6238–6248 *Nucleic Acids Research, 2007, Vol. 35, No. 18* *Published online 13 September 2007*
doi:10.1093/nar/gkm665

Catalytic domain of restriction endonuclease BmrI as a cleavage module for engineering endonucleases with novel substrate specificities

Siu-hong Chan, Yongming Bao, Ewa Ciszak, Sophie Laget and Shuang-yong Xu*

New England Biolabs, Inc. 240 County Road, Ipswich, MA 01938, USA

Received December 14, 2006; Revised and Accepted August 12, 2007

ABSTRACT

Creating endonucleases with novel sequence specificities provides more possibilities to manipulate DNA. We have created a chimeric endonuclease (CH-endonuclease) consisting of the DNA cleavage domain of BmrI restriction endonuclease and C.BclI, a controller protein of the BclI restriction-modification system. The purified chimeric endonuclease, BmrI198-C.BclI, cleaves DNA at specific sites in the vicinity of the recognition sequence of C.BclI. Double-strand (ds) breaks were observed at two sites: 8 bp upstream and 18 bp within the C-box sequence. Using DNA substrates with deletions of C-box sequence, we show that the chimeric endonuclease requires the 5′ half of the C box only for specific cleavage. A schematic model is proposed for the mode of protein–DNA binding and DNA cleavage. The present study demonstrates that the BmrI cleavage domain can be used to create combinatorial endonucleases that cleave DNA at specific sequences dictated by the DNA-binding partner. The resulting endonucleases will be useful *in vitro* and *in vivo* to create ds breaks at specific sites and generate deletions.

INTRODUCTION

Restriction endonucleases (REases), particularly of Type IIP that recognize palindrome sequences and cleave within them, are indispensable tools for DNA manipulation because of their high sequence and cleavage specificity. Substrate specificity, which is always coupled to the catalytic core for Type IIP REases, is known to rely on intricate interactions between the amino acid residues of the REase and the bases and the backbone phosphates of the substrate DNA. However, similar DNA-binding specificity is rarely reflected in amino acid sequence homology among REases. BamHI, for example, recognizes G↓GATCC and cuts between the first two Gs, whereas KpnI, a REase isolated from an evolutionary unrelated bacterium that shares very low sequence similarity with BamHI, recognizes GGTAC↓C and cuts between the last two Cs. This leaves protein engineers no obvious means to identify patterns or recognition modules within the amino acid sequences of Type IIP REases that recognize similar DNA sequences. Engineering Type IIP REases has to resort to genetic screening systems specifically designed for each of the specificities of interest or sophisticated computational design based on the atomic structure of the enzymes concerned. Variants of EcoRV that prefer cleavage sites flanked by AT or GC have been identified by random mutagenesis within specific regions of the EcoRV REase (1). Partial successes have been reported for engineering Type IIP REases that recognizes degenerate sequences. BstYI (R↓GATCY) has been engineered to cleave AGATCT (2) and BsoBI (C↓YCGRG) to cleave CCCGGG preferentially (3). Recently, alternative specificity (GC↓TGCCGC) has been introduced to NotI (GC↓GGCCGC) through genetic screening of a randomized library (4). Computational redesign of homing endonucleases based on their crystal structures had resulted in a variant of I-MsoI with altered sequence specificity (5) and a fusion protein created by swapping domains of I-DmoI and I-CreI and optimization of domain interface (6). Tailored-made specificities had been achieved through genetic screening of I-CreI at specific amino acid residues that make direct or indirect contact with substrate DNA (7–13).

Certain types of REases, such as Types I, IIS, IIG and III endonucleases, cleave substrate DNA outside their recognition sequences. Biochemical and structural studies had shown that these endonucleases consist of separate

*To whom correspondence should be addressed. Tel: +1 978 380 7287; Fax: +1 978 921 1350; Email: xus@neb.com
Present address:
Yongming Bao, Department of Bioscience and Biotechnology, Dalian University of Technology, Dalian, 116024, P. R. China.

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

DNA recognition and cleavage domains. It has been suggested that these endonucleases evolved through recombination of their cleavage domains with DNA-binding domains of different sequence specificity (14,15). This uncoupling of substrate specificity and cleavage activity opens the door to creating combinatorial endonuclease of novel substrate specificity through fusing DNA cleavage domains to DNA recognition proteins or oligonucleotides (16).

FokI is a Type IIS REase that cuts 9 and 13 bp downstream of the recognition sequence of GGATC on the top and bottom strand, respectively (GGATC N_9/N_{13}). Combinatorial endonucleases have been created by fusing the DNA cleavage domain of FokI to a few DNA-binding domains. Fusing the FokI cleavage domain to the *Drosophila Ultrabithorax* (*Ubx*) homeodomain resulted in a chimeric endonuclease that binds and cuts outside the *Ubx*-binding site (17). Chimeric endonucleases were also created by fusing the FokI cleavage domain to the Z-conformation-specific Za domain of human RNA adenosine deaminase (18), Gal4 (19) and natural zinc-finger motifs (20–25), generating novel substrate specificities *in vitro*. Directed-evolution of sequence specificity of Zn fingers has brought us closer toward engineering substrate specificity of endonucleases (26,27). Zn finger/FokI chimeric endonucleases (Zinc finger nucleases, ZFNs) have been shown to generate deletions in target sequences in the germline of *Drosophila in vivo* (28–30). ZFNs injected into nuclei of *Xenopus* oocytes induced efficient insertion of an extra-chromosomal DNA through homologous recombination (28–31).

With increasing number of sequenced and characterized Type IIS REases, we set out to find other cleavage domains that can be used in this combinatorial approach to creating endonucleases with novel substrate specificity. BmrI is a Type IIS REase that recognizes the asymmetric 6-bp sequence ACTGGG and cleaves 5 and 4 bp downstream on the top strand and bottom strand, respectively (ACTGGG N_5/N_4). It is highly homologous to its isoschizomer BfiI in amino acid sequence. BmrI and BfiI are unconventional REases in that they do not require divalent metal ions for DNA cleavage (32). They consist of a DNA-binding domain and a non-specific cleavage domain consisting of a HKD catalytic motif of the phospholipase D family (33). The two domains are joined by a relatively flexible linker sequence (15,34), which is believed to allow domain movements that are needed in the transition from recognition to DNA cleavage activity (35,36). Structural and biochemical studies suggest that BfiI recognizes its target sequence with a single DNA recognition domain and makes ds breaks sequentially (37). The modular structure and sequential cleavage of ds DNA suggest that the DNA cleavage domains of BfiI and BmrI can be connected to other DNA-binding domains to generate novel substrate specificity.

C.BclI is the controller protein of the BclI R-M system. C.BclI binds to a 12-bp inverted repeats upstream of its own open reading frame (ORF) with a dissociation constant in the nanomolar range *in vitro* (38). It represses the expression of the MTases of the BclI R-M system *in vivo* (38). C.BclI contains a helix–turn–helix (HTH) domain for binding to its target DNA sequence as predicted by homology modeling of the crystal structure of C.BclI to that of Cro repressor of phage 434 (38). The 12-bp inverted repeats in the C-box suggests that C.BclI binds to its target sequence as homodimer or tetramer (38).

Here we make use of the modular property of BmrI endonuclease and the DNA-binding specificity of C.BclI to generate a combinatorial endonuclease of novel sequence specificity. A chimeric endonuclease was constructed by linking the cleavage domain of BmrI to C.BclI through a 14-amino acid linker. The chimeric endonuclease requires the 5′ half of the C box only for making specific double-strand (ds) breaks 13–16 bp downstream of the binding site.

MATERIALS AND METHODS

Cloning and expression

The cloning and expression of BmrI R-M system are described in Higgins *et al.* (manuscript submitted). The DNA fragments that encode residues 1–198, 1–204, 1–209 of BmrI were amplified from the cloned BmrI system by PCR and ligated to pET21a (Novagen). The gene of the controller protein of the BclI R-M system (*bclIC*) has been cloned, and the C.BclI protein has been expressed and purified previously (38). The chimeric endonucleases of BmrI-C.BclI were constructed such that two modules were connected by a linker of 14 amino acid residues. Codons of a 6× His tag were added to the 3′ end of the ORF of the fusion proteins to facilitate purification. The coding sequence of the chimeric endonuclease was ligated to pET21a under the control of T7 promoter. All constructs were sequenced to confirm the absence of mutations.

For small-scale expression, *Escherichia coli* strain T7 Express (NEB) was transformed separately by each of the constructs of pET21a carrying fusions between C.BclI and each of the BmrI truncation variants. The transformed cells were cultured in 100 ml of LB with 0.1 mg/ml of ampicillin at 37°C at 200 r.p.m. until OD_{600} reached 0.9. Fusion protein production was induced by adding IPTG to 0.5 mM final concentration and cultured at 16–20°C for 3 h. Ten milliliters of the culture were harvested and cells were lyzed by sonication. Five microliters of the soluble fraction of the lysate were used in cleavage activity assay described below. For large-scale expression and purification, *E. coli* strain T7 Express was transformed by the construct pET21a-BmrI198-Linker-C.BclI (pET-BLC). The transformed cells were cultured in 1 l of LB with 0.1 mg/ml of ampicillin at 37°C at 200 r.p.m. until OD_{600} reached 0.9. Expression was induced by adding IPTG to 0.5 mM and cultured for 3 h under the same condition. The culture was harvested and stored at −20°C until lysed for protein purification.

Purification and refolding

The cell pellet derived from 1 l of IPTG-induced culture (~6 g) was re-suspended in 60 ml of lysis buffer (50 mM NaH_2PO_4, 300 mM NaCl, pH 8.0, 1 mg/ml lysozyme).

The lysate was kept on ice for 30 min and then centrifuged at 15 000*g* for 20 min at 4°C. The fusion protein was mainly found in the insoluble fraction (inclusion bodies). The inclusion bodies were resuspended in 30 ml of washing buffer (2 M urea, 100 mM NaH_2PO_4, 10 mM Tris–HCl, pH 8.0) and kept on ice for 30 min After centrifugation at 15 000*g* for 20 min at 4°C, the inclusion bodies were re-suspended in 30 ml of denaturing buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris–HCl, pH 8.0) and kept on ice for 30 min. The denatured protein solution was centrifuged at 15 000*g* for 20 min at 4°C. The supernatant was loaded onto four 1 ml Ni-NTA columns (Qiagen). After washing with wash buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris–HCl, pH 6.3), the bound protein was eluted using elution buffer (8 M urea, 100 mM NaH_2PO_4, 10 mM Tris–HCl, pH 4.5). A successful refolding condition was found using the Protein Refolding Kit (US Biological). The eluted protein (~12 ml) was then dialyzed in 500 ml of refolding buffer (50 mM Tris–HCl, pH 8.5, 10 mM NaCl, 0.4 mM KCl, 2 mM $MgCl_2$, 2 mM $CaCl_2$, 0.4 M sucrose, 0.5% Triton X-100, 0.05% PEG 3350, 1 mM GSH, 0.1 mM GSSH) at 4°C overnight. The refolded protein was dialyzed against a storage buffer (10 mM Tris–HCl, 250 mM NaCl, 1 mM DTT, 0.1 mM EDTA, 0.5 mg/ml BSA, pH 7.4) and stored at 4°C.

DNA cleavage reactions

Litmus28-*bclIC* (38) that contains the wild-type C box and C.BclI ORF sequences was used as the substrate in DNA cleavage assays. Litmus28 (NEB), from which Litmus28-*bclIC* is derived, was used as a negative control for specific cleavage. 0.4 μg (12 pmol) of purified BmrI198-C.BclI fusion protein was generally used on 125 μg (67 fmol) of substrate DNA in designated buffers and temperatures in 20 μl reaction mixtures. Substrate DNA was either pre-linearized with DraIII or DraIII was added to the cleavage reactions along with the BmrI198-C.BclI fusion protein as indicated. Cleavage products were analyzed by electrophoresis through 1% agarose gels in 1× TBE buffer. The intensities of the DNA bands in ethidium bromide-containing agarose gels were quantified by QuantityOne software (BioRad). Reaction buffers used included: Buffer 1 (10 mM Bis Tris Propane–HCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.0), Buffer 2 (10 mM Tris–HCl, 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9), Buffer 3 (50 mM Tris–HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9), Buffer 4 (20 mM Tris–acetate, 10 mM magnesium acetate, 50 mM potassium acetate, 1 mM DTT, pH 7.9), EDTA buffer (50 mM Tris–HCl, 100 mM NaCl, 10 mM EDTA, 1 mM DTT, pH 7.9), high salt buffer (10 mM Tris–HCl, 150 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9), EcoRI buffer (100 mM Tris–HCl, 50 mM NaCl, 10 mM $MgCl_2$, 0.025% Triton X-100, pH 7.5), Mg^{++} buffer (50 mM Tris–HCl, 100 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT, pH 7.9), Ca^{++} buffer (50 mM Tris–HCl, 100 mM NaCl, 10 mM $CaCl_2$, 1 mM DTT, pH 7.9) and Zn^{++} buffer (50 mM Tris–HCl, 100 mM NaCl, 2 mM $ZnSO_4$, 1 mM DTT, pH 7.9).

To map the cleavage sites, Litmus28-*bclIC* was digested with the indicated REases and 125 μg of the linear DNA was then incubated with 0.4 μg of purified BmrI198-C.BclI fusion protein in Buffer 3 at 37°C for 1 h in 20 μl reaction mixtures, followed by agarose gel electrophoresis. Cleavage sites of BmrI198-C.BclI were determined by incubating Litmus28-*bclIC* with DraIII and BmrI198-C.BclI in Buffer 3 at 37°C for 1 h. The cleaved fragments (~2.1 and ~1.2 kb) were gel-purified and sequenced. The presence of an extra A and a reduction of peak intensity of the proceeding peaks in the electropherograms are indicative of cleavage sites.

The thermal stability of cleavage activity was tested by incubating BmrI198-C.BclI in Buffer 3 at the designated temperatures for 20 min in the absence of substrate DNA. After returning to room temperature, substrate DNA was added and the reaction mixtures were incubated at 37°C for 1 h. To determine the optimal reaction temperature, BmrI198-C.BclI was pre-heated to the designated temperature for 2 min in Buffer 3 before adding substrate DNA. The reaction mixtures were then incubated at 37°C for 1 h.

To determine the mode of binding to the C box, the EcoRI/AccI fragment (5′ half of C box plus 32 bp upstream sequence) and the AccI/EcoNI fragment (3′ half of C box plus 262 bp downstream sequence) were deleted from Litmus28-*bclIC*. The truncation plasmid variants Litmus28-*bclIC*Δ5′ and Litmus28-*bclIC*Δ3′ were used as substrate for cleavage reaction and run-off sequencing as described previously.

RESULTS

Construction and expression of BmrI198-C.BclI fusion protein

CH-endonucleases consisting of N-terminal nuclease domain of BmrI and C.BclI were constructed and sequenced. The CH-endonucleases consist of amino acid residues 1–198, 1–204 and 1–209 of BmrI followed by a 14-amino acid linker and C.BclI sequence (Figure 1A). These fusion proteins contain the HKD motif of BmrI and the natural linker sequence (aa 170–198, colored in yellow; (Figure 1) between the cleavage domain and the DNA recognition domain. Small-scale expression experiments were done and SDS-PAGE analysis showed that all three fusion proteins expressed mostly as inclusion bodies in *E. coli*. DNA cleavage assay showed that the lysate supernatant derived from the clone that expressed the fusion protein consists of amino acid residues 1–198 of BmrI has specific cleavage activity without affecting the growth of the host cells (data not shown). This fusion protein was named BmrI198-C.BclI and was used in the following studies (Figure 1B).

Although low induction temperature at 16–20°C increased the yield of soluble protein, the yield of soluble BmrI198-C.BclI was too low to generate enough protein for further study (data not shown). Therefore, the inclusion bodies were unfolded in a buffer containing 8 M urea, purified through a nickel-charged metal chelation column and then refolded. The best refolding

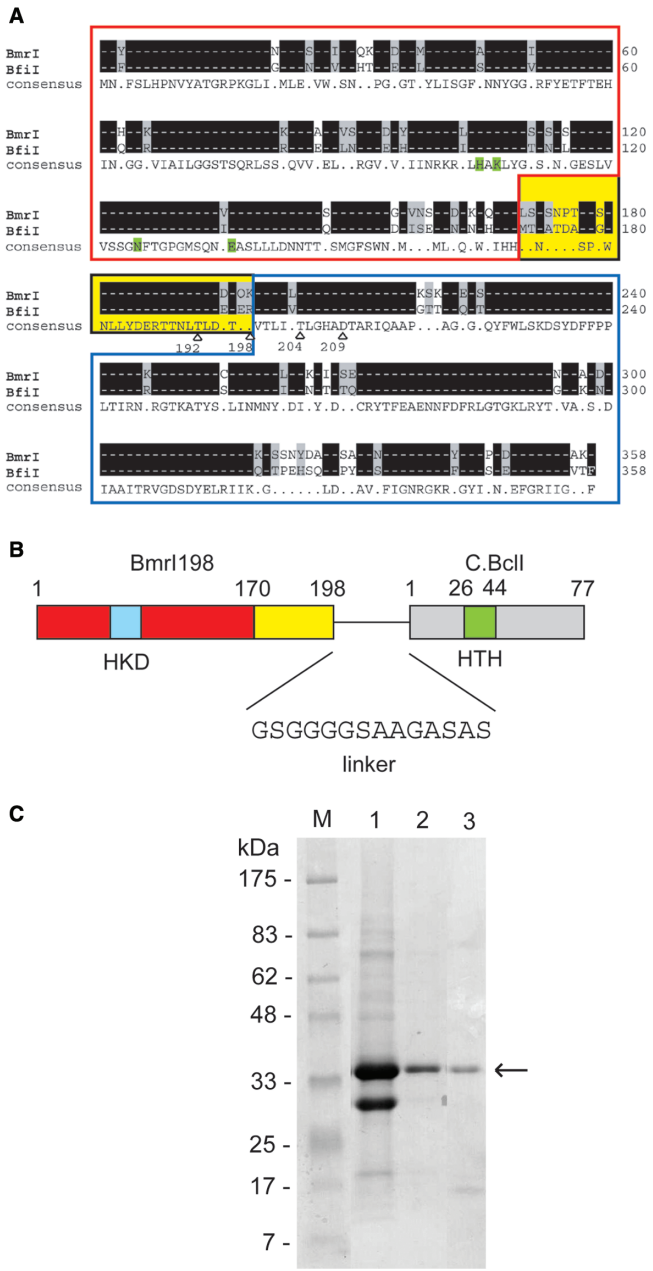

Figure 1. (**A**) Amino acid sequence alignment of BmrI and BfiI endonucleases. Identical residues are shown in the consensus sequence. Conserved residues are colored in gray. The N-terminal cleavage domain is boxed in red, C-terminal DNA-binding domain in blue. The inter-domain intrinsic linker sequence is colored in yellow. Putative catalytic residues of the HKD motif (His105, Lys107, Asn125) are colored in green. Residues 198, 204 and 209 are numbered and marked with triangles. Domain assignment was made according to the crystal structure of BfiI (34). (**B**) Schematic diagram of the CH-endonuclease BmrI198-C.BclI. Peptide sequence containing residues 1–198 of BmrI was linked to C.BclI full-length sequence through a 14-amino acid residues linker sequence that contains Gly, Ala and Ser. The cleavage domain of BmrI is colored in red, the natural inter-domain linker sequence of BmrI in yellow and the HKD catalytic motif in light blue. C.BclI is colored in gray with the helix–turn–helix (HTH) motif highlighted in green. (**C**) Purification of BmrI198-C.BclI. Inclusion bodies isolated from the induced *E. coli* culture was unfolded in a buffer containing 8 M urea (lane 1). The unfolded BmrI198-C.BclI was purified by a Ni-charged metal chelate column (Ni-NTA) (lane 2) and refolded (lane 3).

condition was found by screening 15 refolding buffers. The purified and refolded chimeric endonuclease BmrI198-C.BclI is shown in Figure 1C. The yield of soluble BmrI198-C.BclI was 0.2 mg/g wet cells after refolding.

Sequence-specific cleavage

Plasmid Litmus28-*bclIC* contains the ORF of C.BclI and the C-box sequence (38). BmrI198-C.BclI cleaves DraIII-linearized Litmus28-*bclIC* but not Litmus28 into 2 fragments of ~2.1 and ~1.2 kb (Figure 2A). The sizes of the fragments are consistent with the position of the C box on DraIII-linearized Litmus28-*bclIC*. Restriction mapping using HpaI, ScaI, BsaI and AlwNI further confirmed that the cleavage site is in the vicinity of the C box (Figure 2B). Control experiments using supercoiled DNA as substrates showed that the CH-endonuclease had similar level of nicking activity on specific (Litmus28-*bclIC*) and non-specific (Litmus28) DNA. The CH-endonuclease displayed significantly higher ds cleavage activity on specific DNA than non-specific DNA (data not shown).

The cleavage sites of BmrI198-C.BclI on Litmus28-*bclIC* were determined by sequencing the cleavage products directly. The 2.1 and 1.1 kb cleavage products were gel-purified and subjected to DNA sequencing using primers running from either upstream or downstream of the C box. During sequencing reactions, when the DNA polymerase reaches the end of the template DNA, it 'runs off' from the template. Thus, a sharp decrease in peak intensity is observed in the electropherograms. In addition, an extra A is added to the end of the sequence due to the template-independent terminal transferase activity of the DNA polymerase. Therefore, the presence of an aberrant A accompanied by a sharp drop of peak intensity of proceeding peaks in the electropherograms is interpreted as the end of the template and hence the cleavage site. The height of the aberrant A peak is also suggestive of the population of the template terminated at that site.

DNA sequencing of the cleaved DNA fragments indicated that there is a mixture of molecules cut at two major sites: −8↓−9 and +19↓+20 of C-box sequence (↓ indicating the cleavage site, Figure 2C). On the upper strand, BmrI198-C.BclI makes two major and one minor cuts at −8 to −11 outside the C box and one major cut at +19↓+20 within the C box. On the bottom strand, it makes one major cut at −11↓−12 outside the C box and two major cuts at +17↓+18 and +18↓+19. At the −8 site, blunt ends and 1–3 bp 3′ overhangs are generated whereas at the +20 site, 1–2 bp 3′ overhangs are generated. It is likely that some of the substrate molecules are cleaved at both sites [similar to the BcgI-type REases (39,40)] while some are cleaved at either site. Interestingly, the positions of the two cleavage sites do not align with the symmetry of the C box. The −8↓−9 site lies upstream of the C box whereas the +19↓+20 site lies within the C box.

Substrate sequence requirement for specific cleavage

The asymmetry of the cleavage site relative to the C box suggests that the BmrI198-C.BclI fusion protein

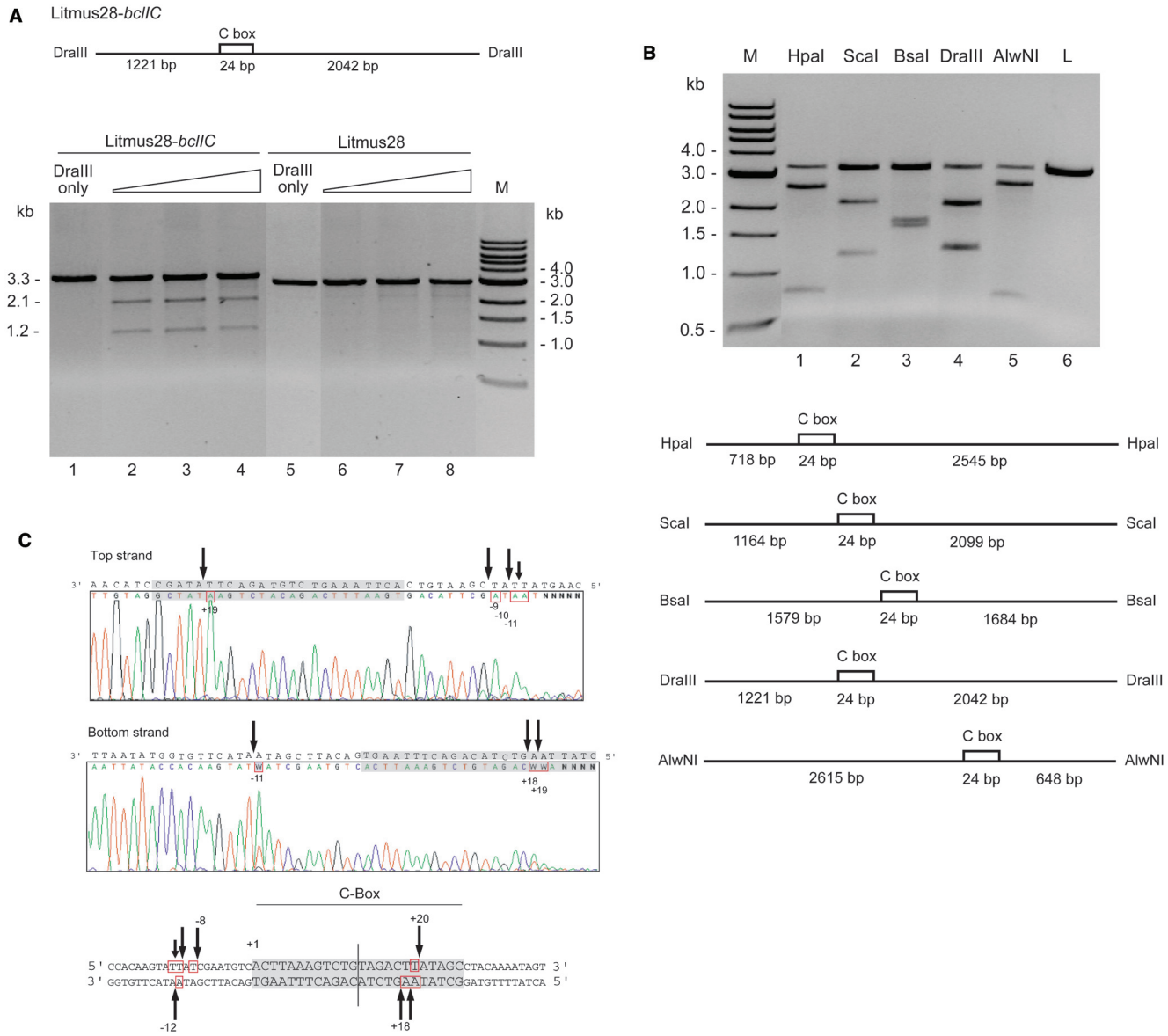

Figure 2. Specific cleavage activity of BmrI198-C.BclI. (**A**) Litmus28-*bclIC* contains the C box and ORF of C.BclI. When linearized, the C box is located 1221 and 2042 bp away from the DraIII site (upper panel). Litmus28-*bclIC* and Litmus28 (67 fmol) were linearized by DraIII and cleaved by 12, 24 or 48 pmol of BmrI198-C.BclI (lanes 2 and 6, 3 and 7, 4 and 8, respectively). Specific cleavage products (2.1 and 1.2 kb) were observed in Litmus28-*bclIC* (lower panel). (**B**) Restriction mapping of BmrI198-C.BclI cleavage site. Litmus28-*bclIC* was pre-cut by the designated REases and then digested by BmrI198-C.BclI (upper panel). The position of the C box with respect to the restriction sites on Litmus28-*bclIC* are shown in the lower panel. In lane 6, the plasmid was incubated with DraIII in the absence of BmrI198-C.BclI (linearized). (**C**) Run-off sequencing of cleavage products. After cleavage by DraIII and BmrI198-C.BclI, the cleavage products were subjected to DNA sequencing from both directions. Electropherograms of sequencing reactions for the top and bottom strands were shown. C-box sequences are boxed in gray. DNA sequences are shown on top of the electropherograms. The nucleotides are numbered relative to the start (+1) of the C box sequence. Down arrows indicate the cleavage sites marked by a drop of the peak intensity and an aberrant A, which are boxed in red line and numbered. Large and small down arrows indicate major and minor cleavage, respectively. The lower panel shows the ds DNA sequence in the vicinity of the C box. The arrows and the red boxes correspond to those in the electropherograms. Nucleotides 3′ to the cleavage sites are numbered.

does not occupy the whole C box. We deleted the 5′ half (Litmus28-*bclIC*Δ5′) or the 3′ half (Litmus28-*bclIC*Δ3′) of BclI C box respectively from Litmus28-*bclIC* and verified the effect of these binding blocks on the cleavage activity of BmrI198-C.BclI. Using the same amount of enzyme, the chimeric endonuclease cleaved Litmus28-*bclIC*Δ3′ into the same patterns as the wild-type substrate Litmus28-*bclIC*, whereas only minor cleavage was observed with the substrate Litmus28-*bclIC*Δ5′ (Figure 3A). This indicates that the nucleotides deleted in Litmus28-*bclIC*Δ3′ are not involved in the specific binding and cleavage for the chimeric endonuclease, and that the nucleotides deleted in Litmus28-*bclIC*Δ5′ are necessary for specific binding.

Figure 3B shows the run-off sequencing of the cleavage products from Litmus28-*bclIC*Δ3′. It shows that

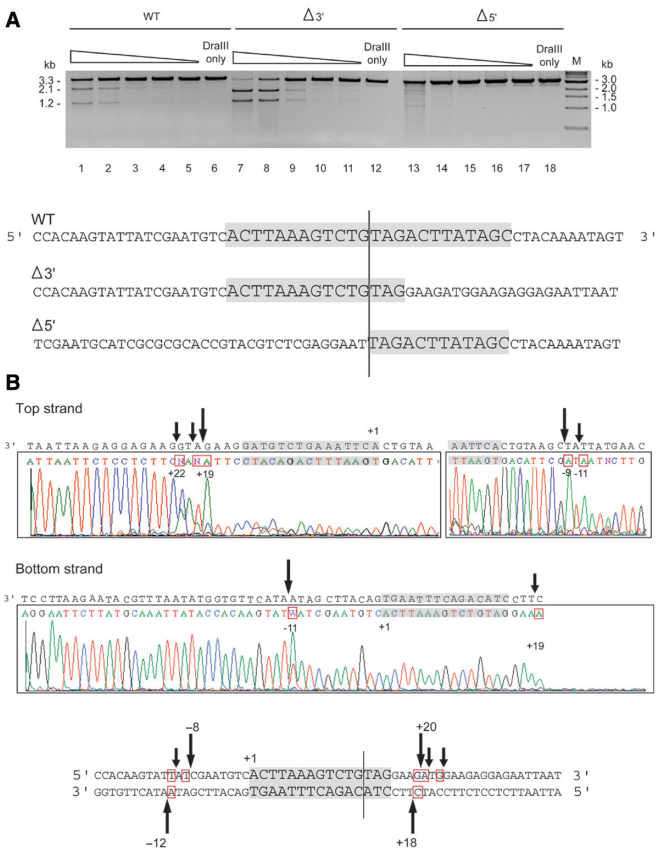

Figure 3. Cleavage of the C-box deletion variants. (**A**) DraIII-linearized WT (lanes 1–4), Δ3′ (lanes 5–8) and Δ5′ (lanes 9–12) variants of Litmus28-*bclIC* were cleaved with increasing amount of BmrI198-C.BclI under the same condition. The lower panel shows the sequences of the C box and flanking sequences for the substrates. (**B**) Run-off sequencing of BmrI198-C.BclI-cleaved Litmus28-*bclIC*Δ3′ (pre-cut by DraIII). Annotation scheme is identical to that of Figure 2C.

BmrI198-C.BclI makes major cuts (large down arrows) on Litmus28-*bclIC*Δ3′ at the same sites (−8↓−9 and +19↓+20 sites) as it cleaves the wild-type C box on Litmus28-*bclIC* (Figure 2C). However, variations of minor cuts (small down arrows) were observed. On the top strand, the minor cut between −10 and −11 (TAT↓TAT) in wild-type substrate was not observed in duplicated sequencing reactions of the variant. The variant has two minor cuts (+20↓+21; +22↓+23) at the top strand of the +20 site, which were not found in wild-type substrate. The fact that BmrI198-C.BclI makes the same major cuts on the wild-type Litmus28-*bclIC* and the deletion variant of Litmus28-*bclIC*Δ3′, and that the chimeric endonuclease does not cut the C box when the first 15 bp of the C box is deleted demonstrate that the first 15 bp of the C box is sufficient for specific cleavage by the chimeric endonuclease.

Kinetics of DNA cleavage

The sequence specificity of BmrI198-C.BclI most likely derives from the DNA-binding HTH motif of C.BclI. The CH-endonuclease appears to possess a low turnover rate on interaction with target site. In fact, titration of the CH-endonuclease against DNA substrate shows that a ∼300-fold molar excess of the fusion protein over DNA substrate is required to make specific cleavage (Figure 3A). Therefore, the kinetics of the DNA cleavage reaction was studied in single-turnover condition. A time course for the cleavage reaction is shown in Figure 4. The intensities of the substrate (3.3 kb) and cleavage products (2.1 and 1.2 kb) were quantified after agarose gel electrophoresis and their quantities were estimated through correlation to that of the input substrate (DraIII only; Figure 4). The amount of the substrate and the sum of that of the cleavage products were plotted against time. In 120 min, ∼80% of the substrate was cleaved and the production of the product leveled off. The increase in the amount of the cleavage products correlated with the decrease in that of the substrate. The sum of the quantity of substrate and products added up to the quantity of input substrate at each time point (data not shown), suggesting that most of the substrate cleaved was transformed into the specific products and that non-specific cleavage activity is not significant within this time frame. Complete cleavage of the substrate was achieved with extended incubation times but at the expense of increasing non-specific cleavage (data not shown).

Ionic strength and magnesium ion requirement for cleavage specificity

Cleavage activity of the fusion enzyme was tested in buffers containing different concentrations of NaCl

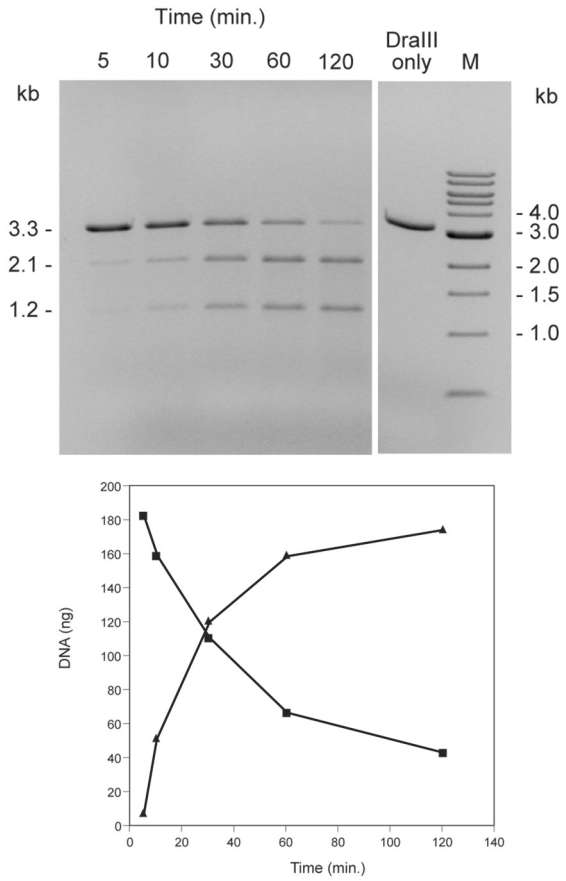

Figure 4. Time course of BmrI198-C.BclI cleavage. Litmus28-*bclIC* was linearized by DraIII and gel-purified. One hundred ninety nanograms (87 fmol) of the DNA were incubated with 24 pmol of BmrI198-C.BclI for 120 min at 37°C. The intensity of the 3.3 kb DraIII-linearized substrate and the 2.1 and 1.2 kb cleavage products were quantified and correlated to the quantity of input DNA (DraIII only). The quantity of the substrate (filled square) and the sum of those of the cleavage products (filled triangle) were plotted against time. The reactions were carried out under single-turnover condition.

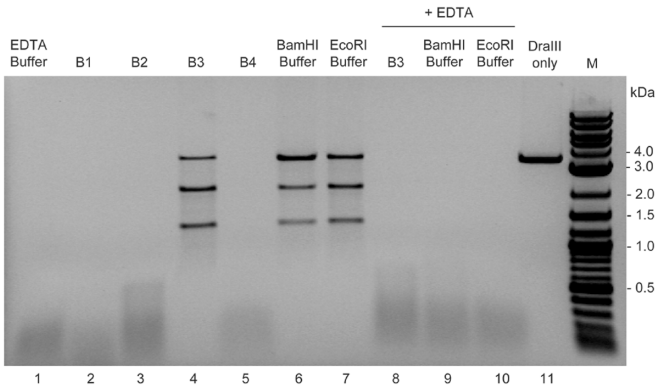

Figure 5. Effect of NaCl and EDTA on specific cleavage activity. Cleavage reactions were carried out using DraIII-linearized Litmus28-*bclIC* in different buffers (lanes 1–7) or in buffer 3, high salt (HS) buffer and EcoRI buffer quenched with 10 mM EDTA (lanes 8–10). B1, Buffer 1; B2, Buffer 2; B3, Buffer3; B4, Buffer 4.

and Tris–HCl (Figure 5). Smears of DNA with ~200 bp and smaller resulting from non-specific cleavage were found in Buffer 1, 2 and 4. Specific cleavage was only observed in Buffer 3, high salt buffer and EcoRI buffer. Buffer 1, 2 and 4 contain 50 mM or less NaCl, whereas Buffer 3, high salt buffer and EcoRI buffer contain 100 mM or higher concentration of chloride ion (NaCl or Tris–HCl). High ionic strength may provide an electrostatic screen for non-cognate hydrogen bonds and/or electrostatic interactions between the HTH DNA-binding motif of C.BclI and the substrate DNA. This is consistent with the buffer condition (10 mM Tris–HCl, 100 mM NaCl, 4 mM $CaCl_2$, 5% glycerol, pH 7.5) with which specific binding of C.BclI to C box was demonstrated (38). The presence of 0.025% Triton X-100 in EcoRI buffer did not increase non-specific activity, suggesting that hydrophobic interactions are not the primary interactions required for cleavage specificity.

Although BmrI and the truncation mutant BmrI198 do not require magnesium ions for cleavage (data not shown), the removal of magnesium ions from the cleavage reaction promotes non-specific cleavage of BmrI198-C.BclI. Quenching Buffer 3, high salt buffer and EcoRI buffer with 10 mM EDTA (Figure 5, lanes 8–10) or replacing 10 mM $MgCl_2$ with 10 mM EDTA in Buffer 3 (EDTA buffer; Figure 5, lane 1) resulted in high non-specific cleavage activity. It is possible that magnesium ions decrease the DNA cleavage activity of the BmrI cleavage domain, or they are required for the interactions between the HTH DNA-binding motif of C.BclI and the target DNA. The absence of divalent ions in the λ repressor-target DNA structure (PDB entry 1LMB) and the decrease in 'star' activity of wild-type BmrI in the presence of magnesium ions (unpublished data) support the notion that magnesium ions increase specific cleavage activity of the fusion protein by inhibiting DNA cleavage activity of the BmrI DNA cleavage domain.

Thermostability and reaction temperature

BmrI is isolated from a mesophilic bacterium (*Bacillus megaterium*). Its cleavage activity decreased ~60% after heating at 55°C for 20 min and was completely destroyed at 65°C (data not shown). Surprisingly, BmrI198-C.BclI exhibited higher thermal stability. Specific cleavage was impaired after pre-incubation at 64°C for 20 min (data not shown). The increased thermostability of the CH-endonuclease may be contributed by the removal of the C-terminal DNA recognition domain, which is probably more susceptible to irreversible thermal denaturation or due to the addition of a thermostable binding partner C.BclI. The BclI producing strain *Bacillus caldolyticus* is a moderately thermophilic strain with growth temperature up to 70°C (REBASE).

The optimal reaction temperature for the CH-endonuclease was determined to be 37–40°C. Non-specific cleavage activity increased above 40°C. Non-specific cleavage activity dominated at 49.2–65°C (data not shown).

DISCUSSION

BmrI and BfiI closely resemble each other in amino acid sequence. They have 358 amino acid residues and share 79.6% sequence identity (Figure 1A). They belong to the phospholipase D family that is characterized by the HKD catalytic motif. Two copies of the HKD motif fold to form a single catalytic site where one of the histidine residues forms the phosphohistidine intermediate and the lysine residues are required for the positioning of the phosphate group being attacked (41). Because one of the histidine residues also acts as a nucleophile, the HKD catalytic motif does not require divalent metal ions for hydrolysis. Unlike most of the phospholipase D family members that contains two copies of the HKD catalytic motif, BfiI and BmrI have only one copy. Presumably, dimerization is required for BfiI and BmrI to form a functional catalytic site at the dimerization interface. Structural and biochemical data suggest that homodimers of BfiI cut the bottom strand first and then undergo conformational rearrangement with respect to the nicked DNA intermediate to cleave the top strand (34,37). Top-strand cleavage activity of BfiI was inhibited at pH 6.5, thereby converting BfiI into a bottom-strand nicking endonuclease (NEase) (37). Experimental evidence suggests that low pH protonates the 5′ phosphate at the new nick in the bottom strand and inhibits the rearrangement and/or chemistry that is required for the sequential top-strand cleavage. Conformational rearrangement and sequential cleavage of the two strands of DNA may be utilized by other Type IIS REases. This makes Type IIS REases an attractive source of modular DNA cleavage domains for creating combinatorial endonucleases.

Here we have successfully created a sequence-specific endonuclease with novel substrate specificity by fusing a DNA recognition domain to the cleavage domain of BmrI. The primary hurdle to cross when creating a combinatorial endonuclease is to attain high specific cleavage and low non-specific cleavage. Factors that can contribute to non-specific cleavage include: (i) the affinity of DNA-binding domain toward target DNA sequence versus non-specific cleavage activity of the cleavage domain and (ii) the position of the cleavage domain relative to the DNA-binding domain. The DNA cleavage domain and the DNA-binding domain can be envisioned as counteracting in terms of specific cleavage: the DNA cleavage domain tends to capture random DNA sequences and make breaks, whereas the DNA-binding domain samples the whole DNA molecule for the target sequence. To make specific cuts, the DNA-binding domain has to find its target sequence before the cleavage domain cuts a random piece of DNA. For BmrI198-C.BclI, the presence of magnesium ions appears to attenuate non-specific cleavage activity of the fusion protein by decreasing the DNA cleavage activity of the BmrI DNA cleavage domain. In order to make cuts at specific sites, the cleavage domain of the chimeric endonuclease has to be in contact with the substrate DNA after the DNA-binding domain binds to the target sequence. In the current study, this was achieved by connecting the two domains with a flexible linker sequence consisting of Gly, Ala and Ser residues (GSGGGGSAAGASAS). The linker is expected to allow the BmrI cleavage domain to adapt a range of orientations such that it can cut the substrate after the C.BclI domain binds to the target sequence.

C.AhdI (PDB entry 1Y7Y), the controller protein of the AhdI R-M system, and C.BclI (PDB entry 2B5A) are highly homologous in 3D structure and their target C-box sequences are also similar. Although with lower sequence homology, C.BclI is also structurally similar to proteins involved in genetic switches, namely, repressor proteins from phage λ (42) and phage 434 (43), and BldD-N, the N-terminal DNA-binding domain of BldD from *Streptomyces coelicolor* repressor (44).

The C box of AhdI R-M system has been proposed to consist of four 5-bp binding blocks: two each in the 5′ half and 3′ half. A homodimer of C.AhdI binds to two binding blocks of either 5′ or 3′ half where each monomer interacts with one of the 5-bp blocks (Figure 6). It has been shown that C.AhdI binds the 5′ half with higher affinity than the 3′ half, and that binding of C.AhdI on the 5′ half induces cooperative binding at the 3′ half by another dimer of C.AhdI (45). Our results show that BclI198-C.BclI cleaves the wild-type C box and the 3′ deletion variant at the same sites, showing that the chimeric endonuclease does not bind to the 3′ half of the C box. This is consistent with the higher binding affinity of C.AhdI toward the 5′ half of its C box. The BmrI198-C.BclI homodimer or tetramer bound to the 5′ half of the C box may exclude another homodimer or tetramer binding to the 3′ half of the C box. This may explain the cleavage sites are predominantly located within the right half of the C box (+19↓+20) instead of being located outside of +24 position.

Based on the published results of C.AhdI and our BclI C-box deletion results, we propose a scheme of DNA binding and cleavage by BmrI198-C.BclI. The BclI C box consists of four binding blocks of AGACTT and its variant sequences TCACTT, AGGCTA (Figure 6, sequences indicated in orange, BLK1 to 4). A 6-bp binding block is considered in contrast to the 5-bp version of C.AhdI. Because BfiI forms stable homodimers *in vitro* (15) at their DNA cleavage domain (34), it is likely that the CH-endonuclease dimerizes before binding to the DNA substrate. Therefore, two homodimers of BmrI198-C.BclI molecules (pairs of homodimer in green or orange, (Figure 6) would bind to the binding block 1 and 2 of substrate DNA (boxed in green or orange lines, respectively) and form a tetramer at the dimerization interface of C.BclI (37,46). Each of the two DNA-bound BmrI198-C.BclI molecules would interact with one of the binding blocks via its HTH motif—the BmrI198-C.BclI subunit (color in green) that binds to BLK1 (in green box) would extend its DNA cleavage domain 13–15 bp downstream to the +20 cleavage site (indicated by green lines) whereas the other subunit (colored in orange) would bind to BLK2 (boxed in orange lines) with its DNA cleavage domain reaching 13–16 bp upstream onto the −8 site (indicated by orange lines) (Figure 6). Thus, a nick is made at the −8 and +20 site, respectively, through the dimeric BmrI cleavage domain at each site. At this point, similar to the mechanism proposed for BfiI, the nicked DNA

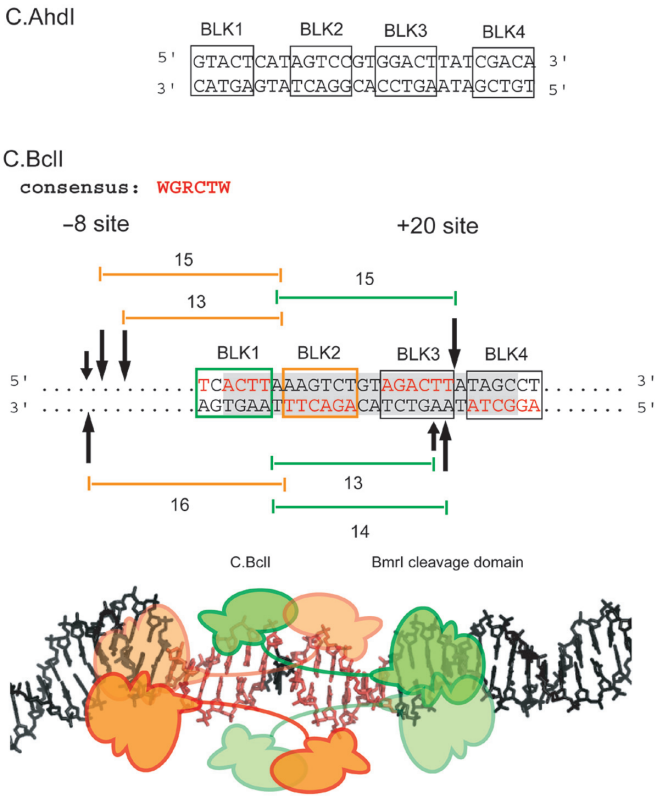

Figure 6. Proposed scheme of DNA binding and cleavage by BmrI198-C.BclI. Upper panel: the C box of AhdI system consists of four 5-bp-binding blocks. Dimers of C.AhdI interact with either BLK1 and BLK2, or BLK3 and BLK4. C.AhdI has a higher affinity towards BLK1 and BLK2, and binding of C.AhdI dimers to BLK1 and BLK2 induces synergetic binding of C.AhdI dimers to BLK3 and BLK4. Middle panel: in C.BclI, the C box can be subdivided into four 6-bp blocks by the conserved and symmetric sequences 5′ WGARTW 3′ (W = A or T; R = A or G; colored in red). Cleavage sites (−8 and +20 sites) depicted in Figure 2C are shown in terms of distance (bp) from BLK 1 and BLK 2. We propose that one of the BmrI198–C.BclI homodimers binds to BLK1 (green box) and cuts 13–15 bp downstream (distance indicated by green lines) and the other homodimer binds to BLK2 (orange box) and cuts 13–16 bp downstream (indicated by orange lines). The C-box sequence is shown in gray box. Lower panel: a cartoon of the proposed scheme of binding. The homodimer of BmrI198–C.BclI that binds BLK1 and cuts at the +20 site is colored in green; the homodimer that binds BLK2 and cuts at the −8 site is colored in orange. Molecules in dark green or orange are in front side of the DNA; those in light green or orange are behind the DNA. The BmrI intrinsic linker plus the 14-amino acid linker connecting the BmrI cleavage domains to C.BclI-binding domains are shown as solid lines. DNA sequences of BLK1 and BLK2 are colored in red.

substrate (37), which is likely to have adopted a different conformation, would induce a conformational change of the DNA-bound chimeric endonuclease, possibly affected by the intrinsic flexible linker within BmrI198 plus the 14-amino acid engineered linker (thin lines connecting the BmrI cleavage domain and C.BclI). This would allow the cleavage domain of the dimeric DNA-bound chimeric endonuclease to make contacts with the opposite strand and repeat the cleavage there. Because the DNA cleavage domain would have to traverse a longer distance around the DNA helix to reach the opposite strand, the cuts at the opposite strand are closer to the binding site. The flexible linker may adopt different conformations when spanning across the ds DNA such that alternative cleavage sites 1–3 bp apart are generated. The enzyme cleaves 1 bp further from the binding site at the −8 site (13–16 bp) than the +20 site (13–15 bp). This is probably caused by bending of the DNA substrate upon binding of the fusion endonuclease, which is often observed in protein–DNA interactions, particularly with those involved in gene regulations (47–49).

Our work demonstrates that the BmrI cleavage domain is useful in the combinatorial approach to creating novel endonucleases that cleave specifically in the vicinity of DNA regulatory region. A fusion of C.BclI to the DNA cleavage domain of BmrI introduces specific ds breaks at the regulatory sequence of the BclI R–M system (C box). Potentially, the BmrI DNA cleavage domain can be fused to other DNA-binding proteins such as transcription factors to cut specifically at the regulatory elements of gene expression. Such cleavage has been demonstrated to induce DNA repair resulting in deletion of genes *in vivo* (28–31). We have also fused the BmrI nuclease domain to a cleavage-deficient variant of NotI such that the fusion protein binds and cleaves outside of NotI site, thus creating a NotI neoschizomer in a related study (Zhang *et al.*, in press, PEDS).

Type IIS REases are useful to create combinatorial endonucleases because they consist of a DNA-binding domain that contributes to their substrate specificity and a separate DNA cleavage domain, which cuts within or outside the recognition sequences. A well-known example making use of this modular property is FokI. The DNA cleavage domain of FokI (GGATC N_9/N_{13}) has been linked to several DNA-binding proteins (18,19) and zinc-finger motifs (24–27) to create endonucleases of novel substrate specificities. In particular, a few ZFNs have been created that possess specific cleavage activity *in vitro* (20–23) and *in vivo* (28–31,50,51). The nuclease domain of BmrI itself can also be used to generate small fragments of DNA (data not shown).

The current version of CH-endonuclease remains to be optimized. The BmrI cleavage domain itself does not fold properly when expressed in *E. coli* and requires refolding. In addition, the presence of two dimerization interfaces, one on the DNA-binding domain and the other on the cleavage domain, may lead to oligomerization that sequesters the catalytic site and attenuates specific activity of the chimeric endonuclease. This may account for the high molar ratio of BmrI198-C.BclI over substrate DNA required for specific cleavage (340:1 for the time course in Figure 4, with 67% of the substrate being cleaved in 1 h). Although complete cleavage of substrate DNA can be achieved by higher concentrations of BmrI198-C.BclI or longer incubation time, these reaction conditions also result in significant non-specific cleavage. Creating a single-chain endonuclease by fusing two identical DNA cleavage domains in tandem can avoid the problems arise from the presence of two different dimerization interfaces. A single-chain variant of PvuII has been created by linking two monomers of the wild-type PvuII through a 4-amino acid linker sequence such that a single molecule of this single-chain PvuII can generate ds breaks

without resorting to dimerization (52). The length and flexibility of the linker sequence connecting the BmrI nuclease domain and the DNA-binding protein may be further optimized for efficient and precise cleavage.

The strategy to couple a DNA-binding protein to a nuclease domain can also be applied to the construction of site-specific nicking endonucleases, for example, by fusing a DNA-nicking domain from I-HmuI to sequence-specific DNA-binding proteins.

ACKNOWLEDGEMENTS

We thank I. Murray and J. Samuelson for critical reading of the manuscript and G.K. Balendiran for discussion. This work was partly supported by an SBIR grant to S.Y - X (1 R43 GM073345-01, NIH). We also thank Z. Zhu for providing the plasmid Litmus28-*bclIC* and Don Comb for support. Funding to pay the Open Access publication charges for this article was provided by New England Biolabs, Inc.

Conflict of interest statement. None declared.

REFERENCES

1. Lanio,T., Jeltsch,A. and Pingoud,A. (1998) Towards the design of rare cutting restriction endonucleases: using directed evolution to generate variants of EcoRV differing in their substrate specificity by two orders of magnitude. *J. Mol. Biol.*, **283**, 59–69.
2. Samuelson,J.C. and Xu,S.Y. (2002) Directed evolution of restriction endonuclease BstYI to achieve increased substrate specificity. *J. Mol. Biol.*, **319**, 673–683.
3. Zhu,Z., Zhou,J., Friedman,A.M. and Xu,S.Y. (2003) Isolation of BsoBI restriction endonuclease variants with altered substrate specificity. *J. Mol. Biol.*, **330**, 359–372.
4. Samuelson,J.C., Morgan,R.D., Benner,J.S., Claus,T.E., Packard,S.L. and Xu,S.Y. (2006) Engineering a rare-cutting restriction enzyme: genetic screening and selection of NotI variants. *Nucleic Acids Res.*, **34**, 796–805.
5. Ashworth,J., Havranek,J.J., Duarte,C.M., Sussman,D., Monnat,R.J., Jr, Stoddard,B.L. and Baker,D. (2006) Computational redesign of endonuclease DNA binding and cleavage specificity. *Nature*, **441**, 656–659.
6. Chevalier,B.S., Kortemme,T., Chadsey,M.S., Baker,D., Monnat,R.J. and Stoddard,B.L. (2002) Design, activity, and structure of a highly specific artificial endonuclease. *Mol. Cell.*, **10**, 895–905.
7. Epinat,J.C., Arnould,S., Chames,P., Rochaix,P., Desfontaines,D., Puzin,C., Patin,A., Zanghellini,A., Paques,F. *et al.* (2003) A novel engineered meganuclease induces homologous recombination in yeast and mammalian cells. *Nucleic Acids Res.*, **31**, 2952–2962.
8. Chames,P., Epinat,J.C., Guillier,S., Patin,A., Lacroix,E. and Paques,F. (2005) In vivo selection of engineered homing endonucleases using double-strand break induced homologous recombination. *Nucleic Acids Res.*, **33**, e178.
9. Perez,C., Guyot,V., Cabaniols,J.P., Gouble,A., Micheaux,B., Smith,J., Leduc,S., Paques,F. and Duchateau,P. (2005) Factors affecting double-strand break-induced homologous recombination in mammalian cells. *Biotechniques.*, **39**, 109–115.
10. Smith,J., Grizot,S., Arnould,S., Duclert,A., Epinat,J.C., Chames,P., Prieto,J., Redondo,P., Blanco,F.J. *et al.* (2006) A combinatorial approach to create artificial homing endonucleases cleaving chosen sequences. *Nucleic Acids Res.*, **34**, e149.
11. Gouble,A., Smith,J., Bruneau,S., Perez,C., Guyot,V., Cabaniols,J.P., Leduc,S., Fiette,L., Ave,P. *et al.* (2006) Efficient *in toto* targeted recombination in mouse liver by meganuclease-induced double-strand break. *J. Gene Med.*, **8**, 616–622.
12. Arnould,S., Chames,P., Perez,C., Lacroix,E., Duclert,A., Epinat,J.C., Stricher,F., Petit,A.S., Patin,A. *et al.* (2006) Engineering of large numbers of highly specific homing endonucleases that induce recombination on novel DNA targets. *J. Mol. Biol.*, **355**, 443–458.
13. Arnould,S., Perez,C., Cabaniols,J.P., Smith,J., Gouble,A., Grizot,S., Epinat,J.C., Duclert,A., Duchateau,P. *et al.* (2007) Engineered I-CreI derivatives cleaving sequences from the human XPC gene can induce highly efficient gene correction in mammalian cells. *J. Mol. Biol.*, **371**, 49–65.
14. Choi,Y.-J., Kim,S.-J., Hwang,H.-Y., Yim,J. and Kim,J.-C. (1994) Characterization of restriction endonuclease EagBI from *Enterobacter agglomerans* CBNU45. *Korean J. Microbiol.*, **32**, 91–95.
15. Zaremba,M., Urbanke,C., Halford,S.E. and Siksnys,V. (2004) Generation of the BfiI restriction endonuclease from the fusion of a DNA recognition domain to a non-specific nuclease from the phospholipase D superfamily. *J. Mol. Biol.*, **336**, 81–92.
16. Eisenschmidt,K., Lanio,T., Simoncsits,A., Jeltsch,A., Pingoud,V., Wende,W. and Pingoud,A. (2005) Developing a programmed restriction endonuclease for highly specific DNA cleavage. *Nucleic Acids Res.*, **33**, 7039–7047.
17. Kim,Y. and Chandrasegaran,S. (1994) Chimeric restriction endonuclease. *Proc. Natl Acad. Sci. USA*, **91**, 883–887.
18. Kim,Y.-G., Kim,P.S., Herbert,A. and Rich,A. (1997) Construction of a Z-DNA-specific restriction endonuclease. *Proc. Natl Acad. Sci. USA*, **94**, 12875–12879.
19. Kim,J.J., Koh,S., Kim,J.S. and Lee,D.-S. (1998) Mode of action on EcoRI restriction endonuclease: EcoRI and EcoRI variant N199H have active monomeric forms. *J. Biochem. Mol. Biol.*, **31**, 149–155.
20. Miller,J.C., Holmes,M.C., Wang,J., Guschin,D.Y., Lee,Y.L., Rupniewski,I., Beausejour,C.M., Waite,A.J., Wang,N.S., *et al.* (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat Biotechnol.* **25**, 778–785.
21. Mani,M., Kandavelou,K., Dy,F.J., Durai,S. and Chandrasegaran,S. (2005) Design, engineering, and characterization of zinc finger nucleases. *Biochem. Biophys. Res. Commun.*, **335**, 447–457.
22. Kandavelou,K., Mani,M., Durai,S. and Chandrasegaran,S. (2005) "Magic" scissors for genome surgery. *Nat. Biotechnol.*, **23**, 686–687.
23. Huang,B., Schaeffer,C.J., Li,Q. and Tsai,M.-D. (1996) A new class IIS zinc finger restriction enzyme with specificity for SP1 binding sites. *J. Protein Chem.*, **15**, 481–489.
24. Kim,Y.-G., Shi,Y., Berg,M. and Chandrasegaran,S. (1997) Site-specific cleavage of DNA-RNA hybrids by zinc finger/FokI cleavage domain fusions. *Gene*, **203**, 43–49.
25. Kaneko,T., Sato,S., Kotani,H., Tanaka,A., Asamizu,E., Nakamura,Y., Miyajima,N., Hirosawa,M., Sugiura,M. *et al.* (1996) Sequence analysis of the genome of the unicellular cyanobacterium Synechocystis sp. strain PCC6803. II. Sequence determination of the entire genome and assignment of potential protein-coding regions. *DNA Res.*, **3**, 109–136.
26. Smith,J., Bibikova,M., Whitby,F.G., Reddy,A.R., Chandrasegaran,S. and Carroll,D. (2000) Requirements for double-strand cleavage by chimeric restriction enzymes with zinc finger DNA-recognition domains. *Nucleic Acids Res.*, **28**, 3361–3369.
27. Smith,J., Berg,J.M. and Chandrasegaran,S. (1999) A detailed study of the substrate specificity of a chimeric restriction enzyme. *Nucleic Acids Res.*, **27**, 674–681.
28. Bibikova,M., Golic,M., Golic,K.G. and Carroll,D. (2002) Targeted chromosomal cleavage and mutagenesis in Drosophila using zinc-finger nucleases. *Genetics*, **161**, 1169–1175.
29. Bibikova,M., Beumer,K., Trautman,J.K. and Carroll,D. (2003) Enhancing gene targeting with designed zinc finger nucleases. *Science*, **300**, 764.
30. Beumer,K., Bhattacharyya,G., Bibikova,M., Trautman,J.K. and Carroll,D. (2006) Efficient gene targeting in Drosophila with zinc-finger nucleases. *Genetics*, **172**, 2391–2403.
31. Bibikova,M., Carroll,D., Segal,D.J., Trautman,J.K., Smith,J., Kim,Y.G. and Chandrasegaran,S. (2001) Stimulation of homologous recombination through targeted cleavage by chimeric nucleases. *Mol. Cell. Biol.*, **21**, 289–297.
32. Sapranauskas,R., Sasnauskas,G., Lagunavicius,A., Vilkaitis,G., Lubys,A. and Siksnys,V. (2000) Novel subtype of type IIs

restriction enzymes. BfiI endonuclease exhibits similarities to the EDTA-resistant nuclease Nuc of *Salmonella typhimurium*. *J. Biol. Chem.*, **275**, 30878–30885.
33. Leiros,I., Secundo,F., Zambonelli,C., Servi,S. and Hough,E. (2000) The first crystal structure of a phospholipase D. *Structure*, **8**, 655–667.
34. Grazulis,S., Manakova,E., Roessle,M., Bochtler,M., Tamulaitiene,G., Huber,R. and Siksnys,V. (2005) Structure of the metal-independent restriction enzyme BfiI reveals fusion of a specific DNA-binding domain with a nonspecific nuclease. *Proc. Natl Acad. Sci. USA*, **102**, 15797–15802.
35. Wah,D.A., Bitinaite,J., Schildkraut,I. and Aggarwal,A.K. (1998) Structure of FokI has implications for DNA cleavage. *Proc. Natl Acad. Sci. USA*, **95**, 10564–10569.
36. Bitinaite,J., Wah,D.A., Aggarwal,A.K. and Schildkraut,I. (1998) FokI dimerization is required for DNA cleavage. *Proc. Natl Acad. Sci. USA*, **95**, 10570–10575.
37. Sasnauskas,G., Halford,S.E. and Siksnys,V. (2003) How the BfiI restriction enzyme uses one active site to cut two DNA strands. *Proc. Natl Acad. Sci. USA*, **100**, 6410–6415.
38. Sawaya,M.R., Zhu,Z., Mersha,F., Chan,S.H., Dabur,R., Xu,S.Y. and Balendiran,G.K. (2005) Crystal structure of the restriction-modification system control element C.BclI and mapping of its binding site. *Structure*, **13**, 1837–1847.
39. Kong,H., Morgan,R.D., Maunus,R.E. and Schildkraut,I. (1993) A unique restriction endonuclease, BcgI, from *Bacillus coagulans*. *Nucleic Acids Res.*, **21**, 987–991.
40. Kong,H., Roemer,S.E., Waite-Rees,P.A., Benner,J.S., Wilson,G.G. and Nwankwo,D.O. (1994) Characterization of BcgI, a new kind of restriction-modification system. *J. Biol. Chem.*, **269**, 683–690.
41. Leiros,I., McSweeney,S. and Hough,E. (2004) The reaction mechanism of phospholipase D from *Streptomyces* sp. strain PMF. Snapshots along the reaction pathway reveal a pentacoordinate reaction intermediate and an unexpected final product. *J. Mol. Biol.*, **339**, 805–820.
42. Beamer,L.J. and Pabo,C.O. (1992) Refined 1.8 A crystal structure of the lambda repressor-operator complex. *J. Mol. Biol.*, **227**, 177–196.
43. Mondragon,A. and Harrison,S.C. (1991) The phage 434 Cro/OR1 complex at 2.5 A resolution. *J. Mol. Biol.*, **219**, 321–334.
44. Kim,I.K., Lee,C.J., Kim,M.K., Kim,J.M., Kim,J.H., Yim,H.S., Cha,S.S. and Kang,S.O. (2006) Crystal structure of the DNA-binding domain of BldD, a central regulator of aerial mycelium formation in *Streptomyces coelicolor* A3(2). *Mol. Microbiol.*, **60**, 1179–1193.
45. McGeehan,J.E., Papapanagiotou,I., Streeter,S.D. and Kneale,G.G. (2006) Cooperative binding of the C.AhdI controller protein to the C/R promoter and its role in endonuclease gene expression. *J. Mol. Biol.*, **358**, 523–531.
46. Lagunavicius,A., Sasnauskas,G., Halford,S.E. and Siksnys,V. (2003) The metal-independent type IIs restriction enzyme BfiI is a dimer that binds two DNA sites but has only one catalytic centre. *J. Mol. Biol.*, **326**, 1051–1064.
47. Kuznetsov,S.V., Sugimura,S., Vivas,P., Crothers,D.M. and Ansari,A. (2006) Direct observation of DNA bending/unbending kinetics in complex with DNA-bending protein IHF. *Proc. Natl Acad. Sci. USA*, **103**, 18515–18520.
48. Sugimura,S. and Crothers,D.M. (2006) Stepwise binding and bending of DNA by *Escherichia coli* integration host factor. *Proc. Natl Acad. Sci. USA*, **103**, 18510–18514.
49. Koudelka,G.B. (1991) Bending of synthetic bacteriophage 434 operators by bacteriophage 434 proteins. *Nucleic Acids Res.*, **19**, 4115–4119.
50. Miller,J.C., Holmes,M.C., Wang,J., Guschin,D.Y., Lee,Y.L., Rupniewski,I., Beausejour,C.M., Waite,A.J., Wang,N.S. *et al.* (2007) An improved zinc-finger nuclease architecture for highly specific genome editing. *Nat. Biotechnol.*, **25**, 778–785.
51. Szczepek,M., Brondani,V., Buchel,J., Serrano,L., Segal,D.J. and Cathomen,T. (2007) Structure-based redesign of the dimerization interface reduces the toxicity of zinc-finger nucleases. *Nat. Biotechnol.*, **25**, 786–793.
52. Simoncsits,A., Tjornhammar,M.L., Rasko,T., Kiss,A. and Pongor,S. (2001) Covalent joining of the subunits of a homodimeric type II restriction endonuclease: single-chain PvuII endonuclease. *J. Mol. Biol.*, **309**, 89–97.