Published online 07 March 2014

Nucleic Acids Research, 2014, Vol. 42, No. 9 ***e77***
doi: 10.1093/nar/gku192

Engineering nicking enzymes that preferentially nick 5-methylcytosine-modified DNA

Alice Gutjahr and Shuang-yong Xu*

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

Received October 8, 2013; Revised February 18, 2014; Accepted February 20, 2014

ABSTRACT

N.ϕGamma is a strand-specific and site-specific DNA nicking enzyme (YCG↓GT or AC↑CGR). Here we describe the isolation of single and double mutants of N.ϕGamma with attenuated activity. The nicking domains (NDs) of E59A and 11 double mutants were fused to the 5mCG-binding domain of MBD2 and generated fusion enzymes that preferentially nick 5mCG-modified DNA. The CG dinucleotide can be modified by C5 methyltransferases (MTases) such as M.SssI, M.HhaI or M.HpaII to create composite sites AC↑YGG $N_{(8–15)}$ 5mCG. We also constructed a fusion enzyme 2xMBD2-ND(N.BceSVIII) targeting more frequent composite sites AS↑YS $N_{(5–12)}$ 5mCG in Mn^{2+} buffer. 5mCG-dependent nicking requires special digestion conditions in high salt (0.3 M KCl) or in Ni^{2+} buffer. The fusion enzyme can be used to nick and label 5mCG-modified plasmid and genomic DNAs with fluorescently labeled Cy3-dUTP and potentially be useful for diagnostic applications, DNA sequencing and optical mapping of epigenetic markers. The importance of the predicted catalytic residues D89, H90, N106 and H115 in N.ϕGamma was confirmed by mutagenesis. We found that the wild-type enzyme N.ϕGamma prefers to nick 5mCG-modified DNA in Ni^{2+} buffer even though the nicking activity is suboptimal compared to the activity in Mg^{2+} buffer.

INTRODUCTION

Restriction and modification systems are one of the bacterial defense mechanisms against invading bacteriophage, transposable elements and conjugative plasmids (reviewed in (1,2)). Bacteria protect their chromosomal DNA by modifying their own sequences, for example, by adding a methyl group to C5 position of cytosine to generate 5-methylcytosine (5mC), resulting in resistance to cleavage by the cognate restriction endonuclease (REase). In contrast, the Type IIM REases DpnI and Type IV REases McrBC, Mrr and GmrSD cleave foreign DNAs only when they are modified (for example, 5mC, 5hmC (5-hydroxylmethylcytosine) or glucosyl-5hmC) (reviewed in (3)). In theory, nicking variants could be isolated from the modification-dependent REases, but no such variants have been successfully engineered.

Strand-specific and site-specific nicking enzymes are found in nature as one subunit of heterodimeric restriction enzymes (e.g. Nb.BsrDI and Nt.BstNBI) (4,5) or as DNA nicking-modification systems (e.g. Nt.CviPII and Nt.CviQII) (6–8). In addition, a number of natural homing endonucleases (HEases) such as I-HmuI and I-BasI recognize and nick a degenerate 16- to 24-bp sequence (9). Nicking enzymes have also been engineered from Type IIS REases and HEases (10–12). Another group of sequence-specific nickases called relaxases is often found in conjugative plasmids that are involved in initiation of DNA replication and DNA strand transfer (13). In the past few years, zinc finger nickases as fusions of zinc finger protein and FokI cleavage domains have been constructed to introduce sequence-specific nicks in genomic DNA for gene targeting (gene addition and gene deletion) (14,15). A nicking domain (ND) variant of a DNA repair enzyme (MutH) has been fused to a TAL effector protein to create a rare nicking enzyme suitable for gene targeting (16). However, no robust 5mC-specific nickases have been found in nature despite extensive enzyme screening in the past 30 years (see REBASE) (17).

We recently discovered a large group of HNH nicking enzymes encoded by phage and prophage; these HNH nicking enzymes are located next to terminase small and large subunits involved in DNA packaging (18). The HNH motif is a small nucleic acid binding and cleavage module with a single metal ion (usually a Zn ion) bound by tetrahedral coordination (19–21). To date, more than 1000 proteins containing this motif have been discovered and they belong to different subgroups of enzymes such as HEases, bacterial toxins (colicins) with non-specific nuclease activity, modification-dependent endonuclease McrA, DNA repair enzymes, putative helicases or reverse transcriptases. It has been proposed that the conserved His residue in the HNH motif

*To whom correspondence should be addressed. Tel: +1 978 380 7287; Fax: +1 978 921 1350; Email: xus@neb.com
Present address:
Alice Gutjahr, University of Strasbourg, Superior School of Biotechnology, ESBS, Parc d'innovation, 300 Boulevard Sébastien Brant, 67412 Illkirch, France.

(HNK or HNN) activates a water molecule for a nucleophilic attack on the sugar phosphate backbone of nucleic acids by behaving like a general base (20,22,23).

The aim of this study is engineering of 5mC-specific nicking endonucleases (NEases or nickases). DNA methylation is an epigenetic process that is involved in the regulation of gene expression and cell differentiation, and the spectrum of methylation levels is broadly variable in mammalian DNA (24). The change in methylation profile is catalyzed by a family of DNA methyltransferases (DNMT1, DNMT3a and DNMT3b) (reviewed in (25)) and also regulated through DNA demethylation by the Tet family enzymes and T/G mismatch repair enzyme TDG (26–28). Methylated cytosines (5mC) are recognized by the methyl-binding domain (MBD) proteins that mediate the recruitment of chromatin remodeling enzymes (24). It has been demonstrated that cancer cells have global alterations in chromatin structure and epigenetic abnormalities play a role in different steps of cancer, from initiation to survival and invasion (reviewed in (29,30)). The analysis of the methylation profile can help to differentiate tumor cells from normal cells and is a useful tool for cancer diagnosis. Both hypo- and hyper-methylations are known to play a role in the onset of oncogenic diseases by either activating oncogenes or by silencing tumor suppressor genes (31–35). Thus, the engineering of a 5mC-dependent NEase has a practical utility by combining C5 MTases with 3–8 bp recognitions and a DNA ND with 2–7 bp specificity. It has been reported that nicking enzymes can be used to nick and label DNA for nicking site profiling (optical mapping), genomic fragment assembly, and ligase-free DNA fragment assembly and cloning (36–38).

The MBD protein family members share a methyl-CpG-binding domain (~80 amino acid (aa) long) that has a specific affinity for methylated CpG sites (5mCG) in double-stranded DNA. The MBD2 of this family appears to bind with a relaxed specificity to methylated DNA at 5mCG. It has been demonstrated that the multimerization of 5mCG-binding domain of MBD2 displays higher binding affinity to 5mCG *in vitro* than the monomeric peptide (39). In this work, two contiguous methyl-CpG-binding domains of MBD2 were constructed and then fused to a DNA ND to create a 5mCG-specific NEase.

The nicking enzyme N.ϕGamma, encoded by *Bacillus anthracis* phage γ, is a small protein with 127-aa residues and its minimal DNA ND of 76-aa has been fused to the zinc finger protein Zif268 to create a rare nickase (18). Structure prediction revealed that the 76-aa catalytic domain contains a β-β-α structure (also known as ββα–metal structural fold). The sequence specificity of this full-length NEase is YCG↓GT (complement strand: AC↑CGR) (Y = C/T; R = A/G) in Mg^{2+} buffer and is reduced to ~3-bp recognition in Mn^{2+} buffer (i.e. one base different from the AC↑CG site, e.g. AC↑TG or AT↑CG) (the down arrow indicates the nicked strand and the up arrow indicates the opposite strand is nicked). A homolog of N.ϕGamma, N.BceSVIII, encoded by a prophage of *Bacillus cereus* ATCC 10987, whose ND was also used in fusion with 2xMBD2. N.BceSVIII, nicks the DNA sites SS↓RT (AY↑SS) (S = G/C) in Mn^{2+} buffer and it has poor nicking activity in Mg^{2+} buffer. Thus, the DNA nicking activity is Mn^{2+}-dependent (18).

In this manuscript, we report the isolation of N.ϕGamma variants with reduced nicking activity, and the construction of fusion nicking enzymes by tethering DNA NDs to the 5mCG-binding domain of MBD2, and characterization of nicking activities in Mg^{2+}, Mn^{2+} and Ni^{2+} buffers. In addition, we demonstrated DNA labeling using the fusion enzyme on 5mCG-modified DNA.

MATERIALS AND METHODS

Bacterial strains, culture media, cloning vector and DNA substrates

The NEB Turbo *Escherichia coli* competent cells were used for cloning and the *E. coli* T7 Express competent cells (New England Biolabs, Inc. (NEB), Ipswich, MA, USA) were used for protein expression. All restriction and modification enzymes and DNA polymerases were from NEB. For protein expression experiments, pTYB1 (7477 bp) was used as a cloning and expression vector (NEB). The stop codon of the target gene was eliminated and the gene was inserted into NdeI and XhoI sites, which allows the expression of target protein as a fusion to an intein-CBD tag (in the C-terminus of the target protein). The intein cleavage is induced by addition of 50 mM Dithiothreitol (DTT) which releases the target protein from the intein-CBD tag that remains bound to the chitin column.

The pBR322 (4361 bp) was used as a DNA substrate for some nicking activity assays. A modified pUC19 plasmid (2686 bp) containing N.ϕGamma nicking sites and the methylation sites of M.HhaI was used. The plasmid substrates were first methylated by M.SssI (CG), M.HhaI (GCGC) or M.HpaII (CCGG) (4 h methylation reactions). The modified DNA substrates were then purified by spin columns (Qiagen, Valencia, CA) and subjected to nicking digestion and run-off sequencing using BigDye® direct cycle sequencing kit (ABI/Life Technologies, Grand Island, NY). To check the extent of 5mC modification of plasmid DNAs, the M.HhaI-, M.HpaII- and M.EagI-modified DNAs were digested by the cognate restriction enzymes HhaI, HpaII and EagI, respectively, to make sure that more than 95% of the modified DNAs were resistant to restriction digestion. For M.SssI-modified plasmid DNAs, we used HpaII digestion to examine the extent of modification (resistance).

The nicking activity assays were performed in different buffers supplemented with metal ions as stated in each experiment. For screening of N.ϕGamma variants, two buffers (10 mM $MgCl_2$ or 1 mM $MnCl_2$ with 50 mM NaCl and 10 mM Tris-HCl, pH 7.5) were used. These two buffer conditions were chosen because N.ϕGamma behaves differently in Mg^{2+} and Mn^{2+} (15). The nicked DNA samples were analyzed by electrophoresis on a 0.8% agarose gel and stained with ethidium bromide.

To enhance 5mCG-binding specificity and nicking, 0.25–0.3 M KCl was supplemented in some nicking reactions with Mg^{2+} or Mn^{2+}. Alternatively, 5mCG-dependent nicking of modified plasmid was achieved in a nickel buffer (50–75 mM NaCl, Tris-HCl, pH 7.5, 0.1–1 mM Ni^{2+}). Five hun-

dred nanograms of plasmid DNA were digested by the fusion enzyme for 1 h at 37°C. The samples were treated with Proteinase K to release the DNA that was tightly bound to the protein and analyzed by gel electrophoresis.

Cloning of nicking domain of N.ϕGamma and 5mCG-binding domain of MBD2

The synthetic gene encoding two consecutive 5mCG-binding domains of MBD2 (2xMBD2) with optimized *E. coli* expression codons were purchased from IDT (Integrated DNA Technologies, Coralville, IA) and subcloned into pTYB1 (NdeI-XhoI fragment). The 2× 5mCG-binding domains of MBD2 (2xMBD2 for abbreviation) contain 235-aa residues and have the following aa sequence:

(NdeI site at the N-terminal coding sequence)

MRATESGKRM DCPALPPGWK KEEVIRKSGL SAGKSDVYYF SPSGKKFRSK PQLARYLGNA VDLSSFDFRT GKMMPSKLQK NKQRLRNDPL NQNKFRLIKK QTLIGL**GASG ASGASGGS**RA TESGKRMDCP ALPPGWKKEE VIRKSGLSAG KSDVYYFSPS GKKFRSKPQL ARYLGNAVDL SSFDFRTGKM MPSKLQKNKQ RLRNDPLNQN KFRLIKKQTL IGL**GASGASG ASGAS**

(NheI-BamHI-XhoI sites after the C-terminal coding sequence; underlined aa sequence, linker between functional domains).

The coding sequence for the minimal nicking domain (76-aa, ND(N.ϕGamma), ND = nicking domain) was amplified by polymerase chain reaction (PCR) and fused to 2xMBD2 (ligation into NheI and XhoI sites) or fusions were constructed by the Gibson assembly method (NEB).

N.BceSVIII ND coding sequence (222 bp, 74-aa) was amplified by PCR and inserted in the pTYBI plasmid containing 2xMBD2 (NheI-XhoI fragment insert). The aa sequence of the fusion enzyme is shown below:

MRATESGKRM DCPALPPGWK KEEVIRKSGL SAGKSDVYYF SPSGKKFRSK PQLARYLGNA VDLSSFDFRT GKMMPSKLQK NKQRLRNDPL NQNKFRLIKK QTLIGL**GASG ASGASGGS**RA TESGKRMDCP ALPPGWKKEE VIRKSGLSAG KSDVYYFSPS GKKFRSKPQL ARYLGNAVDL SSFDFRTGKM MPSKLQKNKQ RLRNDPLNQN KFRLIKKQTL IGL**GASGASG ASGAS**rygra wkrirdsyia ahplceeckr qgkltpanev hhilplargg thdrsnlmal ctpchsaita rdgdrwgtr (the aa residues shown in lower case are the NDs of N.BceSVIII; underlined aa residues, linker sequence).

The recombinant plasmids were transformed into *E. coli* T7 Express competent cells and plasmids were extracted by a mini-prep kit (Sigma-Aldrich, St Louis, USA). Plasmids with correct inserts were screened by restriction digestion and further verified by DNA sequencing.

Site-directed mutagenesis

Site-directed mutagenesis of N.ϕGamma was carried out by inverse PCR. The appropriate mutagenic primers were ordered from IDT and mutagenesis was performed by 25 cycles of PCR (Phusion DNA polymerase, NEB). The template DNA was destroyed by 1 h digestion at 37°C with DpnI. The amplified DNA was purified by spin-column purification (Qiagen), and then phosphorylated and ligated at room temperature with T4 polynucleotide kinase and T4 DNA ligase. The ligated DNA was used in transformation. Mutant alleles were sequenced to confirm the desired mutation(s).

Protein expression and purification

Late-log phase cells were induced with 0.5 mM IPTG and grown in a shaker for 4 h at 37°C for the N.ϕGamma endonuclease and 5 h at 30°C for the 2xMBD2-ND(E59A or other variants of N.ϕGamma) and 2xMBD2-ND(BceSVIII) fusion proteins in Luria broth (LB) medium supplemented with ampicillin. Induced cells were harvested by centrifugation and stored at −20°C. Small-scale purification was performed using 30 ml of IPTG-induced cells. Cell pellets were resuspended in 0.5 ml lysis buffer (0.5 M NaCl, 20 mM Tris-HCl, pH 8.5, 1 mM EDTA, 0.1% Triton X-100) and then subjected to sonication on ice. The cell lysate was centrifuged to eliminate cell debris. Three hundred microliters of chitin beads (NEB) were dispensed into a 1.5 ml Eppendorf tube and washed twice with 1 ml column buffer (0.5 M NaCl, 20 mM Tris-HCl, pH 8.5). The cell extract was added to the resin and agitated for 1 h at 4°C. After this incubation, the resin was washed three times with 1 ml column buffer, and 600 μl of cleavage buffer (0.5 M NaCl, 20 mM Tris-HCl, pH 8.5, 50 mM DTT) was added. DTT cleavage occurred overnight at 4°C. Then, the resin was gently mixed and centrifuged and the supernatant containing the purified protein was removed for analysis. The procedure for chitin column purification was followed based on NEB's purification manual with a minor modification: the cell extract was loaded onto a 10 ml chitin column that was equilibrated with chitin column buffer and the flow-through was reloaded to the resin three times by gravity flow. After the column was washed 10 times with 10 ml column buffer, 10 ml of cleavage buffer (0.5 M NaCl, 20 mM Tris-HCl, pH 8.5, 50 mM DTT) were added and the intein cleavage took place overnight at 4°C. The target protein was then eluted and analyzed by sodium dodecyl sulphate-polyacrylamide gel electrophoresis (SDS-PAGE).

A 5 ml HiTrap heparin HP column (GE Life Sciences) was used for second step purification of the fusion enzyme. The samples eluted from the chitin column were diluted in column buffer (20 mM Tris-HCl, pH 7.5, 1 mM DTT, 1 mM EDTA, 20 mM NaCl, 5% glycerol) to reach a final concentration of 50 mM NaCl. Proteins were loaded onto a heparin Sepharose column using an AKTA FPLC purification system (GE Life Sciences). Elution was carried out using a salt gradient of elution buffer (50 mM to 1 M NaCl, 20 mM Tris-HCl, pH 7.5, 1 mM DTT, 1 mM EDTA, 5% glycerol). The eluted fractions corresponding to UV absorption peaks were analyzed by SDS-PAGE. The fractions containing the desired protein were pooled, concentrated if necessary and stored at −20°C in 50% glycerol.

Gel shift assays

A DNA mobility shift assay in native gel was used to assess DNA binding activity of 2× 5mCG-binding domain of MBD2 and mutants of N.ϕGamma (40). Varying amounts of protein were mixed with oligonucleotides in a binding

buffer (20 mM HEPES pH 7.9, 10% glycerol, 50 mM KCl, 10 mM EDTA). After 30 min incubation at room temperature, 4 μl of glycerol was added and the mixture was loaded onto a 4–20% TB native gel (Life Technologies) and ran 45 min at constant 200 V on ice (pre-ran gel for 30 min, gel box immersed in an ice-water box). The gel was stained with SYBR gold (Life Technologies) and gel imaging was carried out with 526 nm fluorescent emissions in the Typhoon 9400 phosphoimager (GE Life Sciences). The gel images were analyzed using the ImageJ software.

Nicking of 5mCG-containing duplex oligos ($CGN_{14}ACCGGN_{13}CG$)

Duplex oligos of 49 bp were synthesized based on one preferred nicking site found in pBR322, which contains the following sequence:

5′GAGTGG **CG** ATAAGTGGTGTCTT **AC↑CGG** GTTGGACTCAAGA **CG** ATAGTTA3' (top strand).

5′TAACTAT **CG** TCTTGAGTCCAAC **CCG↓GT** AAGACACCACTTAT **CG** CCACTC-FAM 3′ (FAM-labeled bottom strand).

Three CG sites in the duplex oligos were methylated by M.SssI in NEB buffer 2 plus SAM (4 h) and then purified through Bio-Spin 6 spin columns (Bio-Rad, CA). The methylation status was confirmed by HpaII digestion as the modified site C5mCGG confers resistance to HpaII cleavage. Duplex oligos (20–100 ng) were digested with the fusion nicking enzymes or WT N.ϕGamma. Following protease K treatment, the nicked product of FAM-labeled 24-nt oligo was boiled for 5 min and detected by fluorescence imaging after electrophoresis in 4–20% Tris-Borate-EDTA (TBE) gels. For preferential nicking of the modified oligos, a Ni^{2+} buffer (50 mM NaCl, 10 mM Tris-HCl, pH 7.5, 0.5 mM $NiCl_2$) was used in the nicking reactions.

DNA labeling with Cy3-dUTP

For labeling DNA by strand-displacement extension, 1 μg of pBR322 unmethylated or M.SssI-methylated was digested by the 2xMBD2-ND(E59A of N.ϕGamma) fusion enzyme in 0.25 M KCl for 1 h at 37°C. After 20 min heat-inactivation of the enzyme at 80°C, the samples were purified by spin-column purification. Labeling was performed by mixing the purified DNA with 1 mM dTTP, dCTP and dATP (NEB) with 40 pM Cy3 dUTP (GE Life Sciences) and 10 units of the Klenow fragment of *E. coli* DNA polymerase I in 50 μl total volume. The mixture was incubated 5–10 min at room temperature and then analyzed in an agarose gel without ethidium bromide. NIH 3T3 mouse genomic DNA was nicked by 2xMBD2-ND(E59A) in a Mg^{2+} buffer containing 0.3 M KCl for 1 h at 37°C. DNA labeling was performed through 10 min incubation at 25°C with Klenow fragment of *E. coli* DNA polymerase I and a dNTP mix containing Cy3-dUTP. Labeled DNA band images were obtained under 580 nm fluorescent emissions using Typhoon 9400 phosphoimager (GE Life Sciences) and images were edited using the ImageJ software.

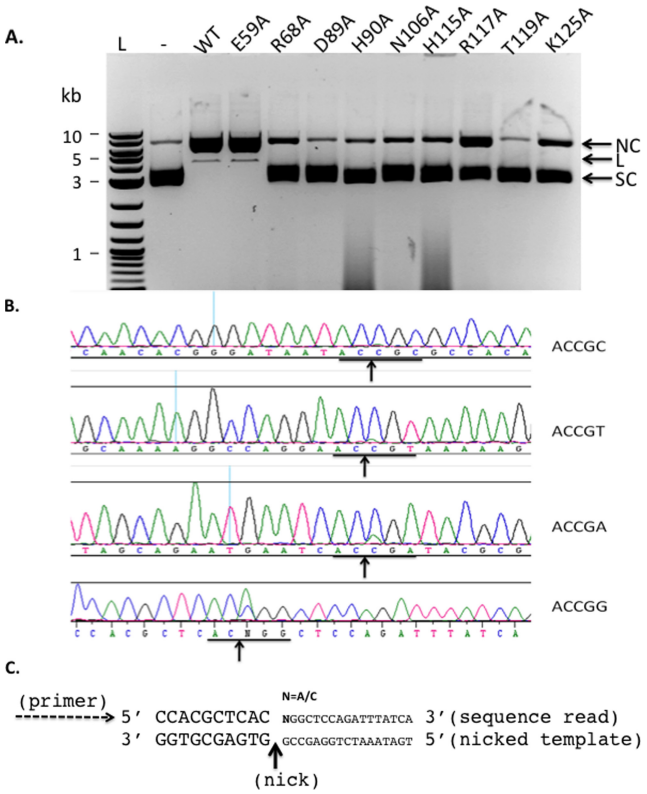

Figure 1. Nicking activity assay for mutants of N.ϕGamma. (**A**) One microgram of pBR322 was digested by purified enzyme (small scale) in a Mg^{2+} buffer (NEB buffer 2). L = 2-log DNA ladder (NEB), WT = wild-type N.ϕGamma. Single mutant aa substitution was indicated on top of each lane. NC: nicked circular; L: linear; SC: supercoiled. (**B**) Run-off sequencing of plasmid DNA nicked by the full-length E59A nickase. Arrows indicate nicked site with double peaks at A/C. (**C**) Schematic diagram of DNA run-off sequencing to determine the nicked site. A doublet peak with extra A indicates the nicked site. The height of base-call peaks following the doublet drops off due to the nicked template.

RESULTS AND DISCUSSION

Isolation of mutants of phage Gamma HNH nicking endonuclease (N.ϕGamma)

The full-length phage Gamma HNH nicking endonuclease (N.ϕGamma, 127-aa) and its minimal ND (76-aa) have been described previously (18). The nicking specificity of N.ϕGamma is YCG↓GT. The 76-aa ND of N.ϕGamma displays more degenerate specificities with one base off from the cognate site CG↓GT under star conditions (high enzyme/glycerol concentration or digestion in Mn^{2+} buffer). In this work, we isolated 54 single mutants and 3 double mutants of N.ϕGamma mainly targeting charged and hydrophilic aa residues or predicted catalytic residues. The goal of this mutagenesis experiment was to confirm the HNH catalytic residues and isolation of variants with attenuated nicking activity (as a result of weakened binding). The mutant proteins were partially purified by batch purification with chitin beads and nicking activity tested on pBR322 (Figure 1 and data not shown). Figure 1A shows the nicking activity of nine variants in Mg^{2+} buffer. The mutants fall into two major groups. Group 1 variants show full or partial nicking activity in Mg^{2+} buffer (e.g. E59A and R117A).

Group 2 variants show poor nicking activity (e.g. N106A, H115A and K125A) or no activity (e.g. D89A and T119A) in Mg^{2+} buffer. In Mn^{2+} buffer, however, some poor activity mutants such as H90A, N106A, H115A and T119A still display partial nicking activities (data not shown). In order to eliminate the nicking activity completely in Mn^{2+} buffer, three double mutants in the HNH motif (H90A/N106A, H90A/H115A and N106A/H115A) were constructed and the mutant proteins were purified and assayed for nicking activity. As expected, the double mutants had minimal detectable nicking activity in Mn^{2+} buffer (data not shown). A number of active variants were further analyzed by run-off sequencing of the nicked sites. One active mutant, E59A, was found to nick the site CCG↓GT preferentially. The range of nicking activity on ACCGN sites is as following: ACCGG>ACCGA>ACCGT>ACCGC as demonstrated by run-off sequencing of the nicked sites in pBR322 (Figure 1B). We choose to work on E59A for the following reasons: it is as active as the WT enzyme, E59A prefers to nick CCG↓GT (ACCGG) site, and the E59 to A substitution is located outside the HNH catalytic core. Later it was shown that the specificity of the 76-aa ND of E59A was confined to ACCGG or ACTGG sites when its ND was fused to a 5mCG-binding domain (ACYGG, Y = C/T) (see below). We have not tested the nicking specificity of the stand-alone 76-aa ND of E59A. A schematic diagram of DNA run-off sequencing is shown in Figure 1C. When Taq DNA polymerase encounters a nick in the template, it adds an adenine base by the template-independent terminal transferase activity, thus creating a doublet peak (A/C).

Table 1 summarizes the single mutant activities and protein expression level. Relative nicking activity (+++, ++, +, +/− or −) was estimated in Mg^{2+} or Mn^{2+} buffer in comparison with the WT enzyme. Most noticeably, 14 out of 15 inactive and poor activity variants (+/− or −) in Mg^{2+} buffer are located in the 76-aa long C-terminal HNH ND with the predicted ββα–metal structure. The protein expression level of three variants (R62A, K77A and K81A) was much reduced, suggesting possible protein folding defect. A negatively charged residue D89 in front of the predicted catalytic residues H90-N106-H115 is conserved in many HNH nucleases and is critical for the catalytic activity. D89A lost nicking activity in both Mg^{2+} and Mn^{2+} buffers. Other poor activity or inactive mutants are K95A, Q108A, N116A, K118A, T119A, K123A and K124A, whose aa substitutions are located in the predicted ββα–metal fold. In structure-guided secondary structure prediction and structure modeling by the Phyre server indicated that N.ϕGamma is similar to a putative DNA repair enzyme Gme HNH endonuclease found in *Geobacter metallireducens* GS-15 (data not shown) (21).

Engineering fusion nicking enzymes targeting 5mCG sites

The 5mCG-binding domain of MBD2 is able to bind to modified 5mCG site. When the 5mCG-binding domain is duplicated or concatenated into three and four copies (2×, 3× and 4×), they show higher binding affinity to 5mCG-modified DNA (39). Therefore, the fusion nicking enzymes we constructed carry two copies of 5mCG-binding domain of MBD2 (2xMBD2) tethered to a DNA nicking domain

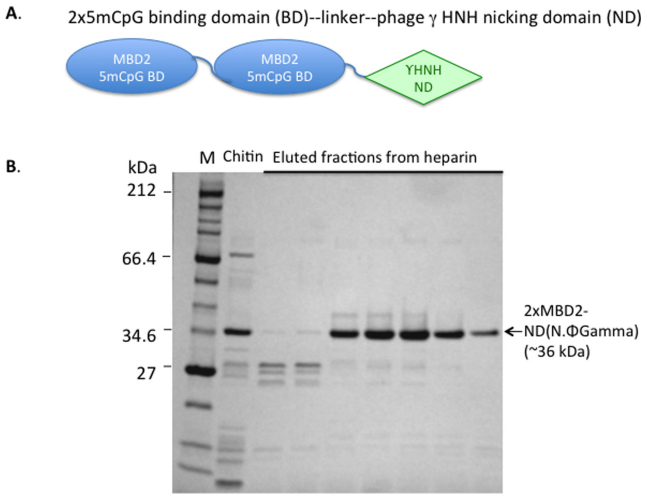

Figure 2. Schematic diagram of 2xMBD2-ND(N.ϕGamma) functional domains and fusion enzyme purification. (**A**) A schematic diagram of the fusion protein 2xMBD2-ND(N.ϕGamma). 5mCG BD, 5mCG-binding domain; 2× = duplicated 5mCG BD; ND, nicking domain. (**B**) SDS-PAGE analysis of partially purified fusion protein. The molecular mass of the fusion protein is predicted to be ∼36 kDa. Lane 2, fusion protein purified from a chitin column after DTT cleavage. Lanes 3–9, eluted fractions from a heparin Sepharose column.

(ND, 76-aa) of N.ϕGamma with a 12-aa linker as diagrammed in Figure 2A. From the mutagenesis experiment described above, we had obtained a small collection of N.ϕGamma variants with full, partial or low nicking activities. We have not constructed 3xMBD2 or 4xMBD2 in fusion with the HNH ND. Such fusion nicking enzymes may nick DNA at farther distance from the 5mCG-binding sites.

A total of 16 fusion enzymes were constructed and partially purified by batch purification. Table 2 shows the relative nicking activity in Mg^{2+} buffer. In general, if the full-length mutant nickase is not very active in Mg^{2+} buffer, the engineered fusion enzyme also has poor nicking activity with the same metal cofactor (e.g. R62A, D89A, H90A, N106A, H115A and T119A in E59A background). That is, the nicking activity of these mutants cannot be rescued (suppressed) by fusion with another DNA binding partner (due to compromised catalytic residues or impaired DNA binding affinity of the ND or some unknown mechanisms). The fusion enzymes are fully active (+++) or partially active (+ or ++) in Mg^{2+} buffer, primarily depending on the 76-aa nicking variant used. However, when the ND of E59A/R68A, E59A/K95A and E59A/K123A was fused to a strong DNA binding partner 2xMBD2, partial nicking activity was restored (rescued). This suggests that the poor nicking activity of the full-length variants of R68A, K95A and K123A is likely due to attenuated DNA binding affinity.

We first characterized the fusion enzyme 2xMBD2-ND(E59A of N.ϕGamma). Figure 2B shows the partially purified fusion enzyme after chromatography through chitin and heparin columns. Figure 3 shows the fusion enzyme nicks both unmodified and M.SssI-modified pBR322 (M.SssI modifies CG sites to yield 5mCG) in NEB buffer 2 or in the presence of 0.1–0.3 M KCl (Figure 3A and B, lanes

Table 1. Nicking activities and protein expression levels of the N.ϕGamma variants

Mutant	Protein expression level	Nicking activity in Mg^{2+}	Nicking activity in Mn^{2+}
WT	+++	+++	+++
R6A	++	+++	+++
H12A	++	++	+++
N15A	++	NT	+++
R18A	++	+++	+++
R20A	++	++	+++
H25A	+++	+++	+++
R26A	+++	+++	+++
D35A	++	+++	+++
R38A	++	+	++
D41A	++	+	++
R42A	+++	+++	+++
R45A	++	–	+
D46A	++	++	+++
R54A	+	+	+++
S55A	+++	+	+++
K56A	++	+	+++
E59A	++	++	+++
R60A	+++	+++	+++
R62A	+	+	+
E63A	+	++	+++
Q64A	+	+	+++
R68A	+	+	+
D69A	++	++	++
K70A	++	+	+++
H75A	+++	++	+++
K77A	–	–	–
N79A	++	NT	+++
R80A	++	+++	+++
K81A	–	–	+
D86A	++	+++	+++
D89A	+++	–	+
H90A	++	+/–	+
K95A	+	–	+
D97A	++	+++	+++
K101A	++	+	+++
E105A	++	++	+++
N106A	+	+/–	+
Q108A	++	+/–	+
N112A	+	+/–	++
H115A	+	+/–	+
N116A	+	+/–	++
R117A	++	+	++
R117C	+	++	+++
R117G	+	++	+++
R117S	++	++	+++
R117V	+	+	++
K118A	+	–	–
T119A	+	–	+
E121A	++	+	+++
D122A	+	+	++
D122G	+	+	++
K123A	+	+/–	++
K124A	+	+/–	++
K125A	+	++	+++

The poor activity and inactive mutants (Mg^{2+} buffer) are shown in bold; the putative catalytic mutants are shown in bold and underlined. Note: the activity of some inactive mutants (Mg^{2+}) can be rescued in Mn^{2+} buffer. NT, not tested. Plasmid pBR322 was used as the nicking substrate. Buffers used in the nicking assay: 50 mM NaCl, 10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$ or 1 mM $MnCl_2$. The relative mutant enzyme activities are defined as follows: +++ good activity, 50–100% of WT; ++ partial activity, 10–49% of WT; +/– or + poor activity, 1–9%; – no activity, background nicked DNA as mock-treated pBR322.

3–5). The fusion enzyme was able to nick 5mCG-modified DNA preferentially in 0.25–0.3 M KCl (lanes 6 and 7), which is consistent with the report that addition of KCl to MBD1 enhances 5mCG binding over non-specific DNA (41). The preferential nicking of 5mCG-modified DNA was also observed for other fusion enzymes carrying two aa substitutions in the ND (Figure 3C; e.g. R68A, N112A, K123A, K124A and R117V in E59A background). Not surprisingly, the catalytic-deficient nicking mutant H90A was inactive when its ND was fused to 2xMBD2; the fusion enzyme from K118A ND was also inactive as the parent (full-length) nicking variant (Figure 3C and data not shown). It

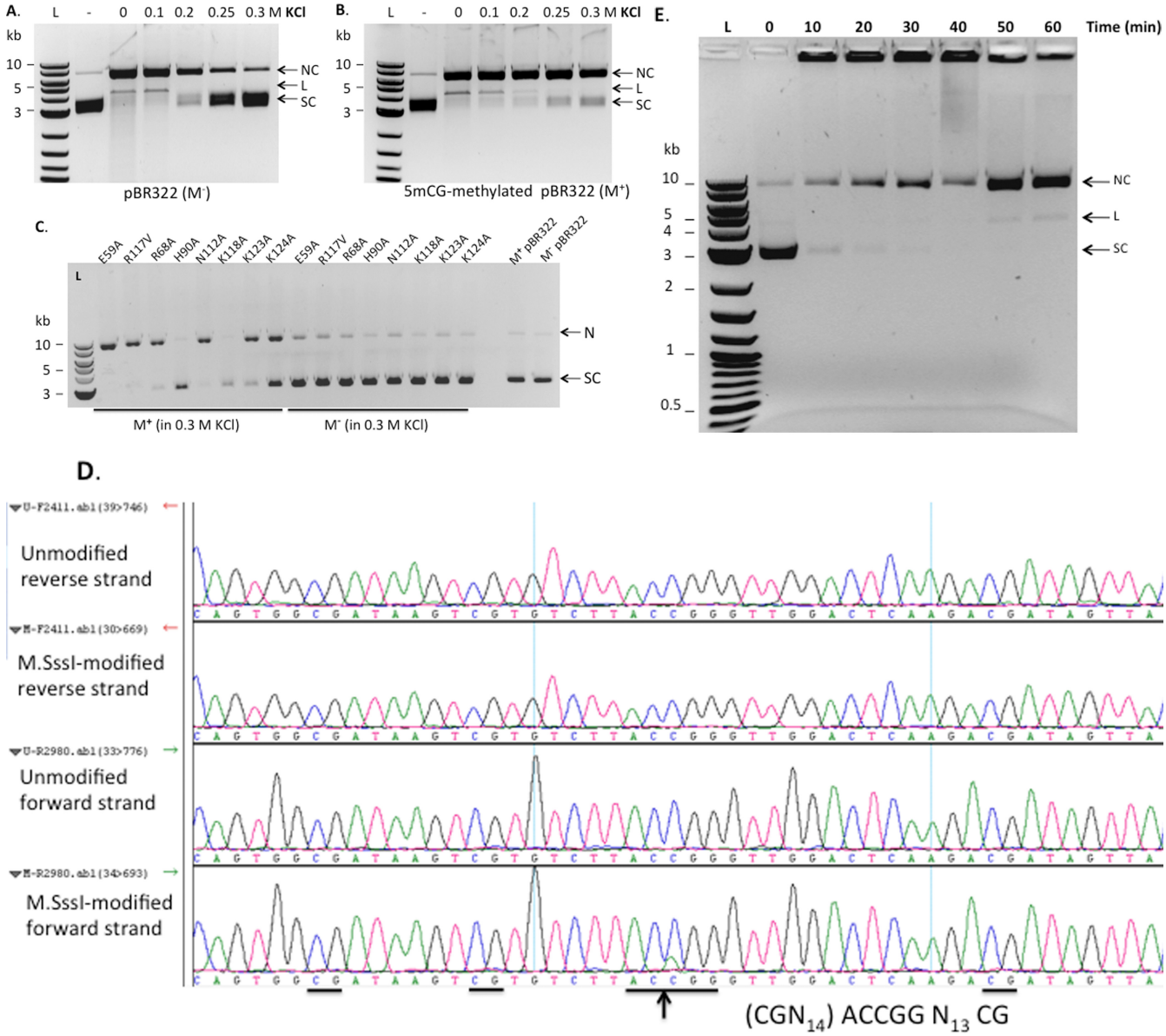

Figure 3. Effect of KCl concentration on the nicking specificity of 2xMBD2-ND(E59A) fusion enzyme. (**A**) Nicking reaction carried out in Mg^{2+} buffer supplemented with 0.1–0.3 M KCl. 0.5 μg of pBR322 (unmodified) was digested by 2.5 μg of the fusion enzyme in NEB buffer 2 (50 mM NaCl, 10 mM $MgCl_2$, 10 mM Tris-HCl, pH 7.8) plus different concentrations of KCl for 1 h at 37°C. NC: nicked circular; L: linear; SC: supercoiled. (**B**) Same digestion condition as in (A), except that 5mCG-modified pBR322 was used. (**C**) Nicking reaction carried out in NEB buffer 2 plus 0.3 M KCl with partially purified E59A and seven nicking domains of double mutants (in E59A background) with modified (M^+) and unmodified (M^-) pBR322. (**D**) Characterization of the nicking site of 2xMBD2-ND(E59A) in Mg^{2+} buffer. Run-off sequencing of modified and nicked pBR322. One of the composite nicking sites is (CG $N_{6–14}$) AC↑CGG N_{13} CG. The arrow indicates the nicking site with A/C doublet. Both strands were sequenced. (**E**) A time course of digestion in 10–60 min with fixed amount of fusion enzyme and DNA.

was concluded that the nicking activity of some mutants could be rescued by fusion with a strong DNA binding partner such as 2xMBD2 and nicking at the modified sites was strongly dependent on the presence of 0.25–0.3 M KCl.

To find out the exact nicked sites, we performed run-off sequencing of the digested pBR322 (unmodified or 5mCG-modified) by 2xMBD2-ND(E59A). The whole plasmid was sequenced on both strands and 10 partially nicked sites were found. Table 3 lists 10 nicked sites in modified pBR322. Eight sites conform to AC↑YGG $N_{(8–15)}$ CG. The run-off sequencing of the one nicked site was shown in Figure 3D (this nicking site was also confirmed on modified synthetic duplex oligos, see below). A doublet peak (A/C) was observed when the template strand was nicked (Taq DNA polymerase adds an adenine base by template-independent nucleotide transferase activity when the template is broken). The nicking specificity of the full-length (127 aa) nicking enzyme variant E59A is ACCGN (N, G>A>T>C) but the specificity of the 76-aa ND of E59A appears to be more relaxed in the third position as ACCGG or ACTGG when it is fused to 2xMBD2 (the fifth position recognition appears to be more specific for G only). The simultaneous effect

Table 2. DNA nicking activity of 2xMBD2-ND(N.ϕGamma) fusion enzymes

Nicking domain variant	Nicking activity in Mg^{2+} buffer
E59A	+++
E59A + R62A	+
E59A + R68A	++
E59A + H90A	–
E59A + K95A	++
E59A + N106A	+/–
E59A + Q108A	+/–
E59A + N112A	++
E59A + H115A	–
E59A + N116A	++
E59A + R117A	++
E59A + R117V	++
E59A + K118A	+/–
E59A + T119A	+/–
E59A + K123A	+
E59A + K124A	+

Relative nicking activity: +++, ++, +, +/–, –, see definition in Table 1 footnote. 0.5 μg of M.SssI-modified pBR322 DNA was digested by the fusion enzyme for 1 h at 37°C in a buffer containing 50 mM NaCl, 0.3 M KCl, 10 mM Tris-HCl, pH 7.5, 10 mM $MgCl_2$.

on recognition on two base pairs was not well understood and remained to be studied. It was concluded that the composite sites in modified pBR322 are confined to AC↑YGG $N_{(8–15)}$ CG. In order to limit the spacer of the composite site to a more defined length, it may be necessary to make the 5mCG-binding domain smaller or make the linker shorter between the 2xMBD2 and the ND or use only one copy of 5mCG-binding domain with a trade-off for lower DNA binding affinity.

A time course of nicking reaction from 10 to 60 min was carried out in order to define the enzyme units (Figure 3E). An extended time course from 10 to 90 min was also performed (data not shown). One unit of the fusion enzyme was defined as the amount of enzyme (~3 μg) required to completely nick 0.5 μg of 5mCG-modified pBR322 DNA in 50 μl of reaction volume at 37°C for 1 h in NEB buffer 2 supplemented with 0.3 M KCl. The specific activity is approximately 333 units/mg protein. Such a low specific activity suggests that the fusion enzyme does not turn over in nicking DNA probably due to the tight binding by 2xMBD2. The other indirect evidence came from the observation that the fusion enzymes remain bound to the DNA. Protease K digestion was necessary to 'release' the DNA after the nicking reactions. Otherwise some of the nicked DNA would be bound to the wells in agarose gels during gel electrophoresis as observed in Figure 3E.

76-aa minimal DNA nicking domain of N.ϕGamma

The 76-aa ND of E59A (N.ϕGamma) can be further minimized to 74-aa by fusion to 2xMBD (by deletion of two more aa residues: FY). Deletion of three aa residues (FYR) from the 76-aa ND abolished the nicking activity (data not shown). A stretch of ~13-aa residues at the beginning of the 76-aa ND is predicted to form an α-helix which is important for the enzyme activity; deletion of this putative α-helix eliminated the nicking activity (18). A large number of homologs (20–100% aa sequence identity) of the 76-aa DNA ND encoded by phage and prophage genomes can be found in GenBank by BlastP search (data not shown), which can be potentially used in engineering fusion nickases. A HK97 phage encoded HNH enzyme with non-specific nicking activity was reported previously and could be used as a non-specific ND with minimal specificity. Interestingly, the HNH endonuclease from ϕHK97 prefers Ni^{2+} for nuclease activity (42).

Fusion enzyme nicking activity in Mn^{2+} buffer

The YCG↓GT (AC↑CGR) nicking specificity of N.ϕGamma can be further relaxed into 3 bp with one base different from the CGGT (ACCG) sites when the nicking reactions are performed in Mn^{2+} buffer (18). Accordingly, the fusion enzyme may also show degenerate composite nicking sites. The fusion enzyme was diluted 2-, 4- and 8-fold and used in digestion of 5mCG-modified pBR322 in Mn^{2+} buffer (supplemented with 0.3 M KCl) for 30, 45 and 60 min. The results are shown in Figure 4A. The 5mCG-modified DNA was preferentially nicked and a small amount of linear DNA was also detected. The nicked sites were analyzed by run-off sequencing. Figure 4B shows one of the partially nicked sites: (CG N_6) ACCC N_{11} CG or (CG N_6) ACCCG N_{10} CG. The 12 nicked sites in Mn^{2+} buffer for the fusion enzyme can be summarized as ASYS $N_{(5–12)}$ CG (S = G/C) (see Table 3). In general, the nicking sites ASYS fall into CG-rich regions (2–5 copies of CG dinucleotide), suggesting that the density of the 5mCG sites is likely important for the nicking activity. The alternative explanation is for the fusion enzyme to dimerize and nick the target DNA or through a looping mechanism.

DNA modification by M.HhaI to create composite nicking sites

In theory, the CG dinucleotide can be modified by C5 methyltransferases (MTases) with 3–8 bp DNA recognition sequence. The combination of C5 MTase modification to create 5mCG and HNH site ACYGG nearby would generate composite rare nicking sites. A different strategy of using DNA MTase, REase and Lac repressor in the Achilles' heel cleavage procedure was used to cleave yeast and *E. coli* genomes at a single site.

M.HhaI modifies the GCGC site to produce G5mCGC (see REBASE). A pUC19 derivative was modified by M.HhaI and then nicked by the fusion enzyme in 0.3 M KCl. The M.HhaI-modified pUC19 derivative was resistant to HhaI digestion (data not shown). Figure 5A shows that the modified DNA was preferentially nicked by the fusion enzyme. Run-off sequencing of the nicked DNA confirmed that a specific nick (a doublet peak A/T indicated by an arrow) was introduced in the composite site AC↑TGG N_{14} GCGC, which is equivalent to a 9-bp nicking site (Figure 5B). Other nicked composite sites are AC↑CGGataagGCGC (ACCGG N_6 CG) and AC↑CGGatacctgtcCGC (ACCGG N_9 CG) (data not shown). CCGC is a possible star site for M.HhaI that could be modified at high enzyme concentration and prolonged *in vitro* modification (4 h).

We also modified pBR322 with M.HpaII and the modified DNA was subjected to nicking by the fusion enzyme.

Table 3. Nicking sites of 5mCG-modified pBR322 in Mg^{2+} and Mn^{2+} buffers as determined by DNA run-off sequencing

Type of modified DNA	Nicking sites in Mg^{2+}	Nicking sites in Mn^{2+}
M.SssI-modified pBR322	ACCGG N_{13} CG: 1 site	ACCG N_{12} CG: 1 site
Same	ACCGG N_{11} CG: 1 site	ACTG N_{12} CG: 1 site
Same	ACTGG N_{10} CG: 2 sites	ACCG N_{11} CG: 1 site
Same	ACTGG N_9 CG: 1 site	ACCC N_{11} CG: 1 site
Same	ACCGG N_6 CG or	ACTG N_{11} CG: 1 site
	ACCGG N_{11} CG*: 1 site	AGTG N_{11} CG: 1 site
Same	ACCGG N_1 CG or	ACCG N_{10} CG: 1 site
	ACCGG N_4 CG or	
	ACCGG N_{13} CG**: 1 site	
Same		AGCG N_7 CG: 1 site
Same		ACCG N_5 CG: 2 sites
		AGCG N_5 CG or
		AGCG N_8 CG: 1 site
		ACGG N_6 CG: 1 site
M.HpaII-modified pBR322	ACCGG N_{11} CG: 1 site	Not tested
M.HhaI-modified pUC19	ACCGG N_8 CG: 1 site	Not tested
M.HhaI-modified pUC19	ACTGG N_{14} CG: 1 site	Not tested
	ACYGG $N_{(8–15)}$ CG	ASYS $N_{(5–12)}$ CG

*AC↑CGGataaggCGcagCG.
**AC↑CGGgCGgCGgccaaagCG.

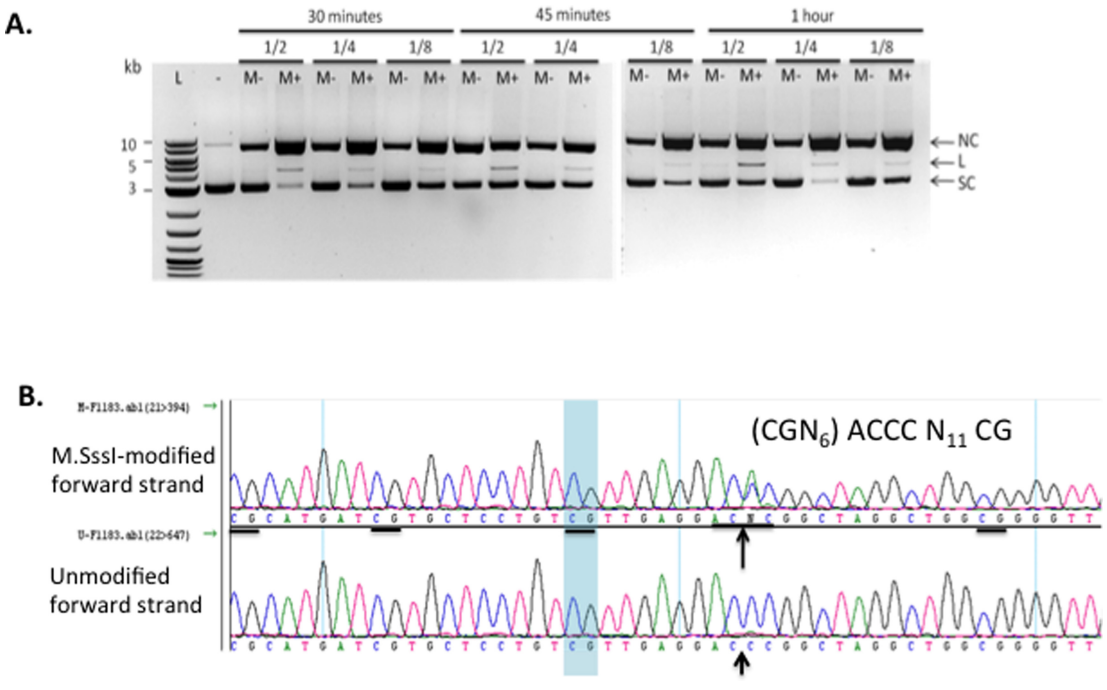

Figure 4. Nicking activity assay for 2xMBD2-ND(E59A) fusion enzyme for methylated DNA in Mn^{2+} buffer. (**A**) 0.5 μg of 5mCG-modified pBR322 was digested by different amounts of enzyme (0.3–1.3 μg) in a buffer containing 10 mM $MnCl_2$, 10 mM Tris-HCl, pH 7.8, 50 mM NCl and 0.3 M KCl in 30 min, 45 min and 1 h at 37°C. M+, 5mCG-modified substrate; M−, unmodified substrate. (**B**) Run-off sequencing of the pre-nicked modified DNA. The composite nicking site is ($CGN_{6–17}$) AC↑CC N_{11} CG. Arrow indicates the nicked site with A/C doublet.

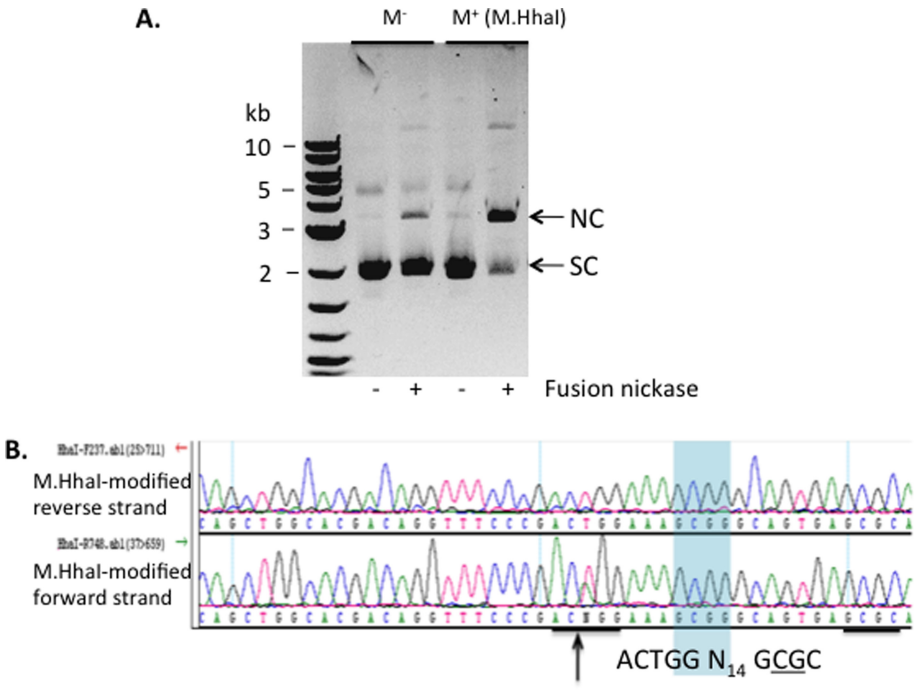

Figure 5. Nicking activity assay of 2xMBD2-ND(E59A) using a pUC19 derivative. (**A**) One microgram of M.HhaI-modified or unmodified pUC19 was digested by 2.5 μg of the fusion nicking enzyme in NEB buffer 2 supplemented with 0.3 M KCl for 2 h at 37°C. M.HhaI modifies the site GCGC to produce G5mCGC, which subsequently serves as the substrate for the 5mCG-binding domain. NC: nicked circular; SC: supercoiled. (**B**) Run-off sequencing of nicked DNA modified by M.HhaI. The composite site ACTGG N_{14} GCGC is shown. Arrow indicates the nicking site with A/T doublet. No nick was found in sequencing of the opposite strand (top panel).

The nicked DNA was analyzed by run-off sequencing. One nicking site was found at CCGG N_4 AC↑CGG N_{10} GCGG (data not shown). In this region, two HpaII sites and one star site (GCGG) are present and are likely modified by M.HpaII. It is known that M.HpaII modified the star site GCGG at high enzyme concentration (M. Zhang and R. Roberts, unpublished results). It is tentatively concluded that M.HpaII can be used to create composite sites containing 5mCG. In a control experiment, a pUC19 derivative with a single EagI site was modified *in vivo* by M.EagI co-expression. M.EagI modifies the site CGGCCG to generate CGG5mCCG (REBASE). The modified plasmid was extracted and digested by EagI, NdeI or the fusion nicking enzyme. The modified and unmodified plasmid DNAs were poorly nicked by the fusion enzyme in 0.3 M KCl, indicating that 5mCC dinucleotide is not a good substrate; the modified plasmid is resistant to EagI digestion and sensitive to NdeI digestion (data not shown).

Fusion enzyme activity in other divalent cations

One fusion enzyme 2xMBD2-ND(E59A/T119A) has poor nicking activity in Mg^{2+} buffer. Since the activity of some of the full-length mutants of N.ϕGamma can be rescued by the presence of Mn^{2+} cofactor, we tested the fusion enzyme activity in Mn^{2+} buffer and in other metal ions. Figure 6 shows that 2xMBD2-ND(E59A/T119A) possesses 5mCG-dependent DNA nicking activity in Mn^{2+} buffer, no activity in buffers with Mg^{2+}, Ca^{2+} or Co^{2+} (the enzyme and DNA formed a brown precipitate in Co^{2+} buffer, data not shown), and a low but detectable nicking activity in 1 mM Ni^{2+} (these reactions were performed in the absence of 0.3 M KCl). It was noted during the mutagenesis experiment that the full-length nicking variant T119A displays poor activity in Mg^{2+} buffer but the activity can be rescued by the presence of Mn^{2+} (see Table 1). It was concluded that the fusion enzyme 2xMBD2-ND(E59A/T119A) displays 5mCG-dependent DNA nicking activity in Mn^{2+} and Ni^{2+} buffers. Another fusion enzyme 2xMBD2-ND (E59/R62A) also displays preferential nicking of modified DNA in nickel buffer (1 mM Ni^{2+}). Nickel ions may enhance 5mCG nicking specificity by binding to the HNH ND (see more Ni^{2+}-dependent nicking below). It is well known that divalent cations can alter restriction enzyme specificities. For example, Mg^{2+} ions enhance EcoRV binding specificity to cognate sites (43). Ca^{2+} ions can enhance KpnI binding specificity to cognate sites (but Mg^{2+} and Mn^{2+} promote star site binding and cleavage). KpnI belongs to the HNH-family endonucleases and high-fidelity (HF) mutants of KpnI have been isolated that increased specificity in Mg^{2+} buffer (reduced star activity) and altered metal ion requirement; KpnI-HF displays no activity in Ca^{2+} buffer and high-fidelity in Mg^{2+} buffer (44). For the HpyAV restriction enzyme (also belonging to the HNH endonuclease family), however, Ni^{2+} is preferred for binding and cleavage of the cognate sites (45). Another HNH endonuclease Sco_McrA requires Mn^{2+} ions for efficient binding and cleavage of 5mC-modified DNA (46). The HNH endonuclease gp74 of phage HK97, however, prefers Ni^{2+} for nicking plasmid and phage DNAs although the exact nicking sites of gp74 have

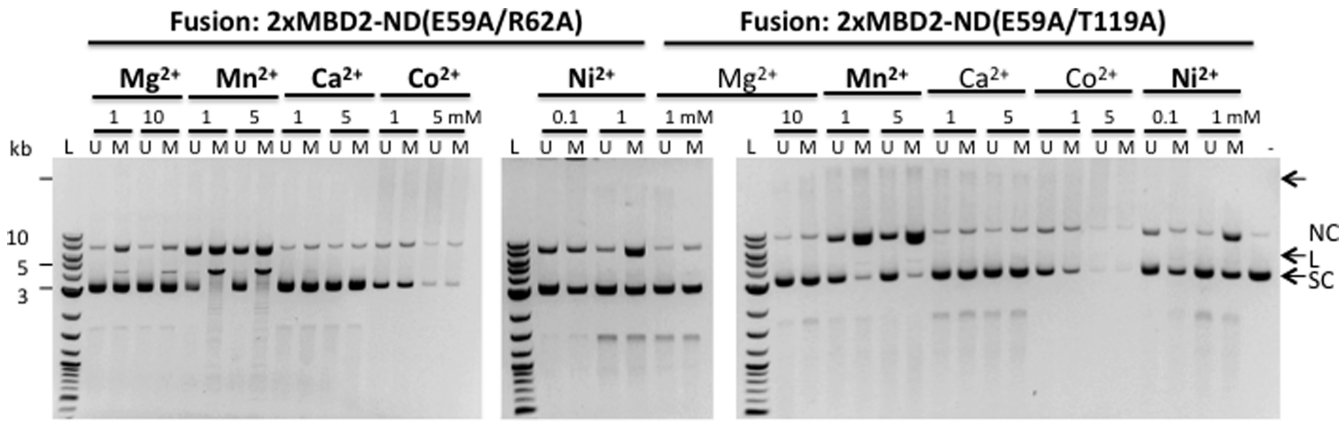

Figure 6. Nicking activity assay in buffers containing different metal ions for the fusion enzymes 2xMBD2-ND(E59A/R62A) and 2xMBD2-ND(E59A/T119A) (nicking domain double mutants). Methylated pBR322 (0.5 μg) by M.SssI (M) or unmodified (U) were digested by 2.5 μg of the fusion enzyme in buffers containing 10 mM Tris-HCl, pH 7.5, 50 mM NaCl and supplemented with different metal ions for 1 h at 37°C. The metal ions and concentration were indicated on top of each lane. Nicking activity was detected in Mg^{2+}, Mn^{2+} and Ni^{2+} buffers for 2xMBD2-ND(E59A/R62A); nicking activity was detected in Mn^{2+} and Ni^{2+} buffers for 2xMBD2-ND(E59A/T119A). These digestions did not include 0.3 M KCl.

not been determined; gp74 of phage HK97 displays poor endonuclease activity in Mg^{2+} buffer (42,47).

Fusion of N.BceSVIII minimal nicking domain to 2xMBD2

The nicking specificity of the wild-type N.BceSVIII is SS↓RT (complement strand: AY↑SS) in Mn^{2+} buffer and it has poor nicking activity in Mg^{2+} buffer (18). By pair-wise aa sequence alignment of N.ϕGamma and N.BceSVIII, the minimal DNA ND of 74-aa was predicted for N.BceSVIII (data not shown). The DNA coding sequence for the 74-aa ND was amplified by PCR and ligated to pTYB1–2xMBD2 (a 10-aa linker was engineered between 2xMBD and the ND). The fusion enzyme was partially purified by chromatography through chitin and heparin columns (Figure 7A). As expected, the fusion enzyme 2xMBD2-ND(N.BceSVIII) nicks DNA poorly in Mg^{2+} buffer (Figure 7B). In the presence of 0.3 M KCl and 10 mM Mn^{2+}, however, the fusion enzyme shows preferential nicking of 5mCG-modified DNA. Some linear DNA was also present after the digestion that may result from two nicks in close proximity on both strands. Figure 7C shows one nicked site in run-off sequencing analysis, indicating the composite site as (CGN_1) ACCC N_5 CG or (CGN_1) ACCCG N_4 CG (a total of five CG sites are present in close proximity).

Nicking of duplex oligos with the composite nicking sites CG N_{14} ACCGG N_{13} CG

FAM-labeled synthetic oligos with the sequence N_6 CG N_{14} ACCGG N_{13} CG N7 was modified by M.SssI and used as the substrate for nicking in Mg^{2+} and Ni^{2+} buffers. The fusion nicking enzymes 2xMBD2-ND(E59A) and 2xMBD2-ND(E59A/R62A) nick modified oligos slightly better in Mg^{2+} buffer supplemented with 0.3 M KCl than unmodified oligos (data not shown). In Ni^{2+} buffer, however, nicking of modified oligos was significantly pronounced than unmodified oligos. Figure 8A and B shows the nicking activity of 2xMBD2-ND(E59A/R62A) on modified (M^+) and unmodified (M^-) duplex oligos in the presence of 0.5 mM Ni^{2+} in a time course (10–60 min, lanes 1–6). The WT enzyme (full-length) N.ϕGamma nicks the oligos more efficiently in Mg^{2+} buffer than in Ni^{2+} buffer regardless of methylation status (Figure 8A, B, lanes 7 and 8). In the control experiment to test modification level, majority of the modified oligos were resistant to HpaII digestion (Figure 8A, lane 10); and the unmodified oligos were digested by HpaII restriction enzyme (Figure 8B, lane 10). Furthermore, in the WT enzyme titration in nicking of modified and unmodified oligos in Ni^{2+} buffer, the same result of preferential nicking of M^+ oligos over M^- oligos was obtained, i.e. WT N.ϕGamma itself prefers to nick 5mCG-modified DNA in Ni^{2+} buffer, a very interesting and unexpected discovery (Figure 8C). Some double-stranded cleavage products were also present (Figure 8A, lane 7; Figure 8C, lanes 2 and 3), suggesting different catalytic mechanisms in the presence of Mg^{2+} and Ni^{2+} ions. Figure 8D (lanes 1–8) shows the enzyme titration (2-fold serial dilution) of 2xMBD2-ND(E59A/R62A) in nicking of modified oligos in Ni^{2+} buffer; the nicked product is proportional to the enzyme concentration. Figure 8E shows that 2xMBD2-ND(E59A/R62A) is active in nicking M^+ oligos in 0.1–5 mM Ni^{2+}. Chelating the divalent cation by EDTA completely inhibited the nicking activity (lane 7, digestion carried out in 10 mM ETDA), while WT N.ϕGamma (full-length) is active in the presence of Mg^{2+}/Ni^{2+} (lane 8) or Mg^{2+} alone (lane 9). In a control for dsDNA cleavage, unmodified oligos were cleaved by HpaII (lane 10). No double-stranded DNA cleavage activity was detected by digestion with fusion enzyme 2xMBD2-ND(E59A/R62A) on modified oligos in the time period tested (30–60 min), although we cannot rule out the possibility that dsDNA breaks may take place in much longer digestion. It was concluded that WT N.ϕGamma prefers to nick/cleave modified DNA in Ni^{2+} buffer. The fusion enzyme 2xMBD2-ND(E59A/R62A) also prefers to nick 5mCG-modified DNA, but lacks dsDNA cleavage activity in Ni^{2+} buffer. Three other fusion enzymes tested (E59A/R68A, E59A/T119A and E59A/K123A) also displayed 5mCG-dependent nicking in Ni^{2+} buffer (data not

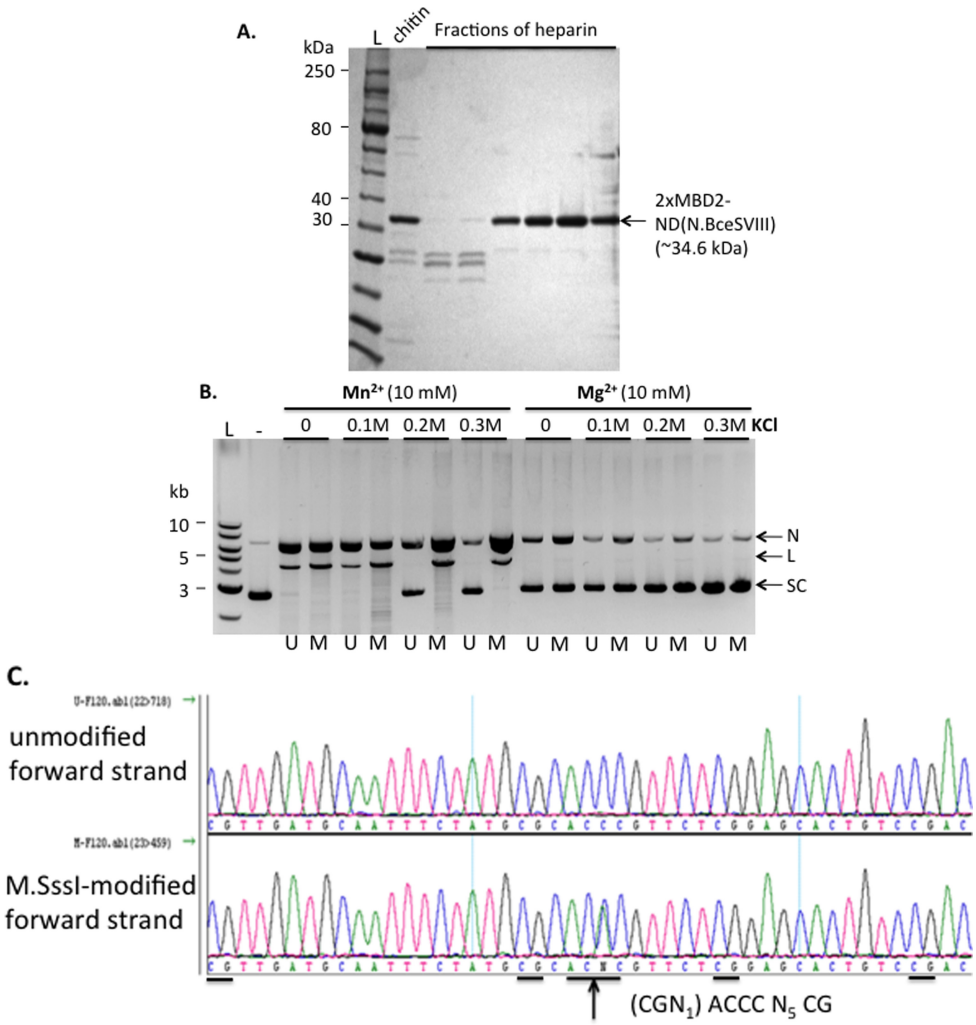

Figure 7. Purification of 2xMBD2-ND(N.BceSVIII) fusion enzyme and nicking activity assay. (**A**) SDS-PAGE analysis of purified fusion enzyme 2xMBD2-ND(N.BceSVIII). The molecular mass of the fusion enzyme is predicted to be ~34.6 kDa. (**B**) Methylated pBR322 DNA (0.5 μg) by M.SssI (M) or unmodified (U) were digested by 2.5 μg of the fusion enzyme in buffers containing 10 mM Tris-HCl, pH 7.8, 50 mM of NaCl supplemented with 10 mM Mn^{2+} or 10 mM Mg^{2+} and 0.1–0.3 M KCl for 1 h at 37°C. NC: nicked circular; L: linear; SC: supercoiled; U: unmodified DNA; M: methylated DNA. (**C**) Run-off sequencing of modified and nicked DNA in Mn^{2+} buffer with 0.3 M KCl. The composite site (CGN_1) ACCC $N_{(5–17)}$ CG is shown. Five CG sites are present in this region.

shown). Further structural and biochemical studies are needed to elucidate the mechanism of 5mCG-dependent endonuclease activity of WT N.ϕGamma and its variant E59A/R62A in buffers with different metal ions. It was found previously that gp74 of phage HK97 nicks plasmid DNA in Ni^{2+} buffer (0.5–10 mM), but prolonged digestion of λ DNA in Ni^{2+} or Zn^{2+} buffers for 2–4 h can completely collapse dsDNA into a smearing pattern, possibly due to introduction of multiple nicks in close proximity (47).

Labeling 5mCG-modified DNA after nicking

DNA synthesis and labeling with fluorescently labeled dNTP can be performed from the nicks on dsDNA (36). M.SssI-modified pBR322 was first nicked by the fusion enzyme 2xMBD2-ND(E59A). Then, strand-displacement synthesis (DNA primer extension) from the nicks was performed by Klenow fragment of *E. coli* DNA polymerase I (or BstI DNA polymerase I large fragment) with dATP, dCTP, dGTP and Cy3-dUTP. The labeled DNA was resolved by agarose gel and detected by fluorescence imaging. Figure 9A shows that the methylated DNA was preferentially nicked and labeled by Cy3-dUMP incorporation into the nascent DNA. The nicked DNA was able to be labeled with Bst DNA pol I large fragment and Cy3-dUTP as well (data not shown). The unmodified DNA was labeled with a lower fluorescence signal probably due to nicks introduced in other ACYGG sites. Overall, the difference in fluorescent signal was significant. When the chromatogram was scanned and digitized, the signal ratio is 9:1 for modified versus unmodified DNA. We also achieved nicking and labeling for mouse genomic DNA since the DNA was naturally modified by DNMT1, DNMT3a and DNMT3b in the cells. But the background DNA labeling was more pronounced, probably due to the random shearing/nicking during genomic DNA preparation (Figure 9B). In a negative control, λ DNA (unmodified at CG sites, modified at

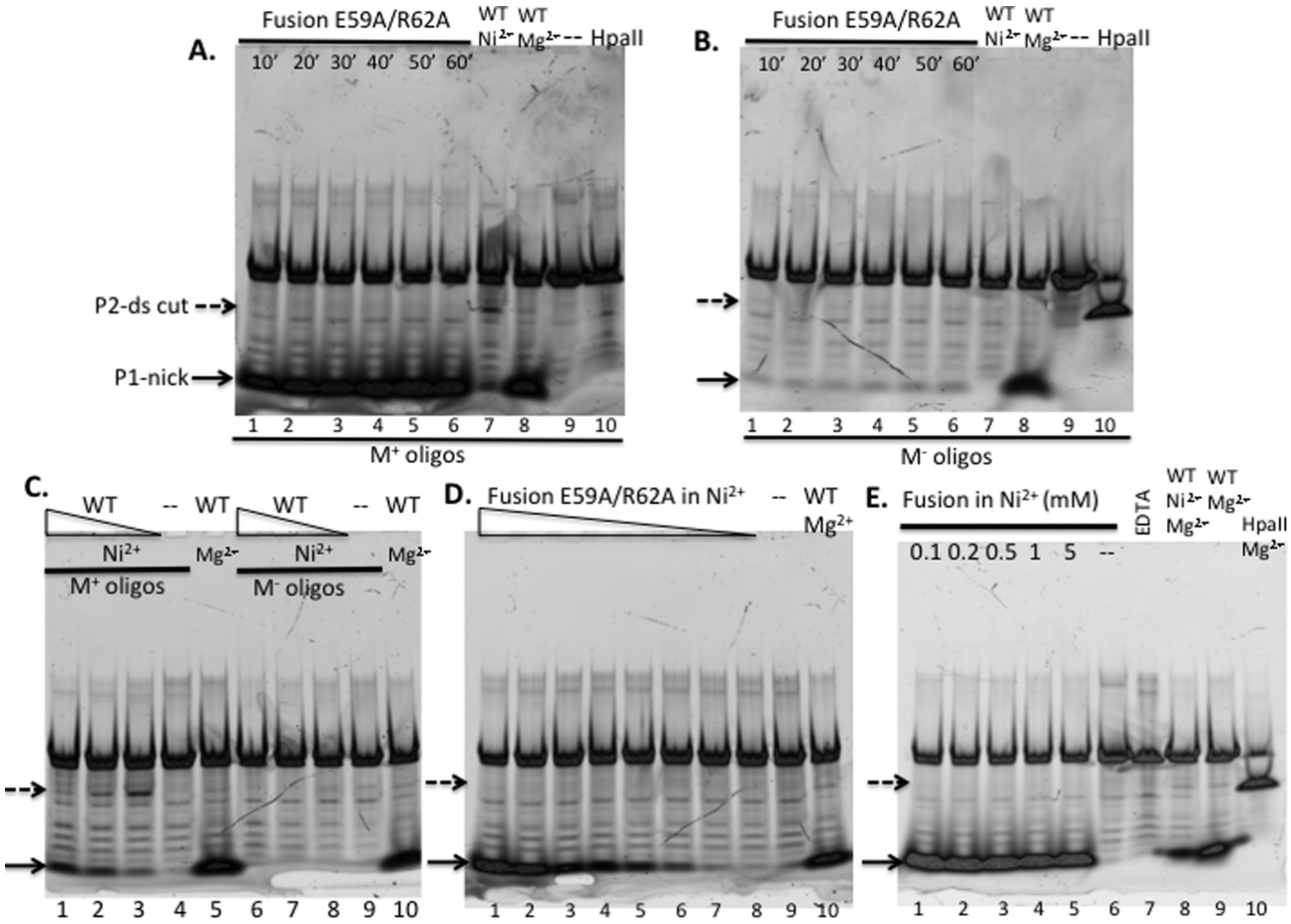

Figure 8. Digestion 5mCG-modified and unmodified duplex oligos in Ni^{2+} buffer by the fusion enzyme 2xMBD2-ND(E59A/R62A), or in Mg^{2+} and Ni^{2+} buffers by the WT nicking enzyme N.ϕGamma (full-length). (**A**) Digestion of M^+ oligos by the fusion enzyme 2xMBD2-ND(E59A/R62A) in Ni^2 buffer, by WT N.ϕGamma in Ni^{2+} or Mg^{2+} buffer, and by HpaII restriction enzyme in buffer 4, respectively. Arrow indicates nicked product and dashed arrow indicates dsDNA cleavage product. (**B**) Digestion of unmodified (M^-) oligos by the fusion enzyme 2xMBD2-ND(E59A/R62A) in Ni^{2+} buffer, by WT N.ϕGamma in Ni^{2+} or Mg^{2+} buffer, and by HpaII restriction enzyme in buffer 4, respectively. (**C**) Digestion of M^+ and M^- oligos by the WT enzyme N.ϕGamma (in 2-fold serial dilution) in Ni^{2+} buffer at 37°C for 30 min. (**D**) Digestion of M^+ oligos by the fusion enzyme in 2xMBD2-ND(E59A/R62A) in 2-fold serial dilution in Ni^{2+} buffer at 37°C for 30 min. (**E**) Digestion of M^+ oligos by the fusion enzyme in 2xMBD2-ND(E59A/R62A) in Ni^{2+} buffer (0.1–5 mM) at 37°C for 30 min. Lane 7 digestion included 1 mM Ni^{2+} and 10 mM EDTA. Lane 10, unmodified oligos cleaved by HpaII restriction enzyme.

CCWGG and GATC by Dcm and Dam methylases, respectively) was nicked by the fusion enzyme and subjected to primer extension/labeling. The nicked λ DNA gave rise to similar level of labeling compared to untreated DNA. In principle, the Cy3-dUMP-incoporated genomic DNA can be stretched in microfluidic-channels and be visualized at the labeled composite sites under fluorescence microscope, a process of optical imaging and mapping of epigenetic marks on normal and cancer DNAs.

Fusion of nicking domain of N.ϕGamma to SRA domain and Zinc finger protein with 5mC preference, and fusion of 2xMBD2 to FokI cleavage domain

We also constructed a fusion enzyme between the SRA domain and the ND E59A (N.ϕGamma). The SRA domain (SET and RING associated (SRA) domain of mouse UHRF1) was reported to preferentially bind to hemi-methylated DNA (48). The fusion nickase was affinity-purified from a chitin column and used to nick hemi-methylated and fully methylated duplex oligos. We found that the fusion enzyme SRA-ND(E59A of N.ϕGamma) was able to nick both hemi-methylated and fully methylated DNA (data not shown). A digestion condition for preferential nicking of hemi-methylated DNA remains to be optimized.

A Zinc finger protein, Zfp57, is able to bind methylated site TGC5mCGC that contains a 5mCG dinucleotide (49). The fusion of Zfp57 DNA binding domain with the NDs of N.ϕGamma or N.BceSVIII or gp74 of phage HK97 can potentially generate 5mC-dependent nickases with rare specificities (6+4.5 = 10.5 bp). TAL effector proteins that can bind to methylated sites have been published previously (50). But construction of fusion nickases targeting methylated sites remains to be carried out.

Another strategy to construct 5mCG-dependent nicking enzyme is to use a catalytic-deficient FokI cleavage domain in fusion with 5mCG-binding domain proteins (51). Such fusion protein can bind to 5mCG sites, but it cannot cleave or nick. FokI mutants that can form transient

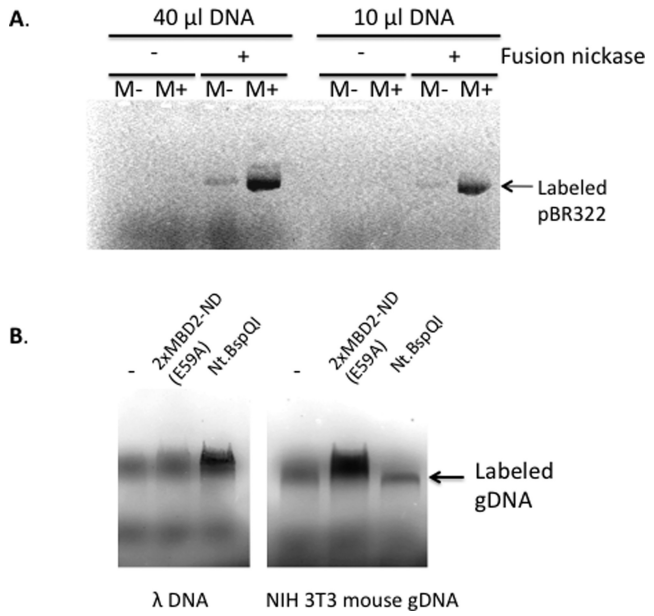

Figure 9. Labeling of nicked plasmid and genomic DNAs by strand-displacement DNA synthesis. (**A**) One microgram of pBR322 methylated by M.SssI (M+) or unmodified (M−) was nicked by 2.5 µg of 2xMBD2-ND(E59A) in Mg^{2+} buffer containing 0.25 M KCl for 1 h at 37°C. The labeling was performed through 5 min incubation at 25°C with 10 units of Klenow fragment of *E. coli* DNA polymerase I and a dNTP mix containing Cy3-dUTP. (**B**) NIH 3T3 mouse genomic DNA was nicked by 2.5 µg of 2xMBD2-ND(E59A) in a Mg^{2+} buffer containing 0.3 M KCl for 1 h at 37°C. DNA labeling was performed through 10 min incubation at 25°C with 10 units of Klenow fragment of *E. coli* DNA polymerase I and a dNTP mix containing Cy3-dUTP. The control enzyme Nt.BspQI nicks the DNA sites GCTCTTCN↓N.

dimer and produce strand-specific nicks have been constructed (12). A catalytic-proficient, dimerization-proficient and DNA binding-deficient full-length FokI mutant can be added and rescue the activity of MBD-FokIR* (cleavage domain R*: catalytic-deficient) by formation of a transient dimer complex: MBD-FokIR*/FokI full-length mutant R^+B^- (cleavage-proficient, binding-deficient). Such engineered enzyme complex may produce 5mCG-specific nicks on DNA. Alternatively, one can engineer a single polypeptide of FokIR* and FokIR$^+$ by fusion of catalytic-deficient and catalytic-proficient cleavage domains. The forced FokIR*/R$^+$ dimer can be subsequently fused to 5mCG-binding domain or 5mC-binding ZFP to create 5mC-specific nicking enzymes. We had constructed a fusion protein of 2xMBD2-FokI* (catalytic-deficient), but the fusion protein was poorly expressed and only small amount of protein can be purified from 2L of IPTG-induced cells. Optimized expression condition and purification method remains to be discovered. In theory, other DNA cleavage/NDs from DNA endonucleases, nickases, integrases, recombinases, transposases, HEases and DNA repair enzymes can be fused to the 5mCG-binding domain of MBD proteins and generate 5mCG-specific cleavage/nicking enzymes.

CONCLUSION

We isolated 54 single mutants and 3 double mutants of N.ϕGamma mainly targeting charged and hydrophilic aa residues or predicted catalytic residues. The importance of the predicted catalytic residues D89, H90, N106 and H115 was confirmed by mutagenesis. The catalytic mutants D89A, H90A, N106A and H115A are inactive in Mg^{2+} buffer, but the activity of H90A, N106A and H115A can be partially restored in Mn^{2+} buffer. Double mutants of catalytic residues of H90, N106 and H115 are inactive in Mn^{2+} buffer, indicating the H90, N106, and H115 and other nearby residues can play an accessory role in Mn^{2+} metal ion binding and activation of water molecule. The NDs of E59A and some double mutants were fused to 5mCG-binding domain of MBD2 and generated fusion enzymes that preferentially nick 5mCG-modified DNA. The CG dinucleotide can be modified by C5 MTases such as M.SssI, M.HhaI or M.HpaII to create composite sites (CG Nx) ACYGG $N_{(8–15)}$ CG. Other C5 MTases with 5–8 bp specificity can potentially be used to create rare composite nicking sites. We found that WT N.ϕGamma and fusion enzymes with ND derived from N.ϕGamma prefer to nick 5mCG-modified oligos in Ni^{2+} buffer despite sub-optimal nicking condition. We also constructed a fusion nicking enzyme 2xMBD2-ND(N.BceSVIII) targeting more frequent composite nicking sites in Mn^{2+} buffer. The fusion enzymes can be used to nick and label 5mCG-modified DNA with fluorescently labeled Cy3-dUTP and potentially be useful for DNA optical mapping and DNA sequencing since a doublet peak would appear at the nicking sites near 5mCG.

ACKNOWLEDGMENTS

We are grateful to Rich Roberts and Bill Jack for critical reading of the manuscript. We thank Nancy Badger for performing DNA sequencing and James Mcfarland for providing the chitin beads. We thank Don Comb and Jim Ellard for support.

FUNDING

New England Biolabs, Inc. Funding for open access charge: New England Biolabs, Inc.

REFERENCES

1. Pingoud,A. and Jeltsch,A. (2001) Structure and function of type II restriction endonucleases. *Nucleic Acids Res.*, **29**, 3705–3727.
2. Roberts,R.J. (2005) How restriction enzymes became the workhorses of molecular biology. *Proc. Natl. Acad. Sci. USA*, **102**, 5905–5908.
3. Loenen,W.A. and Raleigh,E.A. (2014) The other face of restriction: modification-dependent enzymes. *Nucleic Acids Res.*, **42**, 56–69.
4. Xu,S.Y., Zhu,Z., Zhang,P., Chan,S.H., Samuelson,J.C., Xiao,J., Ingalls,D. and Wilson,G.G. (2007) Discovery of natural nicking endonucleases Nb.BsrDI and Nb.BtsI and engineering of top-strand nicking variants from BsrDI and BtsI. *Nucleic Acids Res.*, **35**, 4608–4618.
5. Higgins,L.S., Besnier,C. and Kong,H. (2001) The nicking endonuclease N.BstNBI is closely related to type IIs restriction endonucleases MlyI and PleI. *Nucleic Acids Res.*, **29**, 2492–2501.
6. Chan,S.H., Zhu,Z., Van Etten,J.L. and Xu,S.Y. (2004) Cloning of CviPII nicking and modification system from chlorella virus NYs-1 and application of Nt.CviPII in random DNA amplification. *Nucleic Acids Res.*, **32**, 6187–6199.

7. Xia,Y.N., Morgan,R., Schildkraut,I. and Van Etten,J.L. (1988) A site-specific single strand endonuclease activity induced by NYs-1 virus infection of a Chlorella-like green alga. *Nucleic Acids Res.*, **16**, 9477–9487.
8. Kachalova,G.S., Rogulin,E.A., Yunusova,A.K., Artyukh,R.I., Perevyazova,T.A., Matvienko,N.I., Zheleznaya,L.A. and Bartunik,H.D. (2008) Structural analysis of the heterodimeric type IIS restriction endonuclease R.BspD6I acting as a complex between a monomeric site-specific nickase and a catalytic subunit. *J. Mol. Biol.*, **384**, 489–502.
9. Landthaler,M., Shen,B.W., Stoddard,B.L. and Shub,D.A. (2006) I-BasI and I-HmuI: two phage intron-encoded endonucleases with homologous DNA recognition sequences but distinct DNA specificities. *J. Mol. Biol.*, **358**, 1137–1151.
10. Too,P.H., Zhu,Z., Chan,S.H. and Xu,S.Y. (2010) Engineering Nt.BtsCI and Nb.BtsCI nicking enzymes and applications in generating long overhangs. *Nucleic Acids Res.*, **38**, 1294–1303.
11. McConnell Smith,A., Takeuchi,R., Pellenz,S., Davis,L., Maizels,N., Monnat,R.J. Jr and Stoddard,B.L. (2009) Generation of a nicking enzyme that stimulates site-specific gene conversion from the I-AniI LAGLIDADG homing endonuclease. *Proc. Natl. Acad. Sci. USA*, **106**, 5099–5104.
12. Sanders,K.L., Catto,L.E., Bellamy,S.R. and Halford,S.E. (2009) Targeting individual subunits of the FokI restriction endonuclease to specific DNA strands. *Nucleic Acids Res.*, **37**, 2105–2115.
13. Francia,M.V., Clewell,D.B., de la Cruz,F. and Moncalian,G. (2013) Catalytic domain of plasmid pAD1 relaxase TraX defines a group of relaxases related to restriction endonucleases. *Proc. Natl. Acad. Sci. USA*, **110**, 13606–13611.
14. Kim,E., Kim,S., Kim,D.H., Choi,B.S., Choi,I.Y. and Kim,J.S. (2012) Precision genome engineering with programmable DNA-nicking enzymes. *Genome Res.*, **22**, 1327–1333
15. Ramirez,C.L., Certo,M.T., Mussolino,C., Goodwin,M.J., Cradick,T.J., McCaffrey,A.P., Cathomen,T., Scharenberg,A.M. and Joung,J.K. (2012) Engineered zinc finger nickases induce homology-directed repair with reduced mutagenic effects. *Nucleic Acids Res.*, **40**, 5560–5568.
16. Gabsalilow,L., Schierling,B., Friedhoff,P., Pingoud,A. and Wende,W. (2013) Site- and strand-specific nicking of DNA by fusion proteins derived from MutH and I-SceI or TALE repeats. *Nucleic Acids Res.*, **41**, e83.
17. Roberts,R.J., Vincze,T., Posfai,J. and Macelis,D. (2010) REBASE–a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.*, **38**, D234–D236.
18. Xu,S.Y. and Gupta,Y.K. (2013) Natural zinc ribbon HNH endonucleases and engineered zinc finger nicking endonuclease. *Nucleic Acids Res.*, **41**, 378–390.
19. Kleanthous,C., Kuehlmann,U.C., Pommer,A.J., Ferguson,N., Radford,S.E., Moore,G.R., James,R. and Hemmings,A.M. (1999) Structural and mechanistic basis of immunity toward endonuclease colicins. *Nat. Struct. Biol.*, **6**, 243–252.
20. Cheng,Y.S., Hsia,K.C., Doudeva,L.G., Chak,K.F. and Yuan,H.S. (2002) The crystal structure of the nuclease domain of colicin E7 suggests a mechanism for binding to double-stranded DNA by the H-N-H endonucleases. *J. Mol. Biol.*, **324**, 227–236.
21. Xu,S.Y., Kuzin,A.P., Seetharaman,J., Gutjahr,A., Chan,S.H., Chen,Y., Xiao,R., Acton,T.B., Montelione,G.T. and Tong,L. (2013) Structure determination and biochemical characterization of a putative HNH endonuclease from *Geobacter metallireducens* GS-15. *PLoS One*, **8**, e72114.
22. Galburt,E.A., Chevalier,B., Tang,W., Jurica,M.S., Flick,K.E., Monnat,R.J. Jr and Stoddard,B.L. (1999) A novel endonuclease mechanism directly visualized for I-PpoI. *Nature Struct. Biol.*, **6**, 1096–1099.
23. Shen,B.W., Landthaler,M., Shub,D.A. and Stoddard,B.L. (2004) DNA binding and cleavage by the HNH homing endonuclease I-HmuI. *J. Mol. Biol.*, **342**, 43–56.
24. Bird,A. (2002) DNA methylation patterns and epigenetic memory. *Genes Dev.*, **16**, 6–21.
25. Bestor,T.H. (2000) The DNA methyltransferases of mammals. *Hum. Mol. Genet.*, **9**, 2395–2402.
26. Tahiliani,M., Koh,K.P., Shen,Y., Pastor,W.A., Bandukwala,H., Brudno,Y., Agarwal,S., Iyer,L.M., Liu,D.R., Aravind,L. *et al.* (2009) Conversion of 5-methylcytosine to 5-hydroxymethylcytosine in mammalian DNA by MLL partner TET1. *Science*, **324**, 930–935.
27. He,Y.F., Li,B.Z., Li,Z., Liu,P., Wang,Y., Tang,Q., Ding,J., Jia,Y., Chen,Z., Li,L. *et al.* (2011) Tet-mediated formation of 5-carboxylcytosine and its excision by TDG in mammalian DNA. *Science*, **333**, 1303–1307.
28. Sjolund,A.B., Senejani,A.G. and Sweasy,J.B. (2013) MBD4 and TDG: multifaceted DNA glycosylases with ever expanding biological roles. *Mutation Res.*, **743–744**, 12–25.
29. Warnecke,P.M. and Bestor,T.H. (2000) Cytosine methylation and human cancer. *Curr. Opin. Oncol.*, **12**, 68–73.
30. Jones,P.A. and Baylin,S.B. (2007) The epigenomics of cancer. *Cell*, **128**, 683–692.
31. Gaudet,F., Hodgson,J.G., Eden,A., Jackson-Grusby,L., Dausman,J., Gray,J.W., Leonhardt,H. and Jaenisch,R. (2003) Induction of tumors in mice by genomic hypomethylation. *Science*, **300**, 489–492.
32. Ma,X., Wang,Y.W., Zhang,M.Q. and Gazdar,A.F. (2013) DNA methylation data analysis and its application to cancer research. *Epigenomics*, **5**, 301–316.
33. Delpu,Y., Cordelier,P., Cho,W.C. and Torrisani,J. (2013) DNA methylation and cancer diagnosis. *Int. J. Mol. Sci.*, **14**, 15029–15058.
34. Farkas,S.A., Milutin-Gasperov,N., Grce,M. and Nilsson,T.K. (2013) Genome-wide DNA methylation assay reveals novel candidate biomarker genes in cervical cancer. *Epigenetics*, **8**, 1213–1225.
35. Besaratinia,A., Cockburn,M. and Tommasi,S. (2013) Alterations of DNA methylome in human bladder cancer. *Epigenetics*, **8**, 1013–1022.
36. Zhang,P., Too,P.H., Samuelson,J.C., Chan,S.H., Vincze,T., Doucette,S., Backstrom,S., Potamousis,K.D., Schramm,T.M., Forrest,D. *et al.* (2010) Engineering BspQI nicking enzymes and application of N.BspQI in DNA labeling and production of single-strand DNA. *Protein Express. Purif.*, **69**, 226–234.
37. Hastie,A.R., Dong,L., Smith,A., Finklestein,J., Lam,E.T., Huo,N., Cao,H., Kwok,P.Y., Deal,K.R., Dvorak,J. *et al.* (2013) Rapid genome mapping in nanochannel arrays for highly complete and accurate de novo sequence assembly of the complex Aegilops tauschii genome. *PLoS One*, **8**, e55864.
38. Oster,C.J. and Phillips,G.J. (2011) Vectors for ligation-independent construction of lacZ gene fusions and cloning of PCR products using a nicking endonuclease. *Plasmid*, **66**, 180–185.
39. Jorgensen,H.F., Adie,K., Chaubert,P. and Bird,A.P. (2006) Engineering a high-affinity methyl-CpG-binding protein. *Nucleic Acids Res.*, **34**, e96.
40. Xu,S.-y. and Schildkraut,I. (1991) Isolation of BamHI variants with reduced cleavage activities. *J. Biol. Chem.*, **266**, 4425–4429.
41. Shim,J., Humphreys,G.I., Venkatesan,B.M., Munz,J.M., Zou,X., Sathe,C., Schulten,K., Kosari,F., Nardulli,A.M., Vasmatzis,G. *et al.* (2013) Detection and quantification of methylation in DNA using solid-state nanopores. *Sci. Rep.*, **3**, 1389, doi: 10.1038/srep01389
42. Moodley,S., Maxwell,K.L. and Kanelis,V. (2012) The protein gp74 from the bacteriophage HK97 functions as a HNH endonuclease. *Prot. Sci.*, **21**, 809–818
43. Thielking,V., Selent,U., Kohler,E., Landgraf,Z., Wolfes,H., Alves,J. and Pingoud,A. (1992) Mg2+ confers DNA binding specificity to the EcoRV restriction endonuclease. *Biochemistry*, **31**, 3727–3732.
44. Vasu,K., Nagamalleswari,E., Zahran,M., Imhof,P., Xu,S.Y., Zhu,Z., Chan,S.H. and Nagaraja,V. (2013) Increasing cleavage specificity and activity of restriction endonuclease KpnI. *Nucleic Acids Res.*, **41**, 9812–9824.
45. Chan,S.H., Opitz,L., Higgins,L., O'Loane,D. and Xu,S.Y. (2010) Cofactor requirement of HpyAV restriction endonuclease. *PLoS One*, **5**, e9071.
46. Liu,G., Ou,H.Y., Wang,T., Li,L., Tan,H., Zhou,X., Rajakumar,K., Deng,Z. and He,X. (2010) Cleavage of phosphorothioated DNA and methylated DNA by the type IV restriction endonuclease ScoMcrA. *PLoS Genetics*, **6**, e1001253.
47. Moodley,S. (2010) Biochemical investigation of the bacteriophage protein HK97 gp74. *M.S. Thesis*, University of Toronto, 1–102.
48. Bostick,M., Kim,J.K., Esteve,P.O., Clark,A., Pradhan,S. and Jacobsen,S.E. (2007) UHRF1 plays a role in maintaining DNA methylation in mammalian cells. *Science*, **317**, 1760–1764.
49. Liu,Y., Toh,H., Sasaki,H., Zhang,X. and Cheng,X. (2012) An atomic model of Zfp57 recognition of CpG methylation within a specific DNA sequence. *Genes Dev.*, **26**, 2374–2379.

50. Deng,D., Yin,P., Yan,C., Pan,X., Gong,X., Qi,S., Xie,T., Mahfouz,M., Zhu,J.K., Yan,N. *et al.* (2012) Recognition of methylated DNA by TAL effectors. *Cell Res.*, **22**, 1502–1504.

51. Waugh,D.S. and Sauer,R.T. (1993) Single amino acid substitutions uncouple the DNA binding and strand scission activities of Fok I endonuclease. *Proc. Natl. Acad. Sci. USA*, **90**, 9596–9600.