Published in final edited form as:
Protein Expr Purif. 2010 February ; 69(2): 226–234. doi:10.1016/j.pep.2009.09.003.

Engineering BspQI Nicking Enzymes and Application of N.BspQI in DNA Labeling and Production of Single-strand DNA

Penghua Zhang[1], **Priscilla Hiu-Mei Too**[1], **James C. Samuelson**[1], **Siu-Hong Chan**[1], **Tamas Vincze**[1], **Stephanie Doucette**[1], **Stefan Bäckström**[2], **Konstantinos D. Potamousis**[3], **Timothy M. Schramm**[3], **Dan Forrest**[3], **David C. Schwartz**[3], and **Shuang-yong Xu**[1,*]

[1]New England Biolabs, Inc. 240 County Road, Ipswich, MA 01938

[2]Göteborg University, Department of Medical Biochemistry and Cell Biology, Medicinaregatan 9A, Box 440, SE-40530, Göteborg, Sweden

[3]Laboratory for Molecular and Computational Genomics, Department of Chemistry, Laboratory for Genetics, University of Wisconsin-Madison Biotechnology Center, University of Wisconsin-Madison, 425 Henry Mall, Madison, Wisconsin 53706

Abstract

BspQI is a thermostable Type IIS restriction endonuclease (REase) with the recognition sequence 5' GCTCTTC N1/N4 3'. Here we report the cloning and expression of the *bspQIR* gene for the BspQI restriction enzyme in *E. coli*. Alanine scanning of the BspQI charged residues identified a number of DNA nicking variants. After sampling combinations of different amino acid substitutions, an Nt.BspQI triple mutant (E172A/E248A/E255K) was constructed with predominantly top-strand DNA nicking activity. Furthermore, a triple mutant of BspQI (Nb.BspQI, N235A/K331A/R428A) was engineered to create a bottom-strand nicking enzyme. In addition, we demonstrated the application of Nt.BspQI in optical mapping of single DNA molecules. Nt or Nb.BspQI-nicked dsDNA can be further digested by *E. coli* exonuclease III to create ssDNA for downstream applications. BspQI contains two potential catalytic sites: a top-strand catalytic site (Ct) with a D-H-N-K motif found in the HNH endonuclease family and a bottom-strand catalytic site (Cb) with three scattered Glu residues. BlastP analysis of proteins in Genbank indicated a putative restriction enzyme with significant amino acid sequence identity to BspQI from the sequenced bacterial genome *Croceibacter atlanticus* HTCC2559. This restriction gene was amplified by PCR and cloned into a T7 expression vector. Restriction mapping and run-off DNA sequencing of digested products from the partially purified enzyme indicated that it is an EarI isoschizomer with 6-bp recognition, which we named CatHI (CTCTTC N1/N4).

Keywords

alanine scanning; nicking endonuclease; CatHI; DNA labeling; optical mapping of DNA

*Corresponding author Phone: 978-380-7287, Fax: 978-921-1350, xus@neb.com.

Individual contributions: P. Zhang expressed wt BspQI and engineered Nt.BspQI and Nb.BspQI; P. Too purified Nb.BspQI, performed nicking assays and produced circular ssDNA by exonuclease III; Siu-Hong Chan cloned and purified CatHI; J. Samuelson constructed a BspQI genomic DNA libraries and performed the methylase selection procedure and "endo-blue" screening; T. Vincze analyzed the 454 sequence contigs; S. Doucette purified wt BspQI; S.Bäckström independently tested Nb.BspQI/exonuclease III on different plasmid constructs. K.D. Potamousis, T.M. Schramm, D. Forrest, and D.C. Schwartz at University of Wisconsin-Madison performed the single molecule labeling and optical mapping and wrote the nanocoding section of the paper. SYX contributed to experimental design and wrote the manuscript.

Introduction

Restriction endonucleases (REases) with 4-8 bp recognition sequences are indispensible tools in creating recombinant DNA molecules [1]. Nearly 300 unique Type II specificities have been found so far from bacterial and viral sources. DNA nicking endonucleases (NEases) which bind DNA sequence specifically and cleave only one strand, however, are only available in small numbers. The first two natural NEases, Nt.CviPII and Nt.CviQII, were found from infected cells of Chlorella viruses [2,3]. Other natural NEases are Nt.BstNBI/Nt.BspD6I/N.BstSEI, Nb.BsrDI, and Nb.BtsI, which are the large subunits of their respective REases [4-7] (G.Wilson, unpublished results). NEases have also been engineered from heterodimeric REases BbvCI and Bpu10I by inactivation of the top-strand or bottom-strand catalytic site [8] (Janulaitis A et al. (2005) Strand-specific polynucleotide nickases, US patent number 6,867,028). A number of NEases were derived from protein engineering of Type IIS REases. For example, nicking variants have been engineered from AlwI, BsmAI, BsmBI, BsaI, BsmI, FokI, MlyI, and SapI either by site-directed mutagenesis of a few amino acid residues or by protein domain swapping [9-13] (Nb.BsmI, Z.Zhu and SYX, unpublished results). In addition, nicking variants have been constructed from homing endonucleases with large recognition sequences useful for gene targeting [14,15].

NEases are useful to prepare pre-nicked DNA substrates for DNA repair study [16]. In addition, NEases have been employed in DNA amplification in conjunction with a DNA polymerase with strand-displacement activity. For example, the isothermal exponential amplification reaction (EXPAR) exponentially amplifies short oligonucleotides called “triggers” by primer extension and regeneration through DNA polymerase and NEase activities [17,18]. Nt.CviPII can be used in random DNA amplification (NEMDA, nicking endonuclease mediated DNA amplification) without addition of any primers since DNA can be directly amplified from partially nicked duplexes [19]. Target-specific DNA amplification can also be achieved by a set of specific primers, NEases, single-strand DNA binding protein such as gp32, and a DNA polymerase with strand displacement activity (R.Kucera and E. Raleigh, unpublished results). Another example of NEases in DNA-based diagnostic application is termed nicking endonuclease signal amplification (NESA) [20]. A fluorescent probe and target DNA anneal to create a recognition site for a strand-specific NEase. The NEase cleaves the probe into two pieces while leaving the target intact. Fluorescence coupled with capillary electrophoresis can be used to measure the nicked products.

NEases can also be used to label DNA using the nicked sites for nick-translation with fluorescent-labeled dNTPs and a suitable DNA polymerase. The labeled DNA molecules can be physically stretched out by passing through nanochannels or nanoslit devices and the fluorescent “dots” on single DNA molecules are visualized by fluorescent imaging and optical mapping [21,22] (BioNanomatrix platform technology: www.bionanomatrix.com). The distance of the fluorescent “dots” can be measured to reveal particular kind of genetic rearrangement (deletion or insertion). For DNA nicking and single DNA molecule labeling and detection, it would be useful to have a collection of strand-specific NEases with 5-7 bp specificities. Here we report the successful engineering of strand-specific and sequence-specific BspQI nicking enzymes and the application of Nt.BspQI in DNA nicking and single molecule labeling.

During the cloning and expression of the BspQI R-M system, we also found a putative R-M system from the sequenced bacterial genome *Croceibacter atlanticus* HTCC2559 that shows significant amino acid sequence identity to BspQI and SapI. We report here the characterization of this R-M system.

Materials and Methods

Bacterial strains, plasmid DNA, restriction and modification enzymes

T7 Express (ER2566) was used as the expression strain (New England Biolabs, NEB). BspQI nicking enzymes were expressed in M1.EarI and M2.EarI-modified host ER2566 [pLG339-*earIM1M2*] (EarI recognition sequence 5'CTCTTC3' overlaps with the BspQI site, the plasmid carries Km^R selection marker and pSC101 origin). Site-directed mutagenesis was carried out by inverse PCR using either Vent DNA polymerase or Phusion DNA polymerase (NEB). The T7 expression vector pET21a was purchased from Novagen. Restriction and modification enzymes, pUC19 and pBR322 were from NEB. Plasmid pACYC-T7-ter is a low-copy number expression vector with a T7 promoter and a transcription terminator. IPTG-induced cell extracts were prepared as described [23]. One unit of BspQI REase is defined as the amount of enzyme to completely digest 1 μg of pUC19 DNA in 1 h at 50°C in NEB buffer 3. One unit of Nt.BspQI is defined as the amount of enzyme to completely nick pUC19 (covalently-closed circular form) into nicked circular DNA in 1 h at 50°C in NEB buffer 3.

Construction of genomic DNA libraries and shot-gun sequencing

Bacillus sphaericus genomic DNA was partially digested with BfuCI (5'/GATC3'). Genomic DNA fragments in the range of 1.5 to 10 kb were gel-purified and ligated into pBR322 (BamHI/CIP), pUC19 (BamHI/CIP) or pACYC-T7-ter (BamHI/CIP). Ligated DNA was transferred into *E. coli* ER2683 or ER1992 competent cells (endo-blue indicator strain) by transformation. Plasmid pUC19 or pBR322 with genomic DNA insert library was plated on LB agar plates plus Amp (100 μg/ml) and X-gal (80 μg/ml). Genomic DNA fragments inserted in plasmid pACYC-T7-ter was selected on LB agar plates plus Cm (30 μg/ml) and X-gal (80 μg/ml). Blue colonies were screened for possible clones with *bspQIR* gene insert. The inserts in the dark blue colonies were amplified in PCR by forward and reverse universal primers from pUC19 derivatives or T7 forward and reverser primers from pACYC-T7-ter derivatives. The PCR DNA was sequenced by using the BigDye terminator cycle sequencing kit (Applied Biosystems). In the methylase selection procedure, primary genomic DNA library (pBR322 plus genomic inserts) was challenged by 10-fold over-digestion with concentrated SapI and the digested plasmid DNA was used to transform *E. coli* ER2683 competent cells. Transformants were plated on LB agar plates supplemented with Amp. Plasmid DNA was extracted from overnight culture of individual tranformants and analyzed for resistance to SapI digestion.

Bacillus sphaericus genomic DNA was shot-gun sequenced using the 454 method (454 Inc.) [24]. Mutant *bspQIR* alleles were confirmed by DNA sequencing (Applied Biosystems).

Protein purification of BspQI and BspQI nicking enzymes

T7 Express strains (ER2566) containing the *bspQIR* or *Nt.bspQIR* gene were protected in vivo by the EarI M1 & M2 methyltransferases. Cells were grown to OD_{600} 0.5 and protein production was induced by addition of IPTG to a final concentration of 0.5 mM for 3-4 hours. Harvested cell pellets were resuspended in a buffer (20 mM Tris-HCl, pH 7.5, 100 mM NaCl, 0.1 mM EDTA, 10 mM β–mercaptoethanol) and lysed by sonication. The clarified lysate was loaded onto a Heparin Hyper D column (Pall Corp.). Protein was eluted by applying a NaCl gradient (0.05 -1.0 M). Active fractions were collected and diluted to a final salt concentration of 100 mM. The diluted pool was then passed through a Source TM15Q column (GE Healthcare). The flow-through was collected and loaded onto a Heparin-TSK column (Tosoh Bioscience). The fractions with enzyme activity were collected and dialyzed into a storage buffer (10 mM Tris-HCl, pH 7.4, 300 mM NaCl, 0.1 mM EDTA, 1 mM DTT and 50% glycerol). Nb.BspQI was expressed from a pUC19 derivative (pUC19-*Nb.bspQIR*) in a pre-modified host

and partially purified by chromatography through Heparin Sepharose and SP Sepharose HiTrap columns (GE Healthcare).

Site-directed mutagenesis and molecular biology techniques

Mutagenic primers with desired mutations were used in inverse PCR mutagenesis. Typically, 20-25 cycles of inverse PCR were performed using Vent DNA polymerase or Phusion DNA polymerase (NEB). Following DpnI digestion of the template DNA, the amplified DNA was transferred into M.EarI-modified *E. coli* host by transformation or electroporation. Plasmid extraction kits from Qiagen were used to prepare plasmid DNA.

DNA labeling and single molecule imaging

A DNA labeling procedure and two-color fluorescence imaging followed a previously described protocol [21] with modifications including provisions for mounting labeled molecules on optical mapping surfaces, in place of nanoslits [25]. As such, dialysis steps were obviated for reducing solution ionic strength. T7 DNA (Yorkshire Biosciences Ltd., UK) embedded in gel inserts was nicked with Nt.BspQI (2.5 units) in NEB buffer 3, prior to nick translation with *E. coli* DNA polymerase I (5 units, endonuclease free grade, Roche Applied Sciences) and one fluorochrome label (Alexa Fluor 647-aha-dUTP, Invitrogen) in the reaction mix. Nicked, labeled DNA molecules were then electro-eluted into TE buffer (10 mM Tris-HCl, 1 mM EDTA, pH 8.0) for imaging on optical mapping surfaces after staining with YOYO-I (Invitrogen).

Single-strand DNA preparation

Supercoiled pUC19 DNA was first treated with Nt or Nb.BspQI at 50°C in buffer 3. The nicked DNA was purified by a spin column (Qiagen) and then digested with *E. coli* exonuclease III (NEB, 200 U in buffer 1). DNA products were analyzed on agarose gels. T7 DNA polymerase or Bst DNA polymerase, dNTP, random primers or universal primers, and the ssDNA template were used to synthesize the complementary strand.

CatHI cloning and protein purification

The CatHI open reading frame (ORF, *catHIR* gene, GenBank accession ZP_00949034) was PCR-amplified from genomic DNA of *Croceibacter atlanticus* HTCC2559 (a gift from S.J. Giovannoni of Oregon State University, Corvallis, USA) in PCR [26]. The amplified fragment was ligated to pAII17 at NdeI and BamHI sites and the ligated DNA was used to transform *E. coli* T7 Express expressing EarI methylases. The cloned *catHIR* insert was re-sequenced to confirm the wt sequence. The CatHI expression strain was cultured in 1 L of LB medium supplemented with Ampicillin (100 μg/ml) and Kanamycin (50 μg/ml) at 30°C. IPTG was added to 0.25 mM at mid-log phase and the culture was allowed to grow for 12 h at 25°C. The cells were harvested and stored at -20°C. The cell pellet was resuspended using Buffer A (20 mM potassium phosphate, pH 6.5, 50 mM NaCl, 1 mM EDTA) and then subjected to sonication. After centrifugation, the supernatant was loaded onto a HiTrap SP HP column (5 ml bed volume). Protein was eluted from the column by a linear gradient of 50 mM to 1 M NaCl in Buffer A. Peak fractions displaying specific cleavage activity were pooled and concentrated. To determine the recognition site, 1 μg of plasmid pBC4 (NEB) was cleaved by the partially purified CatHI in a reaction containing NEB Buffer 4 and 0.1 mg/ml BSA at 25° C for 1 h. The cleavage pattern was compared to pBC4 digested by EarI or BspQI. To determine the cleavage site, 0.5 μg of pUC19 was cleaved by CatHI or EarI. The 1.8 kb cleavage product from each reaction was gel-purified and subjected to run-off sequencing.

Results and Discussion

Cloning and expression of *bspQIR* gene for the BspQI REase

The methylase selection method [27] has been widely used in cloning restriction-modification (R-M) systems in heterologous hosts. Expression of a methylase gene cloned into a plasmid with appropriate cognate target sites will render the plasmid resistant to digestion by the cognate REase. Therefore, methylase-positive clones can be selected from a plasmid DNA library by restriction digestion and retransformation. However, this method failed for the BspQI methylase gene: no true positive clones were identified from plasmid DNA libraries after challenge with isoschizomer SapI (data not shown). A second strategy to clone *bspQIR* gene was to use the *dinD::lacZ* indicator strain (in vivo SOS induction, the "endo-blue" method) to screen blue colonies that may carry nuclease genes that induce the SOS DNA repair response [28,29]. In such a screening system, the DNA damage inducible promoter for the *dinD* locus is fused to the *lacZ* gene. Induced expression of *dinD* leads to elevated expression level of β-galactosidase, thus forming blue colonies on X-gal plates [28-30]. When the BspQI genomic DNA fragments ligated into pUC19 and pACYC-T7-ter were transferred into the *dinD::lacZ* indicator strain, we characterized a number of dark blue clones which encoded DNA metabolic enzymes rather than the BspQI endonuclease gene. The DNA inserts encoded putative DNA transposases, a 3'-5' exonuclease domain of a DNA polymerase, and a DNA invertase/integrase (data not shown).

The third strategy used shot-gun sequencing of the entire genome. *Bacillus sphaericus* genomic DNA was sequenced using the 454 method by 454 Inc. (data not shown). One sequence contig contained the BspQI R-M system based on its extensive sequence similarity to SapI R-M genes. Sanger sequencing reads were used to correct a few frame shift errors and confirm the final sequence. The gene organization of the BspQI R-M system and the amino acid sequence alignment of BspQI/SapI/TdeII are shown in Figure 1A and 1B. The BspQI R-M system consists of four genes, *bspQIR*, *bspQIM1*, *bspQIM2*, and *C.bspQI* (C for controller protein). BspQI shows 62% amino acid sequence similarity and 44% sequence identity to SapI REase (5'GCTCTTC3') [31]. The BspQI sequence also shows 46% amino acid similarity and 32% identity to TdeII REase which recognizes a less-specific 6-bp DNA target (5'CTCTTC3') [1] (H, Kong, unpublished results). In addition, BspQI also shares amino acid sequence similarity to two hypothetical proteins from sequenced bacterial genomes. The C-terminus of the hypothetical protein ABOONEI_2170 from *Aciduliprofundum boonei* T469 is very similar to BspQI/SapI/TdeII. It is probably a TdeII-like isoschizomer with a 6-bp recognition sequence since both proteins are similar in size (344-aa vs. 330-aa). Similarity is also found with another hypothetical protein CA2559_00420 (CatHI, 484-aa residues) from *Croceibacter atlanticus* HTCC2559. CatHI was shown to be an isoschizomer of EarI and TdeII (see below). The C-terminal region of BspQI also aligns to a number of putative homing endonucleases that contain the active site motif HNH or HNK (data not shown).

M1.BspQI and M2.BspQI also share significant amino acid sequence identity to M1.SapI and M2.SapI (65% and 64% sequence identity, respectively). In addition, there is a gene encoding a putative transcription regulator protein (C.BspQI) upstream of BspQI endonuclease gene. C.BspQI has 50% amino acid sequence identity to the C.BclI protein of the BclI R-M system [32]. Thus, it may be involved in transcription regulation of the *bspQIR* and *bspQIM1 & M2* genes.

The BspQI REase was over-expressed in the expression strain ER2566 [pLG339-*earIM1M2*, pET21a-*bspQIR*] and the expression yield was 3.3×10^5 units per gram of wet cells following IPTG induction. M1.EarI & M2.EarI modify the target site 5'CTCTTC3', which overlaps with BspQI site GCTCTTC (methylase cross-protection). BspQI was purified to homogeneity by

column chromatography and analyzed by SDS-PAGE (Figure 2A). The specific activity of BspQI was estimated to be 1.1×10^5 units/mg protein.

Site-directed mutagenesis of *bspQIR* gene (Alanine scanning)

The Alanine scanning method [33] was applied to 120 positions within BspQI in order to isolate nicking variants and mutants with altered specificity. A second objective was to determine which residues are important for catalysis. Most of the amino acid substitutions were targeted to C, D, E, H, K, R, N, and Q residues. Cys was chosen as the mutagenesis target because of four predicted zinc fingers in BspQI endonuclease. The amino acid residues D, E, H, K, and Q were selected for mutagenesis because these residues have been proved or implicated as catalytic residues in other REases [5,23,34-37]. The R and N residues may be involved in DNA sequence recognition [36]. Several nicking endonucleases have been successfully engineered from Type IIS REases [9,11,13] or Type IIT REases [5,8] (Strand-specific polynucleotide nickases, US patent number 6,867,028, (2005)). One of our goals was to isolate rare nicking enzymes with 7-bp recognition sequence. When pUC19 with a single BspQI site was used as a substrate in DNA nicking assays, we discovered a number of BspQI nicking variants by detecting the accumulation of nicked circular DNA after incubation with cell extracts. In most cases, both linear and nicked circular DNA bands were detected (data not shown). The nicked DNA species was purified from an agarose gel and subjected to run-off sequencing to determine the strand specificity of each BspQI nicking variant. The following BspQI variants predominantly nicked the top strand: E172A, R254A, and E255K (data not shown). A SapI nicking variant E250K at the corresponding position to BspQI-E255K has been isolated before as a top-strand nicking enzyme [13]. A variety of double and triple mutants were constructed (data not shown) and the best top-strand nicking enzyme was found to be Nt.BspQI E172A/E248A/E255K (E248A eliminated a low level of dsDNA cleavage activity). This nicking enzyme was purified to near homogeneity (Figure 2A) and its nicking activity is shown in Figure 2B. A low level of dsDNA cleavage activity was detected at greater than 32-fold over-digestion (32 units incubated with 1 μg pUC19 DNA substrate for 1 hr). The specific nicking activity is 3×10^5 units/mg protein using pUC19 as the substrate in a reaction conducted for 1 h at 50°C. Following Ala-scanning mutagenesis, the following BspQI variants were found to preferentially nick the bottom strand: E169A, K191A, H214A, Q222A, E224A, K331A, R428A (Table 1). But all the mutants displayed some dsDNA cleavage activity. The K331A and R428A mutations were combined into a double mutant Nb.BspQI (K331A/R428A), which displayed bottom-strand nicking activity and partial dsDNA cleavage activity on pUC19 (data not shown). A third amino acid substitution was introduced to generate a triple mutant Nb.BspQI (N235A/K331A/R428A) in an attempt to further reduce dsDNA cleavage activity. This mutant enzyme was partially purified by chromatography and the nicking activity was tested on pUC19 DNA. Double-strand cleavage activity was not detected at 2-fold over-digestion (Figure 2C). At 10-fold over-digestion with Nb.BspQI, less than 5% of DNA was cleaved on both strands (see below). The bottom-strand nicking specificity was confirmed by run-off sequencing (data not shown). Nb.BspQI (N235A/K331A/R428A) shows a low specific activity (~5,000 units/mg protein). No extensive dsDNA cleavage activity was observed when λ DNA was incubated with the partially purified nicking enzymes Nt.BspQI or Nb.BspQI at less than 5-fold over-digestion (data not shown).

BspQI variants with altered specificity (increased or decreased star activity) will be reported elsewhere.

Nanocoding

Nicking endonucleases, such as Nt.BspQI, may be employed to generate high-resolution restriction maps in ways closely paralleling the use of common type II restriction enzymes for optical mapping of DNA molecules [25]; however, additional steps are required to visualize

their cleavage sites since products are target molecules bearing distinct patterns of nicks and not discrete restriction fragments. The nicked positions are readily identified by optical mapping approaches as labeled gaps that are imaged along fluorochrome stained parental molecules that are stretched on charged glass surfaces [25,38]. As such, nicked sites on DNA molecules are revealed by fluorescence microscopy after the incorporation of fluorochrome labeled nucleotides (Alexa-647-aha-NTPs by nick translation [21]. In this way, the nick sites are contrasted against entire molecules, which are also uniformly stained with an intercalating fluorochrome (YOYO-1). Remarkably, this combination of covalently incorporated label and intercalated fluorochrome supports efficient FRET (fluorescence resonance energy transfer) imaging techniques that offer extreme sensitivity, require a single laser excitation source, and virtually eliminate signals from unincorporated labels. An example of Nt.BspQI nick site detection on a T7 phage DNA molecule is shown in Figure 3. Here, nick translation of a T7 phage DNA molecule cleaved by Nt.BspQI reveals nicked sites as red punctates positioned along a molecular backbone stained with YOYO-1 (green). This labeling method points the way to stretching schemes employing nanoconfinement approaches. Consider that within such nanoscale environments, enzymatic action is difficult to accomplish and control reliably. Although optical mapping, utilizing common Type II restriction enzymes and charged surfaces, produces high- resolution maps of entire genomes [25,39], molecular confinement approaches combined with nicking and labeling [21] offer increased densities of presented molecules for detection techniques capable of very high throughput data acquisition.

Making ssDNA by N.BspQI nicking and exonuclease digestion

There is a single BspQI site in pUC19, which can be nicked by Nt or Nb.BspQI. The open-circular DNA was then subjected to *E. coli* exonuclease III digestion to remove the nicked strand. Figure 4 shows that after Nt.BspQI/exonuclease III or Nb.BspQI/exonuclease treatment, the dsDNA can be converted into ssDNA. The EtBr binding to ssDNA is much weaker than to the dsDNA or nicked open circular DNA. Thus, the intensity of the band is consistent with the above conclusion. The strand-specific nicking by Nb.BspQI and exonuclease III digestion was also applied to a different plasmid construct and ssDNA was successfully obtained (SB, data not shown). In addition, the ssDNA can serve as a template for DNA extension by addition of random primers (dN6), or sequence-specific primers, plus dNTP, and T7 DNA polymerase (data not shown). When Bst DNA polymerase with strand displacement activity was used in the primer extension reaction, a large amount of DNA was amplified from the ssDNA template (isothermal extension/amplification, data not shown). When 5mC dCTP or 5mHC dCTP is added in the extension mix, it is possible to make strand-specific hemi-methylated DNA. It should be feasible to uniformly label DNA with the incorporation of fluorescently labeled dUTP during primer extension.

BspQI catalytic site mapping

The engineering of both top-strand and bottom-strand specific nicking enzymes Nt.BspQI and Nb.BspQI suggests the existence of one catalytic site for cleavage of each strand. E172, E175 and E255 may define the bottom-strand catalytic center (Cb) since mutation of each of these three residues resulted in a nicking variant with top-strand nicking preference. K331 and R428 may constitute the top-strand catalytic center (Ct) along with H306 and N328 as BspQI mutants K331A and R428A are nicking enzymes that prefer to nick the bottom strand. BspQI mutants H306A and N328A possessed little catalytic activity (~1% of wt activity), but retained a low nicking activity (data not shown). H306, N328, and K331 in BspQI align with the HNH (or HNK) catalytic sites found in many HNH-superfamily endonucleases. Two negatively charged residues D305 and D345 flanking the HNK motif are also critical for BspQI catalytic activity. BspQI mutants D305A and D345A showed 1% and 0.1% of wt activity, respectively (data not shown). This catalytic site consisting of D305, H306, N328, and K331 is reminiscent of the catalytic sites of I-HmuI (D74, H75, N96, H56) and T4 endonuclease VII (D40, H41, H43,

N62, H105) [40,41]. These four critical residues in BspQI (D305, H306, N328, and K331) are highly conserved in SapI, TdeII, and CatHI endonucleases (see Fig. 1B). Amino acid sequence analysis and motif mining indicates that 24 out of 289 Type II REases carry the HNH or HNK motif [42]. However, only a few Type II REases with the HNH catalytic motif have been confirmed by site-directed mutagenesis or structural studies [43-45].

Mutagenesis of the Cys-x-x-Cys motif

There are four Cys-x-x-Cys motifs ($C_{192}xxC_{195}$, $C_{225}xxC_{228}$, $C_{290}xxC_{293}$, $C_{324}xxC_{327}$) in the BspQI amino acid sequence, which may constitute two zinc-binding sites. The putative zinc binding sites in SapI endonuclease, a BspQI isoschizomer, had been predicted previously [46]. To determine the importance of these Cys residues, each was mutated to Ala and the resulting mutants were assayed for endonuclease activity. BspQI mutants C192A, C195A, C225A, C290A, C293A, C324A, C327A displayed 0.1% to 3.5% of the wt cleavage activity (Table 1). The C228A variant displayed 12% cleavage activity relative to wt BspQI. The protein expression level of BspQI mutants C225A and C327A was much reduced when cell extracts were analyzed on SDS-PAGE in comparison to the wt enzyme (data not shown), suggesting that residues Cys225 and Cys327 are critical for protein folding and stability. The putative zinc finger $C_{324}xxC_{327}$ is located in close proximity to the catalytic residues (D305, H306, N328, and K331), whose structural role is presumed to aid in folding and positioning of the catalytic center near the DNA. The remaining Cys to Ala mutants have comparable protein expression levels when compared to the wt enzyme (data not shown). There are also two Cys residues (Cys193 and Cys322) outside of the Cys-x-x-Cys motifs. When Cys193 was mutated to Ala, the resulting mutant displayed relaxed specificity (enhanced star activity). Variant C322A displayed activity comparable to the wt enzyme. The crystal structures of BspQI and BspQI nicking variants are currently being investigated (B. Shen and B. Stoddard, personal communication). We expect that the structures will shed light on the role of these Cys residues in metal ion binding and protein folding. Among the restriction enzymes with multiple Cys residues, only three enzymes, BslI, Hpy99I, and KpnI, have been shown experimentally to bind zinc metal ions [45,47,48]. The two zinc metal ions in the KpnI endonuclease seem to play dual roles in both structural folding and specificity [48]. In the NotI:DNA cocrystal structure, however, four Cys residues coordinate the binding of an iron atom in a unique metal binding fold thought to be involved in protein structural stability [49]. Mutation of these individual Cys residues abolishes NotI restriction activity (PZ and SYX, unpublished results).

Expression of the CatHI restriction enzyme in *E. coli*

The *catHIR* gene was cloned and expressed in *E. coli* under the T7 promoter. CatHI was partially purified and used to digest a number of DNA substrates such as pUC19, pBR322, pBC4, and phage λ DNA (data not shown). The DNA banding pattern of pBC4 cleaved by CatHI was identical to EarI digested products (data not shown). To determine the cleavage site, pUC19 DNA was cleaved by CatHI and EarI respectively. The 1.8 kb cleavage product from each reaction was gel-purified and subjected to Sanger sequencing. Figure 5 shows the run-off sequencing results on both strands. The cut site is identical to that of EarI and therefore it was concluded that like TdeII, CatHI is a bona fide EarI isoschizomer. Unlike EarI, the optimal temperature of CatHI is 22°C to 30°C. It shows approximately 50% activity at 37°C.

6 bp (CTCTTC) vs 7 bp (GCTCTTC) recognition

Amino acid sequence alignment in Figure 1B shows that the sequences at the N-terminus of the five proteins are more divergent than that at the C-terminus which would be consistent with their involvement in DNA sequence recognition. Thus, the recognition of the extra base by BspQI and SapI may be mediated by the amino acids within the N-terminus (BspQI residues 1-221; SapI residues 1-217). This prediction may be tested by constructing chimeric enzymes

consisting of the BspQI N-terminal region and the C-terminal region of TdeII or CatHI. The chimeric constructs can be transformed into a *dinD::lacZ* indicator strain to screen blue colonies to identify active chimeric enzymes indicative of restriction activity.

Acknowledgments

We thank Geoff Wilson, Bill Jack, Rich Roberts and Barry Stoddard for critical comments and helpful discussions; Laurie Mazzola, Elizabeth Cantin, and Barton Slatko for DNA sequencing; NEB Organic Synthesis for providing oligos; Don Comb and Jim Ellard for support and encouragement. This work was partially supported by a phase II STTR grant (5R42HG3976-3) from NHGRI to SYX, and by funding from NHGRI (R01 HG000225-13) to DCS, and TS was supported by NHGRI training grant to the Genomic Sciences Training Program (5T32HG002760).

References

1. Roberts R, Vincze T, Posfai J, Macelis D. REBASE--enzymes and genes for DNA restriction and modification. Nucl Acids Res 2007;35:269–270. [PubMed: 17164287]
2. Xia Y, Morgan R, Schildkraut I, Van Etten JL. A site-specific single strand endonuclease activity induced by NYs-1 virus infection of a Chlorella-like green alga. Nucleic Acids Res 1988;16:9477–9487. [PubMed: 3186439]
3. Zhang Y, Nelson M, Nietfeldt J, Xia Y, Burbank D, Ropp S, Van Etten JL. Chlorella virus NY-2A encodes at least 12 DNA endonuclease/methyltransferase genes. Virology 1998;240:366–375. [PubMed: 9454710]
4. Morgan RD, Calvet C, Demeter M, Agra R, Kong H. Characterization of the specific DNA nicking activity of restriction endonuclease N.BstNBI. Biol Chem 2000;381:1123–1125. [PubMed: 11154070]
5. Xu S-Y, Zhu Z, Zhang P, Chan S-H, Samuelson JC, Xiao J, Ingalls D, Wilson GG. Discovery of natural nicking endonucleases Nb.BsrDI and Nb.BtsI and engineering of top-strand nicking variants from BsrDI and BtsI. Nucleic Acids Res 2007;35:4608–4618. [PubMed: 17586812]
6. Yunusova A, Rogulin E, Artyukh R, Zheleznaya L, Matvienko N. Nickase and a protein encoded by an open reading frame downstream from the nickase BspD6I gene form a restriction endonuclease complex. Biochemistry (Mosc) 2007;71:815–820. [PubMed: 16903837]
7. Golobova NS, Okhapkina SS, Abdurashitov MA, Degtiarev S. Determination and analysis of the primary structure of NM.BstSEI operone from Bacillus stearothermophilus SE-589 which produces N.BstSEI site-specific nickase. Mol Biol (Mosk) 2005;39:960–964. [PubMed: 16358732]
8. Heiter DF, Lunnen KD, Wilson GG. Site-specific DNA-nicking mutants of the heterodimeric restriction endonuclease R.BbvCI. J Mol Biol 2005;348:631–640. [PubMed: 15826660]
9. Xu Y, Lunnen KD, Kong H. Engineering a nicking endonuclease N.AlwI by domain swapping. Proc Natl Acad Sci U S A 2001;98:12990–12995. [PubMed: 11687651]
10. Besnier CE, Kong H. Converting MlyI endonuclease into a nicking enzyme by changing its oligomerization state. EMBO Rep 2001;2:782–786. [PubMed: 11520857]
11. Zhu Z, Samuelson JC, Zhou J, Dore A, Xu S-Y. Engineering strand-specific DNA nicking enzymes from the type IIS restriction endonucleases BsaI, BsmBI, and BsmAI. J Mol Biol 2004;337:573–583. [PubMed: 15019778]
12. Sanders KL, Catto LE, Bellamy SR, Halford SE. Targeting individual subunits of the FokI restriction endonuclease to specific DNA strands. Nucleic Acids Res 2009;37:2105–2115. [PubMed: 19223323]
13. Samuelson JC, Zhu Z, Xu S-Y. The isolation of strand-specific nicking endonucleases from a randomized SapI expression library. Nucleic Acids Res 2004;32:3661–3671. [PubMed: 15247348]
14. Niu Y, Tenney K, Li H, Gimble FS. Engineering variants of the I-SceI homing endonuclease with strand-specific and site-specific DNA-nicking activity. J Mol Biol 2008;382:188–202. [PubMed: 18644379]
15. McConnell Smith A, Takeuchi R, Pellenz S, Davis L, Maizels N, Monnat RJ Jr, Stoddard BL. Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. Proc Natl Acad Sci U S A 2009;106:5099–5104. [PubMed: 19276110]

16. Wang H, Hays JB. Mismatch repair in human nuclear extracts. Quantitative analyses of excision of nicked circular mismatched DNA substrates, constructed by a new technique employing synthetic oligonucleotides. J Biol Chem 2002;277:26136–26142. [PubMed: 12006560]
17. Van Ness J, Van Ness LK, Galas DJ. Isothermal reactions for the amplification of oligonucleotides. Proc Natl Acad Sci U S A 2003;100:4504–4509. [PubMed: 12679520]
18. Tan E, Erwin B, Dames S, Ferguson T, Buechel M, Irvine B, Voelkerding K, Niemz A. Specific versus nonspecific isothermal DNA amplification through thermophilic polymerase and nicking enzyme activities. Biochemistry 2008;47:9987–9999. [PubMed: 18729381]
19. Chan S-H, Zhu Z, Van Etten JL, Xu S-Y. Cloning of CviPII nicking and modification system from chlorella virus NYs-1 and application of Nt.CviPII in random DNA amplification. Nucleic Acids Res 2004;32:6187–6199. [PubMed: 15570069]
20. Kiesling T, Cox K, Davidson EA, Dretchen K, Grater G, Hibbard S, Lasken RS, Leshin J, Skowronski E, Danielsen M. Sequence specific detection of DNA using nicking endonuclease signal amplification (NESA). Nucleic Acids Res 2007;35:e117. [PubMed: 17827214]
21. Jo K, Dhingra DM, Odijk T, de Pablo JJ, Graham MD, Runnheim R, Forrest D, Schwartz DC. A single-molecule barcoding system using nanoslits for DNA analysis. Proc Natl Acad Sci U S A 2007;104:2673–2678. [PubMed: 17296933]
22. Xiao M, Phong A, Ha C, Chan TF, Cai D, Leung L, Wan E, Kistler AL, DeRisi JL, Selvin PR, Kwok PY. Rapid DNA mapping by fluorescent single molecule detection. Nucleic Acids Res 2007;35:e16. [PubMed: 17175538]
23. Xu S-Y, Schildkraut I. Isolation of BamHI variants with reduced cleavage activities. J Biol Chem 1991;266:4425–4429. [PubMed: 1999426]
24. Rothberg JM, Leamon JH. The development and impact of 454 sequencing. Nat Biotechnol 2008;26:1117–1124. [PubMed: 18846085]
25. Zhou S, Bechner MC, Place M, Churas CP, Pape L, Leong SA, Runnheim R, Forrest DK, Goldstein S, Livny M, Schwartz DC. Validation of rice genome sequence by optical mapping. BMC Genomics 2007;8:278. [PubMed: 17697381]
26. Cho JC, Giovannoni SJ. Croceibacter atlanticus gen. nov., sp. nov., a novel marine bacterium in the family Flavobacteriaceae. Syst Appl Microbiol 2003;26:76–83. [PubMed: 12747413]
27. Szomolanyi I, Kiss A, Venetianer P. Cloning the modification methylase gene of Bacillus sphaericus R in Escherichia coli. Gene 1980;10:219–225. [PubMed: 6254840]
28. Fomenkov A, Xiao J-P, Dila D, Raleigh E, Xu S-Y. The ‘endo-blue method’ for direct cloning of restriction endonuclease genes in E. coli. Nucleic Acids Res 1994;22:2399–2403. [PubMed: 8036170]
29. Heitman J, Model P. SOS induction as an in vivo assay of enzyme-DNA interactions. Gene 1991;103:1–9. [PubMed: 1908806]
30. Kenyon CJ, Walker GC. DNA-damaging agents stimulate gene expression at specific loci in Escherichia coli. Proc Natl Acad Sci U S A 1980;77:2819–2823. [PubMed: 6771759]
31. Xu S-Y, Xiao J-p, Ettwiller L, Holden M, Aliotta J, Poh CL, Dalton M, Robinson DP, Petronzio TR, Moran L, Ganatra M, Ware J, Slatko B, Benner J. Cloning and expression of the ApaLI, NspI, NspHI, SacI, ScaI, and SapI restriction-modification systems in Escherichia coli. Mol Gen Genet 1998;260:226–231. [PubMed: 9862476]
32. Sawaya MR, Zhu Z, Mersha F, Chan S-H, Dabur R, Xu S-Y, Balendiran GK. Crystal structure of the restriction-modification system control element C.Bcll and mapping of its binding site. Structure 2005;13:1837–1847. [PubMed: 16338412]
33. Brunelle A, Schleif RF. Missing contact probing of DNA-protein interactions. Proc Natl Acad Sci U S A 1987;84:6673–6676. [PubMed: 2958845]
34. Wright DJ, King K, Modrich P. The negative charge of Glu-111 is required to activate the cleavage center of EcoRI endonuclease. J Biol Chem 1989;264:11816–11821. [PubMed: 2745418]
35. Jeltsch A, Alves J, Maass G, Pingoud A. On the catalytic mechanism of EcoRI and EcoRV. A detailed proposal based on biochemical results, structural data and molecular modelling. FEBS Lett 1992;304:4–8. [PubMed: 1618296]
36. Aggarwal AK. Structure and function of restriction endonucleases. Curr Opin Struct Biol 1995;5:11–19. [PubMed: 7773740]

37. Vander Woerd M, Pelletier J, Xu S-Y, Friedman A. Restriction enzyme BsoBI-DNA complex: a tunnel for recognition of degerate DNA sequences and potential histidine catalysis. Structure 2001;9:133–144. [PubMed: 11250198]

38. Dimalanta ET, Lim A, Runnheim R, Lamers C, Churas C, Forrest DK, de Pablo JJ, Graham MD, Coppersmith SN, Goldstein S, Schwartz DC. A microfluidic system for large DNA molecule arrays. Anal Chem 2004;76:5293–5301. [PubMed: 15362885]

39. Reslewic S, Zhou S, Place M, Zhang Y, Briska A, Goldstein S, Churas C, Runnheim R, Forrest D, Lim A, Lapidus A, Han CS, Roberts GP, Schwartz DC. Whole-genome shotgun optical mapping of Rhodospirillum rubrum. Appl Environ Microbiol 2005;71:5511–5522. [PubMed: 16151144]

40. Shen BW, Landthaler M, Shub DA, Stoddard BL. DNA binding and cleavage by the HNH homing endonuclease I-HmuI. J Mol Biol 2004;342:43–56. [PubMed: 15313606]

41. Giraud-Panis MJ, Lilley DM. T4 endonuclease VII. Importance of a histidine-aspartate cluster within the zinc-binding domain. J Biol Chem 1996;271:33148–33155. [PubMed: 8955164]

42. Orlowski J, Bujnicki JM. Structural and evolutionary classification of Type II restriction enzymes based on theoretical and experimental analyses. Nucleic Acids Res 2008;36:3552–3569. [PubMed: 18456708]

43. Saravanan M, Bujnicki JM, Cymerman IA, Rao DN, Nagaraja V. Type II restriction endonuclease R.KpnI is a member of the HNH nuclease superfamily. Nucleic Acids Res 2004;32:6129–6135. [PubMed: 15562004]

44. Kriukiene E, Lubiene J, Lagunavicius A, Lubys A. MnlI--The member of H-N-H subtype of Type IIS restriction endonucleases. Biochim Biophys Acta 2005;1751:194–204. [PubMed: 16024301]

45. Sokolowska M, Czapinska H, Bochtler M. Crystal structure of the beta beta alpha-Me type II restriction endonuclease Hpy99I with target DNA. Nucleic Acids Res 2009;37:3799–3810. [PubMed: 19380375]

46. Bujnicki JM, Radlinska M, Rychlewski L. Polyphyletic evolution of type II restriction enzymes revisited: two independent sources of second-hand folds revealed. Trends Biochem Sci 2001;26:9–11. [PubMed: 11165501]

47. Vanamee ES, Hsieh P, Zhu Z, Yates D, Garman E, Xu S-Y, Aggarwal AK. Glucocorticoid receptor-like Zn(Cys)4 motifs in BslI restriction endonuclease. J Mol Biol 2003;334:595–603. [PubMed: 14623197]

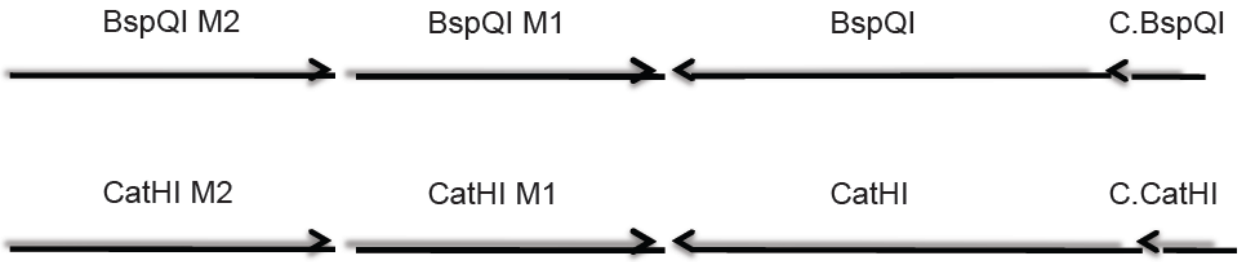
48. Saravanan M, Vasu K, Ghosh S, Nagaraja V. Dual role for Zn2+ in maintaining structural integrity and inducing DNA sequence specificity in a promiscuous endonuclease. J Biol Chem 2007;282:32320–32326. [PubMed: 17785455]

49. Lambert AR, Sussman D, Shen B, Maunus R, Nix J, Samuelson J, Xu S-Y, Stoddard BL. Structures of the rare-cutting restriction endonuclease NotI reveal a unique metal binding fold involved in DNA binding. Structure 2008;16:558–569. [PubMed: 18400177]

Abbreviations

R-M	restriction-modification system
REase	restriction endonuclease
NEase	nicking endonuclease
Nt	top-strand nicking
Nb	bottom-strand nicking

A



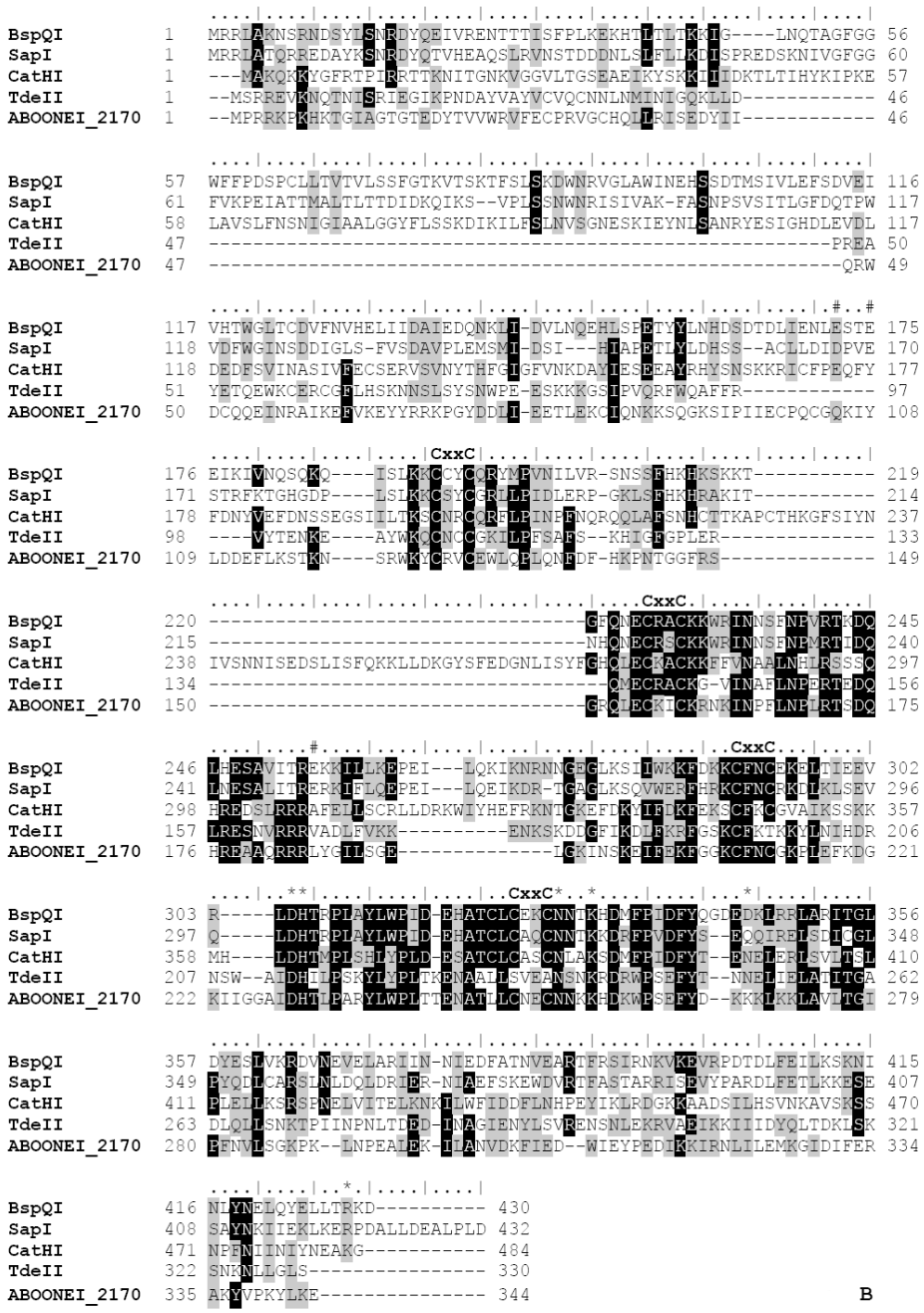


Figure 1.

A. Orientation and relative position of the genes in the BspQI and CatHI R-M systems.

B. Amino acid sequence alignment of BspQI/SapI (5'GCTCTTC3'), TdeII/CatHI (5'CTCTTC3') REases, and one putative endonuclease (ABOONEI_2170) from the sequenced bacterial genome *Aciduliprofundum boonei* T469. The Genbank accession numbers for CatHI (CA2559_00420) and ABOONEI_2170 are ZP_00949034 and EDY36183, respectively. The alignment was made using the BioEdit software-ClustalW Multiple Alignment with the similarity threshold set at 60%. Conserved amino acid residues are shaded. The four putative zinc finger motifs (Cys-x-x-Cys) are shown above the sequence alignment. The BspQI putative top-strand catalytic residues are indicated by a "*" on top of the aligned sequences. The putative

bottom-strand catalytic residues are marked by a "#". The *bspQIR* gene sequence has been deposited in Genbank and assigned the accession number FJ458442.

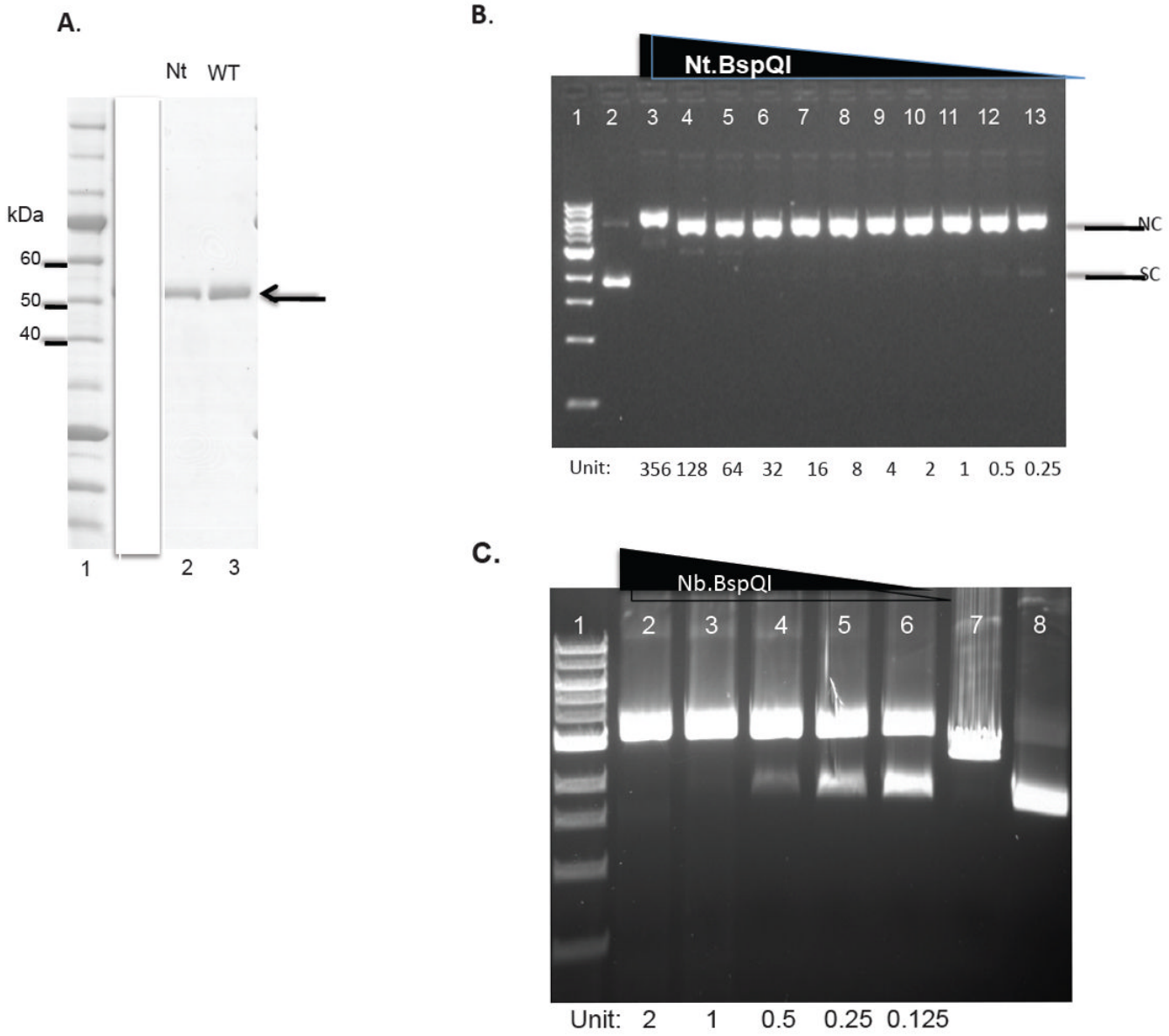


Figure 2.

A. Analysis of the purified Nt.BspQI and wild-type BspQI by SDS-PAGE. Lane 1, protein ladder (NEB). The predicted molecular mass of BspQI is 50.3 kDa.

B. DNA nicking activity assay of Nt.BspQI. Lane 1, 1 kb DNA size marker; lane 2, undigested pUC19 DNA; lane 3-13, pUC19 digested by Nt.BspQI. In this assay, 1 Nt.BspQI enzyme unit is defined as the amount of enzyme required for complete nicking of 1 μg of pUC19 at 50°C for 1 h in buffer 3 (100 mM NaCl, 50 mM Tris-HCl, pH 7.9, 10 mM $MgCl_2$, 1 mM DTT). SC, supercoiled DNA; NC, nicked circular DNA.

C. DNA nicking activity assay of Nb.BspQI. Lane 1, 1 kb DNA size marker; lane 2-6, pUC19 digested (nicked) by the partially purified Nb.BspQI; lane 7, linearized pUC19 by BamHI; lane 8, uncut pUC19.

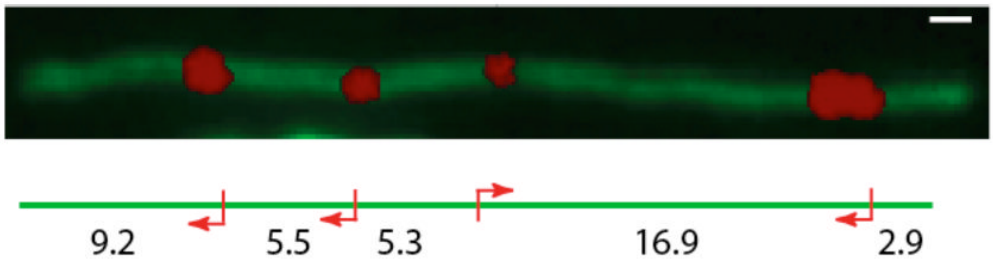


Figure 3.

T7 bacteriophage DNA molecule, nicked by Nt.BspQI and fluorochrome labeled. Two-color fluorescence micrograph showing a T7 bacteriophage DNA molecule stained with YOYO-1 (green) with Nt.BspQI nick sites labeled by nick translation using a mix containing Alexa Fluor 647-aha-dUTP (red punctates). Punctates are revealed by FRET, two-color, imaging techniques using YOYO-1 excitation (green channel) supporting energy transfer to covalently incorporated Alexa fluorophore, which fluoresce for imaging in the red channel. Below, an Nt.BspQI restriction map of T7 with fragment sizes noted. Nick sites are indicated by red lines/arrows indicating strand location and direction of nick translational action. The bar represents 1 micron.

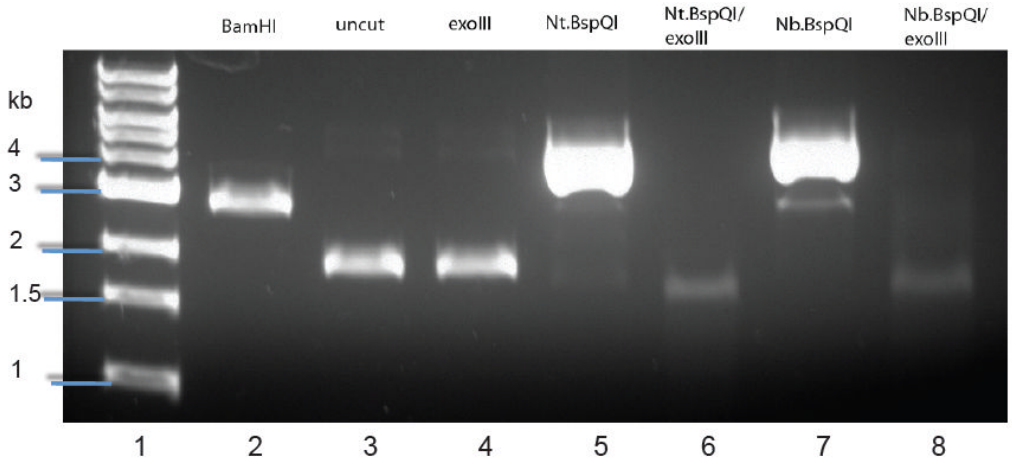


Figure 4.

Preparation of ssDNA by Nt and Nb.BspQI and exonuclease III digestion.

Lane 1, 1 kb DNA ladder; lane 2, BamHI-digested pUC19 (linear); lane 3, pUC19; lane 4, pUC19 treated with *E. coli* exonuclease III (exoIII); lanes 5 and 7, pUC19 digested with Nt.BspQI or Nb.BspQI (nicked DNA); lanes 6 and 8, pUC19 digested with Nt.BspQI/ exonuclease III or Nb.BspQI/exonuclease III (ssDNA).

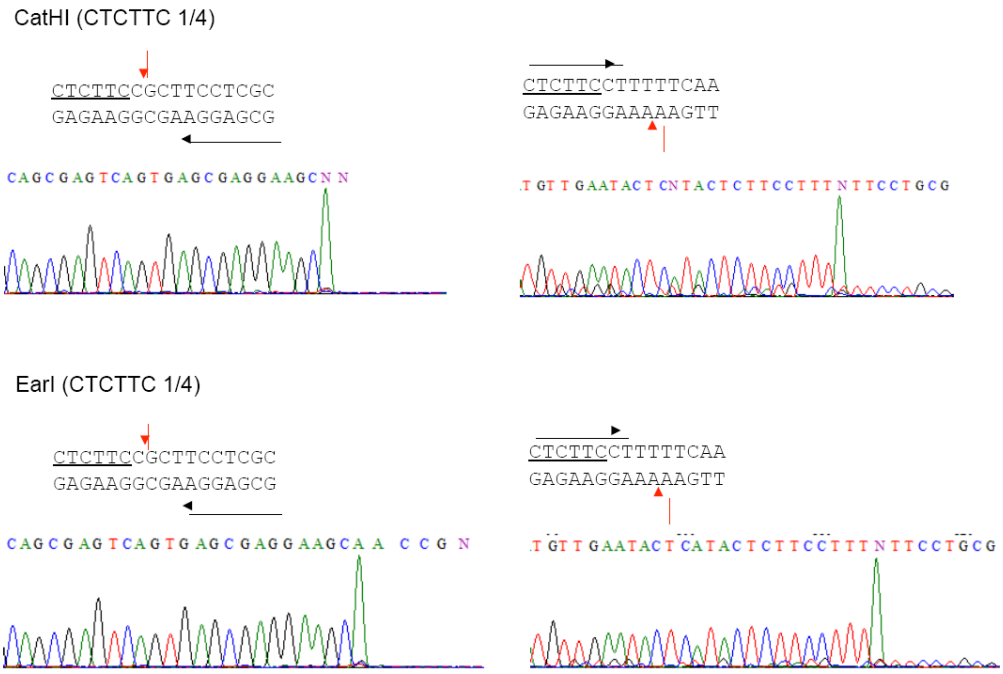


Figure 5.

Run-off sequencing results of CatHI- or EarI-digested pUC19 DNA.

The 1.8 kb CatHI or EarI-digested DNA was gel-purified and subjected to run-off sequencing. An extra adenine (A) base was added by the Taq DNA polymerase after the polymerase "runs-off" from the cleaved template. Following the extra A base, the intensity of the sequence peak signal drops to a lower level. The red arrow indicates the precise cleavage site on top or bottom strand.

Table 1

Listing of BspQI variants

BspQI mutants in the putative zinc fingers (C-x-x-C) C192A, C195A, C225A, C228A, C290A, C293A, C324A, C327A
BspQI top-strand nicking variants (Nt.BspQI) E172A, R254A, E255K E172A/E248A/E255K[a]
BspQI bottom-strand nicking variants (Nb.BspQI) E169A, K191A, H214A, Q222A, E224A, K331A, R428A, K331A/R428A, N235A/K331A/R428A[b]

Note:

[a] improved top-strand nicking enzyme.

[b] improved bottom-strand nicking enzyme