9
RMSD from ideal		
Bond lengths (Å)	0.01	0.01
Bond angles (deg.)	1.3	1.2
Ramachandran plot		
Core region (%)	93.7	90.6
Allowed region (%)	5.3	8.9
Additionally allowed region (%)	1.0	0.0
Disallowed region (%)	0.0	0.5

Values for the highest resolution shell are in parentheses.

Structure determination

All diffraction data were collected at 100K. In-house data were measured on a RUH300 generator with copper anode from MSC/Rigaku equipped with Osmic multi-layer optics and a MAR345 image plate. Synchrotron data were collected at beamline BW6 at DESY, Hamburg. All data were processed with MOSFLM[23] and scaled with SCALA (Table 1).[24]

Crystals without DNA

The BcnI crystals without DNA belonged to space group $P4_12_12$ and contained one monomer in the asymmetric unit. We collected native data and two-wavelength MAD datasets of the selenium variant and of a bromide soak at the selenium and bromide K-edges, respectively. In the selenium MAD data, the anomalous differences at the absorption maximum and inflection point were between 90% and 60% correlated in the resolution range from 20.0–3.0 Å. In the bromide MAD data, the correlations were 73% and 60% in the resolution range from 20.0–2.5 Å. The selenium and bromide anomalous differences were interpreted in terms of the six chemically present Se sites or in terms of 20 bromide sites (the exact cut-off for the number of bromide sites was chosen arbitrarily). For both phase sets, the contrast[25,26] after 20 rounds of SHELXE density modification was significantly higher for the correct space group $P4_12_12$ than for the enantiomorphic alternative (0.54 *versus* 0.19 for the selenium data and 0.53 *versus* 0.11 for the bromide data). The consistency of origin was assured by crossphasing. Selenium and bromide MAD phases were obtained in separate MLPHARE[24] phasing runs, combined with the SIGMAA,[24] improved by density modification, extended to the full resolution, and used as input for automatic model building. The automatically built model was completed manually and refined with the maximum likelihood program REFMAC,[27] using separate TLS parameters for the catalytic lobe and for the recognition lobe (Table 1).

Crystals with DNA

The BcnI crystals with DNA belonged to space group $P2_12_12$ and contained two complexes of a BcnI monomer with a DNA duplex in the asymmetric unit. Surprisingly, molecular replacement with the structure of BcnI taken from the crystals without DNA failed to provide a satisfactory molecular replacement solution. Therefore, we collected a three-wavelength MAD dataset of the selenomethionine variant of the protein in complex with DNA at the K-edge of selenium (at the absorption maximum, inflection point and at a low-energy remote wavelength). Using only the data at the absorption maximum, the SHELXD program[28] readily identified the expected 12 selenium sites (occupancy for correct sites above 0.49 and for incorrect sites below 0.19). Phasing with SHELXE[29] indicated a substantial contrast between the two enantiomorphic alternatives (0.48 *versus* 0.39). The density for the correct enantiomorphic alternative was partially interpretable by automatic model building. The BcnI model was then completed easily using the known structures of the catalytic lobe and the recognition lobe of BcnI from the structure without DNA. Canonical models for the DNA duplexes with the correct sequence were generated with the 3DNA program,[30] and then manually placed and adjusted to the electron density using the modeling program O.[31] The structure was refined with the maximum likelihood program REFMAC,[27] using separate TLS parameters for each catalytic lobe, recognition lobe or DNA duplex (Table 1).

Protein Data Bank accession codes

Structures have been deposited in the Protein Data Bank‖ (and will be available under accession codes 2ODH (apo-BcnI) and 2ODI (BcnI–DNA complex) upon publication.

Acknowledgements

We thank Professor Hans Bartunik for generous allocation of beamtime on BW6 (DESY, Hamburg) and Dr Gleb Bourenkov for collecting MAD data for us. We are grateful to AB "Fermentas" for the BcnI RM system clone. Research in the V.S. laboratory was supported by the Howard Hughes Medical Institute International Research Scholar grant # 55000336. M.B. thanks the European Molecular Biology Organization (EMBO) and HHMI for a Young Investigator award and the UNESCO/Polish Academy of Sciences Cellular and Molecular Biology Network for financial support. H.C. is grateful to the Polish Ministry of Science and Higher Education for a POL-POSTDOC grant (PBZ/MEiN/01/2006/24). This work was supported by the European Commission 5th Framework Programme project "Center of Excellence in Molecular Bio-Medicine Contract no: QLK6-CT-2002-90363" (Warsaw) and "Center of Exellence - Biocell" Contract no: QLK2-CT-2002-30575 (Vilnius). The funds from the Polish Ministry of Science and Information Technology for the purchase of a crystallization robot (Decision Nr. 5210/IA/1789/2005) are gratefully acknowledged.

‖ http://www.rcsb.org

Supplementary Data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.jmb.2007.03.018

References

1. Xu, Q. S., Kucera, R. B., Roberts, R. J. & Guo, H. C. (2004). An asymmetric complex of restriction endonuclease MspI on its palindromic DNA recognition site. *Structure*, **12**, 1741–1747.
2. Yang, Z., Horton, J. R., Maunus, R., Wilson, G. G., Roberts, R. J. & Cheng, X. (2005). Structure of HinP1I endonuclease reveals a striking similarity to the monomeric restriction enzyme MspI. *Nucl. Acids Res.* **33**, 1892–1901.
3. Horton, J. R., Zhang, X., Maunus, R., Yang, Z., Wilson, G. G., Roberts, R. J. & Cheng, X. (2006). DNA nicking by HinP1I endonuclease: bending, base flipping and minor groove expansion. *Nucl. Acids Res.* **34**, 939–948.
4. Petrusyte, M. & Janulaitis, A. (1982). Isolation and some properties of the restriction endonuclease BcnI from *Bacillus centrosporus*. *Eur. J. Biochem.* **121**, 377–381.
5. Janulaitis, A., Povilionis, P. & Sasnauskas, K. (1982). Cloning of the modification methylase gene of *Bacillus centrosporus* in *Escherichia coli*. *Gene*, **20**, 197–204.
6. Janulaitis, A., Klimasauskas, S., Petrusyte, M. & Butkus, V. (1983). Cytosine modification in DNA by BcnI methylase yields N4-methylcytosine. *FEBS Letters*, **161**, 131–134.
7. Vilkaitis, G., Lubys, A., Merkiene, E., Timinskas, A., Janulaitis, A. & Klimasauskas, S. (2002). Circular permutation of DNA cytosine-N4 methyltransferases: *in vivo* coexistence in the BcnI system and *in vitro* probing by hybrid formation. *Nucl. Acids Res.* **30**, 1547–1557.
8. Lebowitz, J., Lewis, M. S. & Schuck, P. (2002). Modern analytical ultracentrifugation in protein science: a tutorial review. *Protein Sci.* **11**, 2067–2079.
9. Kaus-Drobek, M., Czapinska, H., Sokolowska, M., Tamulaitis, G., Szczepanowski, R. H., Urbanke, C., Siksnys, V., & Bochtler, M. (2007). Restriction endonuclease MvaI is a monomer that recognizes its target sequence asymmetrically. *Nucl. Acids Res.* **epub ahead of print**, doi:10.1093/nar/gkm064.
10. Richardson, J. S. (1977). β-Sheet topology and the relatedness of proteins. *Nature*, **268**, 495–500.
11. Harding, M. M. (1999). The geometry of metal-ligand interactions relevant to proteins. *Acta Crystallog. sect. D*, **55**, 1432–1443.
12. Mordasini, T., Curioni, A. & Andreoni, W. (2003). Why do divalent metal ions either promote or inhibit enzymatic reactions? The case of BamHI restriction

endonuclease from combined quantum-classical simulations. *J. Biol. Chem.* **278**, 4381–4384.
13. Pingoud, V., Sudina, A., Geyer, H., Bujnicki, J. M., Lurz, R., Luder, G. *et al.* (2005). Specificity changes in the evolution of type II restriction endonucleases: a biochemical and bioinformatic analysis of restriction enzymes that recognize unrelated sequences. *J. Biol. Chem.* **280**, 4289–4298.
14. Seeman, N. C., Rosenberg, J. M. & Rich, A. (1976). Sequence-specific recognition of double helical nucleic acids by proteins. *Proc. Natl Acad. Sci. USA*, **73**, 804–808.
15. Holm, L. & Sander, C. (1996). The FSSP database: fold classification based on structure-structure alignment of proteins. *Nucl. Acids Res.* **24**, 206–209.
16. Welsh, K. M., Lu, A. L., Clark, S. & Modrich, P. (1987). Isolation and characterization of the *Escherichia coli* mutH gene product. *J. Biol. Chem.* **262**, 15624–15629.
17. Kupper, D., Reuter, M., Mackeldanz, P., Meisel, A., Alves, J., Schroeder, C. & Kruger, D. H. (1995). Hyper-expressed EcoRII renatured from inclusion bodies and native enzyme both exhibit essential cooperativity with two DNA sites. *Protein Expr. Purif.* **6**, 1–9.
18. Zimbro, M. J. & Power, D. A. (2003). *Difco and BBL Manual.* Becton Dickinson and Company, Sparks, MD.
19. Van Duyne, G. D., Standaert, R. F., Karplus, P. A., Schreiber, S. L. & Clardy, J. (1993). Atomic structures of the human immunophilin FKBP-12 complexes with FK506 and rapamycin. *J. Mol. Biol.* **229**, 105–124.
20. Schuck, P. (2000). Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. *Biophys. J.* **78**, 1606–1619.
21. Siksnys, V., Skirgaila, R., Sasnauskas, G., Urbanke, C., Cherny, D., Grazulis, S. & Huber, R. (1999). The Cfr10I restriction enzyme is functional as a tetramer. *J. Mol. Biol.* **291**, 1105–1118.
22. Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D. & A., B. (2005). Protein identification and analysis tools on the ExPASy server. In *The Proteomics Protocols Handbook*, pp. 571–607, Humana Press, Totowa, NJ, USA.
23. Leslie, A. W. G. (1992). Recent changes to the MOSFLM package for processing film and image plate data. *Joint CCP4 + ESF-EAMCB Newsletter Protein Crystallog.* **26**.
24. Collaborative Computational Project Number 4. (1994). The CCP4 suite: programs for protein crystallography. *Acta Crystallog. sect. D*, **50**, 760–763.
25. Terwilliger, T. C. & Berendzen, J. (1999). Evaluation of macromolecular electron-density map quality using the correlation of local r.m.s. density. *Acta Crystallog. sect. D*, **55**, 1872–1877.
26. Terwilliger, T. C. & Berendzen, J. (1999). Discrimination of solvent from protein regions in native Fouriers as a means of evaluating heavy-atom solutions in the MIR and MAD methods. *Acta Crystallog. sect. D*, **55**, 501–505.
27. Murshudov, G. N., Vagin, A. A. & Dodson, E. J. (1997). Refinement of macromolecular structures by the maximum-likelihood method. *Acta Crystallog. sect. D*, **53**, 240–255.
28. Uson, I. & Sheldrick, G. M. (1999). Advances in direct methods for protein crystallography. *Curr. Opin. Struct. Biol.* **9**, 643–648.
29. Sheldrick, G. M. (2002). Macromolecular phasing with SHELXE. *Zeitschrift Kristallog.* **217**, 644–650.
30. Lu, X. J. & Olson, W. K. (2003). 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucl. Acids Res.* **31**, 5108–5121.
31. Kleywegt, G. J. & Jones, T. A. (1998). Databases in protein crystallography. *Acta Crystallog. sect. D*, **54**, 1119–1131.
32. Lee, J. Y., Chang, J., Joseph, N., Ghirlando, R., Rao, D. N. & Yang, W. (2005). MutH complexed with hemi- and unmethylated DNAs: coupling base recognition and DNA cleavage. *Mol. Cell*, **20**, 155–166.
33. Joshi, H. K., Etzkorn, C., Chatwell, L., Bitinaite, J. & Horton, N. C. (2006). Alteration of sequence specificity of the type II restriction endonuclease HincII through an indirect readout mechanism. *J. Biol. Chem.* **281**, 23852–23869.
34. Huai, Q., Colandene, J. D., Topal, M. D. & Ke, H. (2001). Structure of NaeI-DNA complex reveals dual-mode DNA recognition and complete dimer rearrangement. *Nature Struct. Biol.* **8**, 665–669.
35. Vanamee, E. S., Viadiu, H., Kucera, R., Dorner, L., Picone, S., Schildkraut, I. & Aggarwal, A. K. (2005). A view of consecutive binding events from structures of tetrameric endonuclease SfiI bound to DNA. *EMBO J.* **24**, 4198–4208.
36. Horton, N. C. & Perona, J. J. (1998). Role of protein-induced bending in the specificity of DNA recognition: crystal structure of EcoRV endonuclease complexed with d(AAAGAT) + d(ATCTT). *J. Mol. Biol.* **277**, 779–787.
37. Newman, M., Lunnen, K., Wilson, G., Greci, J., Schildkraut, I. & Phillips, S. E. (1998). Crystal structure of restriction endonuclease BglI bound to its interrupted DNA recognition sequence. *EMBO J.* **17**, 5466–5476.
38. Horton, J. R. & Cheng, X. (2000). PvuII endonuclease contains two calcium ions in active sites. *J. Mol. Biol.* **300**, 1049–1056.

Edited by K. Morikawa

(Received 15 January 2007; received in revised form 1 March 2007; accepted 6 March 2007)
Available online 15 March 2007