4296–4308 *Nucleic Acids Research, 2015, Vol. 43, No. 8* *Published online 6 April 2015*
doi: 10.1093/nar/gkv251

Structures of *Escherichia coli* DNA adenine methyltransferase (Dam) in complex with a non-GATC sequence: potential implications for methylation-independent transcriptional repression

John R. Horton[1], Xing Zhang[1], Robert M. Blumenthal[2] and Xiaodong Cheng[1,*]

[1]Department of Biochemistry, Emory University School of Medicine, Atlanta, GA 30322, USA and [2]Department of Medical Microbiology and Immunology and Program in Bioinformatics, The University of Toledo College of Medicine and Life Sciences, Toledo, OH 43614, USA

Received November 21, 2014; Revised March 06, 2015; Accepted March 11, 2015

ABSTRACT

DNA adenine methyltransferase (Dam) is widespread and conserved among the γ-proteobacteria. Methylation of the Ade in GATC sequences regulates diverse bacterial cell functions, including gene expression, mismatch repair and chromosome replication. Dam also controls virulence in many pathogenic Gram-negative bacteria. An unexplained and perplexing observation about *Escherichia coli* Dam (EcoDam) is that there is no obvious relationship between the genes that are transcriptionally responsive to Dam and the promoter-proximal presence of GATC sequences. Here, we demonstrate that EcoDam interacts with a 5-base pair non-cognate sequence distinct from GATC. The crystal structure of a non-cognate complex allowed us to identify a DNA binding element, GTYTA/TARAC (where Y = C/T and R = A/G). This element immediately flanks GATC sites in some Dam-regulated promoters, including the Pap operon which specifies pyelonephritis-associated pili. In addition, Dam interacts with near-cognate GATC sequences (i.e. 3/4-site ATC and GAT). Taken together, these results imply that Dam, in addition to being responsible for GATC methylation, could also function as a methylation-independent transcriptional repressor.

INTRODUCTION

Escherichia coli DNA adenine methyltransferase (EcoDam) methylates the exocyclic amino nitrogen (N6) of the Ade in GATC sequences (1,2). Orthologs of the *dam* gene are widespread among γ-proteobacteria (3), and among a number of their bacteriophages (4). DNA-adenine methylation at specific GATC sites plays a pivotal role in methylation-dependent bacterial gene silencing, DNA replication and DNA mismatch repair (5,6). For example, there is a cluster of GATC sites near the origins of replication of *E. coli* and *Salmonella typhimurium*, each of which is conserved between the two species. This is not limited to the order Enterobacteriales, as there is evidence that Dam methylation controls replication of *Vibrio cholerae* chromosome II (7) and possibly chromosome I (8).

The Dam-dependent control of the various affected cell functions is in response to hemimethylated GATC sites, produced immediately following DNA replication (9). Specifically, Dam activity is relatively low (10), so there is a delay between chromosome replication and methylation of the new daughter strand. This delay is essential to the post-replicative mismatch repair system, where the methylation directs repairs to the new daughter strand (11).

Dam methylation also regulates the expression of specific genes in *E. coli* (12,13). For example, the expression of pyelonephritis-associated pili (Pap) in uropathogenic *E. coli* is epigenetically controlled by the methylation state of the two GATC sites in the Pap regulon (14). As another example, in *Salmonella enterica serovar* Typhimurium, Dam methylation modulates expression and translocation of the secreted *Salmonella* effector protein SopB (15). The interaction of EcoDam with its GATC target sequences has been studied structurally (16,17) and functionally (see below). For example, GATC sites preceded by an A_4 tract are poorly methylated, while the most-efficiently methylated GATC sites have no obvious flanking sequence pattern (18). However, GATC sites may not be associated with all EcoDam effects.

Puzzlingly, studies of global gene expression changes in *dam* mutant *E. coli* suggest that EcoDam can regulate gene expression in a methylation- and GATC-independent manner. Some of this is very likely due to indirect effects, but not necessarily all cases. There have been at least three stud-

*To whom correspondence should be addressed. Tel: +1 404 727 8491; Fax: +1 404 727 3746; Email: xcheng@emory.edu

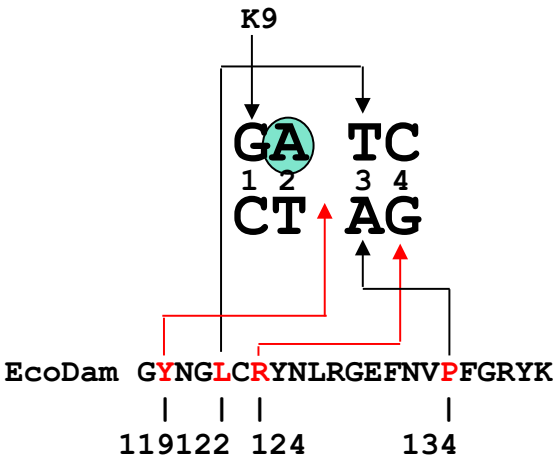


Figure 1. Schematic summary of EcoDam-DNA base contacts in the specific complex. The flipped target adenine base of the top strand is shaded. The R124–Gua4 interaction is conserved in cognate (GATC) and non-cognate (GTYTA/TARAC) complexes. Y119 intercalates between the inner AT base pairs.

ies of global gene expression comparing *dam*$^+$ and *dam E. coli* strain pairs; these involve different strains and media, but it is nevertheless striking that there is very little overlap among the genes affected in the three different microarray studies (12,13,19). The exception to this lack of overlap is the up-regulation of LexA-regulated SOS genes, because recombination protein RecA is chronically activated in *dam* cells (20) by the single-stranded DNA resulting from DNA damage, and facilitates autocatalytic cleavage of the LexA repressor (21). More strikingly, gene expression changes associated with *dam* mutation are not uniformly correlated with the presence of Dam methylation sites (GATC) (22).

We have reported crystal structures for *E. coli* Dam, and its ortholog bacteriophage T4 Dam (23,24). We have characterized these enzymes in both binary complexes with AdoHcy alone, and in ternary complexes with bound cognate GATC-containing DNA and reaction product AdoHcy (16,17). These structural studies, along with biochemical analysis by us and others, yielded four major conclusions. First, Dam is a processive enzyme that slides along the DNA (though see (25)) and switches enroute between nonspecific and specific interactions. Second, there is a consistent temporal order for the formation of specific contacts, with the first contact made between R124 of EcoDam and the Gua of the GATC site in the non-target strand (23) (Gua4 in Figure 1). The R124–Gua4 interaction has been examined by base-pair substitutions at the fourth position and by mutating R124 to alanine (R124A) (23). Wild-type EcoDam methylation of three near-cognate sites (GATN) was at least three orders of magnitude slower than methylation of GATC, whereas R124A methylated GATG and GATT sites 2–3-fold faster than the canonical GATC site. Therefore, R124A has lost the discriminatory requirement for a C:G base pair at the fourth position of GATC. Third, EcoDam Y119 intercalates between the inner AT base pairs (GA^TC), resulting in a local doubling in the DNA helical rise equivalent, overall, to a 5-base pair (bp) length of DNA. The catalytic activity of Y119A mutant was the second most strongly affected by the alanine substitution, after R124A (23). Fourth, the first G:C base pair is recognized less accurately than the third and fourth pair pairs (16,23,26). Relative to GATC, three near-cognate substrates that carry a base-pair substitution at the first position were still methylated by EcoDam, although at rates reduced by 100- (AATC) or 1000-fold (CATC) (23). Interestingly, the contact to the first base pair is not conserved among members of the Dam family, and it has been suggested that Dam evolved from an ancestral protein that recognized ATC (27). Finally, in crystal complexes Dam molecules preferentially bind at the joint between two linearly-aligned (head-to-tail) DNA duplexes, which mimics a damaged or altered B-DNA conformation.

Here, we report the structures of EcoDam in complex with non-cognate DNA, lacking any GATC sequences. These structures allowed us to identify an apparent 5-bp DNA binding element, GTYTA/TARAC (Y = C/T and R = G/A). We subsequently found this 5-bp element (or its variations) in the Pap operon adjacent to both GATC sites, and at or near the −10 promoter elements upstream of the transcription starts of many Dam-responsive promoters. In solution, EcoDam binds to short oligonucleotides containing GTTTA equally well (or even slightly better) than GATC. These results suggest a mechanism of negative regulation by Dam in a methylation- and GATC-independent fashion, and may help explain the heretofore puzzling gene expression patterns associated with EcoDam mutation.

MATERIALS AND METHODS

His-tagged EcoDam was expressed in HMS174(DE3) cells using autoinduction procedures (28), and purified on Ni^{2+}-affinity, UnoS and S75 Sepharose sizing columns (GE Healthcare) as previously described (16,17). A 0.5 liter autoinduced culture yielded ∼7 mg purified enzyme. In the last purification step and during concentration, AdoMet, AdoHcy or sinefungin was added to the protein at ∼2:1 molar ratio. Concentrated binary complexes were mixed with oligonucleotide duplex (synthesized by New England Biolabs, Inc.) at a protein to DNA ratio of ∼2:1 and allowed to stand on ice for at least 2 h before crystallization. Final protein concentration for crystallization trials was ∼15–30 mg/ml.

Initially, in sitting drop crystallization screens, showers of small needles appeared in 5–10% of 2-methyl-2,4-pentanediol (MPD) and polyethylene glycol (PEG) 4000-8000 as precipitants, but it appeared that larger, more single crystals could be grown with low molecular weight PEGs. Ternary complexed crystals utilized for data collection were best grown in hanging drops, with well solution containing 100 mM KCl, 10 mM $MgCl_2$ and with variation of PEG 200 from 5 to 30% and 100 mM MES or HEPES (pH 6.4-7.0). The PEG 200 concentration was increased to ∼40% in the crystallization drop before picking single crystals in cryoloops and froze and stored in liquid nitrogen until data collection. In addition, a few larger crystals appeared in PEG 4000 conditions and 25% ethylene glycol was added to mother liquor before freezing.

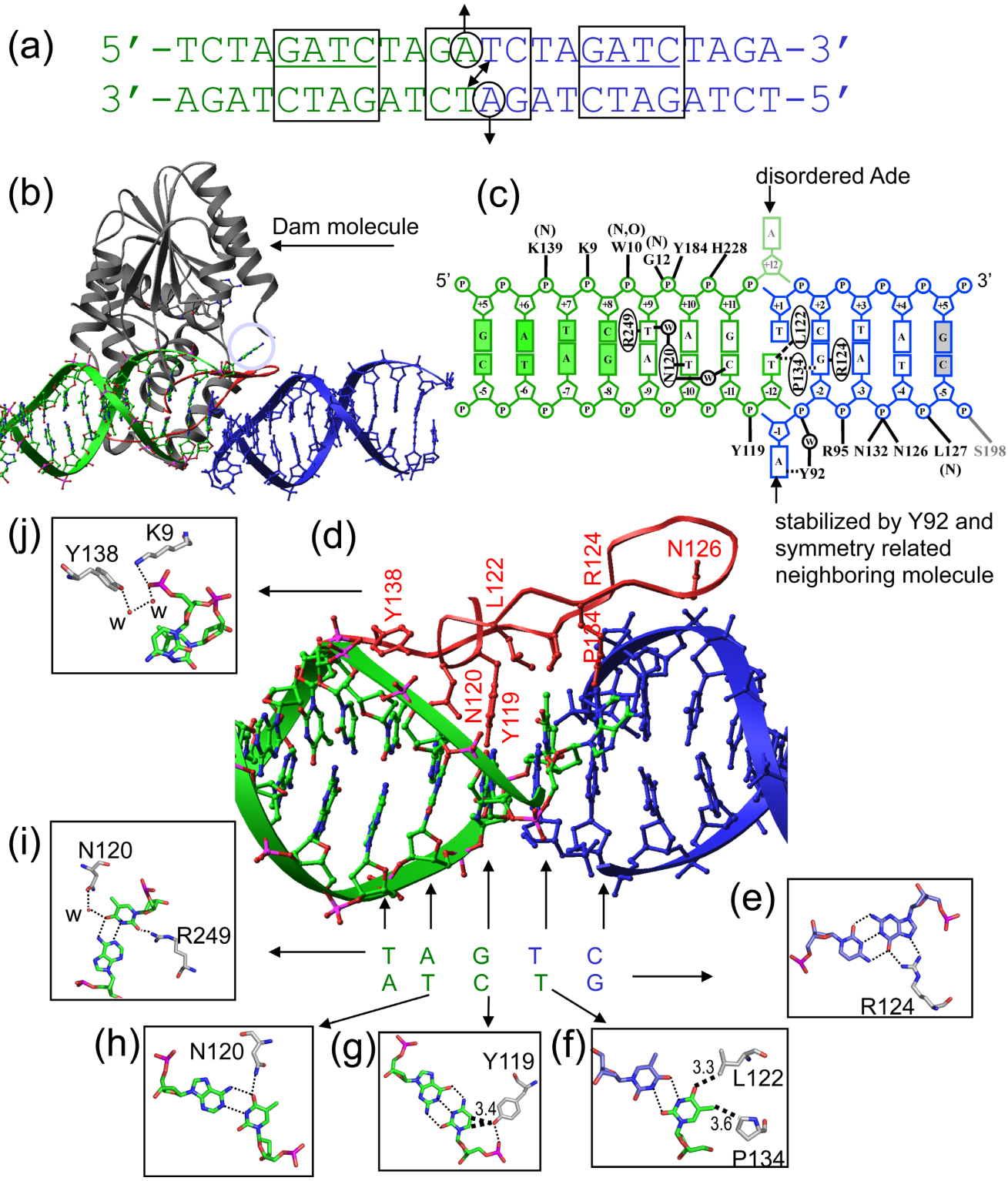


Figure 2. Structure of EcoDAM in complex with a non-canonical site. (**a**) DNA sequence used in the crystallization containing one GATC in the middle region of the duplex. The end sequence of the duplex was chosen such that the sequence at the joint of two molecules (green and blue) mimics a GATC site if the DNA duplexes are stacked head-to-tail. (**b**) EcoDam binds at the joint of two DNA duplexes. The hairpin recognition loop is colored in red. (**c**) Schematic summary of the protein-DNA contacts. Two DNA duplexes are stacked head-to-tail with one T:T mispair in the joint of two duplexes. Backbone-mediated interactions are indicated with main chain amine nitrogen (N) or carbonyl oxygen (O). (**d**) EcoDam-DNA interactions involve a 5-bp non-canonical site (5′-GTCTA-3′). (**e**) R124 interacts with the first G:C pair. (**f**) L122 and P134 interact with the T:T mispair. (**g**) Y119 interacts with the third C:G pair. (**h**) N120 interacts with the fourth T:A pair. (**i**) R249 interacts with the fifth A:T pair. (**j**) K9 and Y138 are involved in phosphate interactions.

Structures were determined by molecular replacement with the program REPLACE (29), using an EcoDam protomer from the previously determined cognate complex structure (PDB: 2G1P) (16) as a search model and DNA was manually built into its obvious electron density. Most refinement was completed using the program CNS (30) and manual manipulation by the program O (31). The last rounds of refinement utilized the PHENIX package (32) and final maps and models were visualized and manual manipulation was completed with COOT (33,34). Molecular graphics were generated using PyMol (DeLano Scientific LLC).

Dam DNA binding was measured by two assays. For electrophoretic mobility shift assays, EcoDam (20 μM) was pre-incubated with the AdoMet analog sinefungin (Sigma) at a 1:1.5 molar ratio in 20 mM HEPES–HCl (pH 7.5), 100 mM NaCl, 5% glycerol, 1 mM dithiothreitol (DTT). The binary complex (at the indicated amount) was mixed with a 32-bp 6-carboxy-fluorescein (FAM)-labeled double-stranded (ds) DNA (5 nM) for 30 min at room temperature (~25°C) in a 20 μl reaction containing 0.1 mg/ml bovine serum albumin (New England Biolabs), with or without 10 μg/ml salmon sperm genomic DNA (Rockland Inc.). Samples were loaded onto a 10 cm × 10 cm 6% native acrylamide gel in 0.5× Tris/Borate/EDTA (TBE) buffer and run 45 min at 80 V. FAM labeled DNA was visualized by Typhoon Trio+ (GE Healthcare).

Fluorescence polarization measurements were carried out at room temperature on a Synergy 4 microplate reader (BioTek). The FAM-labeled dsDNA (5 nM) was incubated for 30 min with increasing amounts of EcoDam–sinefungin complex in 20 mM HEPES–HCl (pH 7.5), 200 mM NaCl, 10% glycerol, 1 mM DTT. Curves were fit individually using GraphPad Prism 5.0 software (GraphPad Software, Inc.). Binding constants (K_D) were calculated as [mP] = [maximum mP] × [C]/(K_D + [C]) + [baseline mP], and saturated [mP] was calculated as saturation = ([mP] − [baseline mP])/([maximum mP] − [baseline mP]), where mP is millipolarization and [C] is protein concentration. Averaged K_D and its standard error are reported.

RESULTS

EcoDam interaction with a non-canonical site

We previously crystallized a complex containing EcoDam and a 12-bp DNA duplex containing a single centrally-located GATC target site (16). The end sequence of the duplex was chosen such that the sequence at the joint of two molecules mimics a GATC target site, if the DNA duplexes are stacked head-to-tail (Figure 2a). Indeed, two Dam molecules were bound to each duplex, one at the central GATC and one at the joint. However, a second crystal form (space group: $P2_12_12_1$) was also produced using the same 12-bp blunt-end DNA duplex (Figure 2a; Table 1).

There were three unexpected observations in the second crystal form. First, while an EcoDam molecule was again bound to the joint between neighboring DNA duplexes, no EcoDam molecule was bound to the unbroken GATC site in the middle of the duplex (Figure 2b). Second, only 11 of the 12 base pairs in each DNA duplex are stacked head-to-tail along the crystal *a*-axis with the length of ~36 Å (average helical rise per base pair of ~3.3 Å). The electron density maps indicate that the two 3′ Ade bases at the ends of each DNA duplex were flipped out, with one being disordered and the other stabilized by Dam (Figure 2, panels b and c). This is surprising because, in the cognate complex, only the target (methylatable) Ade is flipped out and into the active site (16), whereas in this non-cognate complex both Ade are flipped and neither one is in the active site. Third, the two 5′ Thy bases formed a T:T mismatch at the joint of the two DNA molecules (Figure 2c).

EcoDam is thus in contact with five base pairs spanning the joint, which constitutes a non-canonical site. Three contacted bases are from the green/left duplex, one is from the T:T mispair, and one is from the blue/right duplex (using the colors and orientation of Figure 2, panels c and d). In this particular complex, the 5′ Gua (blue DNA) interacts with R124 (Figure 2e), and the interactions of its four 5′ phosphates are identical to those of EcoDam interacting with cognate DNA (16). One thymine of the T:T mispair makes a van der Waals contact with P134 and a C–H···O type of hydrogen bond with L122 (Figure 2f) (35). Other residues, previously identified in cognate complexes as being involved in intercalation (Y119), base-amino acid hydrogen bonding (N120), and first base pair recognition (K9 and Y138), are in the case of this non-cognate complex located in the major groove of the green DNA. It is as if they were positioned for invasion into the DNA at a GATC sequence, but then switched their roles to making phosphate contacts (Y119 and K9), base contacts (Y119 and N120), or water-mediated DNA interactions (Y138) (Figure 2g–j). An additional base contact is formed in the minor groove of the green DNA by R249 (Figure 2i). Taken together, these interactions suggest EcoDam recognizes and binds to a non-canonical 5-bp sequence of 5′-GTCTA-3′ / 5′-TAGTC-3′. This non-canonical binding might be to a more degenerate sequence, but the structure provides direct evidence for recognition of this particular sequence.

Comparing the cognate complex (16) and the non-canonical complex, the protein components are structurally similar (root mean squared deviation = 0.4 Å across 245 Cα atoms; Figure 3a), as are the analogous cofactors (AdoHcy in the cognate complex and sinefungin in the non-canonical complex). The associated active-site residues are in equivalent positions (Figure 3b). In addition, there is one particularly well-conserved protein–DNA base interaction: R124–Gua (Figure 3c). The interactions with GTCTA/TAGTC (Figure 2, panels d–j) suggest that the Dam binding is likely to be specific for the 5′ G bp, but could be degenerate at the other positions. This partially resembles a 1/4-site recognition of GATC, in which EcoDam traps the sequence that mimics part (1/4) of the GATC sequence.

Interactions with *pap* promoter sequences

As mentioned in the introduction, the Pap operon contains two GATC sites (Figure 4a). In contrast to most GATC sites in the *E. coli* genome, the *pap*-associated sites are not always completely re-methylated after DNA replication, and their methylation state determines in part the phase variation of pilus formation, which is under epigenetic control (14). The failure to methylate these sites is due in part to the bind-

Table 1. Statistics of X-ray diffraction and refinement

Crystal	#1	#2	#3	#4	#5	#6	#7	#8	#9	#10
Name	Non-cognate	Pap (dist)			Pap (prox)			GTTTA (prox)		GTCTA (dist)
Cofactor	Sinefungin	AdoHcy	Sinefungin	AdoMet	AdoHcy	Sinefungin	AdoMet	AdoHcy	AdoMet	AdoMet
DNA sequence	TCTAGATCTAGA AGATCTAGATCT		TCTAAAGATCG GATTTCTAGCA			TTTAAAGATCG AATTTCTAGCA		ACTTAAACTTAA GAATTTGAATTT		ACTTAGACTTAG GAATCTGAATCT
PDB code	4RTJ	4RTK	4RTL	4RTM	4RTN	4RTO	4RTP	4RTQ	4RTR	4RTS
Data collection										
Cell dimensions (Å)	**a** = 36.3	36.3	38.3	36.4	36.2	35.9	35.8	44.0	43.7	39.6
(α=β=γ=90°)	**b** = 63.1	63.0	66.3	63.1	63.5	63.3	62.2	62.9	62.9	63.8
	c = 145.0	142.2	149.2	143.2	144.0	143.2	144.1	136.6	136.3	155.0
[a]Resolution (Å)	32.51–1.99	26.22-1.96	27.61-2.19	32.43-2.49	29.04-2.59	34.85-2.69	34.76-2.39	30.02-1.99	29.96-2.39	33.60-2.49
	(2.06–1.99)	(2.03-1.96)	(2.27-2.19)	(2.58-2.49)	(2.68-2.59)	(2.79-2.69)	(2.48-2.39)	(2.06-1.99)	(2.48-2.39)	(2.58-2.49)
Measured reflections	181,130	102,658	175,481	144,853	98,237	52,728	79,618	199,845	65,792	53,993
Unique reflections	23,469	23,847	20,122	12,050	10,911	9,617	13,573	26,232	14,421	13,946
Redundancy	7.7	4.3	8.7	12	9	5.5	5.9	7.6	4.6	3.9
[b]<I/σI>	20.9	20.4	24.4	10.7	12.0	11.4	17.3	15.9	12.2	10.7
Completeness (%)	98.5(98.0)	98.0(99.6)	99.3(98.5)	100(100)	100(100)	99.7(99.6)	99.4(99.8)	99.4(98.3)	93.1(90.5)	96.4(98.4)
[c]R_{merge}	0.051(0.374)	0.057(0.297)	0.056(0.287)	0.113(0.439)	0.083(0.503)	0.095(0.470)	0.069(0.311)	0.111(0.220)	0.086(0.471)	0.135(0.454)
Refinement										
[d]R_{work} / [e]R_{free}	0.214 / 0.236	0.210 / 0.223	0.221/ 0.251	0.208 / 0.237	0.205 / 0.240	0.218 / 0.250	0.228 / 0.255	0.160 / 0.198	0.177 / 0.227	0.174 / 0.214
Non-hydrogen atoms										
Protein	2016	1973	2043	1969	1968	1958	1937	1991	1992	1954
DNA	486	444	445	445	445	445	445	468	468	469
Cofactor	27	26	27	27	26	27	27	26	27	27
Solvent	101	105	88	50	34	19	25	232	106	43
B-factors (Å^2)										
Protein	54.3	39.9	46.6	43.7	46.5	66.3	61.8	35.9	42.8	48.2
DNA	67.8	48.2	53.6	58.2	56.0	68.1	74.4	48.1	58.7	69.7
Cofactor	41.6	38.5	32.9	42.6	31.0	63.3	68.7	34.2	49.1	63.8
Solvent	49.1	42.2	42.5	41.4	38.5	58.1	50.2	44.5	45.0	51.5
Room-mean-square deviation from ideality										
Bond lengths (Å)	0.003	0.002	0.011	0.003	0.003	0.002	0.002	0.003	0.003	0.004
Bond angles (°)	0.6	0.6	1.1	0.6	0.6	0.5	0.5	0.8	0.6	0.8
[f]Estimated error										
Coordinate (Å)	0.17	0.19	0.30	0.29	0.26	0.30	0.26	0.17	0.13	0.29
Phase (°)	24.0	23.5	27.2	23.3	25.3	26.2	25.2	18.1	21.8	21.0

Note: Wavelength=1 Å; Space group $P2_12_12_1$; One ternary complex per asymmetric unit; synchrotron beamline APS 22-BM or 22-ID

[a]Values in parenthesis correspond to highest resolution shell;

[b]<I/σI> = averaged ratio of the intensity (I) to the error of the intensity (σI);

[c]$R_{merge} = \Sigma \mid I - \langle I \rangle \mid / \Sigma I$, where I is the observed intensity and <I> is the averaged intensity from multiple observations;

[d]$R_{work} = \Sigma \mid F_{obs} - F_{cal} \mid / \Sigma \mid F_{obs} \mid$, where F_{obs} and F_{cal} are the observed and calculated structure factors, respectively;

[e]R_{free} was calculated using a randomly chosen subset (5%) of the reflections not used in refinement.

[f]maximum-likelihood based

ing of regulatory proteins (Lrp, PapI) that block access of EcoDam (14,36,37), but may also reflect reduced EcoDam methylation activity due to their adjacent DNA sequences. Interestingly, methylation of these two sites is nonprocessive in the absence any regulators, which suggested that sequences flanking these two GATC sites might prevent the expected processivity of EcoDam (38). This is consistent with the observation that the ability of EcoDam to methylate a particular GATC depends on the immediate flanking DNA sequences (25,39).

By inspecting the Pap sequences flanking GATC (Figure 4a box), we discovered sequence elements (GTYTA/TARAC) that include the 5-bp non-canonical site identified in our crystal structure, located immediately adjacent to both GATC sites. The two GTYTA elements are in opposite orientations, and they differ at the third base pair. Furthermore, 16 out of the 17 bp flanking the two GATC sites are identical, differing only at a Y (Y = C in one and T in the other) – TAAAA**GATC**GTYTAAAT (Figure 4a). This observation, together with the demonstrated binding to GTYTA/TARAC motifs in the structural studies, suggests that the different behaviors of the Dist and Prox sites are associated with the difference in GTCTA/TAGAC and GTTTA/TAAAC sequence or orientation.

The 5-bp elements might bind EcoDam before or after it binds the GATC site, thus affecting processivity or local Dam concentration (or both) and modulating the regulation of *pap* expression. Accordingly, we next used X-ray crystallography to characterize the interactions between EcoDam and the two GATC regions (Dist and Prox) of the *pap* regulatory region (Figure 4b and c). We designed a series of oligonucleotides representing the Dist and the Prox sites. One of the 11 bp duplexes that contains a Prox site sequence (Figure 4c) crystallized in complex with EcoDam, in over one-third of the 48 conditions screened, in the presence of cofactor AdoMet, its analog Sinefungin, or its product AdoHcy (Table 1). The influence of the AdoMet or its proxy on EcoDam interactions with non-GATC elements is limited, as the structures are highly similar to each other. The 11-bp DNA duplexes are stacked head-to-tail with 5′ overhangs A and T forming a base pair, resulting in a sequence representing 14 of the 17 conserved base pairs between the Dist and Prox sites (underlined in Figure 4b and c). Although the canonical GATC sequence is present, the EcoDam molecule does not occupy it. Instead, EcoDam occupies the joint between two neighboring DNA duplexes, with R124 making direct base contacts with the 5′ Gua of GTTTA of the Prox site (Figure 4d, top left). In addition, L122 and P134 contact the second base pair (Figure 4d, top middle), Y119 contacts the third base pair with weak carbon–carbon interaction between the methyl group of thymine and a ring carbon atom of Y119 (Figure 4d, top right), and N124 contacts both the fourth and the fifth base

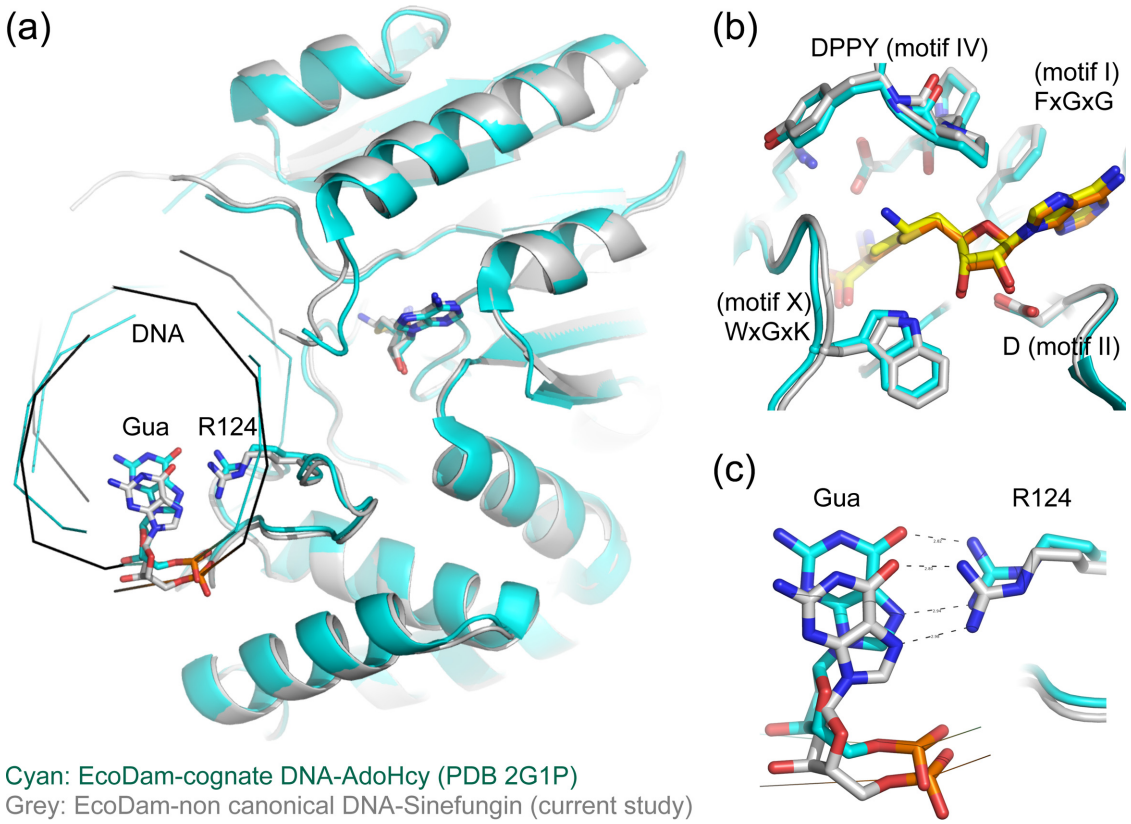


Figure 3. Structural comparison of cognate and non-canonical complexes. (**a**) Superimposition of EcoDam–cognate DNA–AdoHcy ternary complex (PDB: 2G1P in cyan) and the non-canonical complex (in gray). (**b**) The active sites containing Sinefungin (yellow) and AdoHcy (orange) and conserved sequence motifs (defined in (57)). (**c**) Conserved R124–Gua interaction.

pairs, via a water-mediated interaction (base pair 4) or a weak interaction with the methyl group of thymine of base pair 5 (Figure 4d, bottom).

For the Dist site, we suspected that the same R124 could interact with the 5′ Gua of GTCTA on the bottom strand (see Figure 5). Surprisingly, we also obtained crystals in which R124 was interacting with the 5′ Gua of GACGA of the top strand in the joint between two neighboring DNA duplexes (Figure 4b). This complex resembles a 1/2-site recognition, in which the 5′ Gua forms hydrogen bonds with R124 while the second-position Ade interactions with P134 and L122 (Figure 4e, top left and middle) are identical to those of EcoDam interacting with cognate GATC sequence (16). It seems that the interactions for the first 2-bp of the GACGA are specific, while the interactions for the last 3-bp are less so and those bp could be replaced with alternatives (comparing Figure 4d and e). In this orientation, the Dam molecule occupying the Dist site is in the same direction as the one in Prox site (i.e. both R124–Gua interactions occur for the top strand (Figure 4b and c).

EcoDam alters DNA conformation in non-cognate complexes

We next designed oligos containing the two repeated 5-bp GTYTA elements, from either Dist or Prox site (Figure 5a and b), with one repeat within the DNA duplex and the other at the joint of two neighboring DNA molecules. Again, we observed that a Dam molecule occupies the 5-bp element formed in the joint between two neighboring DNA duplexes (Figure 5). Interestingly, we observed two distinct Y119 interactions with DNA at the Prox and Dist sequences. In the Prox site, Y119 intercalates into the DNA between the two thymine bases of GT^TTA (Figure 5c), similarly to the cognate complex where Y119 inserts into the DNA between Ade and Thy of GA^TC (16). In the Dist site, Y119 stacks with the thymine base at position 2, GTCTA, but pushes the rest of DNA away, resulting in two DNA molecules shifted relative to one another perpendicularly to the DNA axis (Figure 5d). This second phenomenon has been observed previously with the phage T4 Dam (23).

We have observed previously that a mammalian SRA domain protein could bind the junction between the two DNA duplexes (40). In addition, protein–DNA complex crystals of some DNA repair glycosylases were obtained in which the enzyme did not bind to the middle region of the duplex containing the lesion, but bound at the joint (41–43). The examples mentioned here, Dam, SRA, and DNA glycosylases, all use a base-flipping mechanism (44) to access their target base, whether for the purpose of mismatch excision, damage repair or generating/recognizing modification. It is possible that the junction between the two DNA duplexes mimics the altered B-DNA conformation generated during or after base flipping. In fact, HhaI methyltransferase binds more tightly to DNA sequences containing a mismatch at the target base (45).

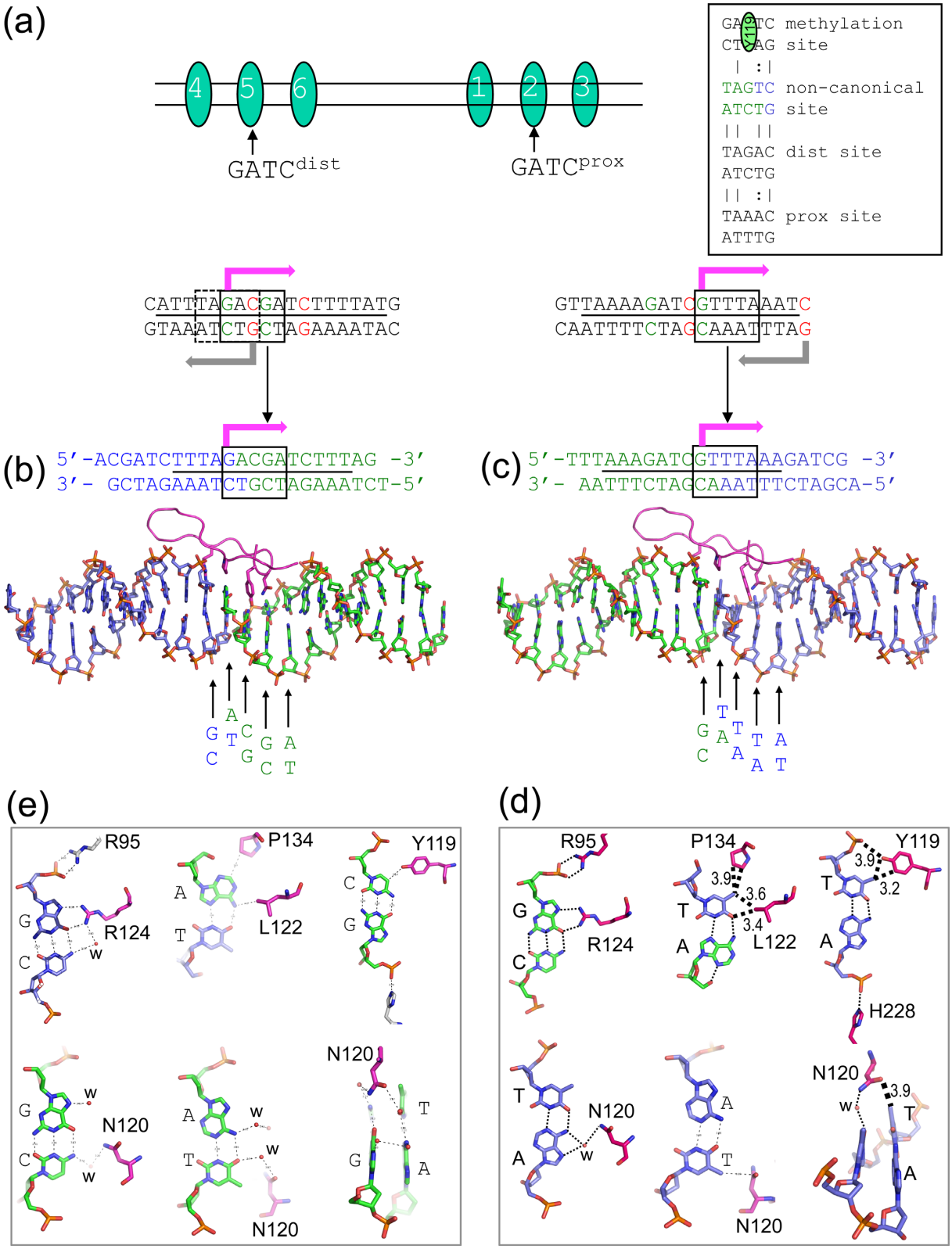


Figure 4. Structure of EcoDam in complex with *pap* promoter sequences. (**a**) Organization of *pap* regulatory sequence: numbers (1–6) indicate six leucine-responsive-regulatory-protein (Lrp) binding sites (14). Among the six Lrp binding sites, sites 2 (Prox) and 5 (Dist) contain a GATC sequence. The Pap GATC flanking sequences, shown underneath, share sequence similarity with the non-canonical site (inset box). (**b** and **c**) The 11-mer DNA duplexes are stacked head-to-tail with the 5′ overhangs A and T forming a base pair, resulting in a sequence representing the Dist site or Prox site (underlined). The Dam molecule is trapped at the 5′ Gua of GACGA (Dist) or GTTTA (Prox). (**d** and **e**) EcoDam-DNA interactions involves 5-bp at the Prox site (GTTTA) or Dist site (GACGA).

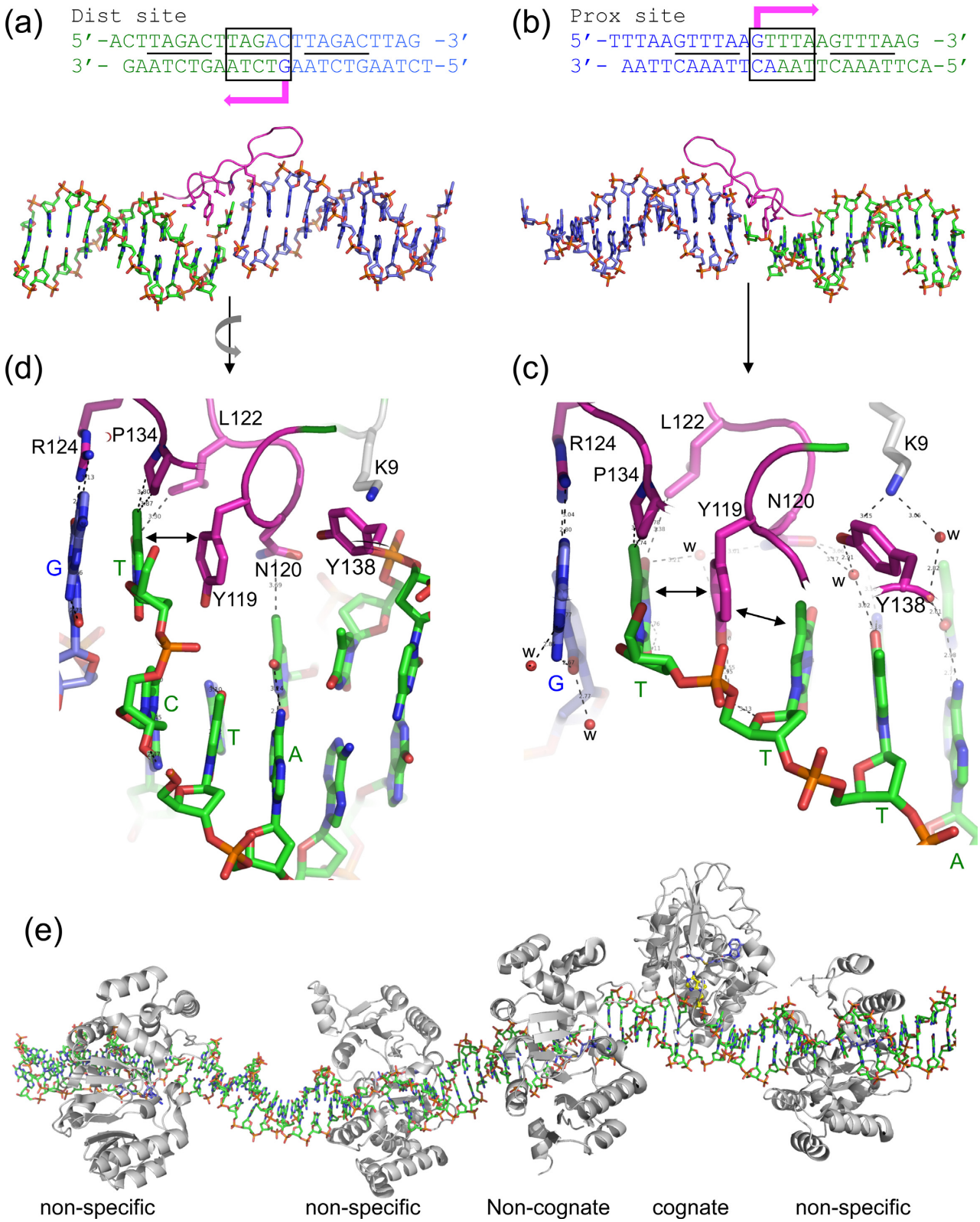


Figure 5. Structure of EcoDam in complex with the 5-bp GT(C/T)TA element. (**a** and **b**) The repeated 5-bp elements used in the crystallization. The organization of the DNA sequences mimics what appears in the *pap* regulatory sequence (see Figure 4). (**c**) Y119 intercalates into DNA resembling the Prox site, and stacks between two thymine bases. (**d**) Y119 pushes the green DNA, resulting in the two DNA molecules (blue and green, resembling the Dist site) being shifted relative to one another perpendicularly to the DNA axis. (**e**) A hypothetical model of EcoDam molecules sliding along DNA and binding at adjacent noncognate and cognate sites. The model was generated based on T4Dam structure (PDB: 1YFJ) (23).

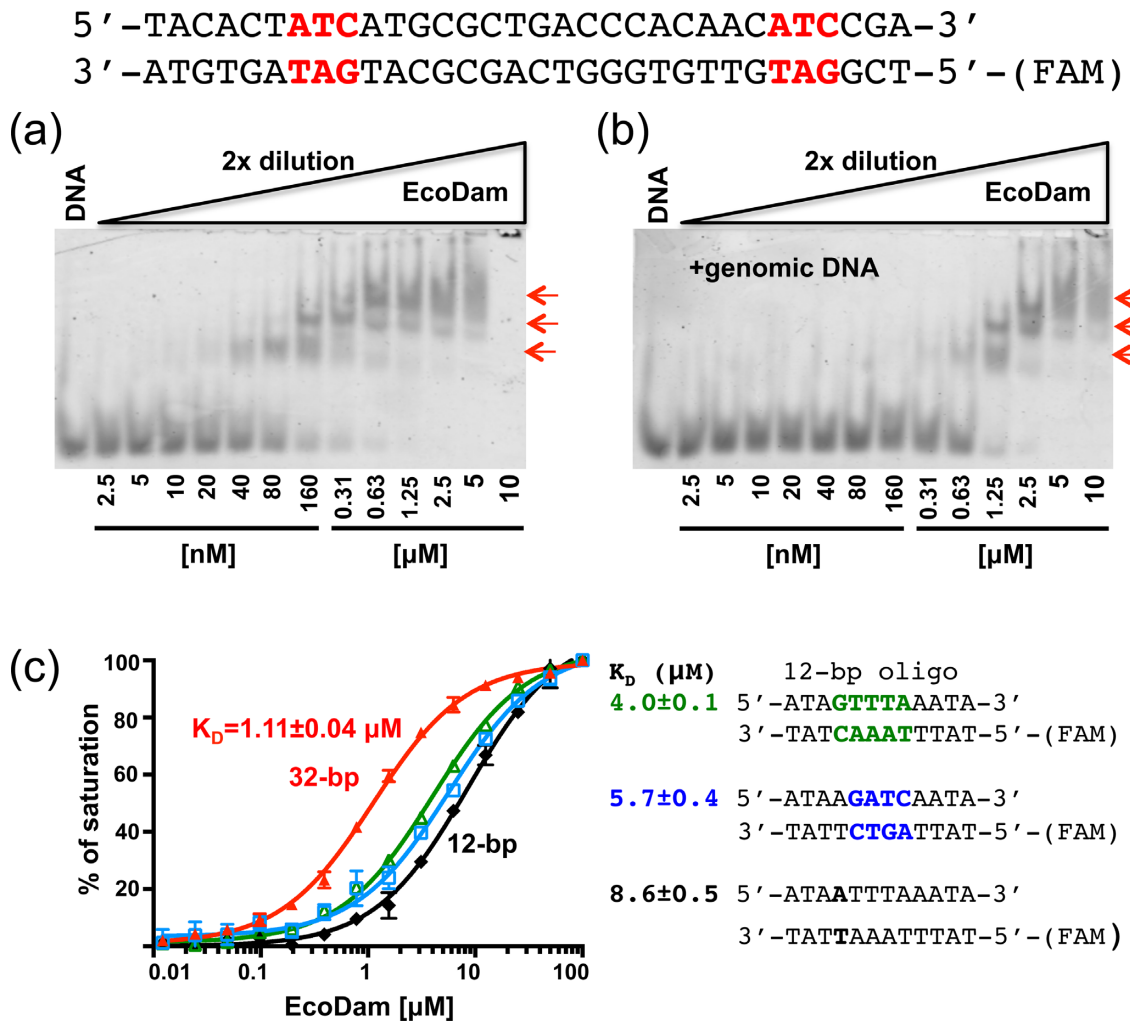


Figure 6. EcoDam–DNA interactions in solution. (**a**) DNA binding assays were performed by incubating 32-bp FAM labeled oligonucleotides with increasing amount of EcoDam. Specific band shifts (indicated by red arrows) were observed at ∼40, 160 and 630 nM of EcoDam, respectively. (**b**) Detection of specific EcoDam–DNA complexes in the presence of salmon sperm genome DNA. The band shifts were observed at higher EcoDam concentrations at approximately 0.63, 1.25 and 2.5 μM, respectively. (**c**) Binding affinities of EcoDam with oligonucleotides (32-bp or 12-bp) containing two GAT/ATC sites (red), one GTTTA (green), one GATC (blue) or no G:C base pair (black).

EcoDam–DNA interaction in solution

To explore the effect of DNA sequence variation on EcoDam binding, we first measured the binding activity of EcoDam using a 32-bp DNA oligonucleotide containing two GAT/ATC sites, but no GATC (Figure 6). In electrophoretic mobility-shift assays, specific band shifts were observed at ∼40, 160 nM, and became a smear at 630 nM or higher concentrations of EcoDam (Figure 6a). EcoDam maintains the specific band shifts in the presence of salmon sperm DNA, though the shifts occur at higher EcoDam concentrations of approximately 0.63, 1.25 and 2.5 μM, respectively (Figure 6b). This observation of non-GATC specific interactions was in agreement with a previous study using a long 150-bp DNA fragment containing a single GATC site and six 3/4 sites (ATC or GAT) (27), where the appearance of a ladder of bands indicated multiple Dam molecules binding to non-GATC sites. In contrast, when the same long DNA molecule, which contained a single EcoRV binding site (GATATC), was incubated with EcoRV methyltransferase (a restriction-modification enzyme which is related to EcoDam by protein sequence homology and overlapping substrate specificities), only a single specifically-shifted band resulted (46).

We next measured the dissociation constants (K_D) between EcoDam and a 12-bp oligonucleotide containing either a GATC, GTTTA or no G:C base pair using fluorescence polarization analysis (Figure 6c). Surprisingly, the binding affinity for GTTTA is slightly stronger than that of GATC under the assay conditions (Figure 6c). While the difference is only ∼40%, it is clear that the Dam affinity for GTTTA is no lower than that for GATC. Changing the single G:C base pair of GTTTA to A:T reduced binding affinity by a factor of >2 (Figure 6c), indicating nonspecific DNA binding, which might account for the smear observed in the electrophoretic mobility-shift assays. The 32-bp oligonucleotide exhibited the strongest binding at K_D of ∼1.1 μM, probably because it allowed binding of multiple EcoDam molecules at the same time.

Table 2. Examples of GTYTA/TARAC elements in the GATC-regulated promoters

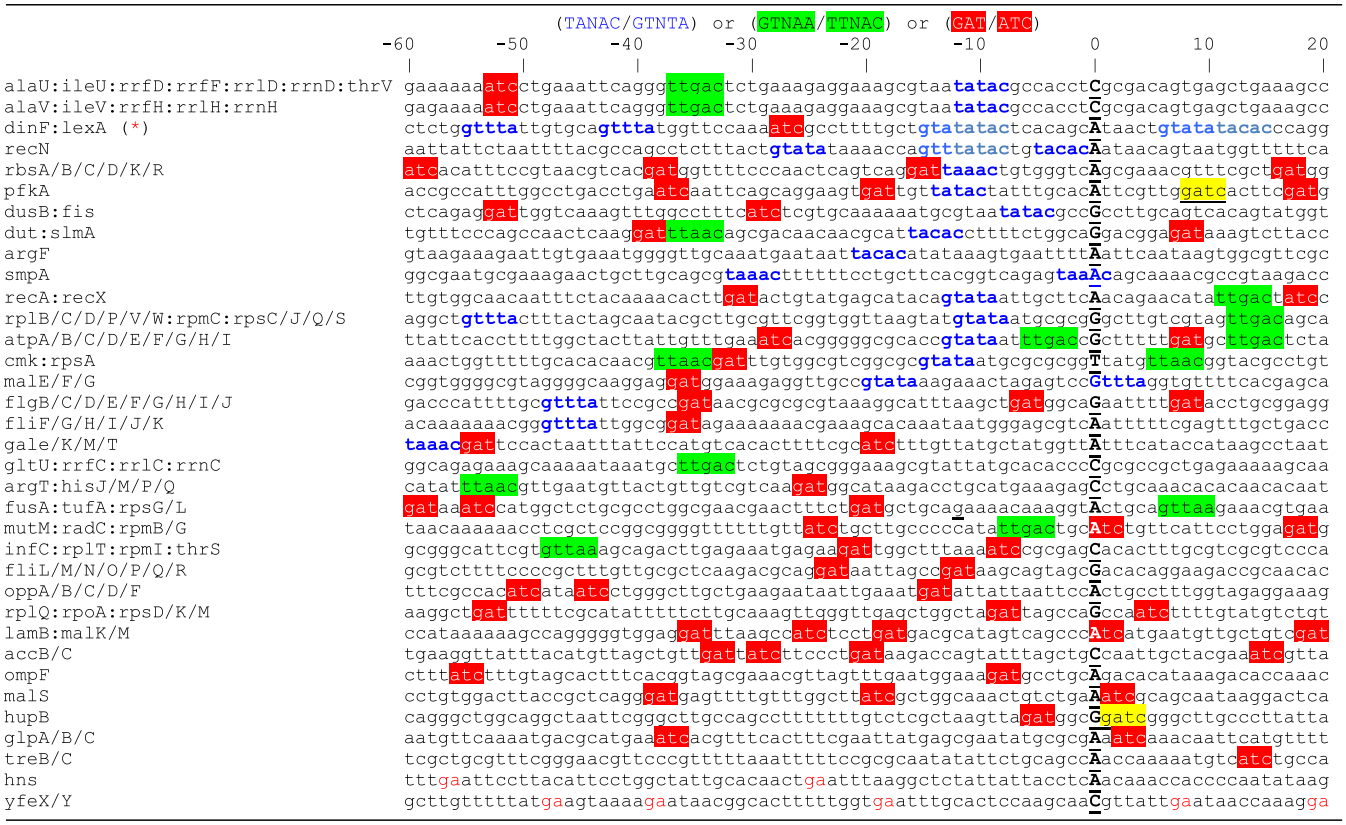
Note: GATC sites are in lower case and boxed. Variations of GATC sequence (GAT or GA or their complements) are in red. Motif GTYTA/TARAC is in blue and its variation GTYAA/TTRAC in green.

DISCUSSION

A new EcoDam binding site on DNA

We have demonstrated, by means of X-ray crystallography, that EcoDam binds to a 5-bp sequence distinct from the cognate GATC. This is different from the previously described preferences for certain sequences flanking GATC sites (18), or from the sometimes substantial binding preference for particular repeat symmetries in non-cognate DNA sequences (47). However, the new type of EcoDam binding site may affect binding to an adjacent GATC. The new sequence, GTYTA/TARAC (Y = C/T, R = A/G), immediately flanks the proximal and distal GATC sites in the Pap operon. These sites control the expression of pyelonephritis-associated pili. In the Dist site, we also found that EcoDam binds GACGA, the sequence overlapping with the 5-bp element and the cognate GATC (Figure 4a), which share two base pairs. The two elements near the Dist GATC, GTCTA (top strand) and GACGA (bottom strand), are oriented in opposite directions, suggesting EcoDam linear diffusion along DNA in either direction will result in the enzyme passing over a non-cognate site and thus affecting processivity of GATC methylation. The same could be true for the Prox site, where the top strand GTTTA and the bottom strand contains a 3/4-site (GAT) (Figure 4a). These non-cognate sites contain at least one G:C base pair (recognized by R124) and may also affect Dam intrasite hopping (48). A previous study also suggested that a GATC site with immediate 5′ polyA-tract (as occurred in the Pap operon) was methylated at a lower rate (18). The GATC site itself is palindrome, but the neighboring sequences at the Dist and the Prox sites are inverted (Figure 4a) as well as differing in a C:G versus A:T base pair, contributing to and possibly explaining the different methylation rates of the two target Ade on the same strand. T4Dam, the Dam ortholog of phage T4, has a similar ability to bind DNA sequences containing part (1/4-, 1/2- or 3/4-site) of GATC sequences (23). In T4Dam, we previously observed five different modes of T4Dam–DNA interaction ranging from nonspecific, through noncognate to specific (23). We suggest that EcoDam might be able to bind neighboring cognate and noncanonical sites in *pap* (Figure 5e), providing a potential Dam–Dam contact mechanism (either recruiting, or clashing with the second Dam molecule if the two sites are too close) that would work at adjacent cognate–noncognate DNA sites.

Table 3. Experimentally verified transcription units containing the genes that are differentially expressed in a *dam* mutant (adopted from (22))



*For LexA-regulated SOS genes, recombination protein RecA is chronically activated in *dam* cells (20) by the single-stranded DNA resulting from DNA damage, and facilitates autocatalytic cleavage of the LexA repressor (21).
Note that only two of these sequences, which include RpoD (σ^{70}) binding promoter sites, have full GATC Dam substrate sites (highlighted in yellow). For each 80nt line, a random sequence would be expected to contain 2.5 (GAT or ATC) motifs [2*80/(4^3)].

Regulatory implications of non-GATC EcoDam binding sites

To assess the possible role of EcoDam as a methylation-independent regulator, we searched the literature for Dam-regulated promoters. Table 2 shows promoters confirmed to be Dam responsive where GATC methylation plays a role, while Table 3 shows Dam-responsive promoters (whether due to direct or indirect effects) where methylation is not known to play a role—in fact only 2/35 of these Table 3 promoter regions have a GATC at all, and in those two cases the GATC is downstream of the transcription start. Focusing first on Table 2, and the methylation-dependent promoters, we found that two copies of the motif GTYTA are present in the regulatory region for another pilus type, *foo* (49), where the GTYTA and GATC sites are equivalent to the way they appear in the Pap operon. One copy of GTCTA is upstream of the protease gene *clp* (50), while a second copy contains a variation at fourth position of the element, GTTAA. Two copies of GTYAA are present in *sfa* (51) and *fot* (49). In addition, antigen 43 (52) and *fae* (53) contain one copy of GTTTA or GTTAA. In addition, variations of cognate GATC sequence (i.e. 3/4-site GAT or 1/2-site GA) can be found in *agn43* (the gene for antigen 43) and *fae*. For the division inhibitor *sulA*, the GATC overlaps the putative −35 hexamer, with an overlapping upstream GTTGA.

Table 3 was assembled because multiple earlier analyses of gene expression changes in *dam vs.* WT *E. coli* indicated a surprising lack of correlation between gene expression level and the presence of GATC sites upstream of the Dam-responsive transcriptional start sites (19,22). To further assess the role of the 5-bp element as a regulatory motif at the transcriptional level, we analyzed the upstream sequences of 35 transcription units that have been experimentally-verified as being differentially expressed in a *dam* mutant. Of these, 18 (51%) have TANAC or GTNTA (Table 3). In addition, 28 (80%) have ATC or GAT (i.e. the 3/4-site of cognate GATC sequence). Combining these two together, either the 5-bp element or GAT or both, it accounts for 33 (94%). The shorter or more degenerate sequences will, of course, occur more often in random sequences so, while they may nonetheless show significant non-cognate EcoDam binding (particularly where such binding is cooperative), it is necessary to be more cautious in interpreting those subsites. In particular, the promoter regions for *hns* and *yfeXY* have only 2 bp (1/2-site) sequences, while the single 3/4 site associated with *treBC* is downstream of the transcript start, and it may be that these three promoters are Dam-responsive only via indirect effects.

As mentioned above, EcoDam has low but measurable activity *in vitro* on the 3/4-sites that carry a base-pair substitution at the first position of NATC (23), implying that

binding of EcoDam occurs at near-cognate sites (Figure 6). A recent study of genome-wide mapping of methylated adenine residues in a pathogenic *E. coli* strain detected adenine-specific methylation of GATC sites as well as ATC/GAT sites, when the *dam* homologs were expressed in a plasmid system (54).

Based on our evidence for specific Dam binding to non-GATC sequences, and the occurrence of such sequences in Dam-responsive promoters, we suggest that Dam may function as a transcriptional repressor, in a methylation- and GATC-independent manner. It would be informative to know if a non-catalytic mutant of Dam still binds the non-cognate sites and has regulatory effects.

Finally, we note one technical implication of these findings. EcoDam has been used as a tool to assess mammalian chromatin structure by determining relative accessibility of GATC sites for methylation (55), or preferential methylation of specific GATC sites when EcoDam is fused to a mammalian regulatory protein (DamID) (56). The binding of Dam to noncanonical sites might subtly bias these uses of EcoDam.

ACCESSION NUMBERS

Protein Data Bank: the coordinates and structure factors of EcoDam–DNA complexes have been deposited (see Table 1 for accession numbers).

ACKNOWLEDGEMENTS

We thank Brenda Baker at the organic synthesis unit of New England Biolabs for synthesizing the oligonucleotides for crystallization. X.C. is a Georgia Research Alliance Eminent Scholar.

Author contributions: J.R.H. performed all crystallographic experiments and DNA binding assays. X.Z performed sequence analyses. R.M.B. participated in discussion and assisted in preparing the manuscript. X.C. and all authors were involved in analyzing data and preparing the manuscript.

FUNDING

The U.S. National Institutes of Health (NIH) [GM049245-21]; The Department of Biochemistry at the Emory University School of Medicine supported the use of the Southeast Regional Collaborative Access Team (SERCAT) synchrotron beamlines at the Advanced Photon Source of Argonne National Laboratory; The U.S. Department of Energy, Office of Science, Office of Basic Energy Sciences [W-31-109-Eng-38]; Georgia Research Alliance Eminent Scholar (to X.C.). Funding for open access charge: The open access publication charge for this paper has been waived by Oxford University Press - *NAR* Editorial Board members are entitled to one free paper per year in recognition of their work on behalf of the journal.

Conflict of interest statement. None declared.

REFERENCES

1. Lacks,S. and Greenberg,B. (1977) Complementary specificity of restriction endonucleases of Diplococcus pneumoniae with respect to DNA methylation. *J. Mol. Biol.*, **114**, 153–168.
2. Hattman,S., Brooks,J.E. and Masurekar,M. (1978) Sequence specificity of the P1 modification methylase (M.Eco P1) and the DNA methylase (M.Eco dam) controlled by the Escherichia coli dam gene. *J. Mol. Biol.*, **126**, 367–380.
3. Brooks,J.E., Blumenthal,R.M. and Gingeras,T.R. (1983) The isolation and characterization of the Escherichia coli DNA adenine methylase (dam) gene. *Nucleic Acids Res.*, **11**, 837–851.
4. Hattman,S. and Malygin,E.G. (2004) Bacteriophage T2Dam and T4Dam DNA-[N6-adenine]-methyltransferases. *Prog. Nucleic Res. Mol. Biol.*, **77**, 67–126.
5. Lobner-Olesen,A., Skovgaard,O. and Marinus,M.G. (2005) Dam methylation: coordinating cellular processes. *Curr. Opin. Microbiol.*, **8**, 154–160.
6. Marinus,M.G. and Casadesus,J. (2009) Roles of DNA adenine methylation in host-pathogen interactions: mismatch repair, transcriptional regulation, and more. *FEMS Microbiol. Rev.*, **33**, 488–503.
7. Val,M.E., Kennedy,S.P., Soler-Bistue,A.J., Barbe,V., Bouchier,C., Ducos-Galand,M., Skovgaard,O. and Mazel,D. (2014) Fuse or die: how to survive the loss of Dam in Vibrio cholerae. *Mol. Microbiol.*, **91**, 665–678.
8. Koch,B., Ma,X. and Lobner-Olesen,A. (2010) Replication of Vibrio cholerae chromosome I in Escherichia coli: dependence on dam methylation. *J. Bacteriol.*, **192**, 3903–3914.
9. Messer,W. and Noyer-Weidner,M. (1988) Timing and targeting: the biological functions of Dam methylation in E. coli. *Cell*, **54**, 735–737.
10. Campbell,J.L. and Kleckner,N. (1988) The rate of Dam-mediated DNA adenine methylation in Escherichia coli. *Gene*, **74**, 189–190.
11. Modrich,P. and Lahue,R. (1996) Mismatch repair in replication fidelity, genetic recombination, and cancer biology. *Annu. Rev. Biochem.*, **65**, 101–133.
12. Oshima,T., Wada,C., Kawagoe,Y., Ara,T., Maeda,M., Masuda,Y., Hiraga,S. and Mori,H. (2002) Genome-wide analysis of deoxyadenosine methyltransferase-mediated control of gene expression in Escherichia coli. *Mol. Microbiol.*, **45**, 673–695.
13. Lobner-Olesen,A., Marinus,M.G. and Hansen,F.G. (2003) Role of SeqA and Dam in Escherichia coli gene expression: a global/microarray analysis. *Proc. Natl. Acad. Sci. U.S.A.*, **100**, 4672–4677.
14. Hernday,A.D., Braaten,B.A. and Low,D.A. (2003) The mechanism by which DNA adenine methylase and PapI activate the pap epigenetic switch. *Mol. Cell*, **12**, 947–957.
15. Giacomodonato,M.N., Llana,M.N., Castaneda,M.D., Buzzola,F., Garcia,M.D., Calderon,M.G., Sarnacki,S.H. and Cerquetti,M.C. (2014) Dam methylation regulates the expression of SPI-5-encoded sopB gene in Salmonella enterica serovar Typhimurium. *Microb. Infect.*, **16**, 615–622.
16. Horton,J.R., Liebert,K., Bekes,M., Jeltsch,A. and Cheng,X. (2006) Structure and substrate recognition of the Escherichia coli DNA adenine methyltransferase. *J. Mol. Biol.*, **358**, 559–570.
17. Liebert,K., Horton,J.R., Chahar,S., Orwick,M., Cheng,X. and Jeltsch,A. (2007) Two alternative conformations of S-adenosyl-L-homocysteine bound to Escherichia coli DNA adenine methyltransferase and the implication of conformational changes in regulating the catalytic cycle. *J. Biol. Chem.*, **282**, 22848–22855.
18. Coffin,S.R. and Reich,N.O. (2008) Modulation of Escherichia coli DNA methyltransferase activity by biologically derived GATC-flanking sequences. *J. Biol. Chem.*, **283**, 20106–20116.
19. Robbins-Manke,J.L., Zdraveski,Z.Z., Marinus,M. and Essigmann,J.M. (2005) Analysis of global gene expression and double-strand-break formation in DNA adenine methyltransferase- and mismatch repair-deficient Escherichia coli. *J. Bacteriol.*, **187**, 7027–7037.
20. Peterson,K.R., Wertman,K.F., Mount,D.W. and Marinus,M.G. (1985) Viability of Escherichia coli K-12 DNA adenine methylase (dam) mutants requires increased expression of specific genes in the SOS regulon. *Mol. Gene Genet.*, **201**, 14–19.
21. Simmons,L.A., Foti,J.J., Cohen,S.E. and Walker,G.C. (2008) The SOS regulatory network. *EcoSal Plus*, doi:10.1128/ecosalplus.5.4.3.
22. Seshasayee,A.S. (2007) An assessment of the role of DNA adenine methyltransferase on gene expression regulation in E coli. *PLoS ONE*, **2**, e273.
23. Horton,J.R., Liebert,K., Hattman,S., Jeltsch,A. and Cheng,X. (2005) Transition from nonspecific to specific DNA interactions along the

substrate-recognition pathway of dam methyltransferase. *Cell*, **121**, 349–361.
24. Yang,Z., Horton,J.R., Zhou,L., Zhang,X.J., Dong,A., Zhang,X., Schlagman,S.L., Kossykh,V., Hattman,S. and Cheng,X. (2003) Structure of the bacteriophage T4 DNA adenine methyltransferase. *Nat. Struct. Biol.*, **10**, 849–855.
25. Pollak,A.J., Chin,A.T., Brown,F.L. and Reich,N.O. (2014) DNA looping provides for "intersegmental hopping" by proteins: a mechanism for long-range site localization. *J. Mol. Biol.*, **426**, 3539–3552.
26. Liebert,K., Hermann,A., Schlickenrieder,M. and Jeltsch,A. (2004) Stopped-flow and mutational analysis of base flipping by the Escherichia coli Dam DNA-(adenine-N6)-methyltransferase. *J. Mol. Biol.*, **341**, 443–454.
27. Elsawy,H., Podobinschi,S., Chahar,S. and Jeltsch,A. (2009) Transition from EcoDam to T4Dam DNA recognition mechanism without loss of activity and specificity. *Chembiochem*, **10**, 2488–2493.
28. Studier,F.W. (2005) Protein production by auto-induction in high density shaking cultures. *Protein Expr. Purif.*, **41**, 207–234.
29. Tong,L. and Rossmann,M.G. (1997) Rotation function calculations with GLRF program. *Methods Enzymol.*, **276**, 594–611.
30. Brunger,A.T., Adams,P.D., Clore,G.M., DeLano,W.L., Gros,P., Grosse-Kunstleve,R.W., Jiang,J.S., Kuszewski,J., Nilges,M., Pannu,N.S. *et al.* (1998) Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.*, **54**, 905–921.
31. Jones,T.A., Zou,J.Y., Cowan,S.W. and Kjeldgaard,M. (1991) Improved methods for building protein models in electron density maps and the location of errors in these models. *Acta Crystallogr. A*, **47**, 110–119.
32. Adams,P.D., Afonine,P.V., Bunkoczi,G., Chen,V.B., Davis,I.W., Echols,N., Headd,J.J., Hung,L.W., Kapral,G.J., Grosse-Kunstleve,R.W. *et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 213–221.
33. Emsley,P., Lohkamp,B., Scott,W.G. and Cowtan,K. (2010) Features and development of Coot. *Acta Crystallogr. D Biol. Crystallogr.*, **66**, 486–501.
34. Emsley,P. and Cowtan,K. (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.*, **60**, 2126–2132.
35. Horowitz,S. and Trievel,R.C. (2012) Carbon-oxygen hydrogen bonding in biological structure and function. *J. Biol. Chem.*, **287**, 41576–41582.
36. Peterson,S.N. and Reich,N.O. (2008) Competitive Lrp and Dam assembly at the pap regulatory region: implications for mechanisms of epigenetic regulation. *J. Mol. Biol.*, **383**, 92–105.
37. Kawamura,T., Vartanian,A.S., Zhou,H. and Dahlquist,F.W. (2011) The design involved in PapI and Lrp regulation of the pap operon. *J. Mol. Biol.*, **409**, 311–332.
38. Mashhoon,N., Carroll,M., Pruss,C., Eberhard,J., Ishikawa,S., Estabrook,R.A. and Reich,N. (2004) Functional characterization of Escherichia coli DNA adenine methyltransferase, a novel target for antibiotics. *J. Biol. Chem.*, **279**, 52075–52081.
39. Bergerat,A., Kriebardis,A. and Guschlbauer,W. (1989) Preferential site-specific hemimethylation of GATC sites in pBR322 DNA by Dam methyltransferase from Escherichia coli. *J. Biol. Chem.*, **264**, 4064–4070.
40. Hashimoto,H., Horton,J.R., Zhang,X., Bostick,M., Jacobsen,S.E. and Cheng,X. (2008) The SRA domain of UHRF1 flips 5-methylcytosine out of the DNA helix. *Nature*, **455**, 826–829.
41. Bowman,B.R., Lee,S., Wang,S. and Verdine,G.L. (2008) Structure of the E. coli DNA glycosylase AlkA bound to the ends of duplex DNA: a system for the structure determination of lesion-containing DNA. *Structure*, **16**, 1166–1174.
42. Wibley,J.E., Waters,T.R., Haushalter,K., Verdine,G.L. and Pearl,L.H. (2003) Structure and specificity of the vertebrate anti-mutator uracil-DNA glycosylase SMUG1. *Mol. Cell*, **11**, 1647–1659.
43. Barrett,T.E., Savva,R., Barlow,T., Brown,T., Jiricny,J. and Pearl,L.H. (1998) Structure of a DNA base-excision product resembling a cisplatin inter-strand adduct. *Nat. Struct. Biol.*, **5**, 697–701.
44. Roberts,R.J. and Cheng,X. (1998) Base flipping. *Annu. Rev. Biochem.*, **67**, 181–198.
45. Klimasauskas,S. and Roberts,R.J. (1995) M.HhaI binds tightly to substrates containing mismatches at the target base. *Nucleic Acids Res.*, **23**, 1388–1395.
46. Jurkowski,T.P., Anspach,N., Kulishova,L., Nellen,W. and Jeltsch,A. (2007) The M.EcoRV DNA-(adenine N6)-methyltransferase uses DNA bending for recognition of an expanded EcoDam recognition site. *J. Biol. Chem.*, **282**, 36942–36952.
47. Afek,A., Schipper,J.L., Horton,J., Gordan,R. and Lukatsky,D.B. (2014) Protein-DNA binding in the absence of specific base-pair recognition. *Proc. Natl. Acad. Sci. U.S.A.*, 17140–17145.
48. Pollak,A.J. and Reich,N.O. (2012) Proximal recognition sites facilitate intrasite hopping by DNA adenine methyltransferase: mechanistic exploration of epigenetic gene regulation. *J. Biol. Chem.*, **287**, 22873–22881.
49. Daigle,F., Forget,C., Martin,C., Drolet,M., Tessier,M.C., Dezfulian,H. and Harel,J. (2000) Effects of global regulatory proteins and environmental conditions on fimbrial gene expression of F165(1) and F165(2) produced by Escherichia coli causing septicaemia in pigs. *Res. Microbiol.*, **151**, 563–574.
50. Crost,C., Garrivier,A., Harel,J. and Martin,C. (2003) Leucine-responsive regulatory protein-mediated repression of clp (encoding CS31A) expression by L-leucine and L-alanine in Escherichia coli. *J. Bacteriol.*, **185**, 1886–1894.
51. van der Woude,M.W. and Low,D.A. (1994) Leucine-responsive regulatory protein and deoxyadenosine methylase control the phase variation and expression of the sfa and daa pili operons in Escherichia coli. *Mol. Microbiol.*, **11**, 605–618.
52. Waldron,D.E., Owen,P. and Dorman,C.J. (2002) Competitive interaction of the OxyR DNA-binding protein and the Dam methylase at the antigen 43 gene regulatory region in Escherichia coli. *Mol. Microbiol.*, **44**, 509–520.
53. Huisman,T.T. and de Graaf,F.K. (1995) Negative control of fae (K88) expression by the 'global' regulator Lrp is modulated by the 'local' regulator FaeA and affected by DNA methylation. *Mol. Microbiol.*, **16**, 943–953.
54. Fang,G., Munera,D., Friedman,D.I., Mandlik,A., Chao,M.C., Banerjee,O., Feng,Z., Losic,B., Mahajan,M.C., Jabado,O.J. *et al.* (2012) Genome-wide mapping of methylated adenine residues in pathogenic Escherichia coli using single-molecule real-time sequencing. *Nat. Biotechnol.*, **30**, 1232–1239.
55. Bulanenkova,S.S., Kozlova,A.A., Kotova,E.S., Snezhkov,E.V., Azhikina,T.L., Akopov,S.B., Nikolaev,L.G. and Sverdlov,E.D. (2011) Dam methylase accessibility as an instrument for analysis of mammalian chromatin structure. *Epigenetics*, **6**, 1078–1084.
56. van Steensel,B., Delrow,J. and Henikoff,S. (2001) Chromatin profiling using targeted DNA adenine methyltransferase. *Nat. Genet.*, **27**, 304–308.
57. Malone,T., Blumenthal,R.M. and Cheng,X. (1995) Structure-guided analysis reveals nine sequence motifs conserved among DNA amino-methyltransferases, and suggests a catalytic mechanism for these enzymes. *J. Mol. Biol.*, **253**, 618–632.