Structural basis of asymmetric DNA methylation and ATP-triggered long-range diffusion by EcoP15I

Yogesh K. Gupta1, Siu-Hong Chan2, Shuang-yong Xu2 & Aneel K. Aggarwal1

Type III R-M enzymes were identified >40 years ago and yet there is no structural information on these multisubunit enzymes. Here we report the structure of a Type III R-M system, consisting of the entire EcoP15I complex (Mod2Res1) bound to DNA. The structure suggests how ATP hydrolysis is coupled to long-range diffusion of a helicase on DNA, and how a dimeric methyltransferase functions to methylate only one of the two DNA strands. We show that the EcoP15I motor domains are specifically adapted to bind double-stranded DNA and to facilitate DNA sliding via a novel ‘Pin’ domain. We also uncover unexpected ‘division of labour’, where one Mod subunit recognizes DNA, while the other Mod subunit methylates the target adenine—a mechanism that may extend to adenine N6 RNA methylation in mammalian cells. Together the structure sheds new light on the mechanisms of both helicases and methyltransferases in DNA and RNA metabolism.
We provide here a basis for two prevailing questions in the study of DNA helicases and DNA methyltransferases. First, how is ATP hydrolysis coupled to long-range diffusion of a helicase on DNA? And second, how does a dimeric methyltransferase methylate one DNA strand but not the other? Both of these activities are embodied in the Type III restriction–modification (R–M) systems in bacteria and archaea, which play roles analogous to the innate immune system in higher eukaryotes2–4. R–M systems are categorized in four groups (Type I–IV) based on their subunit assembly, cofactor requirements and associated cleavage patterns5.

EcoP15I is a prototype of the Type III R–M family and consists of two methyltransferase (Mod) and one (or two) restriction (Res) subunits6, resulting in a Mod1Res1 or Mod1Res2 complex7–10. The Mod subunits are responsible for DNA recognition and methylation, while the Res subunits are responsible for ATP hydrolysis and cleavage. Cleavage only occurs if the two recognition sites (CAGCAG) are in an inverted-repeat orientation, arranged either as ‘head-to-head’ or ‘tail-to-tail’11,12. The sites can be separated by thousands of base pairs but ATP hydrolysis is absolutely required for cleavage8.

According to classical helicase mechanisms, motion of EcoP15I should require one ATP per base pair moved on the DNA. However, even though EcoP15I contains the standard helicase motifs found in members of superfamily 2 (SF2) helicases and translocases13,14, the enzyme communicates over thousands of base pairs by consuming only a few ATP15,16. A number of competing models have been proposed for this long-range communication, including classical translocation and three-dimensional DNA looping17,18. More recently, single-molecule and ensemble fluorescence measurements have shown that EcoP15I undergoes random one-dimensional diffusion along the DNA16,19. The diffusion coefficient is one of the largest yet measured for a DNA sliding process \(D = 0.92 \pm 0.06 \mu m^2/s\)19,20, with DNA cleavage only occurring when the freely diffusive EcoP15I collides with a stationary enzyme bound to the second DNA site19 (Supplementary Fig. 1). Together, these results posit a new functionality for helicases, as molecular switches for long-lived DNA sliding (rather than conventional DNA/RNA unwinding or stepwise translocation)21. EcoP15I hydrolyses \(\sim 30\) ATP molecules in two steps (a fast consumption of \(\sim 10\) ATP molecules followed by a slower consumption of \(\sim 20\) ATP molecules), which switches the enzyme into a distinct structural state that can diffuse on DNA over long distances19. A similar sliding-based mechanism has been proposed for the mismatch repair protein MutS (and its eukaryotic homologue), where nucleotide exchange (rather than hydrolysis) triggers sliding on DNA after mismatch recognition22–24. Altogether, ATP-triggered sliding is an emerging theme in helicase-like enzymes, but questions about the mechanism remain unanswered due to the lack of structural data.

Intriguingly, the EcoP15I Mod subunit is also unusual in that it functions as a dimer25 as compared to monomeric methyltransferases26,27. In this respect, it is similar to several DNA methyltransferases in mammals (Dnmt3a/3b/3L) and plants (Domains rearranged methyltransferase 2, DRM2) that also function as dimers or other higher order oligomers in processes ranging from de novo DNA methylation to RNA-directed DNA methylation28,29. Similar to DRM2, for example, EcoP15I methylates one DNA strand but not the other. The EcoP15I mode of action also extends to the methylation of adenine N6 in mammalian mRNAs, mediated by a heterodimer of methyltransferase-like 3 and 14 (METTL3 and METTL14), which has been shown to be critical for cellular homeostasis and stem cell commitment and differentiation30,31. Like the Mod subunit, METTL3 and METTL14 belong to the \(\beta\) class of amino methyltransferases and may operate in a similar dimer mode as EcoP15I. Despite the availability of several crystal structures of dimeric DNA methyltransferases (all in DNA-free form), the structural basis for asymmetric DNA (and RNA) methylation has remained mysterious.

Although Type III R–M enzymes were identified \(\sim 40\) years ago1–4,32, there is still no structural information on these multisubunit enzymes that encompass DNA methylation, DNA translocation and DNA cleavage activities all within the same complex. Here we report the structure of the entire EcoP15I complex (Mod1Res1) bound to its DNA substrate. The structure is only the second (after \(S.\) solfataricus RAD54) of an SF2 helicase with a duplex DNA bound to the motor domains33. As such, it provides new insights into the mechanisms of DNA translocation and the nature of the conformational changes that switches EcoP15I into a long-lived sliding machine. The structure is also the first, to our knowledge, of a dimeric DNA methyltransferase bound to DNA. It reveals a remarkable division of labour, where one Mod subunit recognizes the DNA, while the other Mod subunit methylates the target adenine base. Together these structural features shed new light on the diversity of helicases and methyltransferases in DNA and RNA metabolism.

Results

Structure determination. The EcoP15I holoenzyme was co-crystallized with a 20-mer DNA duplex containing a single EcoP15I recognition site (CAGCAG). The best co-crystals were obtained in the presence of AMP and diffracted to \(\sim 2.6\) Å resolution with synchrotron radiation. The co-crystals belong to space group \(P4_1_2_1_2\) with unit cell dimensions of \(a = b = 101\) Å, \(c = 533\) Å and \(\alpha = \beta = \gamma = 90^\circ\) and contain one EcoP15I/DNA/AMP complex in the crystallographic asymmetric unit (Table 1). The structure was determined by the multiple isomorphous replacement with anomalous scattering (MIRAS) method and refined to \(2.6\) Å resolution (Table 1). The refined model consists of two Mod subunits (ModA, residues 13–644; ModB, residues 2–644), one Res subunit (residues 7–810), 20-mer DNA (nucleotides 1–20 on each strand), one AMP molecule, 3 ions and 103 solvent molecules. Regions of protein with no electron density were omitted, and amino acids with weak side chain densities were modelled as alanines (Supplementary Table 1). The current model lacks theendonuclease portion of the Res subunit due to the lack of electron density for this region.

Overall architecture. The EcoP15I Mod1Res1 heterotrimer embraces the DNA duplex and makes extensive protein–DNA contacts. The Mod subunits engage the upstream portion of the DNA duplex that contains the CAGCAG recognition sequence (Fig. 1). The target adenine (CAGCAG) rotates out of the DNA helix and enters the catalytic pocket of one of the two Mods, ModB. The other Mod, ModA, makes the majority of base-specific contacts with the CAGCAG recognition sequence. The Res subunit interacts with the downstream portion of the DNA, approximately one half-turn away from recognition site (Fig. 1). Only the helicase core of the Res subunit is visible in the electron density map; the cleavage domain is disordered and may only become ordered when the enzyme collides with another EcoP15I complex and becomes cleavage competent. AMP lies in a cleft in the helicase core.

Each Mod subunit is composed of four domains, an amino-terminal domain (NTD, aa 14–60), a central methyltransferase domain (MTase, aa 62–262, 390–516), a target recognition domain (TRD, aa 263–384) and a carboxy-terminal domain (CTD, aa 539–644; Fig. 2a,b, Supplementary Fig. 2). The MTase domain contains nine motifs (I–VIII and X) characteristic of
The helicase core of the Res subunit is composed of tandem RecA-like domains\(^3\) (Supplementary Fig. 4b), N-terminal RecA1 (aa 7–269) and C-terminal RecA2 (aa 366–594), followed by a helical spacer (aa 620–810; Fig. 2b). The spacer connects to the endonuclease domain (disordered in the structure). Each RecA-like domain consists of a central β sheet of six to seven parallel β-strands sandwiched by helices. AMP binds to the ‘bottom’ side of the cleft at the confluence of two domains, while the DNA duplex is accommodated on the ‘top’ side of the cleft (Fig. 2b). The AMP is highly mobile in the structure (B-factor of 139 Å\(^2\)).

The AMP is highly mobile in the structure (B-factor of 139 Å\(^2\)).
are located on loops that line the cleft (Fig. 2b). Altogether, RecA1 contains the classical motifs Q, I (or Walker A), Ia, Ib, Ic, II (or Walker B), IIa and III, whereas RecA2 contains the motifs IV, IVa, V and VI (Fig. 2b and Supplementary Fig. 5). The specificity of helicases and translocases for different substrates is dictated to a large extent by accessory domains derived from ‘inserts’ in RecA1 or RecA2, or from the N- and C-terminal flanking sequences. In DNA and RNA helicases, for example, an accessory domain can act as a ‘wedge’ to disrupt base pairing for the unwinding reaction. In EcoP15I, we identify three new substructures, namely a loop after motif Ic (‘Ic-extension’; aa 198–211), a β-hairpin-like ‘Q-arm’, formed by an ~50aa insertion (aa 28–77) in RecA1, and a more elaborate substructure, ‘Pin’ domain, formed by an ~77aa insertion (aa 288–365) in RecA2 (Supplementary Table 2). The Pin domain adopts a β-sandwich-like tertiary structure with two overlaid β-sheets that extends towards the ModA TRD and interacts with the translocating strand of the DNA duplex (Figs 1 and 2b). The Pin domain is highly mobile (B-factor of 89 Å²), but we could assign the main chain and some of the side chains.

**DNA conformation.** The DNA is severely distorted from B-form at two sites along its axis (Fig. 2c and Supplementary Fig. 6). First,
the site where the target adenine is ejected from the recognition sequence (CAGCAG), and second, near the ModA–Res interface (Fig. 2c and Supplementary Fig. 6). For convenience, we refer to the DNA strand containing the target adenine as the ‘methylating’ strand, and the complementary DNA strand as the ‘translocating’ strand (which makes the majority of contacts with the motor domains—described later; Fig. 2c). The distortions around target adenine and the recognition sequence are mainly induced by the intrusion of ModA TRD in the DNA major groove (Fig. 1). At the ModA–Res junction the DNA is bent ∼24° towards the minor groove, in the direction of the ModA TRD and the Res Pin domain (Figs 1 and 2c). At the site of bending, the torsion angles ε (C3′–O3′) and ζ (O3′–P) are gauche−, trans rather than more characteristic trans, gauche− conformation found in B-DNA35. Analogous deviations in torsion angles were observed for inner thymines in DNA bound to BglII, where the DNA experiences an overall bend of ∼23° (ref. 36). Most importantly, the ∼24° bend in the EcoP15I DNA reduces the distance between the ModA TRD and the Res Pin domain to <14 Å and may facilitate an interaction between the two domains when EcoP15I assumes a diffusive or sliding state on DNA (Fig. 1).

Division of labour: DNA recognition and methylation. The EcoP15I structure is the first of a β-class of an amino DNA methyltransferase bound to DNA and it suggests a fundamentally different mechanism of methylation. There is not only a division of labour between two Mod subunits in terms of DNA recognition (ModA) and methylation (ModB), but also the methylating subunit (ModB) binds to DNA in a radically different manner from other methyltransferases.

ModA makes the majority of base-specific contacts, via the bilobed TRD that tracks the DNA major groove and interacts with bases over the entire length of the recognition sequence
In contrast, ModB makes only a few contacts to bases and its role is mainly to methylate the target adenine (CAGCAG). The adenine rotates ~180° out of DNA helix and enters the ModB catalytic cleft (Fig. 2a). In contrast, the ModA catalytic cleft is empty and lies 30 Å away from the DNA (Fig. 2a). Compared with the other amino methyltransferases, ModB pivots around the aspartate/asparagine of the conserved D/NPY catalytic sequence (motif IV) by ~140° so that the 'PPY' sequence runs along the Watson–Crick edge of the extrahelical adenine base rather than the Hoogsteen edge (Fig. 3b and Supplementary Fig. 8), and the conserved tyrosine (PPY) stacks on opposite face of the base (Supplementary Fig. 8c–e). This unusual mode of DNA docking is a consequence of Mod2 dimerization and interactions with the Res subunit, whereby if ModB were to assume the same orientation as say in the monomeric M.TaqI/DNA complex then ModA would not be in the correct position to recognize the CAGCAG sequence. In addition, the ModA TRD would directly clash with the Res subunit bound to downstream portion of the DNA (Supplementary Fig. 8f).

**The RecA motor domains.** EcoP15I is only the second SF2 translocase (after ssRad54) to be crystallized with a duplex DNA bound to the RecA motor domains. As with ssRad54 (ref. 33), the EcoP15I motor domains interact predominantly with one strand of the DNA duplex—the translocating strand (Figs 1 and 4). However, the motor domains in ssRad54 adopt an unusual arrangement, in which RecA2 is flipped 180° with respect to RecA1, limiting the number of possible interactions with the DNA (Supplementary Fig. 9a)33. The EcoP15I motor domains adopt a more canonical configuration, which is intermediate...
between the fully ‘closed’ configuration observed in the SF2 RNA helicase VASA/ssDNA/AMPPNP complex and the ‘open’ configuration observed in zebrafish Rad54 (zRad54) (Fig. 4). This ‘semi-closed’ configuration (16° outward motion of RecA2 when compared with VASA) reinforces the notion that contacts primarily to one DNA strand is a conserved feature in different helicase structures. In monomeric M.HhaI (class) and M.TaqI (γ class) for example, appears to represent an intermediate state, following the position of the translocating DNA strand in the EcoP15I structure overlays with the ssRNA in VASA and NS3 RNA helicase structures, reinforcing the notion that contacts to DNA from RecA2 are from the mains chain amides of residues on the accessory Pin domain (Gly352, Ile353, Glu354 and Ser359).

Both RecA1 and RecA2 interact with the DNA duplex (Figs 1 and 4). The RecA1 residues Thr116 and Leu117 of motif Ia, Ser151 of motif Ib, and Asn187, Met190, Ser193 and Lys194 of motif Ic interact with successive phosphates on the translocating DNA strand, while residues Lys235, Lys236 and Thr237 on the switch II region interact with the opposite methylating DNA strand (Fig. 4). RecA2 interacts with the more downstream portion of the translocating DNA strand via amino acids Thr503 and Arg505 of motif V. In particular, the Pin domain in RecA2 is involved in a number of hydrogen bonds with DNA via the main chain amides of Gly354 and Lys356 and the side chain of Ser359, as well as hydrophobic contacts via Gly352 and Ile353 (Fig. 4). Altogether, the EcoP15I motor domains are specifically adapted to bind double-stranded (ds) DNA. Importantly, there is no equivalent of a ‘wedge’ to separate DNA strands of the DNA duplex, but instead the Pin domain in RecA2 augments interactions with the backbone of the translocating DNA strand and which is more apt for diffusion on ds DNA (Fig. 4).

**Discussion**

The multisubunit Type I and Type III enzymes are exceptional in their dependency on ATP for restriction activity. By contrast, the Type II enzymes do not require ATP and majority of them harbour functionally independent R and M subunits with the exception of a few enzymes like Type IIG BpuSI or Type III Mmef40,41. Type IV enzymes only cleave modified DNA substrates. The Type III R–M enzymes have defined structural interpretation for >40 years. We report here the first structure of a Type III R–M system, consisting of the entire EcoP15I complex (Mod1,Res1) bound to its DNA substrate. The structure provides unprecedented new insights into the molecular underpinnings of asymmetric DNA methylation and ATP-triggered DNA diffusion.

The early structures of DNA methyltransferases with DNA revealed monomeric enzymes with the ability to both recognize and methylate DNA. This feature extended to both cytosine C5 and adenine N6 methyltransferases26,27. As such, much of the subsequent data on DNA methyltransferases have been interpreted in a context of a monomer, even in cases where they were observed as dimers42–44. All of the current structural information on dimeric DNA methyltransferases is limited to crystal structures in the absence of DNA. The EcoP15I structure provides a mechanistic basis for the action of β amino methyltransferases, which are observed primarily as dimers in solution or in crystals42–44. Indeed, it is conceivable that this entire subfamily of DNA methyltransferases works in the same manner as EcoP15I, wherein one subunit recognizes the DNA while the other subunit methylates the target base. The β-amino methyltransferases differ from methyltransferases in other subfamilies in how the TRD is positioned with respect to the MTase domain. In monomeric M.HhaI (α class) and M.TaqI (γ class), for example, the TRD is adjacent to the active site cleft and in a direction that permits it to enter the DNA major groove next to the flipped target base. By contrast, in EcoP15I, the TRD lies far off from the active site cleft and in a direction that makes it geometrically impossible for a single Mod subunit to both methylate a target base and recognize the DNA sequence; instead it is reliant on the TRD of the second Mod subunit.

Strikingly, this division of labour may also extend to RNA methylation in mammalian cells40. In particular, adenine N6 methylation is the most prevalent modification in the body of nuclear and cytosolic RNAs in mammals and is implicated in processes ranging from mRNA splicing to translation regulation45,46. Intriguingly, adenine N6 methylation of mRNA has also been shown recently to be critical for...
stem cell commitment and differentiation\textsuperscript{31,47,48}. The modification often occurs in the context of a G(G/A)ACU sequence and the enzyme(s) responsible have recently been identified as the METTL3/METTL14 heterodimer\textsuperscript{49–51}. Intriguingly, both METTL3 and METTL14 contain motifs (X, I–VIII) characteristic of amino methyltransferase\textsuperscript{52}, including the equivalent of the ‘DPY’ sequence in motif I (DPYW in METTL3 and EPPI in METTL14). The linear order of these motifs suggests that both METTL3 and METTL14 belong to the β-amino class of methyltransferases and—based on the EcoP15I structure—it is conceivable that one methyltransferase (METTL3 or METTL14) plays a more dominant role in recognition of the G(G/A)ACU sequence (and perhaps also the surrounding RNA secondary structure), while the other plays a more central role in adenine methylation.

The EcoP15I methyltransferase mechanism may also extend to other subfamilies of DNA methyltransferases, such as the plant de novo DNA methyltransferase DRM2, which functions as a homodimer\textsuperscript{28}. On the basis of the EcoP15I structure, one can envisage a mechanism where one DRM2 monomer recognizes the DNA sequence context (CG/CHG/CHH, where H = A, T, or C) while the other methylates the target cytosine. The dimeric RNA MTases from SPOUT family display another form of division of labour in which the RNA binds in a cleft between the two monomers, whereas the target RNA base for methylation resides in the catalytic pocket of one monomer\textsuperscript{53–55}. Altogether, a division of labour between two or more methyltransferase subunits appears to be a more general mechanism in DNA and RNA methylation. The EcoP15I structure provides a framework for beginning to understand the interplay between different methyltransferase subunits.

As recently as 1993, helicases were considered as DNA- or RNA-unwinding machines that couple ATP hydrolysis for the unwinding reaction\textsuperscript{56}. The discovery that many helicases are actually translocases (especially those in the SF2 superfamily) has changed this view\textsuperscript{14}; however, even this view has been found wanting with the discovery that several helicases or translocases behave as molecular switches\textsuperscript{21}. These molecular switches have also been referred to as pseudo-helicases\textsuperscript{19}, wherein ATP hydrolysis is coupled to a conformational change in the enzyme for thermally driven diffusion on the DNA or RNA. The EcoP15I structure uncovers a helicase motor that is generally similar to that observed in classical helicases and translocases, composed of tandem RecA-like domains with Walker A and B motifs and an arginine finger, among other classical motifs\textsuperscript{14}.

What is the nature of the ATP-triggered conformational change for diffusion on DNA? This structure here provides some interesting clues. In particular, proximity of the ModA TRD to the Pin domain in RecA2 suggests a model in which the TRD may switch its location from the DNA major groove to the Pin domain and hence, adopt a ‘nonspecific’ conformation more amenable to DNA sliding\textsuperscript{19} (Fig. 5). The TRD is joined to the MTase domain by a flexible linker, and a simple rotation of ~40° about this linker puts the ModA TRD in direct contact with the Pin domain—sequestered away from the DNA (Fig. 5a). Moreover, the DNA duplex is bent by ~24° at this precise ModA–Res nexus, which reduces the distance between the TRD and the Pin domain to <14 Å. The Pin domain is highly mobile and may only become fully ordered when it recruits the ModA TRD. The asymmetric nature of ModA and ModB DNA binding seems to ensure that only one TRD (and not both) needs to be drawn away from the DNA. Overall, the structural model is in accord with single-molecule studies, which suggest that the entire Mod\textsubscript{A}Res\textsubscript{A} complex diffuses on the DNA (and not just the Res subunit) until it collides with another complex to become cleavage competent (Fig. 5b, Supplementary Fig. 1)\textsuperscript{19}.

In conclusion, we present here the first structure of a Type III R–M system, consisting of the entire EcoP15I complex (Mod\textsubscript{A}Res\textsubscript{A}) bound to its DNA substrate. Asymmetric methylation and ATP-triggered DNA diffusion are emerging themes in the study of methyltransferases and helicases but the mechanisms remain unclear. Plant DRM2 homodimer\textsuperscript{28} and the mammalian METTL3/METTL14 heterodimer\textsuperscript{50,45,46,49–51}, for example, may operate in a similar manner to EcoP15I, where one monomer recognizes the DNA or RNA sequence context, while the other methylates the target base. Neither DRM2 nor METTL3/METTL14 possesses helicase activity. Furthermore, an EcoP15I type DNA sliding-based mechanism has also been proposed for the mismatch repair protein MutS (and its eukaryotic homologue), but where nucleotide exchange (rather than hydrolysis) triggers sliding on DNA after mismatch recognition\textsuperscript{22–24}. Similarly, the loading of processivity clamps at replication forks by clamp loaders occurs via a two-step conformational change mediated by ATP binding\textsuperscript{27}. Altogether, the EcoP15I structure proffers unprecedented new insights into the molecular underpinnings of asymmetric DNA/RNA methylation and ATP-triggered thermal diffusion in broad array of DNA and RNA metabolism.

**Methods**

**Expression and purification.** The genes encoding Res and Mod subunits of EcoP15I were subcloned from a plasmid kindly provided by Dr D.N. Rao (Indian Institute of Science) into an expression vector pRRS\textsuperscript{58}. E. coli expression host NEB Express (NEB) was transformed and was grown in LB medium containing 100 µg ml\textsuperscript{−1} of ampicillin. Protein expression was carried out for 18 h at 30°C.
The harvested cells from 6 l of culture were lysed and the derived cell pellet was centrifuged at a final concentration of >10 mg ml⁻¹.

**Crystallization.** We co-crystallized EcoP15I complex in presence of a 20-mer DNA duplex and AMP. The crystals were obtained in a hanging drop set up by mixing 1 l of EcoP15I/DNA/AMP complex with 1 l of precipitant solution containing 10% PEG 5000 monomethyl ether, 0.1 M HEPES pH 7.5, 0.2 M potas-
sium acetate and 15 mM MnCl₂ at 20 °C, and were cryoprotected by serial transfer into mother liquor containing 30% PEG 5000 MME and 10% PEG400 before plunging them into liquid N₂. Neither S-adenosyl methionine (AdoMet) nor its analogue AdoHcy was included during purification or crystallization. The crystals belong to the space group P4₁2₁2 with unit cell dimensions of a = b = 101 Å, c = 533 Å and α = β = γ = 90°. X-ray diffraction data were measured at beamlines NECAT-24IDC at Advanced Photon Source (APS), and X4A, X25 and X29 at NSLS of Brookhaven National Laboratory (BNL; Table 1).

**Structure determination.** To calculate the experimental phases for structure determination, we used X-ray data from native crystals and seven heavy atom derivatives (Se, Br, I, Ta, Sm, Co, Ho). The phases were calculated by the MIRAS method, using the program TRUESHARP59. The bromine and iodine derivatives were prepared by soaking native crystals into the mother liquor containing crystals was measured at a longer wavelength (2.07 Å) on five of such crystals and processed, merged, and scaled using the XDS programme package64. Crystals and derivatives, the Se-Met data set gave the best phases (anomalous phasing power of 26.5 and 21.9%, respectively (Table 1).

**References.**

1. Haberman, A. The bacteriophage P1 restriction endonuclease. *J. Mol. Biol.* 89, 545–563 (1974).
2. Haberman, A., Heywood, J. & Meselson, M. DNA modification methylase activity of Escherichia coli restriction endonucleases K and P. *Proc. Natl Acad. Sci. USA* 69, 3318–3314 (1972).
3. Meselson, M. & Yuan, R. DNA restriction enzyme from *E. coli*. *Nature* 217, 1110–1114 (1968).
4. Arber, W. & Wauters-Willems, D. Host specificity of DNA produced by *Escherichia coli* XII. The two restriction and modification systems of *Escherichia coli*. *J. Mol. Biol.* 80, 203–217 (1974).
5. Roberts, R. I., Vincze, T., Posfai, J. & Macelis, D. RELBASE—a database for DNA restriction and modification: enzymes, genes and genomes. *Nucleic Acids Res.* 43, D298–D299 (2015).
6. Raghavendra, N. K., Bheemanaik, S. & Rao, D. N. Mechanistic insights into type III restriction enzymes. *Roadmaps* (2011), 1079–1083 (2013).
7. Gupta, Y. K. et al. Structural insights into the assembly and shape of type III restriction modification (R-M) EcoP15I complex by small-angle X-ray scattering. *J. Mol. Biol.* 420, 261–268 (2012).
8. Janscak, P., Sandmeier, U., Szczelkun, M. D. & Bickle, T. A. Subunit assembly and mode of DNA cleavage of the type III restriction endonucleases EcoP11 and EcoP15I. *J. Mol. Biol.* 306, 417–431 (2001).
9. Wollmert, K. H. & Posfai, J. Type III restriction endonuclease EcoP15I is a heterotrimetric complex containing one Res subunit with several DNA-binding regions and ATPase activity. *Nucleic Acids Res.* 40, 3610–3622 (2012).
10. Butterer, A. et al. Type III restriction endonucleases are heterotrimetric: comprising one helicase-nuclease subunit and a dimeric methytransferase that binds only one specific DNA. *Nucleic Acids Res.* 42, 5139–5150 (2014).
11. Eng, A., Bickle, T. A., Kruger, D. H. & Schroeder, C. Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. *Nature* 355, 467–469 (1992).
12. van Aelst, K. et al. Type III restriction enzymes cleave DNA by long-range interaction between sites in both head-to-head and tail-to-tail inverted repeat. *Proc. Natl Acad. Sci. USA* 107, 9123–9128 (2010).
13. Gorbalenya, A. E. & Koonin, E. V. Endonuclease (R) subunits of type-I and type-III restriction-modification enzymes contain a helicase-like domain. *J. Mol. Biol.* 289, 277–291 (1999).
14. Singleton, M. R., Dillingham, M. S. & Wigley, D. B. Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* 76, 23–50 (2007).
15. Reiser, J. & Yuan, R. Purification and properties of the P1 specific restriction endonuclease from *Escherichia coli*. *J. Biol. Chem.* 252, 451–456 (1977).
16. Ramanathan, S. P. et al. Type III restriction enzymes communicate in 1D without looping between their target sites. *Proc. Natl Acad. Sci. USA* 106, 1748–1753 (2009).
17. Meisel, A., Mackeldanz, P., Bickle, T. A., Kruger, D. H. & Schroeder, C. Type III restriction endonucleases translocate DNA in a reaction driven by recognition of specific ATP. *Nucleic Acids Res.* 23, 11147–11166 (1995).
18. Crampton, N. et al. DNA looping and translocation provide an optimal cleavage mechanism for the type III restriction enzymes. *EMBO J.* 26, 3815–3825 (2007).
19. Schwarz, F. W. et al. The helicase-like domains of type III restriction enzymes trigger long-range diffusion along DNA. *Science* 340, 353–356 (2013).
20. Blainey, P. C. et al. Nonspecifically bound proteins spin while diffusing along DNA. *Nat. Struct. Mol. Biol.* 16, 1224–1229 (2009).
21. Szczelkun, M. D. Roles for helicases as ATP-dependent molecular switches. *Adv. Exp. Med. Biol.* 767, 225–244 (2013).
22. Qiu, R. et al. Large conformational changes in MutS during DNA scanning, mismatch recognition and repair signalling. *EMBO J.* 31, 2528–2540 (2012).
23. Gorman, J. et al. Single-molecule imaging reveals target-search mechanisms during DNA mismatch repair. *Proc. Natl Acad. Sci. USA* 109, E3074–E3083 (2012).
24. Cho, W. K. et al. ATP alters the diffusion mechanics of MutS on mismatched DNA. *Structure* 20, 1264–1274 (2012).
25. Ahmad, I., Krishnamurthy, V. & Rao, D. N. DNA recognition by the EcoP15I and EcoP11 modification methyltransferases. *Gene* 157, 143–147 (1995).
26. Durr, H., Korner, C., Muller, M., Hickmann, V. & Hopfner, K. P. X-ray scattering studies of the N6-adenine DNA methyltransferase M.TaqI in complex with DNA and a cofactor analog. *Nat. Struct. Biol.* 8, 121–125 (2001).
27. Zhong, X. et al. Molecular mechanism of action of plant DRM de novo DNA methyltransferases. *Cell* 157, 1050–1060 (2014).
28. Jia, D., Jurkowski, R. Z., Zhang, X., Jeltsch, A. & Cheng, X. Structure of Dnmt3a bound to Dnmt3l suggests a model for de novo DNA methylation. *Nature* 449, 248–251 (2007).
29. Meyer, K. D. & Jaffrey, S. R. The dynamic epitranscriptome: N6-methyladenosine and gene expression control. *Nat. Rev. Mol. Cell Biol.* 15, 313–326 (2014).
30. Stemmerborg, H. G., Vermulen, M. & Atlass, Y. Developmental biology. A MetAge for pluripotency. *Science* 347, 614–615 (2015).
31. Rao, D. N., Dryden, D. T. & Bheemanaik, S. Type III restriction modification enzymes: a historical perspective. *Nucleic Acids Res.* 42, 45–55 (2013).
32. Durr, H., Korner, C., Muller, M., Hickmann, V. & Hopfner, K. P. X-ray structures of the Sulfobolus solfatarius SW12/SNF2 ATPase core and its complex with DNA. *Cell* 121, 363–373 (2005).
33. Malone, T., Blumenthal, R. M. & Cheng, X. 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 (1995).
34. Schneider, B., Neidle, S. & Berman, H. M. Conformations of the sugar-phosphate backbone in helical DNA crystal structures. *Biopolymers* 42, 113–124 (1997).
35. Luksas, C. M., Kucera, R., Schildkraut, I. & Aggarwal, A. K. Understanding the immutability of restriction enzymes: crystal structure of BglII and its DNA substrate at 1.5 Å resolution [see comments]. *Nat. Struct. Biol.* 7, 134–140 (2000).
37. Sengoku, T., Nureki, O., Nakamura, A., Kobayashi, S. & Yokoyama, S. Structural basis for RNA unwinding by the DEAD-box protein Drosophila Vasa. Cell 125, 287–300 (2006).
38. Thoma, N. H. et al. Structure of the SWI2/SNF2 chromatin-remodeling domain of eukaryotic Rad54. Nat. Struct. Mol. Biol. 12, 350–356 (2005).
39. Kim, J. L. et al. Hepatitis C virus NS3 RNA helicase domain with a bound oligonucleotide: the crystal structure provides insights into the mode of unwinding. Structure 6, 89–100 (1998).
40. Morgan, R. D., Dwimell, E. A., Bhatta, T. K., Lang, E. M. & Luyten, Y. A. The Mmef family: type II restriction-modification enzymes that employ single-strand modification for host protection. Nucleic Acids Res. 37, 5208–5221 (2009).
41. Shen, B. W. et al. Characterization and crystal structure of the type IIG restriction endonuclease Rm.BpuSI. Nucleic Acids Res. 39, 8223–8236 (2011).
42. Scavetta, R. D. et al. Structure of RsrI methyltransferase, a member of the N6-adenine beta class of DNA methyltransferases. Nucleic Acids Res. 28, 3950–3961 (2000).
43. Osipiuk, J., Walsh, M. A. & Joachimski, A. Crystal structure of MboIIA methyltransferase. Nucleic Acids Res. 31, 5440–5448 (2003).
44. Thomas, C. B. & Gumport, R. I. Dimerization of the bacterial Rsr1 N6-adenine DNA methyltransferase. Nucleic Acids Res. 34, 806–815 (2006).
45. Meyer, K. D. et al. Comprehensive analysis of mRNA methylation reveals enrichment in 3′ UTRs and near stop codons. Cell 149, 1635–1646 (2012).
46. Dominissini, D. et al. Topology of the human and mouse m6A RNA methylomes revealed by m6A-seq. Nature 485, 201–206 (2012).
47. Batista, P. J. et al. m(6)A RNA modification controls cell fate transition in mammalian embryonic stem cells. Cell Stem Cell 15, 707–719 (2014).
48. Geula, S. et al. m6A mRNA methylation facilitates resolution of naive pluripotency toward differentiation. Science 347, 1002–1006 (2015).
49. Wang, Y. et al. N6-methyladenosine modification destabilizes developmental regulators in embryonic stem cells. Nat. Cell Biol. 16, 191–198 (2014).
50. Ping, X. L. et al. Mammalian WTAP is a regulatory subunit of the N6-methyladenosine methyltransferase. Cell Res. 24, 177–189 (2014).
51. Liu, J. et al. A METTL3-METTL14 complex mediates mammalian nuclear RNA N6-methyladenosine methylation. Nat. Chem. Biol. 10, 93–95 (2014).
52. Bujnicki, J. M., Feder, M., Radlinska, M. & Blumenthal, R. M. Structure prediction and phylogenetic analysis of a functionally diverse family of proteins of eukaryotic Rad54. J. Mol. Evol. 55, 431–442 (2002).
53. Michel, G. et al. The structure of the RlmB 23 S rRNA methyltransferase reveals a new methyltransferase fold with a unique knot. Structure 10, 1303–1315 (2002).
54. Thomas, S. R., Keller, C. A., Szyk, A., Cannon, J. R. & Laronde-Leblanc, N. A. Structural insights into the functional mechanism of Npl1/Emg1 N1-specific pseudouridine methyltransferase in ribosome biogenesis. Nucleic Acids Res. 39, 2445–2457 (2011).
55. Schubert, H. L., Blumenthal, R. M. & Cheng, X. Many paths to methyltransfer: a chronicle of convergence. Trends Biochem Sci. 28, 329–335 (2003).
56. Gorbalenya, A. E. & Koonin, E. V. Helicases: amino acid sequence comparisons and structure-function relationships. Curr. Opin. Struct. Biol. 3, 419–429 (1993).
57. Kelch, B. A., Makino, D. L., O’Donnell, M. & Kuriyan, J. How a DNA polymerase clamp loader opens a sliding clamp. Science 334, 1675–1680 (2011).
58. Skoglund, C. M., Smith, H. O. & Chandrasegaran, S. Construction of an efficient overproducer clone of Hinfl restriction endonuclease using the polymerase chain reaction. Gene 88, 1–5 (1990).
59. Fortelle, d. L. & Bricogne, G. Maximum-likelihood heavy atom parameter refinement for multiple isomorphous replacement and multil wavelength anomalous diffraction methods. Methods Enzymol. 276, 472–494 (1997).
60. Hendrickson, W.A. Determination of macromolecular structures from anomalous diffraction of synchrotron radiation. Science 254, 51–58 (1991).
61. Vournhein, C. et al. Data processing and analysis with the autoPROC toolbox. Acta Crystallogr. D Biol. Crystallogr. 67, 293–302 (2011).
62. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132 (2004).
63. Blanc, E. et al. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. Acta Crystallogr. D Biol. Crystallogr. 60, 2210–2221 (2004).
64. Kabsch, W. Xds. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132 (2010).
65. McCoy, A. J., Grosse-Kunstleve, R. W., Storoni, L. C. & Read, R. J. Likelihood-enhanced fast translation functions. Acta Crystallogr. D Biol. Crystallogr. 61, 458–464 (2005).

Acknowledgements
We thank David Hough for expert technical assistance with protein purification and Jim Samuelson for development of the overexpressing strain. We also thank Rajasankar (NECAT-24ID, APS) and Qun Liu (X4A, BNL) for facilitating data collection and useful suggestions, and M. Szczeklik and R. Seidel for helpful discussions. We thank Y. Vanamee for help during early stages of the project. We are thankful to the staff of NECAT 24ID, GM-CAT 23ID, SBC-CAT 19ID (APS, ANL), X25, X29 and X4A (NSLS, BNL), and A1 and F1 beams at MacCHESS for the provision of synchrotron beamtime. All structure figures were generated using PyMol. This work was supported by funding from NIH (R01 GM111507) and partially by New England Biolabs Inc.

Author contributions
Y.K.G. and A.K.A. designed the crystallographic studies; Y.K.G. performed the crystallographic studies; S.-H.C. and S.-y.X. performed the cloning and protein purification; A.K.A. and Y.K.G. wrote the paper.

Additional information
Accession codes: Atomic coordinates and structure factors have been deposited in the Protein Data Bank under accession codes 4ZCF.

Supplementary Information accompanies this paper at http://www.nature.com/naturecommunications

Competing financial interests: S.-H.C. and S.-y.X. are employees and shareholder of New England Biolabs Inc. The remaining authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/

How to cite this article: Gupta, Y. K. et al. Structural basis of asymmetric DNA methylation and ATP-triggered long-range diffusion by EcoP1571. Nat. Commun. 6:7363 doi: 10.1038/ncomms8363 (2015).