Supplementary Material for

Activity and structure of EcoKMcrA

Honorata Czapinska[1], Monika Kowalska[1], Evelina Zagorskaitė[2], Elena Manakova[2], Anton Slyvka[1], Shuang-yong Xu[3], Virginijus Siksnys[2], Giedrius Sasnauskas[2,#], Matthias Bochtler[1,4,#]

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

[2]*Institute of Biotechnology, Vilnius University, Saulėtekio av. 7, 10257 Vilnius, Lithuania*

[3]*New England Biolabs, Inc. 240 County Road, Ipswich, MA 01938, USA*

[4]*Institute of Biochemistry and Biophysics PAS, Pawinskiego 5a, 02-106 Warsaw, Poland*

Correspondence to mbochtler@iimcb.gov.pl or gsasnaus@ibt.lt

Supplementary methods

Analytical gel filtration: The gel filtration was carried out at room temperature on an AKTA FPLC system using a Superdex 200 10/300 GL column (GE Healthcare) pre-equilibrated with 20 mM Tris-HCl (pH 8.0 at 25 °C), 0.5 M KCl and 5 mM 2-mercaptoethanol. The samples (2 μM wt EcoKMcrA dimer or 4 μM EcoKMcrA-N) were prepared in 0.1 ml of the same buffer. Elution from the column (flow rate 0.5 ml/min) was monitored by measuring A_{220}. The calibration line was generated by measuring the elution volumes of a series of standards of known molecular mass (Gel Filtration Calibration Kit from GE Healthcare). The molecular weights of proteins were calculated by interpolating its elution volumes onto the calibration line.

Plasmid and phage DNA cleavage assays. The mixtures contained 0.05 or 0.5 μM wt EcoKMcrA or its H229A variant and 0.2 or 0.5 μg DNA (reaction volume 25 μl). Reactions were carried out using supercoiled pACYC184 plasmid; supercoiled pACYC184_M.HpaII plasmid (M.HpaII methylated DNA); single-stranded M13mp18 DNA (all isolated from either a dcm+ or dcm– *E. coli* strain); supercoiled plasmid pBR322; or phage lambda DNA (isolated from either Dam+/Dcm+ or Dam-/Dcm- *E. coli* strains). Cleavage was stopped by phenol extraction. Samples were analyzed by agarose electrophoresis and visualized by ethidium bromide staining.

Digestion of mixed DNA substrates. The reaction mixture contained 10 μl mixed DNA (~1 μg total DNA): 0.36 μg phage XP12 DNA, 0.34 μg 5hmC containing PCR DNA, 0.35 μg cytosine containing PCR DNA and 0, 0.5, 1, 2, 4 μl of EcoKMcrA stock (175 μM). Digestion was performed in a buffer containing 50 mM KCl, 10 mM DTT, 50 mM Tris-HCl pH 8.0 supplemented with 1 mM $MnCl_2$ at 37 °C for 1.5 h. The reaction was stopped by the addition of 5 μl of Protease K for 15 min at 37 °C and then stop solution. Cleavage products were visualized on 1% agarose gel.

Digestion of oligonucleotides containing multiple modified CCGG sites. The 37 and 52 bp long oligonucleotides containing two unmodified, hemi- or fully methylated CCGG target sites were purchased from Genomed (37_X/37_X and 52_X/52_X in Table S1). The reactions were carried out for 3 h at 37 °C in the buffer containing 50 mM Tris-HCl, pH 8.0, 50 mM KCl, 0.5 mM $MnCl_2$ and 1 mM DTT. After the incubation samples were alkali treated, mixed with formamide and loaded on 20% Urea-PAGE gel. The DNA size marker was created by the controlled DNase I digestion of the 37 and 52 bp oligoduplexes.

Phage restriction assay: C2566 *E. coli* cells carrying either wild type EcoKMcrA or its H228A, H229A or N-terminal variants (residues 1-174), or REM14, were cultured in 6 ml of Phage broth+Amp to log-phase and concentrated 10-fold in phage broth+Amp. 0.2 ml cells and 4 ml soft top agar were used to pour plates on Rich+Amp. The plates were air dried for 15 min. 8 μl of diluted phage Lambda, T4gt (5hmC) and T4 (g5hmC) were spotted onto the cells lawns, air dried for 10 min and incubated overnight at 37 °C.

Run-off sequencing: Run-off sequencing was carried out as previously described (1), using the BigDye™ Terminator v3.1 Cycle Sequencing Kit (ThermoFisher/Applied Biosystems), containing an engineered Taq DNA polymerase.

Fluorescence measurements: Pyrrolocytosine steady state fluorescence measurements in solution were performed at 25 °C on a Fluoromax-3 spectrofluorimeter. Excitation wavelength was 350 nm, emission wavelength was 450 nm (5 nm slits). The samples contained 0.2 μM hemimodified 12 bp oligoduplex with or without the proteins (1 μM EcoKMcrA (dimer), 2 μM EcoKMcrA-N monomer) in 20 mM Tris-HCl, pH 8.0, and 100 mM KCl. Measurements were also made with 0.2 μM of the single-stranded pyrrolocytosine-containing oligonucleotide (corresponding to the modified strand of the 12 bp oligoduplex).

Cloning, expression and purification of the EcoKMcrA - N.ϕGamma nickase fusion. The expression vector for chitin-binding domain (CBD) tagged EcoKMcrA-gHNH and EcoKMcrA(H252A)-gHNH fusion proteins was made by ligating the synthetic genes into pTYB1 plasmid (NdeI and XhoI cut) using the NEB Gibson assembly mix. The C-terminal gHNH domain corresponded to the previously described attenuated variant of the 76 amino acid nuclease domain of the N.ϕGamma nickase of *Bacillus anthracis* (2), which was connected to EcoKMcrA or EcoKMcrA(H252A) with a six amino acid (GASGAS) linker. Cells were grown at 37 °C to OD600 of 0.7, induced by the addition of 0.5 mM IPTG. Protein expression was carried out overnight at 20 °C. The EcoKMcrA-gHNH fusion proteins were purified by affinity chromatography via a chitin column. Elution was after DTT cleavage at 4 °C overnight. Proteins were concentrated by ultrafiltration, and stored in 20 mM Tris-HCl, 200 mM KCl, 1 mM DTT, and 50% v/v glycerol, pH 7.5.

Supplementary Figures

Fig. S1

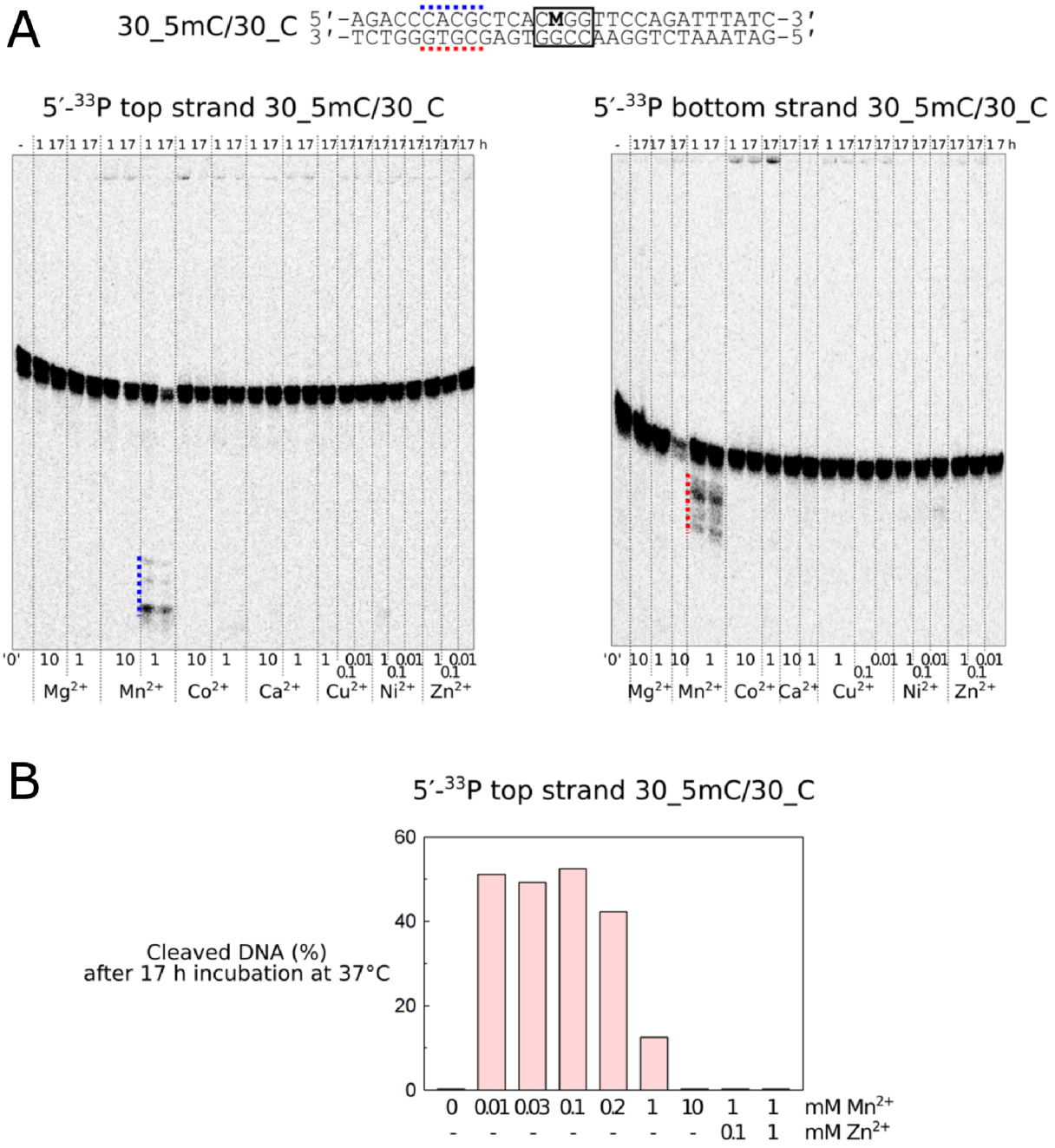

Fig. S1: Catalytic activity of EcoKMcrA in the presence of divalent metal ions. **(A)** Reactions of wt EcoKMcrA (0.5 µM dimer) on 0.2 µM radiolabeled oligoduplex DNA 30_5mC/30_C (DNA sequence shown at the top, 'M' designates 5-methylcytosine). Incubation times in hours are indicated above the gel lanes, divalent metal ions and their concentrations in mM are shown below gel lanes. Sample '0' is untreated DNA. Approximate cleavage positions observed in the presence of Mn^{2+} are marked by blue (top strand) or red (bottom strand) dotted lines. **(B)** Quantification of 30_5mC/30_C DNA top strand cleavage by wt EcoKMcrA at various Mn^{2+} and Zn^{2+} concentrations.

Fig. S2

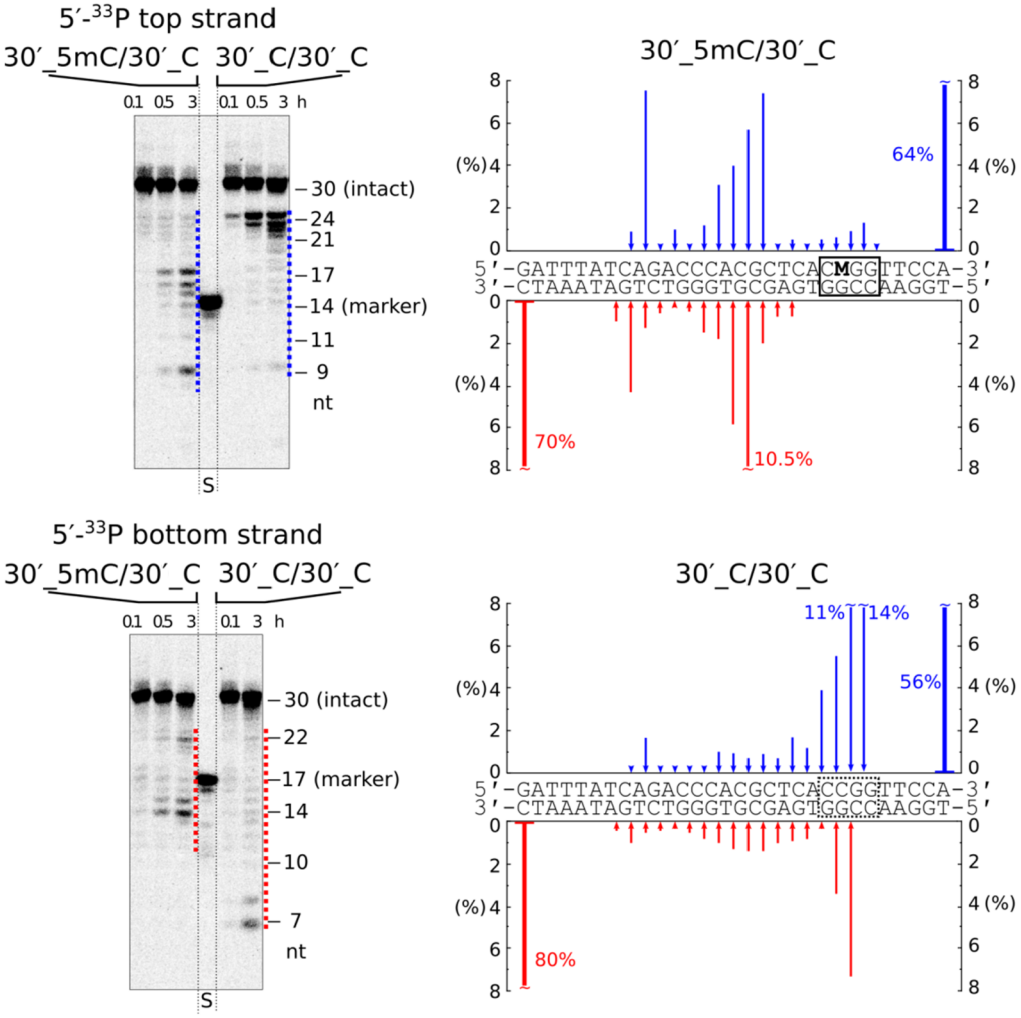

Fig. S2: Cleavage of hemi-methylated and non-modified oligoduplex: Reactions were performed at 37 °C with 0.5 µM wild type EcoKMcrA (dimer) and 10 nM hemimethylated (30′_5mC/30′_C) or unmethylated (30′_C/30′_C) oligoduplex DNA in a buffer supplemented with 0.1 mM Mn^{2+}. Top and bottom strand cleavage positions on the gels are marked by blue and red dotted lines, respectively. Gel lanes 'S' contained radiolabeled single-stranded oligonucleotides that correspond to the 5′-terminal fragments of the respective DNA strands (sizes in nucleotides are shown on the sides of the gels). The amounts of the top and bottom strand cleavage products after 3 hours of cleavage are plotted as blue/red arrows along the oligoduplex sequences.

Fig. S3

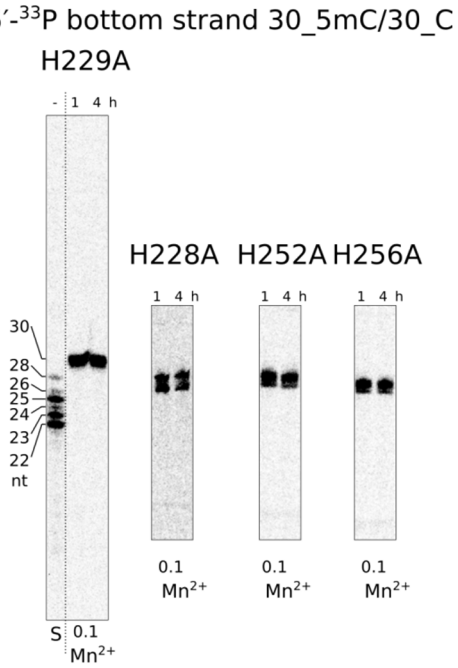

Fig. S3: Inactivity of EcoKMcrA mutants on oligoduplex DNA. The reactions contained 0.5 µM (dimer) EcoKMcrA mutants and 0.2 µM 30_5mC/30_C DNA (radiolabel on the bottom strand) in a reaction buffer supplemented with 0.1 mM $MnCl_2$. Reactions were performed for up to 4 hours at 37 °C.

Fig. S4

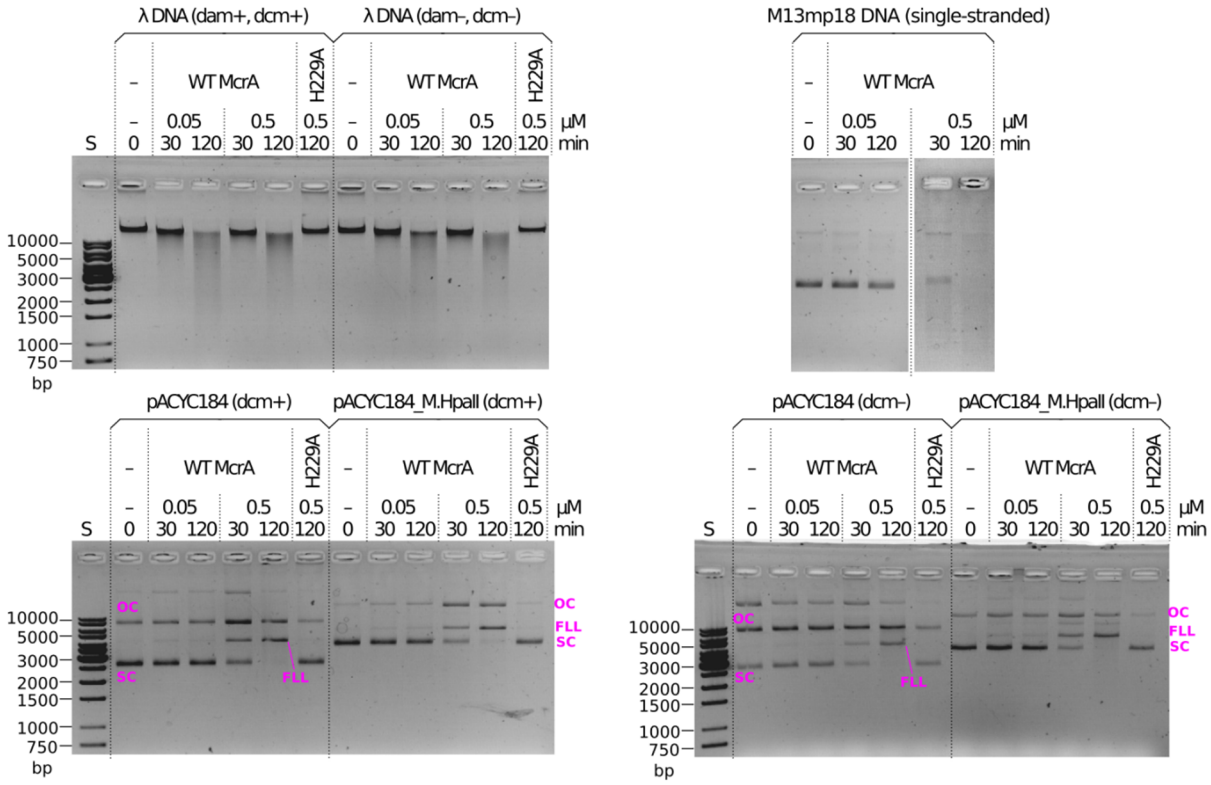

Fig. S4: ***In vitro*** **phage and plasmid DNA cleavage by EcoKMcrA.** The reactions were performed on phage λ DNA (isolated from either a (dam+, dcm+) or (dam–, dcm–) *E. coli* strain), single-stranded M13mp18 DNA, supercoiled plasmid pACYC184 and supercoiled plasmid pACYC184_M.HpaII, each plasmid isolated from either a dcm+ or dcm– *E. coli* strain. Gel lane 'S' contained DNA size marker. The positions of supercoiled, nicked (one or multiple nicks), and linear (a single double-strand break) DNA forms are marked as 'SC', 'OC' and 'FLL'. pACYC184 (dcm+ and dcm–) contained a detectable amount of heavier, presumably dimeric DNA forms. The reactions contained 0.05 or 0.5 µM wild type EcoKMcrA, 0.5 µM H229A EcoKMcrA variant and 0.5 µg / 25 µl DNA in a reaction buffer supplemented with 0.1 mM $MnCl_2$. The incubations were performed at 37 °C.

Fig. S5

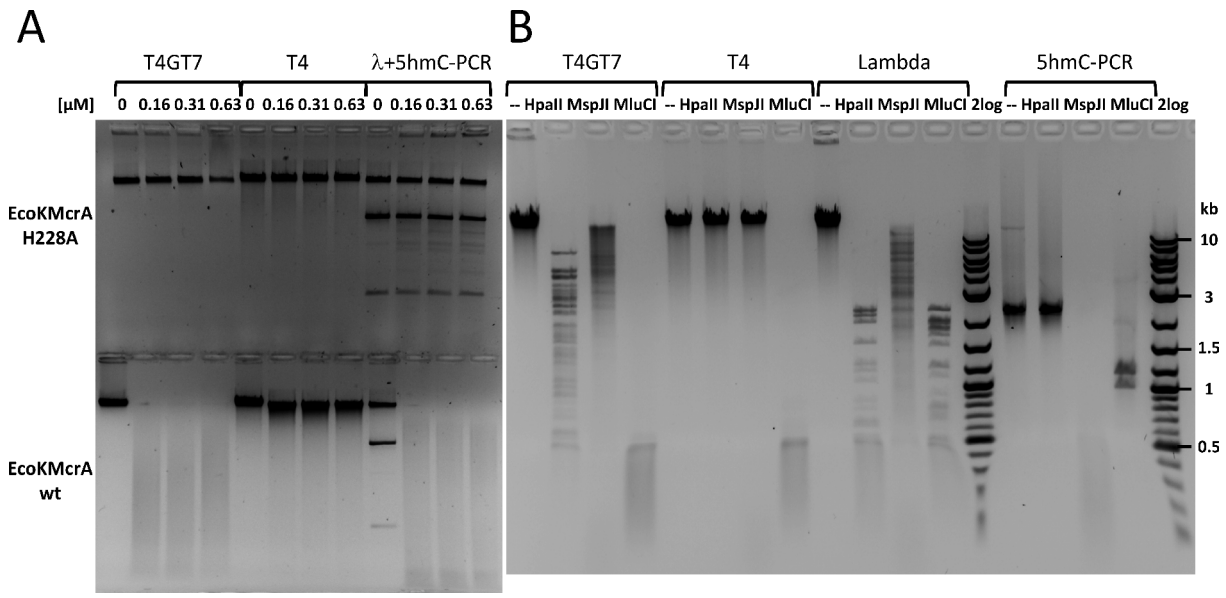

Fig. S5: ***In vitro*** **phage and PCR DNA cleavage by EcoKMcrA (wt and catalytic mutant) and control enzymes.** The DNA of T4GT7, T4 and lambda phage together with the 5hmC containing PCR products was digested by **(A)** EcoKMcrA or its active site variant; **(B)** control HpaII, MspJI and MluCI enzymes. HpaII REase is blocked by 5hmC modification in the C/CGG context, MspJI is selective for 5mC modified DNA (5mCNNRN9/) and MluCI (/AATT) is unaffected by the presence of cytosine modifications. T4GT7 and lambda DNA is cleaved by all enzymes which indicates that although the two phages do not by themselves introduce cytosine modifications, their DNA may contain some cytosines modified in the C5mCWGG context due to a passage in a dcm+ host. T4 DNA contains glucosyl-5hmC residues that block the activity of EcoKMcrA, MspJI and HpaII and thus is cleaved only by the cytosine modification insensitive MluCI enzyme. 5hmC containing PCR DNA is cleaved by all enzymes but HpaII, as predicted. The EcoKMcrA active site variant retained residual activity only.

Fig. S6

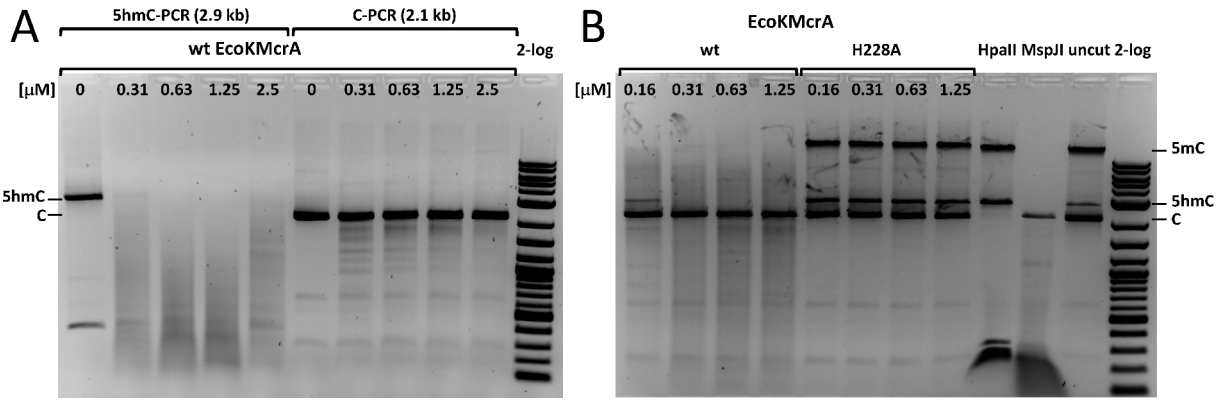

Fig. S6: *In vitro* cleavage of (A) separate and (B) mixed DNA substrates by EcoKMcrA (wild type and catalytic mutant) and control enzymes. Unmodified PCR DNA (2.1 kb), 5hmC containing PCR DNA (2.9 kb) and 5mC containing XP12 phage DNA were digested by EcoKMcrA. Digestion was performed in 50 mM KCl, 10 mM DTT, 50 mM Tris-HCl, pH 8.0 supplemented with 1 mM $MnCl_2$ at 37 °C for 1.5 h. Wild type EcoKMcrA cleaved 5mC and 5hmC containing DNA but not the unmodified PCR product. The H228A active site variant of EcoKMcrA did not exhibit any activity. The activity of the control enzymes agreed with their substrate preferences: HpaII endonuclease only cleaved unmodified DNA, whereas MspJI had very similar cleavage properties to wt EcoKMcrA and cleaved only modified DNAs.

Fig. S7

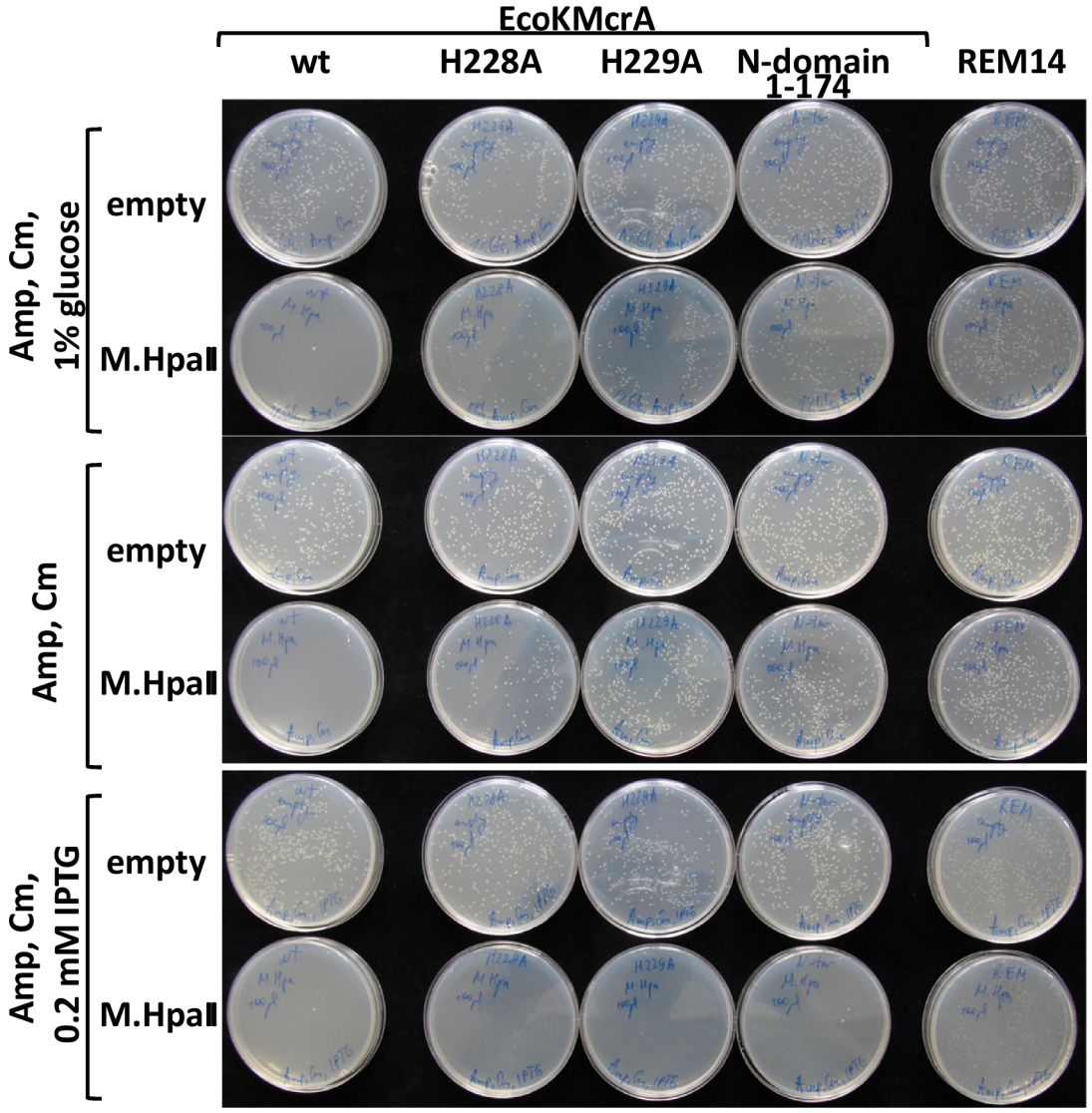

Fig. S7: Plasmid restriction assay. One of the repeats of the restriction assay of pACYC184 plasmid carrying a chloramphenicol resistance gene was assayed in BL21(DE3) (McrA-) *E. coli* cells expressing wt EcoKMcrA, catalytic variants of the enzyme, its N-terminal fragment lacking the nuclease domain or an unrelated protein (REM14). The plasmid was either empty or carried a gene for M.HpaII methyltransferase and thus was 5mC modified in the C5mCGG sequence context.

Fig. S8

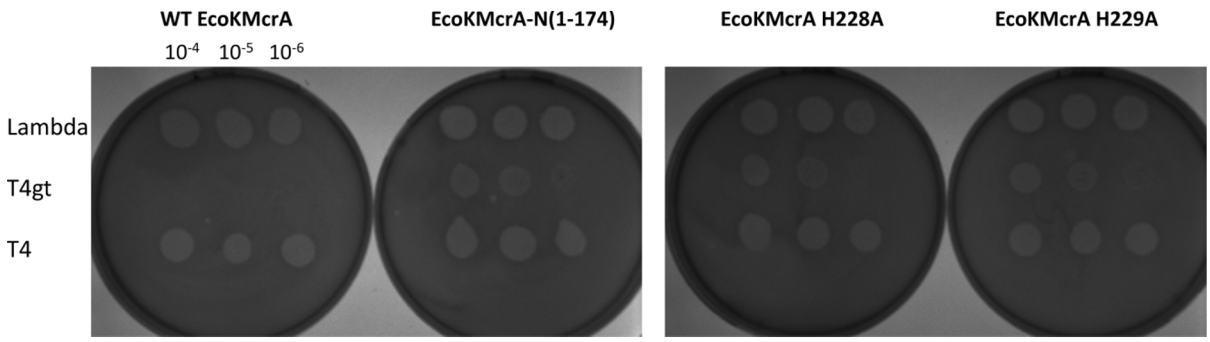

Fig. S8: Phage restriction assay. 8 µl of Lambda (C), T4gt (5hmC) and T4 (g5hmC) phages were plagued onto the lawns of C2566 (McrA-) *E. coli* cells expressing either wild type EcoKMcrA or its catalytic variants (without IPTG induction). Under low expression conditions catalytically competent EcoKMcrA restricted the 5hmC containing T4gt phage much more efficiently than its inactive variants. The phages containing unmodified or g5hmC modified DNA were unaffected by the presence of the enzyme.

Fig. S9

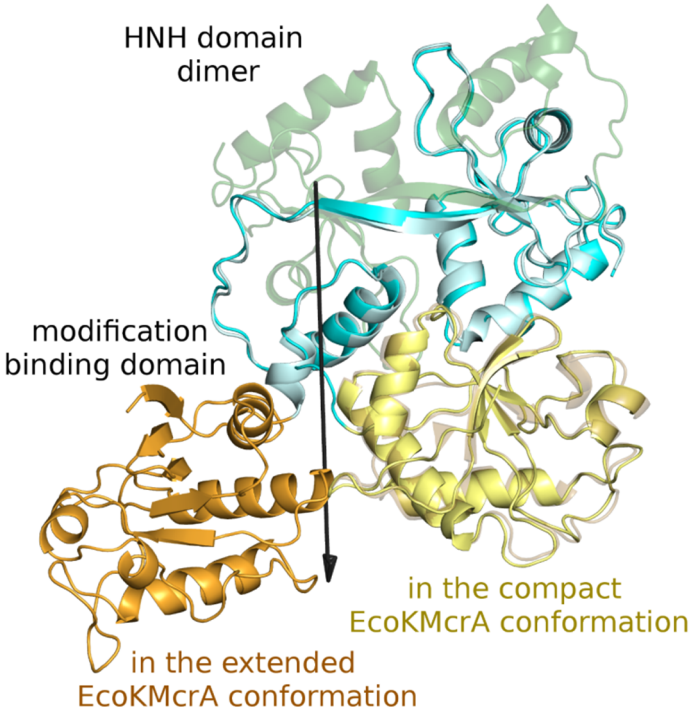

Fig. S9: EcoKMcrA conformational flexibility. The EcoKMcrA crystals contain a full dimer in the asymmetric unit. When the C-terminal HNH domains are overlaid (cyan), the N-terminal domains (orange and yellow) adopt different orientations. According to DynDom (3) a 150° rotation is required to map one N-terminal domain onto the other. The N-terminal domains were also separately superposed to show that they adopt the same conformation (faint yellow).

Fig. S10

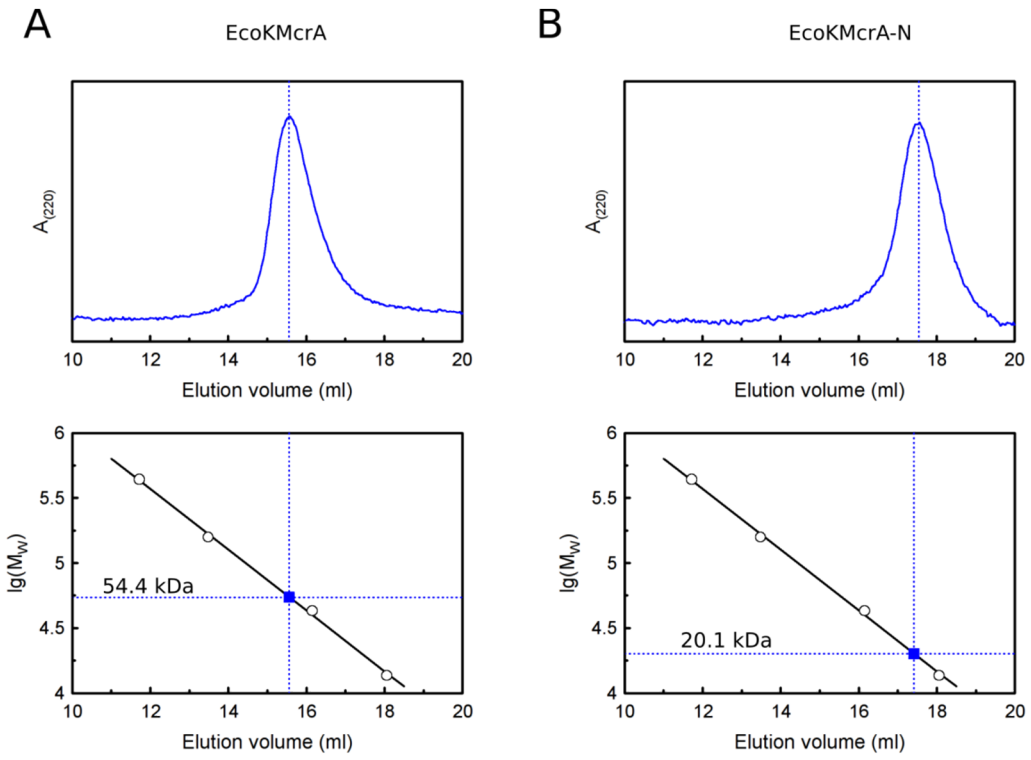

Fig. S10: Gel filtration of wt EcoKMcrA and its N-terminal fragment EcoKMcrA-N. (A) The apparent MW of EcoKMcrA determined by gel-filtration (15.55 ml elution volume, equivalent to 54.4 kDa, blue square) is close to the theoretical mass of EcoKMcrA dimer (64.5 kDa). (B) The apparent MW of the N-terminal EcoKMcrA fragment EcoKMcrA-N (17.41 ml elution volume, equivalent to 20.1 kDa, blue square), is a close match to the theoretical mass of the EcoKMcrA-N monomer (20.6 kDa). The proteins used as MW standards were ribonuclease A (13.7 kDa), ovalbumin (43 kDa), ferritin (440 kDa), aldolase (158 kDa).

Fig. S11

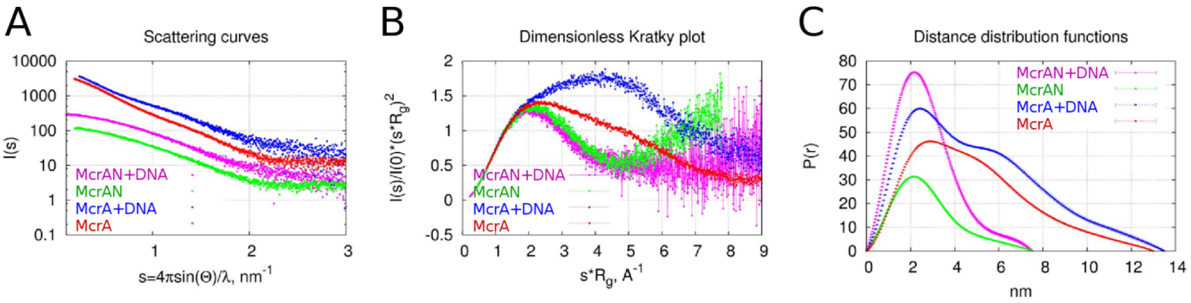

Fig. S11: Small-angle X-ray scattering data for apo-EcoKMcrA (red), EcoKMcrA+DNA (blue), apo-EcoKMcrA-N (green) and EcoKMcrA-N with DNA (magenta). (A) Scattering data; (B) Kratky plot, normalized by Rg and I(0) parameters (4); (C) distance distribution function.

Fig. S12

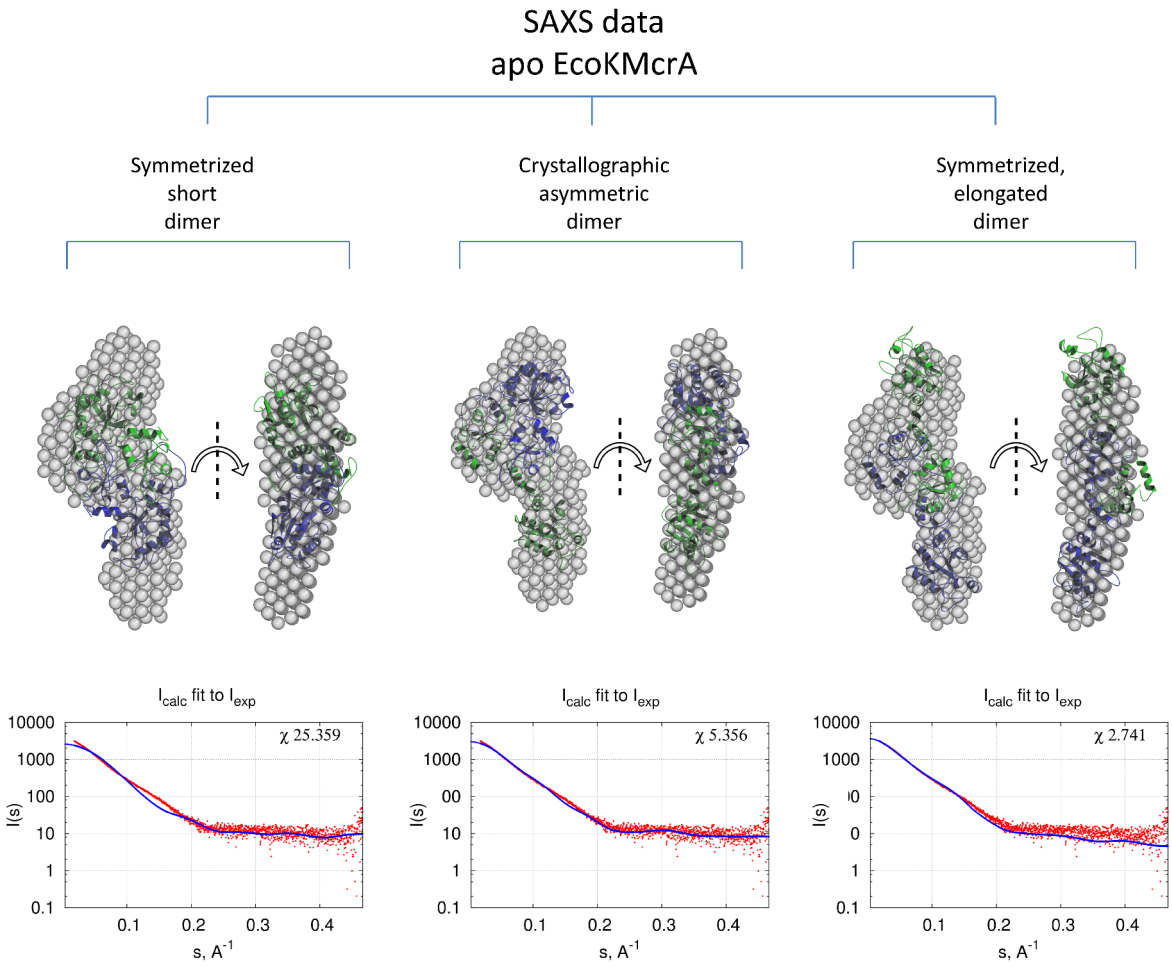

Fig. S12: Comparison of small-angle X-ray scattering data with structure based predictions for the dimer observed in the crystal or for two-fold symmetric models based on the conformation of either subunit in the asymmetric unit. The top panel illustrates the agreement in real space, the bottom panel shows the agreement in reciprocal space. The *ab initio* SAXS model of the EcoKMcrA dimer was calculated without imposing two-fold symmetry restraints.

Fig. S13

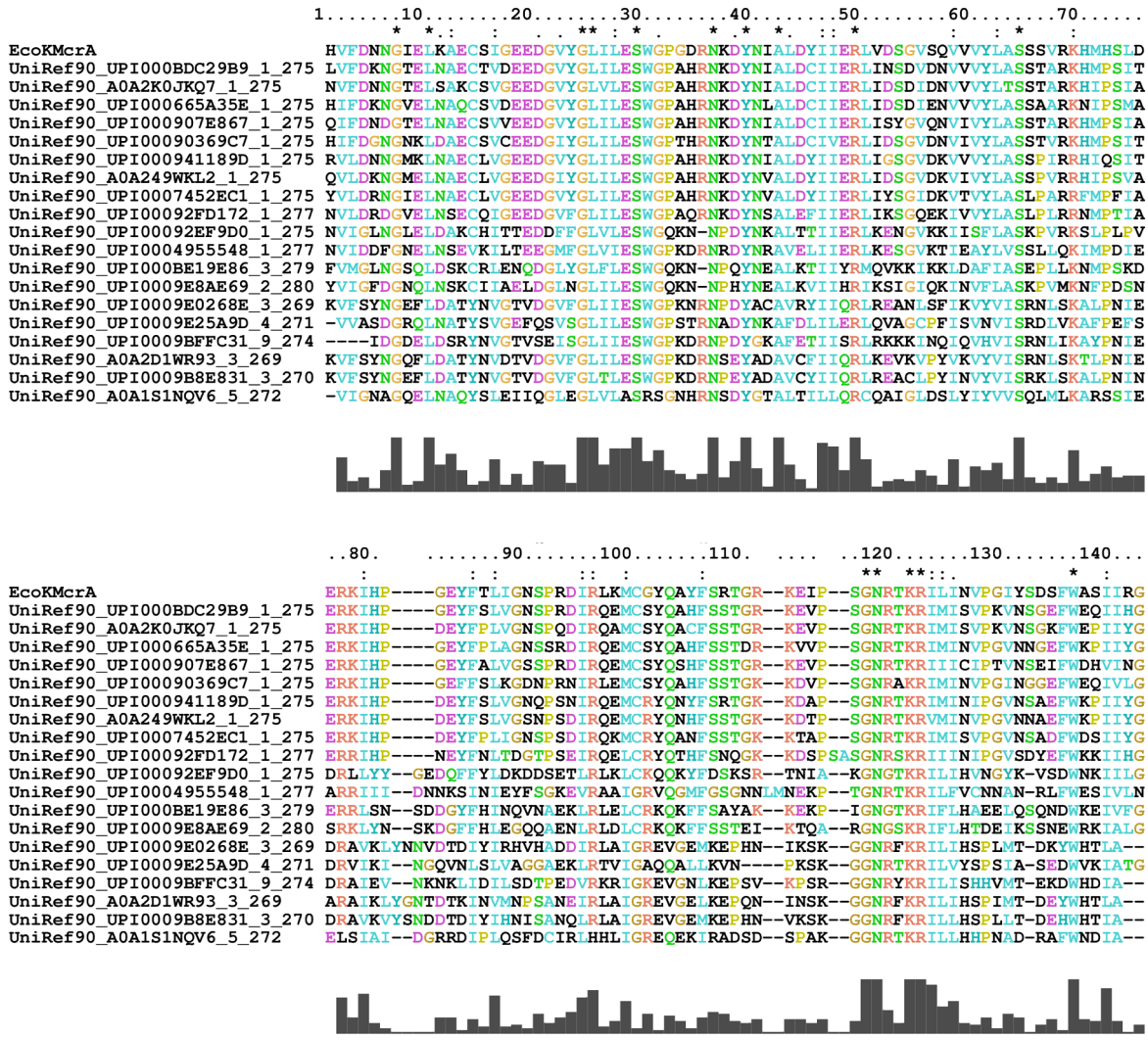

Fig. S13: Sequence alignment of the EcoKMcrA N-terminal domain used for the calculation of the conservation scores presented in Fig. 3. The standard ConSurf (5) parameters were used to generate the alignment. The conservation scores were very similar irrespective of whether the full length EcoKMcrA or just its N-terminal domain was used as a query.

Fig. S14

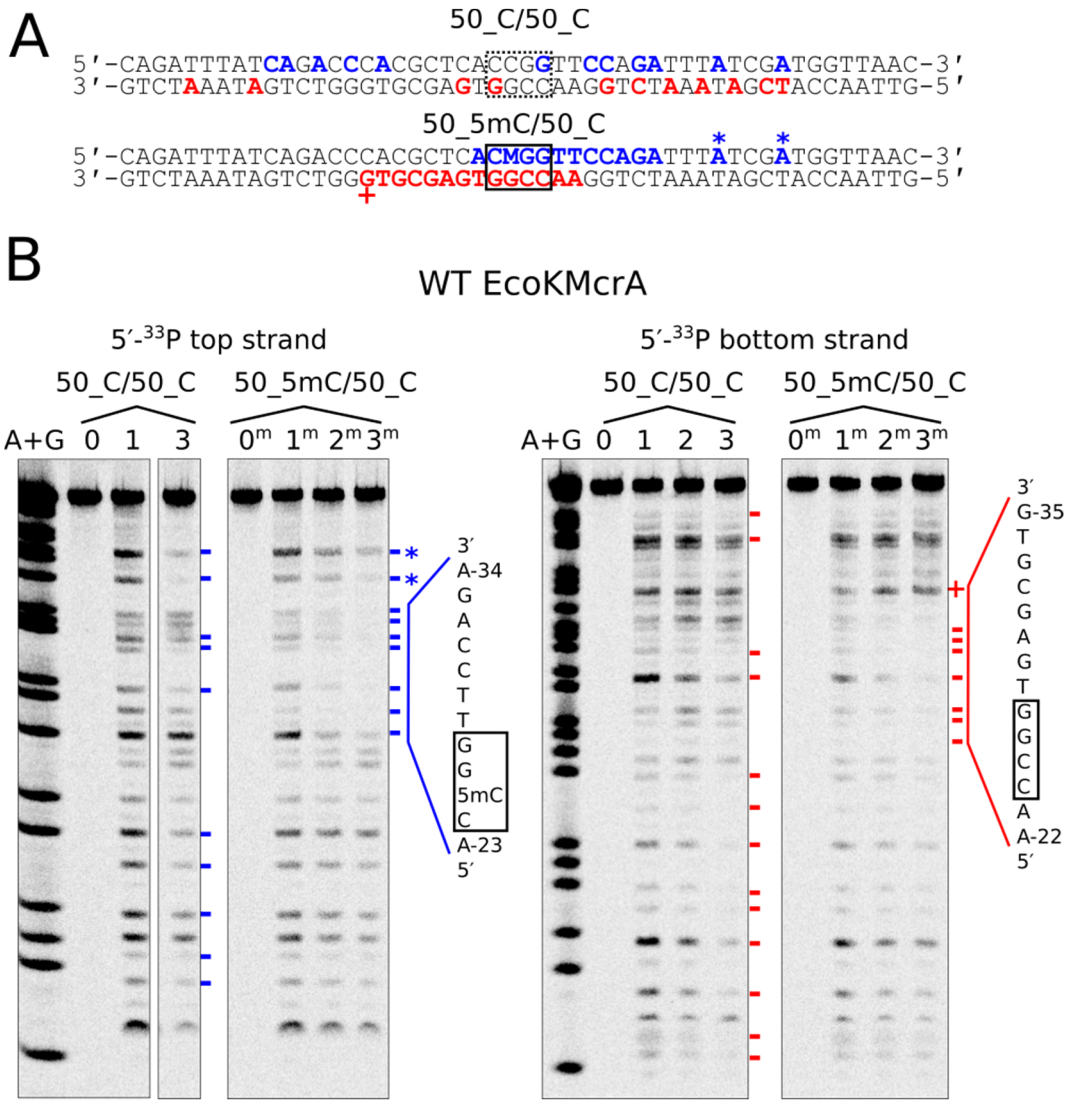

Fig. S14: DNase I footprint of wt EcoKMcrA. (A) The unmethylated and hemimethylated DNA oligoduplexes used for DNase I footprinting experiments. 'M' marks 5-methylcytosine. The EcoKMcrA recognition sequence in hemimethylated DNA is marked by a black box; an equivalent region in unmethylated DNA is marked by a dotted line. The positions protected from DNase I cleavage by EcoKMcrA are shown in bold font. **(B)** DNA protection by wt EcoKMcrA. DNase I footprinting experiment with radiolabeled 5′-terminus of either top (left) or bottom (right) DNA strand. Gel lanes '0' contained untreated unmethylated DNA, '1' – the DNA treated with DNase I in the absence of EcoKMcrA, lanes '2' and '3' – the DNA treated with DNase I in the presence of 0.25 μM and 0.5 μM EcoKMcrA homodimer. Gel lanes '0m', '1m', '2m' and '3m' contained analogous samples prepared with hemimethylated DNA. Positions protected from DNase I cleavage upon EcoKMcrA binding are marked by blue (top strand) or red (bottom strand) dashes. The sequences of the protected regions in the hemimethylated DNA are shown on the right-hand side of the gels. Asterisks (*) mark top strand positions distal from the methylated site that are protected on both unmethylated and methylated DNAs. The bottom strand position G-35, which in hemimethylated DNA becomes more susceptible to DNase I treatment upon EcoKMcrA binding, is marked by a '+' sign.

Fig. S15

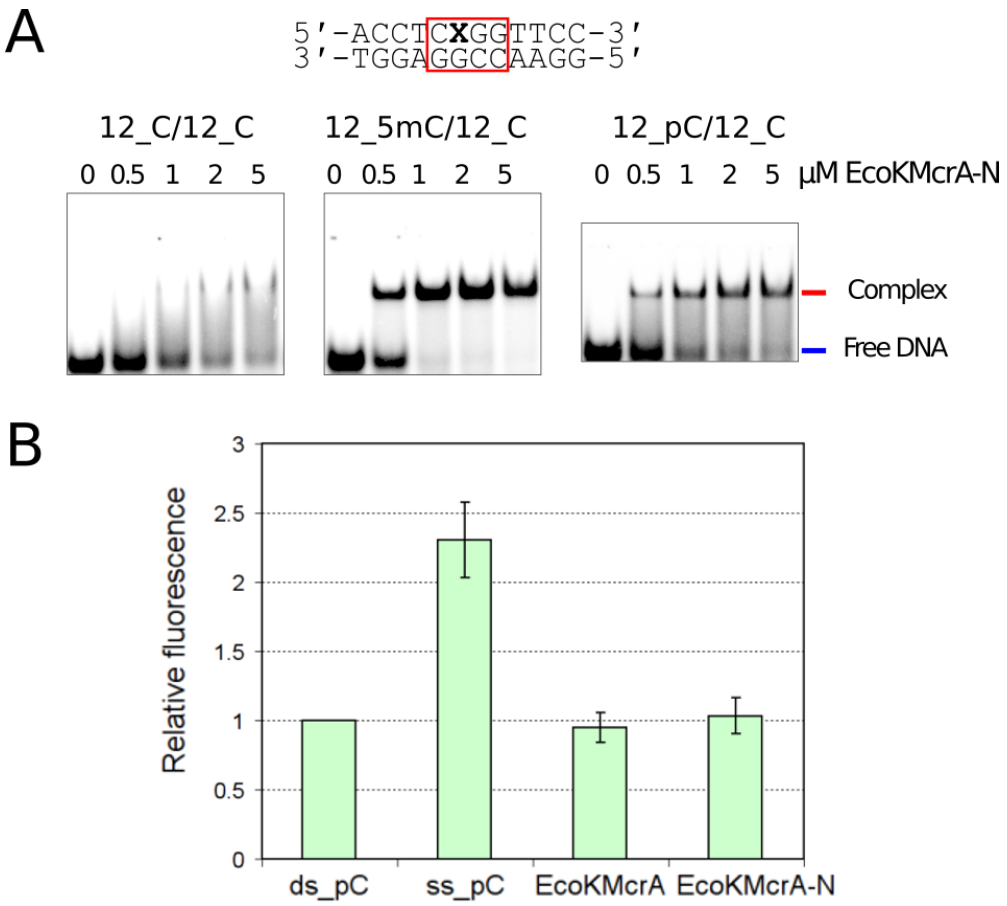

Fig. S15: EcoKMcrA-N interaction with pyrrolocytosine containing DNA. (A) Electrophoretic mobility shift assay performed with unmodified and hemimodified 12 bp DNA oligoduplexes. The DNA sequence is shown at the top of the panel, 'X' designates unmodified C (oligoduplex 12_C/12_C), 5mC (12_5mC/12_C), or pyrrolocytosine (12_pC/12_C). DNA concentration was 0.5 µM, protein concentrations are indicated above gel lanes. Experiments were performed in a 40 mM Tris-acetate pH 8.3 buffer as described in Materials and Methods. Positions of free DNA and the specific protein-DNA complex are marked. **(B)** Pyrrolocytosine steady state fluorescence measurements in solution. The 'EcoKMcrA-N' and 'EcoKMcrA' samples contained the respective protein (1 µM EcoKMcrA dimer, 2 µM EcoKMcrA-N monomer) and 0.2 µM oligoduplex 12_pC/12_C. Control sample 'ds_pC' contained 0.2 µM of the double-stranded oligonucleotide 12_pC/12_C, the sample 'ss_pC' contained 0.2 µM of the modified (top) 12_pC strand. The emission intensities of the 'EcoKMcrA-N', 'EcoKMcrA' and 'ss_pC' samples were normalized against the 'ds_pC' sample and are presented as the average value of 3 independent measurements ±SD.

Fig. S16:

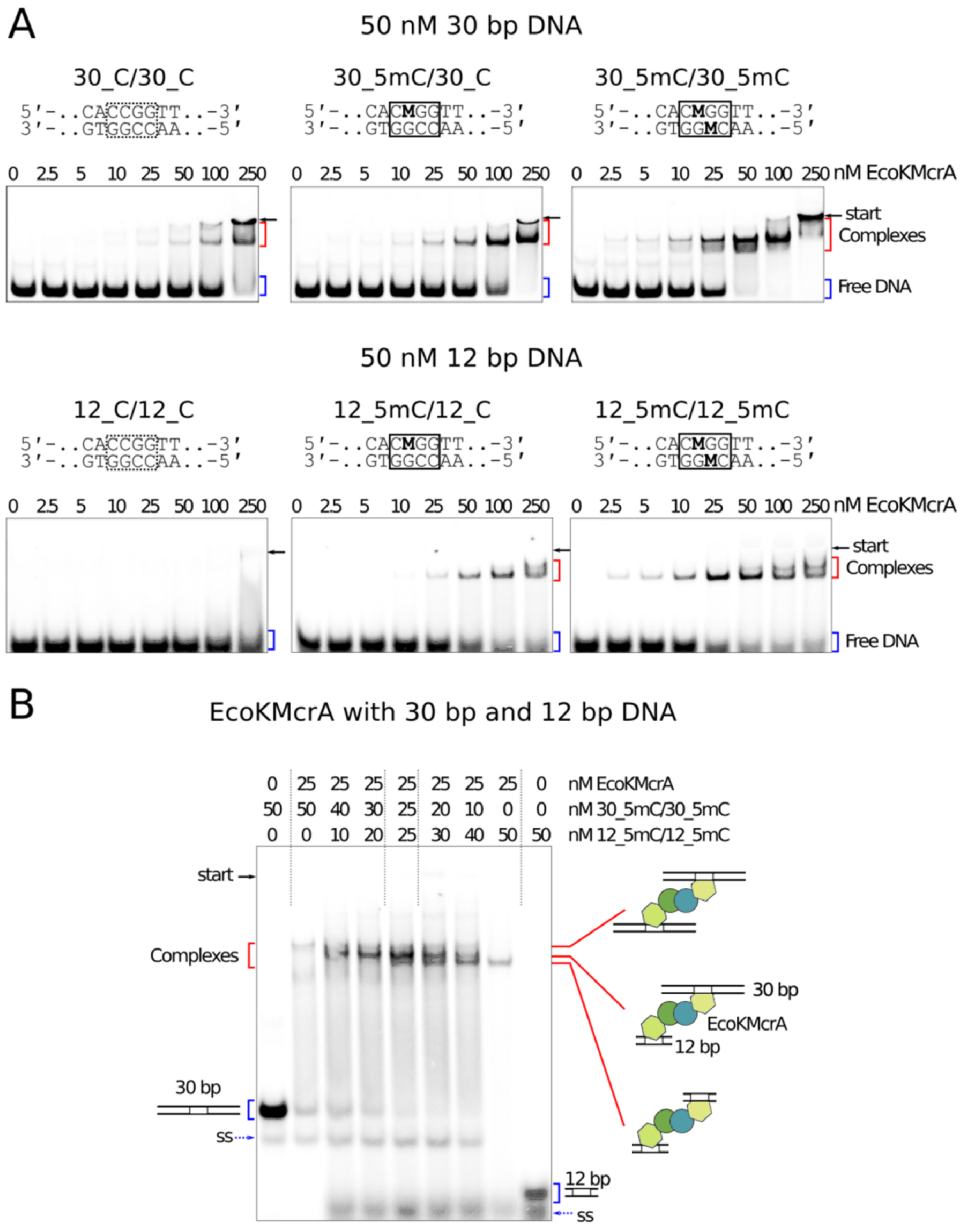

Fig. S16: Electrophoretic mobility shift assay of DNA binding by EcoKMcrA. (A) EMSA with unmodified, hemi- or fully methylated 30 or 12 bp DNA (central sequences shown on top; 'M' denotes 5mC), DNA concentration was 50 nM, concentrations of EcoKMcrA dimer are indicated above gel lanes. Positions of free DNA and protein-DNA complexes are marked by blue and red brackets, respectively. The gels were run for 1 h at 5 V/cm. **(B)** EMSA with a mixture of radiolabeled 30 bp (30_5mC/30_5mC) and 12 bp (12_5mC/12_5mC) DNA oligoduplexes. Concentrations of EcoKMcrA and DNAs are indicated above gel lanes. The gel was run for 3 h at 5 V/cm. Cartoons depict two types of unbound oligoduplexes (12 and 30 bp) and three types of protein-DNA complexes, containing two 30 bp, two 12 bp, or one of each DNAs. EcoKMcrA N- and C-terminal domains are depicted as blue hexagons and red circles, respectively. Bands of the single-stranded forms of the respective oligonucleotides are marked by blue dotted lines.

Fig. S17:

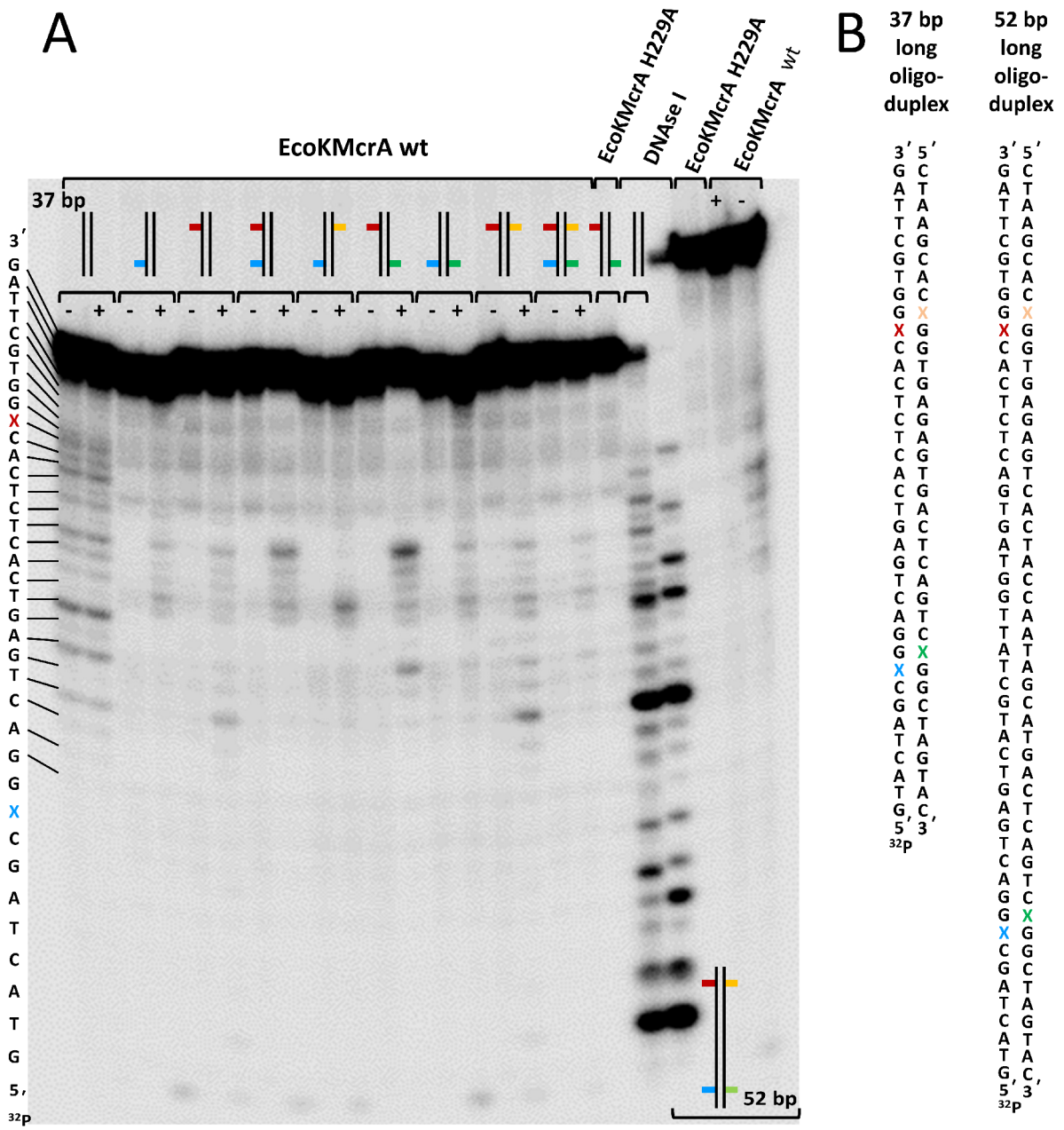

Fig. S17: Cleavage of substrates with either one or two target sites for EcoKMcrA modification binding domain. (A) 37 or 52 (last four lanes) base pair long oligoduplex DNAs were cleaved by wild type EcoKMcrA or its H229A catalytic mutant. The oligoduplexes carried two CCGG sites that were either unmodified, hemi- or fully methylated as indicated on the diagrams above each pair of lanes. One DNA strand was labelled with ^{32}P on the 5′ end. The unmodified duplex was not HPLC purified after synthesis and therefore shorter DNA oligonucleotides were also present. **(B)** The sequences of the oligoduplexes used in (A).

Fig. S18:

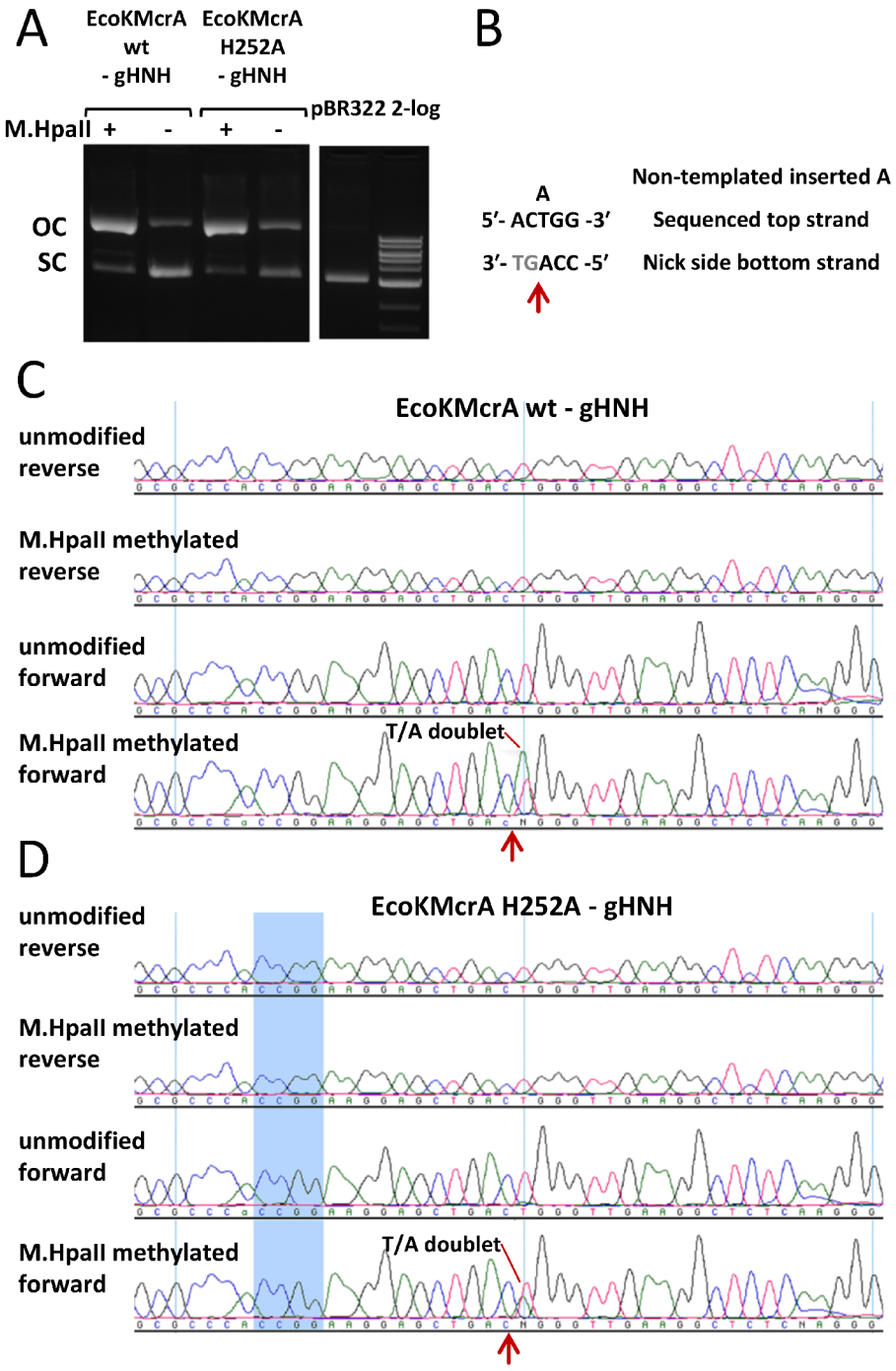

Fig. S18: Methylation dependent cleavage of YCR/GT consensus sequences by the EcoKMcrA-gHNH and EcoKMcrA(H252A)-gHNH fusion proteins. (A) The nicking activity of the fusion proteins on supercoiled pBR322. (B) Principle of run-off sequencing to map cleavage sites in a DNA template. The polymerase inserts an adenine 2′-deoxynucleotide downstream of the last inserted nucleotide when it encounters a strand break. The presence of an A/T doublet in the sequenced strand indicates a strand break in the bottom strand in the 5′-CCA/GT-3′ position. (C) and (D) Actual sequence traces demonstrating methylation dependent bottom strand cleavage in the vicinity of an M.HpaII methylated CCGG target site (blue bars). The 5′-CCA/GT-3′ cleavage site is compatible with the previously reported YCG/GT nicking sequence logo (2). OC-nicked, SC-supercoiled plasmid.

Fig. S19

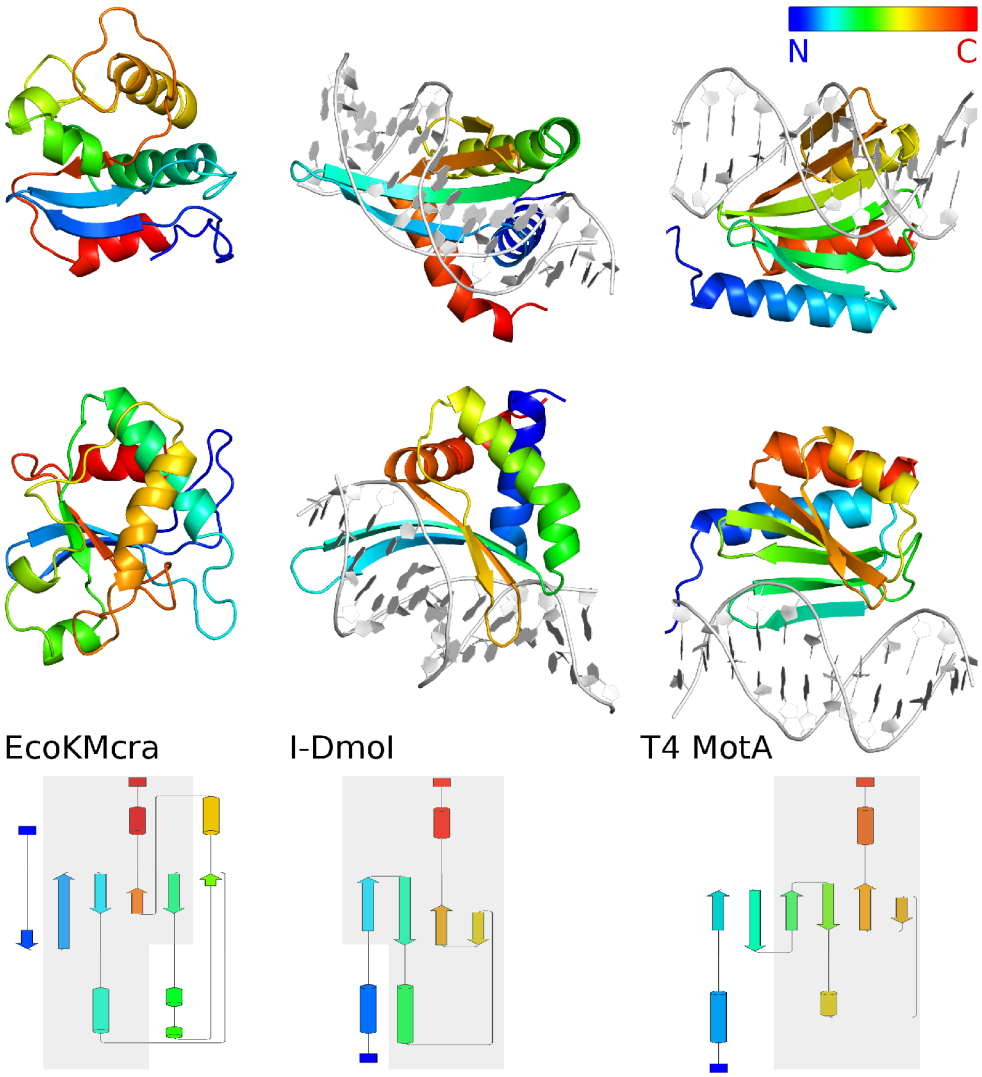

Fig. S19: Comparison of EcoKMcrA and structurally "similar" I-DmoI (PDB: 1mow)(6) and T4 MotA (PDB: 5jlt)(7). The top two rows show the structures in two different orientations, with bound nucleic acids where possible. The meganuclease I-DmoI is a dimer, but only one protomer is shown. The bottom row illustrates the folds using diagrams generated by the Pro-origami server (8).

Fig. S20

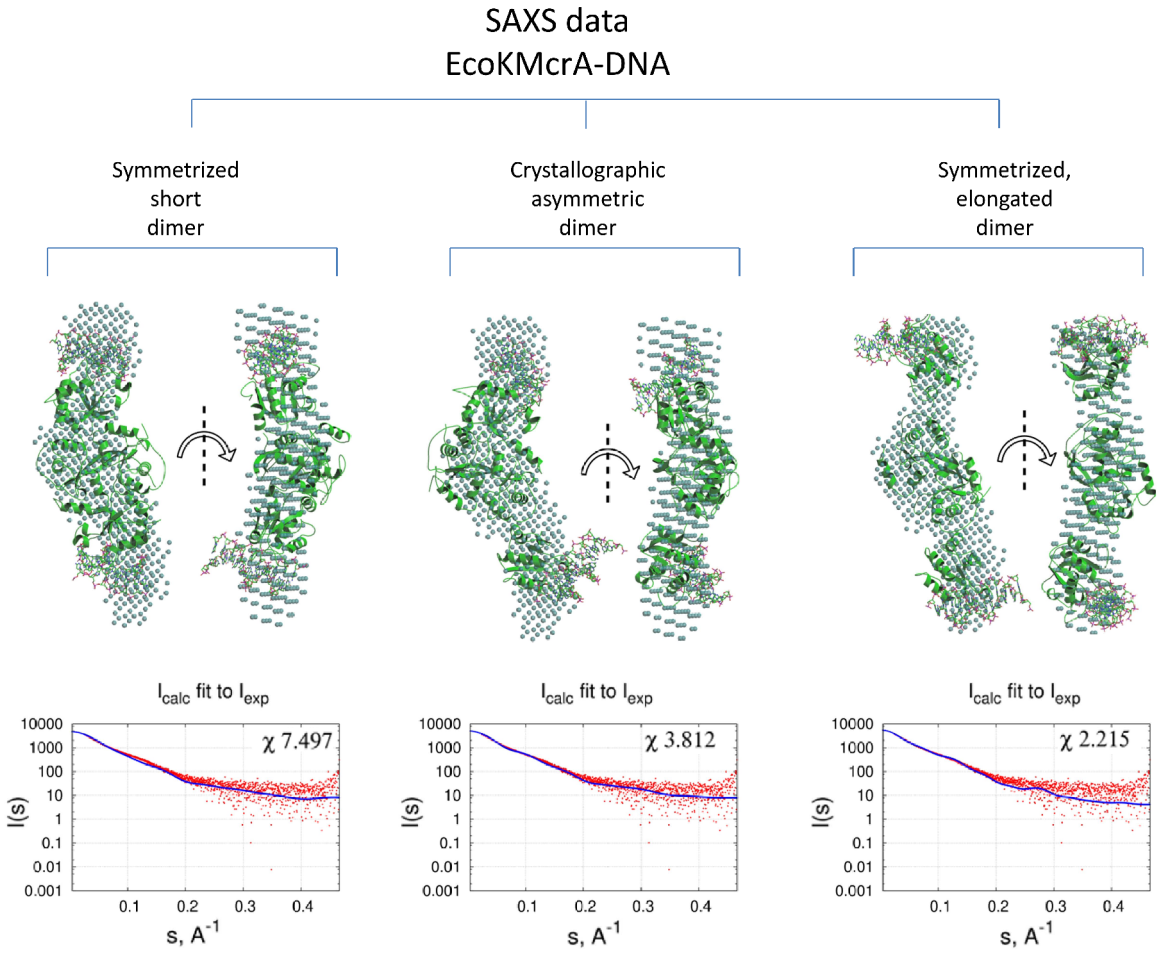

Fig. S20: Comparison of small-angle X-ray scattering data for the EcoKMcrA - hemimethylated DNA mixture with the predicted complex structure. The protein was mixed with DNA in one dimer per two DNA duplexes ratio. The top panel illustrates the agreement in real space, the bottom panel shows the agreement in reciprocal space. The *ab initio* SAXS model of the EcoKMcrA-DNA complex was calculated without imposing two-fold symmetry restraints.

Supplementary Tables

Table S1: Oligonucleotides used in this study. The CCGG sequence that, if modified, is recognized by EcoKMcrA is underlined**.**

Name	Sequence	Specification
12_C/12_C 12_5mC/12_C 12_pC/12_C	5′-ACCTC**X**GGTTCC-3′ 3′-TGGAGGCCAAGG-5′	12 bp oligoduplexes; 'X' marks unmodified C, 5mC or pyrrolocytosine. Used in EMSA and fluorescence measurements.
30_C/30_C 30_5mC/30_C 30_5mC/30_5mC	5′-AGACCCACGCTCAC**X**GGTTCCAGATTTATC-3′ 3′-TCTGGGTGCGAGTGG**X**CAAGGTCTAAATAG-5′	30 bp oligoduplexes; 'X' marks unmodified C or 5mC. Used for EMSA and DNA cleavage experiments.
T4, T5, T6, T7	5′-AGAC-3′ 5′-AGACC-3′ 5′-AGACCC-3′ 5′-AGACCCA-3′	Used as size markers to monitor top strand cleavage of the 30-mer oligoduplexes.
B22, B23, B24, B25, B26, B28	3′-TGGGTGCGAGTGGCCAAGGTCTAAATAG-5′ 3′-GGTGCGAGTGGCCAAGGTCTAAATAG-5′ 3′-GTGCGAGTGGCCAAGGTCTAAATAG-5′ 3′-TGCGAGTGGCCAAGGTCTAAATAG-5′ 3′-GCGAGTGGCCAAGGTCTAAATAG-5′ 3′-CGAGTGGCCAAGGTCTAAATAG-5′	Used as size markers to monitor bottom strand cleavage of the 30-mer oligoduplexes.
30′_C/30′_C 30′_5mC/30′_C	5′-GATTTATCAGACCCACGCTCAC**X**GGTTCCA-3′ 3′-CTAAATAGTCTGGGTGCGAGTGGCCAAGGT-5′	Alternative 30 bp oligoduplexes used for DNA cleavage experiments. 'X' marks unmodified C or 5mC.
T14′	5′-GATTTATCAGACCC-3′	Used as a size marker to monitor top strand cleavage of the 30′-mer oligoduplexes
B17′	3′-GTGCGAGTGGCCAAGGT-5′	Used as a size marker to monitor bottom strand cleavage of the 30′ -mer oligoduplexes
50_C/50_C 50_5mC/50_C	5′-CAGATTTATCAGACCCACGCTCAC**X**GGTTCCAGATTTATCGATGGTTAAC-3′ 3′-GTCTAAATAGTCTGGGTGCGAGTGGCCAAGGTCTAAATAGCTACCAATTG-5′	50 bp oligoduplex; 'X' marks unmodified C or 5mC. Used in DNase I footprinting experiments
37_X/37_X	5′-CTAAGCAC**X**GGTGAGAGTGACTCAGTC**X**GGCTAGTAC-3′ 3′-GATTCGTGG**X**CACTCTCACTGAGTCAGG**X**CGATCATG-5′	37 bp oligoduplex, 'X' marks unmodified C or 5mC. Used for the cleavage assay of substrates with either one or two modification sites
52_X/52_X	5′-CTAAGCAC**X**GGTGAGAGTCACTACCAATAGCATGACTCAGTC**X**GGCTAGTAC-3′ 3′-GATTCGTGG**X**CACTCTCAGTGATGGTTATCGTACTGAGTCAGG**X**CGATCATG-5′	52 bp oligoduplex, 'X' marks unmodified C or 5mC. Used for the cleavage assay of substrates with either one or two modification sites

Table S2: Parameters of the SAXS data

Samples	EcoKMcrA, merged data	EcoKMcrA + DNA	EcoKMcrA-N, merged data	EcoKMcrA-N + DNA
Concentration range, mg/ml	2, 5.2 and 9	1.7	1.3, 3.5 and 11.2	1.14
Guinier range (point number)	42-110	62-100	45-200	12-200
s range used in GNOM, nm^{-1}	0.1848-2.5753	0.2380-2.8944	0.1928-2.3174	0.1050-2.7614
Rg reciprocal, nm (from Guinier approximation)	4.19	3.88	3.13	2.09
Rg real, nm GNOM	3.835 ± 0.01259	4.044 ± 0.02507	2.184 ± 0.01286	2.128 ± 0.009277
Dmax, nm DATGNOM	18.7	13.6	9.1	8.1
Dmax, as parameter of GNOM	13.0	13.5	7.5	7.5
I(0) GNOM	3578 ± 11.82	4860 ± 28.59	1247 ± 6.513	2944 ± 7.980

Table S3: Molecular mass determination from SAXS data using various methods

	DATVC (9)	Porod volume (DATPOROD) (10)	SAXSMoW (11)	DAMMIF volume (10) (20 averaged models)	DAMMIN volume (10) (20 averaged models)
EcoKMcrA	64.7	58.7	77.6	57.43±0.47	74.67±0.61
EcoKMcrA +DNA	51.2	41.2	59.1	59.14±0.39	55.98±0.42
EcoKMcrA-N	21.7	16.8	-	23.68±0.23	22.24±0.15
EcoKMcrA-N +DNA	19.3	16.8	20.8	21.62±0.13	20.08±0.22

All molecular masses are given in kDa.

Table S4: Data collection and refinement statistics

Data collection statistics	
Space group	P2(1)2(1)2(1)
Cell dimensions a (Å) b (Å) c (Å)	 82.04 90.38 95.41
Wavelength (Å)	1.22021
Resolution range (Å) Highest shell Lowest shell	42 - 2.85 42 - 8.4 3.02 - 2.85
Total reflections	196068
Unique reflections	17013
Completeness (%)*	99.1 (95.2, 96.4)
Multiplicity *	11.5 (9.2, 11.4)
Mean $I/\sigma I$ *	21.0 (50.4, 1.82)
R (%)*	6.9 (3.8, 135.7)
R(meas) (%)*	7.2 (4.0, 141.9)
Solvent content (%)	56
B(iso) from Wilson (Å^2)	99.0
Refinement statistics	
Protein atoms excluding H	4577
Solvent molecules	51
R_{cryst} (%)	21.76
R_{free} (%)#	28.27
RMSD bond lengths (Å)	0.007
RMSD angles (°)	1.13
Ramachandran favored region (%)	100.0
Ramachandran allowed region (%)	96.5
Molprobity clashscore	0.7

* Lowest and highest shell in brackets

#5% of reflections were set aside randomly

References

1. 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.
2. Gutjahr, A. and Xu, S.Y. (2014) Engineering nicking enzymes that preferentially nick 5-methylcytosine-modified DNA. *Nucleic Acids Res*, **42**, e77.
3. Taylor, D., Cawley, G. and Hayward, S. (2014) Quantitative method for the assignment of hinge and shear mechanism in protein domain movements. *Bioinformatics*, **30**, 3189-3196.
4. Durand, D., Vives, C., Cannella, D., Perez, J., Pebay-Peyroula, E., Vachette, P. and Fieschi, F. (2010) NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. *J Struct Biol*, **169**, 45-53.
5. Ashkenazy, H., Erez, E., Martz, E., Pupko, T. and Ben-Tal, N. (2010) ConSurf 2010: calculating evolutionary conservation in sequence and structure of proteins and nucleic acids. *Nucleic Acids Res*, **38**, W529-533.
6. Chevalier, B.S., Kortemme, T., Chadsey, M.S., Baker, D., Monnat, R.J. and Stoddard, B.L. (2002) Design, activity, and structure of a highly specific artificial endonuclease. *Mol Cell*, **10**, 895-905.
7. Cuypers, M.G., Robertson, R.M., Knipling, L., Waddell, M.B., Moon, K., Hinton, D.M. and White, S.W. (2018) The phage T4 MotA transcription factor contains a novel DNA binding motif that specifically recognizes modified DNA. *Nucleic Acids Research*, **46**, 5308-5318.
8. Stivala, A., Wybrow, M., Wirth, A., Whisstock, J.C. and Stuckey, P.J. (2011) Automatic generation of protein structure cartoons with Pro-origami. *Bioinformatics*, **27**, 3315-3316.
9. Rambo, R.P. and Tainer, J.A. (2013) Accurate assessment of mass, models and resolution by small-angle scattering. *Nature*, **496**, 477-481.
10. Petoukhov, M.V., Franke, D., Shkumatov, A.V., Tria, G., Kikhney, A.G., Gajda, M., Gorba, C., Mertens, H.D., Konarev, P.V. and Svergun, D.I. (2012) New developments in the ATSAS program package for small-angle scattering data analysis. *J Appl Crystallogr*, **45**, 342-350.
11. Fischer, H., Polikarpov, I. and Craievich, A.F. (2004) Average protein density is a molecular-weight-dependent function. *Protein Sci*, **13**, 2825-2828.