Article No. mb982250 *J. Mol. Biol.* (1998) **284**, 937–948

JMB

The DNA Recognition Subunit of the Type IB Restriction-Modification Enzyme *Eco*AI Tolerates Circular Permutions of its Polypeptide Chain

Pavel Janscak and Thomas A. Bickle*

Department of Microbiology Biozentrum, University of Basel, Klingelbergstrasse 70 CH-4056 Switzerland

The DNA specificity subunit (HsdS) of type I restriction-modification enzymes is composed of two independent target recognition domains and several regions whose amino acid sequence is conserved within an enzyme family. The conserved regions participate in intersubunit interactions with two modification subunits (HsdM) and two restriction subunits (HsdR) to form the complete endonuclease. It has been proposed that the domains of the HsdS subunit have a circular organisation providing the required symmetry for their interaction with the other subunits and with the bipartite DNA target. To test this model, we circularly permuted the HsdS subunit of the type IB R-M enzyme *Eco*AI at the DNA level by direct linkage of codons for original termini and introduction of new termini elsewhere along the N-terminal and central conserved regions. By analysing the activity of mutant enzymes, two circularly permuted variants of HsdS that had termini located at equivalent positions in the N-terminal and central repeats, respectively, were found to fold into a functional DNA recognition subunit with wild-type specificity, suggesting a close proximity of the N and C termini in the native protein. The wild-type HsdS subunit was purified to homogeneity and shown to form a stable trimeric complex with HsdM, M_2S_1, which was fully active as a DNA methyltransferase. Gel electrophoretic mobility shift assays revealed that the HsdS protein alone was not able to form a specific complex with a 30-mer oligoduplex containing a single *Eco*AI recognition site. However, addition of stoichiometric amounts of HsdM to HsdS led to efficient specific DNA binding. Our data provide evidence for the circular organisation of domains of the HsdS subunit. In addition, they suggest a possible role of HsdM subunits in the formation of this structure.

Keywords: restriction-modification; DNA binding; circularly permuted protein; HsdS polypeptide; subunit complex

**Corresponding author*

Introduction

Restriction-modification (R-M) systems protect a bacterial cell against invasion of foreign DNA by endonucleolytic cleavage of DNA that lacks a site-specific modification. The host genome is protected from cleavage by methylation of specific nucleotides in the target sites. In type I R-M systems, both restriction and modification activities are present in one heteromeric enzyme complex composed of one DNA specificity subunit HsdS, two modification subunits HsdM and two restriction subunits HsdR (for reviews, see Bickle & Krüger, 1993; Redaschi & Bickle, 1996). The enzyme complex recognises an asymmetric DNA sequence that is composed of two specific 3-4 bp components split by a non-specific spacer of 6-8 bp (e.g. GAGNNNNNNNGTCA for *Eco*AI). If the recognition site is not modified, the enzyme subsequently translocates DNA past the DNA-enzyme

Abbreviations used: R-M, restriction-modification; TRD, target recognition domain; cp, circularly permuted; MTase, methyltransferase; ENase, endonuclease; AdoMet, *S*-adenosylmethionine; EtBr, ethidium bromide; Caps, 3-(cyclohexylamino)-1-propanesulphonic acid; GEMSA, gel electrophoretic mobility shift assay; e.o.p., efficiency of plating.

E-mail address of the corresponding author: bickle@ubaclu.unibas.ch

0022–2836/98/490937–12 $30.00/0

complex using free energy derived from Mg^{2+}-dependent ATP hydrolysis (Yuan *et al*., 1980; Endlich & Linn, 1985; Szczelkun *et al*., 1996). DNA is cleaved at a random site, distant from the recognition sequence, by a mechanism involving an *in cis*-collision of two translocating enzyme complexes (Studier & Bandyopadhyay, 1988; Dreier *et al*., 1996). The type I enzymes cleave DNA only once, but ATP hydrolysis continues long after DNA cleavage (Eskin & Linn, 1972; Yuan *et al*., 1972). If the target site is hemimethylated, or when ATP is absent in the case of non-modified DNA, the enzyme complex exhibits N6-adenine DNA methyltransferase activity and methylates a specific adenine residue in each strand of its recognition sequence using *S*-adenosylmethionine (AdoMet) as a methyl donor (Burckhardt *et al*., 1981; Powell *et al*., 1993; Taylor *et al*., 1993). The HsdR subunit is not essential for methylation and a monofunctional DNA methyltransferase with a subunit composition of M_2S_1 can be formed (Dryden *et al*., 1993; Taylor *et al*., 1992).

The type I R-M enzymes are divided into four families (IA, IB, IC and ID) on the basis of the genetic complementation, amino acid conservation and biochemical properties of the enzymes. While HsdM and HsdR subunits are highly conserved within a family, an intra-family alignment of HsdS polypeptides reveals two large regions that are very variable (Gough & Murray, 1983; Kannan *et al*., 1989; Gubler *et al*., 1992). Swapping these variable regions between members of the same family generates a new DNA specificity, so that each half of the new recognition sequence is identical with one parental half-site (Fuller-Pace *et al*., 1984; Nagaraja *et al*., 1985; Gubler *et al*., 1992; Thorpe *et al*., 1997). Thus, each variable sequence constitutes a domain that is responsible for recognition of one specific half of the target site (TRD, target recognition domain). Amino acid sequences of TRDs from different families sometimes show a significant level of similarity if these TRDs recognise the same sequence (Cowan *et al*., 1989). The TRDs are separated by a region that is conserved amongst family members, and the length of this region determines the length of the spacer in the recognition sequence (Gubler & Bickle, 1991). Two additional conserved regions within the HsdS polypeptide are located at the N and C termini. The length of these regions varies among individual families. Since the HsdS subunits within a family are interchangeable, the conserved regions are believed to be involved in protein-protein interactions with the other subunits.

It has been proposed that the N and C-terminal conserved regions of HsdS are in close proximity and together form a ''split'' domain linker between the two inversely oriented TRDs, which makes contact with one HsdM subunit (Kneale, 1994). The second HsdM subunit would interact with the central conserved region of HsdS to form a structure with a 2-fold rotational symmetry (Kneale, 1994; Figure 1). The ''circular'' model for the domain

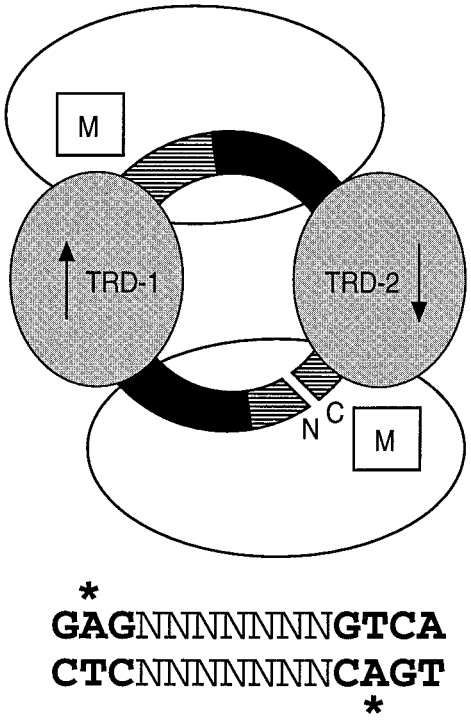

Figure 1. A model for domain organisation of the *Eco*AI DNA methyltransferase. The HsdS subunit adopts a circular structure that is mantained by interaction of one HsdM subunit (open oval) with the terminal, conserved regions of HsdS. The second HsdM subunit makes a symmetrical contact with the central conserved region of HsdS. Internally duplicated portions of the conserved regions of HsdS polypeptide are shown in black and the non-repeated part of conserved regions is hatched. The two target recognition domains of HsdS (TRD-1 and TRD-2) are represented by grey ovals. The DNA recognition sequence of *Eco*AI is shown. The specific nucleotides of the recognition sequence are in bold and the two targets for N6-adenine methylation are marked by asterisks (*). The GAG-half site is recognised by TRD-1 and the GTCA-half site is recognised by TRD-2. Upon interaction of the TRDs with DNA, the methylation moieties of HsdM subunits (open squares) are located in the vicinity of the adenine targets.

organisation of HsdS subunit is supported by a number of findings: (i) intramolecular alignment of an HsdS polypeptide reveals significant homologies between the central and terminal conserved regions. In the HsdS subunits of type IC R-M enzymes, one repeated sequence is split between the polypeptide termini, suggesting their close proximity (Tyndall *et al*., 1994; Kneale, 1994; MacWilliams & Bickle, 1996). (ii) Truncated forms of HsdS polypeptides of the type IC R-M systems, *Eco*R124I and *Eco*DXXI, which contain a single TRD flanked by the conserved regions can still function as DNA recognition subunits. The mutant R-M enzyme recognises a novel sequence that is an interrupted palindrome of the sequence recognised by the TRD present (Meister *et al*., 1993; Abadjieva *et al*., 1993; MacWilliams & Bickle, 1996). This

suggests that two inversely oriented copies of the deleted subunit are present in the enzyme complex. (iii) The amino-TRD of a member of the IB family, *Sty*SKI, and the carboxy-TRD of IC family member, *Eco*R124I, show 36% identity and recognise complementary sequences, which suggests that these TRDs are inversely oriented on their targets (Thorpe *et al.*, 1997).

We decided to test the model proposed for domain organisation of HsdS by circularly permuting the HsdS subunit of the type IB R-M enzyme *Eco*AI at the DNA level. A circularly permuted protein variant is produced by covalent linking of the original N and C termini, and introducing new termini in an alternative region. Circularly permuted variants of many proteins have been created to study the relationship between primary protein sequence and folding (Goldenberg & Creighton, 1983; Luger *et al.*, 1989; Yang & Schachman, 1993; Koebnik & Kramer, 1995; Johnson & Raushel, 1996). A requirement for the formation of a functional circularly permuted protein is the close proximity of the N and C termini in the folded protein. New termini can be readily introduced into solvent-exposed loops; however, permuted variants with termini within helices of the wild-type protein have been generated (Graf & Schachman, 1996). Here, we report that the HsdS subunit of the *Eco*AI restriction enzyme indeed tolerates circular permutations within both N-terminal and central conserved regions. In addition, we show that the purified HsdS subunit was incapable of specific binding to a synthetic oligoduplex containing a single *Eco*AI site in the absence of HsdM. This suggests a role of the HsdM subunits in DNA recognition by HsdS, perhaps by maintaining a circular organisation of HsdS domains.

Results

Design of circularly permuted variants of the HsdS polypeptide

Circularly permuted variants of the *Eco*AI *hsd*S were constructed by PCR amplification of two specific regions of the native gene and subsequent recombination of these gene fragments in reverse order, as described in Materials and Methods. For most proteins that have been subjected to permutation analysis, the crystal structure was already known, enabling linkers to be designed for connecting the original termini and to place new termini into the solvent-exposed loops, turns or non-ordered regions. Since no such structural information is available for any HsdS polypeptide, we decided to link the *Eco*AI HsdS ends directly to each other and to introduce new termini either into the N-terminal or the central conserved regions of the original polypeptide. In addition to a central conserved region (130 amino acid residues), the *Eco*AI HsdS subunit has a long conserved region at the N terminus (107 residues) and a very short conserved region at the C terminus (16 residues). The amino acid sequence of more than one half of the N-terminal conserved region (64 residues) is duplicated in the central conserved region (Figure 2).

Four circularly permuted variants of HsdS were produced (Figure 2). One mutant (cp91) had termini in the N-terminal repeat at Glu91, since the sequence in the vicinity of this residue was reminiscent of a turn (PLPEIS). This manipulation split this repeat between the N and C ends, producing an organisation similar to that of type IC HsdS subunits. A deletion derivative lacking 90 N-terminal amino acid residues (Δ90) was also constructed as a control. The termini of the second permuted protein (cp108) resided on the border

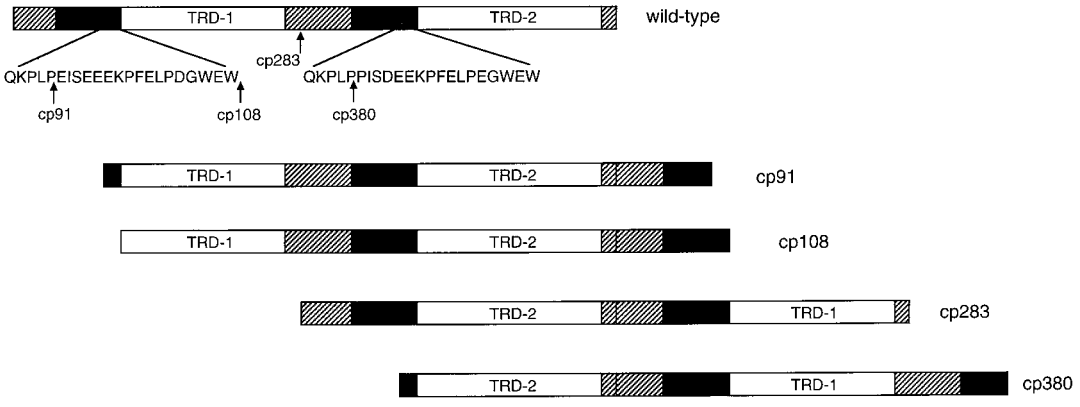

Figure 2. Circularly permuted variants of the *Eco*AI HsdS polypeptide constructed in this study. The wild-type HsdS polypeptide of *Eco*AI R-M system (569 amino acid residues) contains two target recognition domains (TRD-1 and TRD-2) and three regions conserved amongst members of the type IB family. A stretch of 64 residues in the N-terminal conserved region is repeated in the central conserved region (shown in black). Non-repeated parts of the conserved regions are hatched. A part of the amino acid sequence of the repeats is shown by one-letter symbols. Circularly permuted variants of the HsdS polypeptide were constructed genetically by direct linkage of the original N and C termini and introduction of new termini at positions indicated by arrowheads (see Materials and Methods). Numbers in names of the HsdS variants refer to the positions of their N-terminal amino acid residues in the wild-type sequence.

between the N-terminal conserved region and the TRD-1 at Thr108. The third permuted variant (cp283) had termini in the non-repeated part of the central conserved region at Asp283. This organisation is similar to that of the wild-type in terms of the lengths of the conserved regions at the N and C termini. The termini of the fourth permuted variant (cp380) were located in the central repeat at Pro380. The position of this residue in the central repeat sequences is equivalent to that of Glu91, which was chosen as the N-terminal residue for cp91 (Figure 2). If HsdS has a circular organisation, a reverse order of TRDs produced by circular permutations in the central conserved region should not change the order of specific halves in the recognition sequence. The circularly permuted *hsd*S genes as well as the wild-type gene were cloned under the control of the phage T7 promoter in various pET-expression vectors (Materials and Methods).

In vivo functional analysis of circularly permuted variants of HsdS

To analyse the activity of the individual circularly permuted variants of the *Eco*AI HsdS polypeptide, the corresponding permuted *hsd*S genes were expressed from an appropriate pET-construct in *Escherichia coli* BL21(DE3) along with the *hsd*M and *hsd*R genes provided *in trans* on pJP28. The cell cultures were tested for the presence of an active restriction enzyme by analysing their ability to restrict growth of unmodified bacteriophage λ. Various dilutions of λ*vir* phage lysate were plated on BL21(DE3) expressing a particular set of *hsd* genes as well as the same cells lacking the R-M system. The resulting efficiency of plating of the phage (e.o.p. is the ratio of the phage titre on tested host to the titre on non-restricting host) was used as a measure of the restriction activity of the cells. Cultures of uninduced cells could be used for the phage infection assay, because there was a sufficient level of constitutive expression of the wild-type and mutant HsdS proteins in the pET system. The results from the restriction assay (Figure 3(a)) showed that of all mutant R-M systems, only that expressing $hsdS_{cp91}$ exhibited restriction activity (e.o.p. 2.1×10^{-2}); the level of restriction was similar to that produced by the wild-type system (e.o.p. 2.7×10^{-2}). The relative inefficiency of *Eco*AI restriction is presumably a consequence of the presence of only one *Eco*AI recognition site in λ DNA.

The mutant cells were tested for the presence of an active methylase. First, λ*vir* phage was propagated in host cells expressing the R-M system to be tested. The lysate of this phage was plated on non-restricting DH5α and on DH5α expressing the wild-type *Eco*AI R-M system (DH5α/pFFP30). The resulting e.o.p. served as a measure of *Eco*AI-specific modification activity. The modification assay (Figure 3(b)) revealed that, in addition to the cp91 system, the cp380 system protected λ*vir* phage against restriction by *Eco*AI, showing that

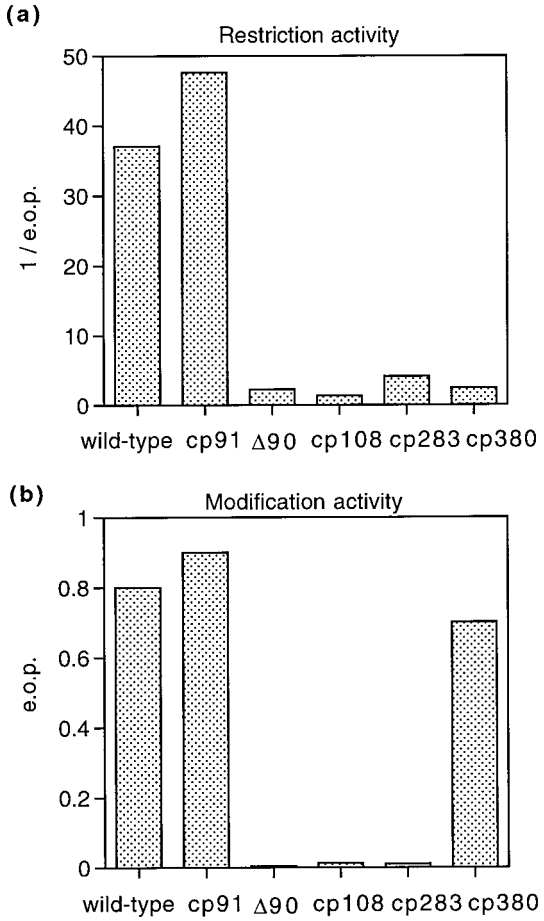

Figure 3. Restriction and modification phenotypes of mutant R-M systems expressing circularly permuted variants of the *Eco*AI HsdS polypeptide. (a) Restriction assay. A reciprocal plot of efficiency of plating of λ*vir* on *E. coli* BL21(DE3) expressing wild-type and mutant R-M systems. (b) *Eco*AI-specific methylation. A plot of efficiency of plating of λ*vir* propagated in *E. coli* BL21(DE3) expressing wild-type and mutant R-M systems on *E. coli* DH5α expressing wild-type *Eco*AI R-M system from the plasmid pFFP30. The phage infection assay was performed as described in Materials and Methods. The e.o.p. values are averages of at least three independent measurements. Wild-type *hsd*S, cp91, Δ90, cp108, cp283 and cp380 genes were expressed from plasmids pJP26, pJP30, pJP31, pJP32, pJP33 and pJP34, respectively. The *hsd*R and *hsd*M genes were provided *in trans* on the plasmid pJP28. Uninduced cells were used for the assays.

both mutants exhibit *Eco*AI site-specific modification activity. λ*vir* phage modified by passage through cells expressing the wild-type *Eco*AI system was not restricted by cp91 system, confirming that this system has only *Eco*AI DNA specificity. The cp380 system could not be tested by this assay, due to the lack of restriction activity. The remaining mutants, $hsdS_{cp108}$, $hsdS_{cp283}$ and $hsdS_{\Delta 90}$, showed no activity when expressed in the two-plasmid system.

These results demonstrate that the $hsdS_{cp91}$ and $hsdS_{cp380}$ genes code for functional DNA recog-

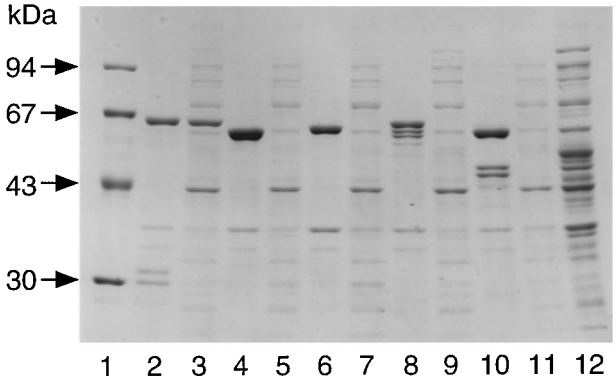

Figure 4. SDS/polyacrylamide gel showing high-level expression of circularly permuted variants of the HsdS subunit of *Eco*AI using the pET system. The cells of *E. coli* BL21(DE3) transformed with a particular plasmid were grown to the mid-log phase ($A_{600} = 0.4$) and then induced with 0.4 mM IPTG for three hours at 37°C. Harvested cells were broken by sonication in buffer A (Materials and Methods), and soluble and insoluble portions of the cell extract were separated by centrifugation. Lane 1, molecular mass markers; lanes 2, 4, 6, 8 and 10, pellet fractions; lanes 3, 5, 7, 9 and 11, soluble fractions from BL21(DE3) harbouring pJP26 (wild-type), pJP30 (cp91), pJP32 (cp108), pJP33 (cp283) and pJP34 (cp380); lane 12, whole-cell extract from BL21(DE3).

nition subunits exhibiting wild-type DNA specificity. Thus, the DNA recognition subunit of the *Eco*AI R-M system tolerates circular permutations within both the N-terminal and central conserved regions of its polypeptide chain.

Expression of circularly permuted proteins

To analyse the expression of circularly permuted variants of *hsd*S gene at the protein level, BL21(DE3) cells harbouring an appropriate plasmid were induced at mid-log phase of growth with IPTG, and both soluble and insoluble portions of crude extract from these cells were subjected to SDS/polyacrylamide gel electrophoresis (Figure 4). All permuted proteins were found to be insoluble when produced in the absence of HsdM and HsdR subunits, while the wild-type HsdS protein was soluble. In addition, $HsdS_{cp91}$, $HsdS_{cp108}$ and $HsdS_{cp380}$ proteins were smaller than the wild-type HsdS, suggesting that they had undergone a limited proteolysis in the cell. $HsdS_{cp91}$ and $HsdS_{cp380}$ proteins appeared to be the same size, while the $HsdS_{cp108}$ protein was slightly larger. In the insoluble part of the crude extract from the cells that produced $HsdS_{cp380}$, two additional smaller protein species were identified by SDS-PAGE, suggesting further processing of this mutant. The $HsdS_{cp283}$ protein had the same size as wild-type, although two additional smaller proteins were revealed by the SDS-PAGE that may have resulted from partial proteolysis of cp283. Estimation of apparent molecular masses of individual proteins from their mobility in SDS/polyacrylamide gel revealed that the size difference between the wild-type HsdS and either $HsdS_{cp91}$ or $HsdS_{cp380}$ would correspond to the loss of a 3.8 kDa peptide, while the size of $HsdS_{cp108}$ corresponds to the loss of a 2.4 kDa peptide. N-terminal sequencing of the $HsdS_{cp91}$, $HsdS_{cp108}$ and $HsdS_{cp380}$ proteins revealed that they all have the predicted N-terminal sequence, suggesting that processing had occurred at the C terminus. To investigate whether coproduction of circularly permuted HsdS proteins along with HsdM and HsdR subunits would prevent the observed proteolysis, the crude extracts from cell cultures used for testing restriction and modification phenotypes (Figure 3) were subjected to SDS-PAGE analysis followed by immunoblotting. This experiment showed that all permuted proteins underwent the same processing as they did when produced independently (not shown).

Purification and properties of DNA methyltransferases containing the functional circularly permuted variants of HsdS

To produce DNA methyltransferases (MTase) containing the functional circularly permuted variants of the HsdS subunit, $HsdS_{cp91}$ and $HsdS_{cp380}$, the corresponding genes were placed under control of the the *tac* promoter in plasmids pCYP3 and pCYP2, respectively. These constructs were coexpressed with the *hsd*M gene provided on the plasmid pJP29 (T7-*hsd*M) in BL21(DE3). With this plasmid combination, both HsdM and HsdS proteins were located in the soluble fraction with a large excess of HsdM over HsdS (not shown). The wild-type methylase was produced using the same two-plasmid system as that used for the mutants.

The wild-type and mutant methylases were purified to homogeneity as described in Materials and Methods. SDS-PAGE analysis of the final preparations again revealed that the $HsdS_{cp91}$ and $HsdS_{cp380}$ subunits were smaller than the wild-type subunit (Figure 5). The difference in size between the wild-type and mutant subunits appeared to be similar to that obtained when the mutant proteins were produced independently (Figure 4). N-terminal sequencing of the permuted HsdS subunits again confirmed that they start with the predicted sequence. Densitometric scanning of the SDS/polyacrylamide gel (stained with Coomassie brilliant blue) showed that the mutant methylases have the same subunit stoichiometry as the wild-type, which appeared to be M_2S_1 as judged from the ratio of densitometer traces of individual subunits (not shown).

To test whether the purified mutant methylases were functional *in vitro*, they were assayed together with the wild-type enzyme for their ability to transfer radiolabelled methyl groups from AdoMet to a synthetic 30-mer DNA oligoduplex containing a single *Eco*AI site. This experiment indicated that both $MTase_{cp91}$ and $MTase_{cp380}$ were active *in vitro*; however, their methylation activities

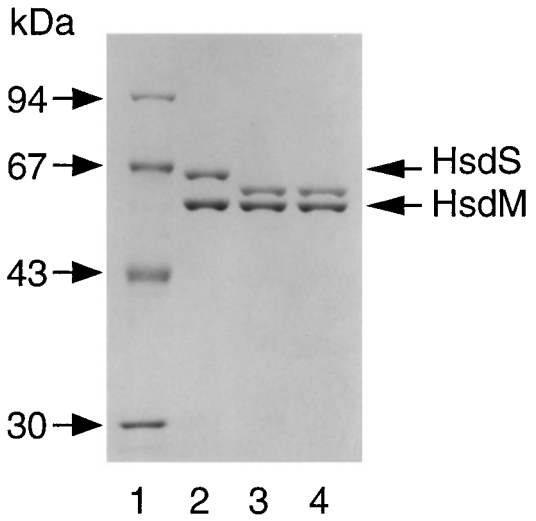

Figure 5. SDS/polyacrylamide gel showing purified wild-type and mutant DNA methyltransferases. Lane 1, molecular mass markers; lane 2, wild-type MTase; lane 3, $MTase_{cp91}$; lane 4, $MTase_{cp380}$. Positions of the HsdM and wild-type HsdS subunits are indicated.

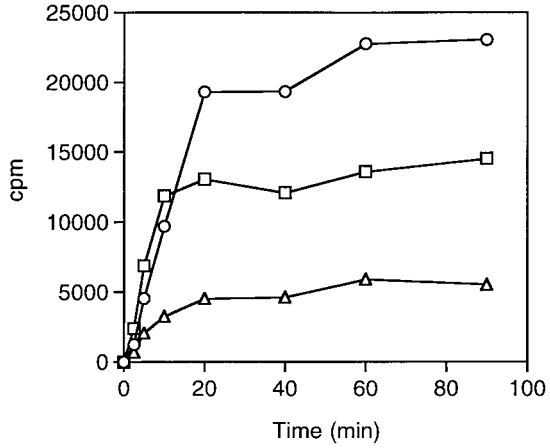

Figure 6. *In vitro* DNA methylation activity of mutant *Eco*AI methylases containing circularly permuted variants of the HsdS subunit. Purified wild-type (○), cp91 (△) and cp380 (□) methylases (50 nM each) were incubated with 1 μM 30-mer DNA oligoduplex containing a single *Eco*AI site in the presence of 1 μM [^{3}H]AdoMet as described in Materials and Methods. Aliquots (20 μl) removed at various time-points were subjected to DE81 filter-binding assay to determine the amount of tritiated methyl groups (cpm) incorporated into DNA.

were reduced in comparison to that of the wild-type enzyme (Figure 6).

To test whether the mutant methylases were capable of interacting with purified HsdR subunit *in vitro* to produce a functional restriction endonuclease (ENase), a mixture of HsdR and each methylase was incubated with the plasmid pJP25 containing a single *Eco*AI recognition site in the presence of ATP, as described in Materials and Methods. This experiment showed that addition of the HsdR subunit to both mutant methylases in a molar ratio of 8 to 1 (HsdR to MTase) yielded fully functional restriction enzymes that, when present in twofold excess over DNA, cleaved the supercoiled form of the plasmid to its linear form in a period of five minutes (Figure 7). Neither mutant endonucleases nor the wild-type enzyme showed restriction activity towards pDRM.1R (Janscak *et al*., 1996), a plasmid derived from the same vector as pJP25 but lacking an *Eco*AI site (Figure 7). The *in vitro* restriction activity of the cp380 mutant contrasts with its restriction minus phenotype (Figure 3). To investigate the possibility that the lack of *in vivo* activity of $ENase_{cp380}$ was due to weaker binding of HsdR to MTase, we compared plasmid DNA cleavage for wild-type and both mutant enzymes at various MTase to HsdR ratios. A 1:1 mixture of appropriate methylase and pJP25 DNA (each at a concentration of 40 nM) was incubated with increasing concentrations of HsdR, and the relative amount of linear DNA product for each HsdR to MTase ratio was measured as described in Materials and Methods. No difference between the wild-type and mutant enzymes was observed in this experiment (Figure 8). Thus, the reason for the defective restriction phenotype of cp380 is not clear. As seen in Figure 8, saturation of HsdR-binding sites on the MTase, even for the wild-type enzyme, was reached at higher HsdR to MTase ratios than in the $R_2M_2S_1$ cleavage-competent complex. This agrees with previously published data showing that $R_2M_2S_1$ is a weak complex (Suri *et al*., 1984).

Purification and *in vitro* properties of the wild-type HsdS subunit

The finding that the HsdS subunit of *Eco*AI, unlike other HsdS proteins so far investigated, was soluble when expressed in the absence of HsdM subunit (Figure 4, lane 3), encouraged us to purify this protein and investigate its DNA-binding properties *in vitro*. The HsdM protein was also overproduced and purified. Analytical ultracentrifugation revealed that both HsdS and HsdM proteins are present as monomers in solution and, when mixed together, they form a stable complex with a molecular mass of approximately 180 kDa. This value corresponds to the expected subunit stoichiometry of M_2S_1. The *in vitro* assembled M_2S_1 complex was found to be as active as the *in vivo* assembled enzyme (not shown).

DNA binding of HsdS was examined by gel electrophoretic mobility shift assays using a 30-mer oligoduplex containing a single *Eco*AI recognition site. We observed that HsdS protein alone did not interact with 85 nM DNA probe when present in concentrations below 4 μM (Figure 9(a)). At higher protein concentrations, binding occurred but the shifted DNA did not enter the gel, suggesting aggregation rather than specific binding. In contrast, when *in vitro* assembled M_2S_1 complex was

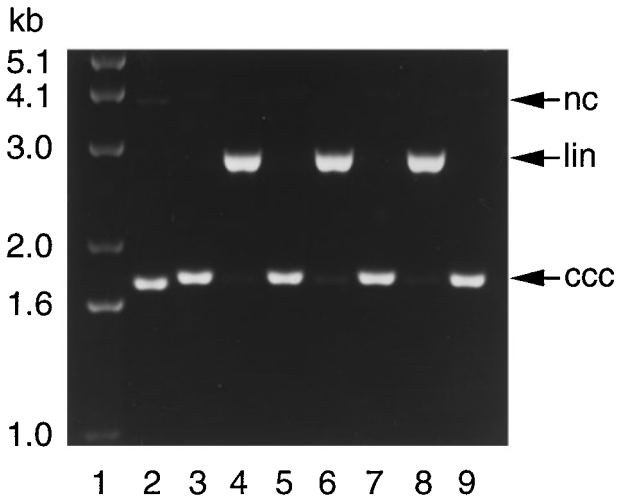

Figure 7. *In vitro* restriction activity of reconstituted endonucleases containing circularly permuted variants of the *Eco*AI HsdS subunits on plasmid DNA substrate. Wild-type and mutant ENases were prepared by addition of purified HsdR subunit (300 nM) to the appropriate methylase (40 nM) directly in the cleavage reaction. Plasmid DNA was present at a concentration of 20 nM and the concentration of AdoMet was 0.1 mM. Reactions were started by addition of ATP to a final concentration of 5 mM and incubated for five minutes at 37°C. Following heat-inactivation of enzymes, reaction mixtures were electrophoresed in a 0.9% agarose gel. DNA was visualised by EtBr staining. Lane 1, molecular mass markers. Lanes 2, 4, 6 and 8 show plasmid pJP25 (one *Eco*AI site) incubated with no enzyme, wild-type ENase, $ENase_{cp91}$ and $ENase_{cp380}$, respectively. Lanes 3, 5, 7 and 9 show plasmid pDRM.1R (no *Eco*AI sites) incubated with no enzyme, wild-type ENase, $ENase_{cp91}$ and $ENase_{cp380}$, respectively. Positions of covalently closed circular DNA (ccc), nicked circular DNA (nc) and linear DNA product (lin) are indicated.

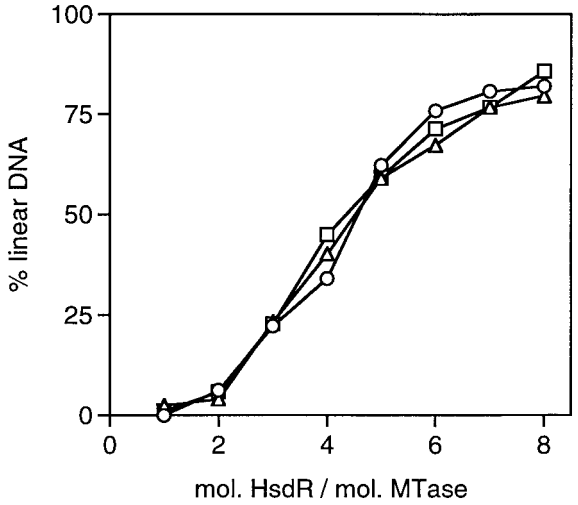

Figure 8. Titration of wild-type and mutant methylases of *Eco*AI with HsdR as monitored by DNA cleavage assay. Purified wild-type (○), cp91 (△) or cp380 (□) methylases (40 nM each) were incubated with increasing concentrations of HsdR subunit for two minutes at room temperature in the presence of 40 nM pJP25 DNA (a single *Eco*AI site). DNA cleavage reactions were initiated by addition of ATP and AdoMet to final concentrations of 5 mM and 0.1 mM, respectively. Following incubation at 37°C for ten minutes, the individual reactions were electrophoresed in a 0.9% agarose gel. The percentage of DNA cleaved to linear DNA was determined for each HsdR to MTase ratio by gel scanning as described in Materials and Methods.

assayed, a discrete DNA-protein complex was detected at much lower protein concentrations (Figure 9(a)). This complex was specific, because it was still formed in the presence of a 1000-fold excess (w/w) of poly(dI:dC) (not shown). HsdM protein alone was not able to shift the oligoduplex, even at very high protein concentrations (Figure 9(a)). Addition of HsdR protein to HsdS protein did not promote DNA binding (not shown). Titration of a mixture of 200 nM HsdS and 85 nM DNA with HsdM revealed two shifted species (Figure 9(b)). The slower-migrating complex (complex I) was predominantly present at all HsdM to HsdS ratios tested and had the same mobility as the DNA-protein complex containing the M_2S_1 species. The faster-migrating complex (complex II) appeared only at HsdM to HsdS ratios of less than 2, suggesting that it contains the assembly intermediate M_1S_1. The low intensity of the band corresponding to this complex suggests that DNA binding of M_1S_1 is weak, or that this species by itself is not stable. These data clearly demonstrate that HsdM is required for specific DNA recognition by HsdS.

Discussion

The HsdS subunit of type I restriction-modification enzymes has two main functions. It interacts with the other subunits to form the enzyme complex (M_2S_1 or $R_2M_2S_1$) and it is responsible for specific binding of this complex to DNA. The dual role of HsdS is reflected in its primary structure. The HsdS polypeptide contains two extensive variable regions that constitute two independent target recognition domains. Each TRD recognises one specific half of a non-palindromic DNA sequence. The TRDs are flanked by regions, conserved within an enzyme family, that are thought to mediate interactions with the other subunits. Although the action of the enzyme on DNA requires a symmetrical subunit complex, the location of the subunit-interaction regions in the HsdS polypeptide with respect to TRDs is asymmetric. It was proposed that the required symmetry may be provided by a circular organisation of HsdS domains (Kneale, 1994). According to this model, the methylase (M_2S_1) is a 2-fold rotationally symmetrical complex in which one HsdM subunit interacts with conserved regions at the N and C termini, and the

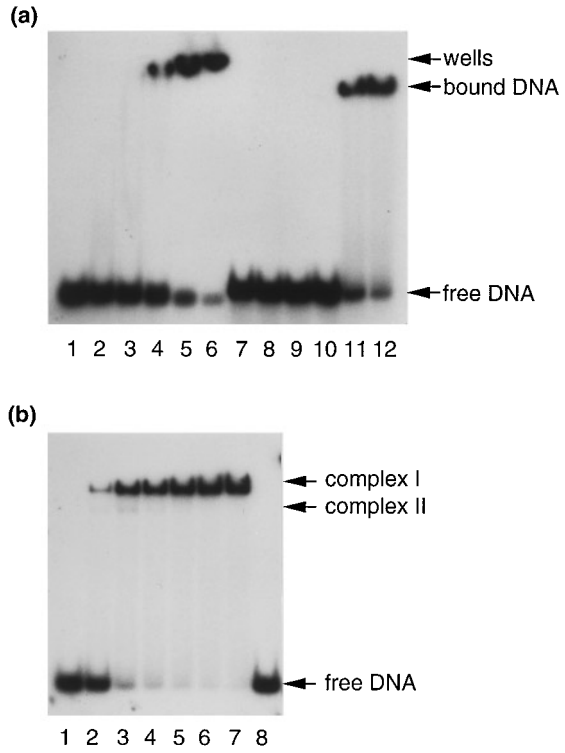

Figure 9. DNA binding properties of purified *Eco*AI HsdS subunit. (a) Gel electrophoretic mobility shift assay of purified HsdS subunit. The 5′-end labelled 30-mer oligoduplex containing a single *Eco*AI site was present at a concentration of 85 nM. Lane 1, free DNA; lanes 2-6, DNA incubated with HsdS protein at concentrations of 2, 4, 8, 16 and 32 μM, respectively; lanes 7-10, DNA incubated with HsdM protein at concentrations of 1.5, 3, 6 and 12 μM, respectively; lanes 11 and 12, DNA incubated with *in vitro* assembled M_2S_1 complex (prepared as described in Materials and Methods) at concentrations of 0.06 and 0.12 μM, respectively. (b) Titration of 200 nM HsdS with HsdM monitored by GEMSA. The 30-mer oligoduplex was present at a concentration of 85 nM throughout. Lanes 1-6 correspond to the following HsdM to HsdS ratios, 0, 0.5, 1, 1.5, 2 and 4; lane 7, DNA incubated with 200 nM *in vitro* assembled M_2S_1; lane 8, free DNA.

other HsdM subunit makes contact with the central conserved region (Figure 1). Upon binding of TRDs to the recognition sequence, the proposed organisation of the enzyme complex would enable the methylation moieties of the HsdM subunits to be positioned near the adenine targets (Willcock *et al*., 1994; Dryden *et al*., 1995). Our finding that the HsdS subunit of *Eco*AI R-M enzyme tolerates circular permutations of its polypeptide chain in both N-terminal and central conserved regions without change in DNA specificity indicates that the N and C termini in the native protein lie in close proximity and thus provides strong evidence for the ''circular'' model.

Gel electrophoretic mobility shift assays with purified HsdS revealed that the protein alone was not able to form a specific complex with a synthetic oligoduplex containing a single *Eco*AI site. However, addition of stoichiometric amounts of HsdM subunit to HsdS, which resulted in formation of a tight M_2S_1 complex, led to efficient, specific DNA binding. Similar results were obtained with the *Eco*R124I HsdS-GST fusion protein (Mernagh *et al*., 1997). These results suggest that the proposed ''circularisation'' of the HsdS polypeptide may be mediated by interaction with HsdM. This conformational change in the HsdS protein would then correctly position the TRDs for binding the two specific halves of the DNA target. From our data, we cannot rule out the possibility that HsdM participates directly in DNA binding.

In the permuted polypeptide chains of *Eco*AI HsdS, the original termini were linked directly without the aid of a peptide linker. No linker was needed to generate a circularly permuted variant of a β-glucanase that has N and C termini located on adjacent β-strands (Hahn *et al*., 1994). In contrast, a linkage of α-helical N and C-terminal structural units of a β-lactamase required a longer peptide insertion (Pieper *et al*., 1997). It is also possible that the region(s) at one or both termini of the HsdS protein are unstructured and this would provide flexibility for direct linkage of the polypeptide ends.

Our results with the cp380 mutant demonstrate that the central conserved region that links the two TRDs of *Eco*AI can be disconnected by a circular permutation without loss of activity or change in DNA specificity. This is in agreement with previously published data showing that coexpression of N-terminal and C-terminal truncated variants of the *Eco*DXXI HsdS polypeptide, with termini in about the same positions within the central conserved repeat, along with HsdM and HsdR subunits regenerates the wild-type specificity (MacWilliams & Bickle, 1996). On the other hand, separation of the two TRDs in *Eco*R124 methylase by limited proteolysis at the beginning of TRD-2 did not impair stability of the complex but it dramatically reduced DNA binding specificity (Webb *et al*., 1995). Thus, it appears that an HsdM subunit acts as a ''crosslink'' between the TRDs by interacting with flanking conserved regions. Interruption of the HsdS polypeptide at positions outside of the conserved regions would prevent this connection and result in an ''open'' complex with impaired DNA-binding properties. This would explain why circular permutation at the boundary between the N-terminal conserved region and TRD-1 of *Eco*AI HsdS (cp108) completely abolished activity. It is possible that this permuted subunit is still capable of forming the complex with two HsdM subunits but this complex may have similar properties to those observed with the proteolysed *Eco*R124I methylase. Introduction of new termini into the non-repeated part of the central conserved region of *Eco*AI HsdS at Asp283 (cp283) also abolished activity. This loss of activity may indicate that the

original polypeptide chain was broken within an important secondary structure element.

Interestingly, all circularly permuted variants of HsdS polypeptide were proteolytically processed even when coproduced with HsdM. In both functional permuted subunits, $HsdS_{cp91}$ and $HsdS_{cp380}$, a 3.8 kDa peptide was removed from the C terminus without loss of activity. This reduction in size would correspond to a deletion of approximately 35 amino acid residues. In agreement with these results, deletion analysis of the *hsd*S_{cp91} gene demonstrated that the codons for the last 20 amino acid residues could be removed without loss of function (K. Endele, unpublished results). In the wild-type HsdS polypeptide, the regions proteolytically removed from $HsdS_{cp91}$ and $HsdS_{cp380}$ are located within the internal repeats (Figure 2) and may form connecting structural elements that have no additional function. When one of these regions was moved to the C terminus by circularly permuting the polypeptide chain, it became sensitive to proteases. Proteolysis at the C terminus was observed also with truncated variants of type IC HsdS subunits (K. Firman, unpublished results; M.P. MacWilliams, unpublished results).

In contrast to the wild-type HsdS, all permuted variants were insoluble when expressed independently of HsdM. This may be due the observed C-terminal processing of these proteins. Alternatively, since $HsdS_{cp283}$ was proteolysed only partially and the remaining full-length protein was still insoluble, one can speculate that the intramolecular rearrangements of the HsdS polypeptide by a circular permutation of primary sequence affects protein folding and thus leads to protein precipitation. Interaction with the HsdM subunit overcame this problem. The HsdS subunits of the type IA and type IC R-M enzymes characterised to date are insoluble when expressed in the absence of the HsdM subunit (Dryden *et al*., 1993; Patel *et al*., 1992). These subunits differ from that of *Eco*AI in the length of the conserved regions at the N and C termini. The *Eco*AI HsdS polypeptide has a long conserved region at the N terminus and only a short conserved region at the C terminus (Figure 2). The type IA and type IC HsdS subunit organisations resemble those of the *Eco*A circularly permuted variants produced by transposition of various parts of the N-terminal conserved region to the C terminus. It would be interesting to examine the effect on protein solubility of transposition of the C-terminal conserved regions of these HsdS polypeptides to their N termini.

In conclusion, the HsdS subunit of the type IB restriction-modification enzyme *Eco*AI tolerates circular permutations of its polypeptide chain, indicating close proximity of the N and C termini in the native protein. This result lends credence to the proposed circular organisation of the HsdS domains. In addition, both HsdS and HsdM subunits are required for specific DNA binding, which suggests a role for HsdM in maintaining the circular domain organisation of HsdS.

Materials and Methods

Bacterial strains and bacteriophages

Plasmid constructions were carried out using *E. coli* DH5α (Woodcock *et al*., 1989). *E. coli* BL21(DE3) was utilized for protein expression by the pET system (Studier *et al*., 1990; Novagen). Bacteriophage λ*vir* (Jacob & Wollman, 1954) was used for phage infection assays.

Plasmid constructions

The plasmid pFFP30 (Fuller-Pace *et al*., 1985) was the source of *Eco*AI *hsd* genes for their cloning in various pET plasmids (Novagen) under control of bacteriophage T7 transcription and translation signals. The plasmid pJP21, expressing the *Eco*AI methylase, was constructed by subcloning the *Stu*I-*Bst*BI fragment from pFFP30 (*hsd*M and *hsd*S genes) into the expression vector pET32a cut with *Nde*I and *Nco*I after the cohesive ends produced by *Bst*BI, *Nde*I and *Nco*I were made blunt with Klenow fragment. The plasmid pJP22, expressing the *hsd*R gene, was constructed by cloning the *Nco*I-*Stu*I fragment from pFFP30 into pET15b cut with *Nco*I and *Nde*I. The *Nde*I site was first made blunt with Klenow to join it to the *Stu*I site of the fragment. The plasmid pJP23, expressing only the *hsd*M gene, is a *Nde*I/*Not*I deletion derivative of pJP21. To express the *hsd*S gene independently, an *Nco*I site was introduced into the initiation codon of this gene in the plasmid pJP21 by site-directed mutagenesis and the *Nco*I-*Xho*I fragment from the resulting construct was subcloned into a *Nco*I-*Xho*I digestion of the vector pET15b. The final construct was named pJP26. Circularly permuted variants of the *Eco*AI *hsd*S gene were constructed by exchanging the order of two portions of the gene at various codons. This was achieved by the following procedure. First, two PCR fragments were generated using $Vent_R$ DNA polymerase (New England Biolabs). One PCR product, coded by the proximal part of the *hsd*S gene, started directly with the codon for the original N-terminal amino acid (Ser). The other end of this product contained a new stop codon followed by a *Xho*I cloning site. The second PCR product, coded by the distal part of the *hsd*S gene, started with a sequence containing either *Nco*I or *Nde*I cloning sites in which the ATG was followed by the codon for the new N-terminal amino acid. The second PCR product ended with the codon for the original C-terminal amino acid (Asn). Following digestion with *Nco*I (*Nde*I) and *Xho*I, respectively, and phosphorylation of the opposite blunt ends, each pair of PCR products was cloned into a pET-protein expression vector cleaved with corresponding restriction enzymes. In these constructs, the codons for the original N and C-terminal amino acids were directly linked and the new polypeptide termini were formed by the amino acids present at the point of disruption of the wild-type polypeptide. Four plasmids containing circularly permuted *hsd*S genes were constructed using the above procedure. The plasmids pJP30 and pJP33, derivatives of pET15b, contain circularly permuted variants of *hsd*S coding for polypeptides with the N terminus at Glu91 (cp91) and Asp283 (cp283), respectively, cloned *via Nco*I and *Xho*I sites. In the plasmids pJP32 and pJP34, derivatives of the pET21b, the permuted genes start at the codons for Thr108 (cp108) and Pro380 (cp380), respectively, and are cloned *via Nde*I and *Xho*I sites. The plasmid pJP31 codes for a truncated mutant of the HsdS polypeptide that lacks the first 90 amino acid residues. This plasmid is a derivative of pJP30, in which the *Afl*II-

*Xho*I fragment was replaced by the *Afl*II-*Xho*I fragment from pJP21. To construct a plasmid expressing only *hsd*R and *hsd*M genes for *in vivo* functional analysis of the circularly permuted variants of *hsd*S, the *hsd*S gene in the plasmid pFFP30 was inactivated by a frame-shift mutation in the *Nde*I site located at the codon for Met10. The *Hin*dIII fragment from this pFFP30 derivative containing the entire *hsd* locus was cloned into the plasmid pACYC184 (Chang & Cohen, 1978), which is compatible with plasmids containing permuted variants of *hsd*S. The resulting plasmid was named pJP28. The plasmid pJP29 was constructed by subcloning the *Sph*I-*Hin*dIII fragment from the plasmid pJP23 into the plasmid pACYC184 cut with *Sph*I and *Hin*dIII. Plasmids pJP36 and pJP37 were constructed by subcloning the *Nco*I-*Xho*I fragment from the plasmids pJP26 and pJP30, respectively, into the plasmid pCYP3 (New England Biolabs) cut with *Nco*I-*Xho*I. The plasmid pJP38 is a derivative of pCYP2 (New England Biolabs) containing *Nde*I-*Xho*I fragment from pJP34. In these plasmids, the *hsd*S gene or its permuted variants $hsdS_{cp91}$ and $hsdS_{cp380}$ are placed under control of the *tac* promoter. The plasmid pJP25 was constructed by cloning the *Kpn*I-*Sal*I fragment from a synthetic 30-mer oligoduplex containing a single *Eco*AI site (shown below) into the vector pTZ18R (Pharmacia).

Phage infection assays

The cells were grown in LB medium containing, when required, 200 μg/ml ampicillin and 20 μg/ml chloramphenicol to select for the various plasmids harbouring wild-type or mutant *hsd* genes. At $A_{600} = 1$, the cells were infected with serial dilutions of bacteriophage λ*vir* and phage titre was determined as described (Gubler & Bickle, 1991).

Protein expression and purification

To produce wild-type and mutant *Eco*AI DNA methyltransferases, *E. coli* BL21(DE3) was transformed with the plasmid pJP29 (HsdM) in combination with pJP36 (wild-type HsdS), pJP37 ($HsdS_{cp91}$) or pJP38 ($HsdS_{cp380}$). HsdS, HsdM and HsdR proteins were also produced in *E. coli* BL21(DE3) from the plasmids pJP26, pJP23 and pJP22, respectively. The cells harbouring the recombinant plasmid(s) were grown in four litres of LB medium containing the appropriate antibiotic(s). At $A_{600} = 0.4–0.5$, protein synthesis was induced by addition of IPTG to a final concentration of 0.4 mM and the cultures were incubated for an additional three hours at 37°C for the methylase, HsdM and HsdS proteins or at 30°C for HsdR protein. The first steps of purification were the same for all proteins: the harvested cells were resuspended in 40 ml of buffer A (10 mM Tris-HCl (pH 8), 150 mM NaCl, 5 mM $MgCl_2$, 2.5 mM EDTA, 3 mM DTT) and broken down by two passages through a French press cell at 1000 psi (1 psi ≈ 6.9 kPa). The cell extracts were clarified by ultracentrifugation at 100,000 *g* for two hours. The NaCl concentration in the supernatant was increased to 0.4 M and nucleic acids were precipitated by addition of neutralised polyethyleneimine to a concentration of 0.4% (v/v). The proteins in the supernatant were precipitated with ammonium sulphate at 70% saturation, redisolved in 20 ml of buffer B (10 mM Tris-HCl (pH 8), 50 mM NaCl, 0.1 mM EDTA, 1 mM DTT) and dialysed against the same buffer. The dialysate was loaded onto a 70 ml DEAE Sepharose FF column (Pharmacia) equilibrated with buffer B and bound proteins were eluted with a linear gradient of NaCl (0.05 M–0.45 M, 855 ml). Individual purification steps were monitored by SDS-PAGE analysis. HsdS protein, wild-type and mutant methylases were further purified on a 40 ml heparin Sepharose CL6B column (Pharmacia) equilibrated with buffer B. The proteins were eluted with a 0.05 M–0.55 M linear gradient of NaCl in buffer B (530 ml). Both HsdM and HsdR proteins were further purified by gel-filtration chromatography on a 200 ml Sephacryl S200 column (Pharmacia) run in buffer C (20 mM Tris (pH 8), 150 mM NaCl, 0.1 mM EDTA, 1 mM DTT). Finally, gel-filtration fractions containing HsdM or HsdR were loaded onto a Mono Q 10/10 column (Pharmacia) and bound proteins were eluted with a 0.05 M–0.45 M linear gradient of NaCl (150 ml) in buffer B. The fractions containing pure proteins were concentrated using Centricon-30 devices (Amicon) and stored at −20°C when supplemented with 50% (v/v) glycerol.

For preparation of reconstituted *Eco*AI methylase (M_2S_1 complex), purified HsdM and HsdS subunits were mixed in a ratio of 2.5:1 (w/w) in buffer B and than passed through a 5 ml HighTrap heparin cartridge (Pharmacia) to remove excess HsdM (HsdM is not bound to heparin). The bound methylase was eluted with a 0.05 M–0.40 M linear gradient of NaCl in buffer B (50 ml). Pooled fractions containing the methylase were concentrated using a Centricon-30 device.

Protein analysis

Protein concentration was determined from the UV absorption at 280 nm using an extinction coefficient derived from the aromatic amino acid composition in the predicted amino acid sequence. Protein samples for ultracentrifugation analysis were dialysed against a buffer containing 20 mM Tris (pH 8), 150 mM NaCl, 1 mM DTT. SDS-PAGE was carried out according to Laemmli (1970). Proteins for N-terminal sequencing were electroblotted on to Immobilon-P membrane (Millipor) in a buffer containing 10 mM Caps (pH 11) and 10% (v/v) methanol. For immunoblotting, polyclonal antibodies raised against a mixture of purified *Eco*AI methylase and HsdR subunit were used (Suri & Bickle, 1985).

Gel electrophoretic mobility shift assay

Complementary synthetic 30-mer oligonucleotides containing a single *Eco*AI site used for the assay were purchased, gel-purified, from Pharmacia. The DNA concentration was determined from measurement of absorption at 260 nm. One of the oligonucleotides was 5′-end labelled using phage T4 polynucleotide kinase and [^{32}P]γATP and then annealed to the complementary strand to generate the duplex shown below. The *Eco*AI recognition sequence including the non-specific spacer (underlined) is shown in bold.

5′-GGGGTACC**GAGGATATCGGTCA**GTCGACGG-3′
3′-CCCCATGG**CTCCTATAGCCAGT**CAGCTGCC-5′

The end-labelled oligoduplex, typically at a concentration of 85 nM, was incubated with various amounts of HsdS, HsdM or M_2S_1 complex in a buffer containing 50 mM Tris (pH 8), 50 mM NaCl, 10% (v/v) glycerol, 10 mM $MgCl_2$, 1 mM DTT. The incubations were carried out at room temperature for 15 minutes. The samples were electrophoresed in non-denaturing 8% polyacrylamide gel run in TBE buffer at 100 V for 1.5 hours. The gel was dried and subjected to autoradiography.

Enzyme assays

DNA methylation and DNA cleavage assays were performed in a buffer containing 50 mM Tris (pH 8), 50 mM NaCl, 10 mM $MgCl_2$, 1 mM DTT. The synthetic 30-mer oligoduplex containing one *Eco*AI site shown above was used as a substrate for *in vitro* methylation assay of wild-type and mutant *Eco*AI methylases. Typically, 1 μM substrate DNA was incubated with 50 nM methylase in the presence of 1 μM [^{3}H]*S*-adenosylmethionine (2.77 TBq/mmol; Amersham) at 37°C. The 20 μl reaction aliquots were withdrawn at timed intervals and the reaction was terminated by heating at 65°C for ten minutes. The amount of incorporated label was measured by scintillation counting of labelled DNA bound to DE81 filters (Whatman) after removal of unincorporated label by washing filters three times in 0.5 M phosphate buffer (pH 6.8).

The plasmid pJP25 containing a single *Eco*AI site was used as a substrate for DNA cleavage assay. Wild-type and mutant restriction endonucleases were reconstituted by mixing the corresponding MTase with the HsdR subunit directly in DNA cleavage reactions. Reactions were started by addition of ATP to a final concentration of 5 mM and incubated at 37°C for, usually, five minutes. Protein and DNA concentrations are indicated in the Figure legends. AdoMet was added to a final concentration of 0.1 mM. Following heat inactivation at 65°C for ten minutes, reactions were electrophoresed in a 0.9% agarose gel run in TBE buffer supplemented with ethidium bromide (EtBr) at a concentration of 0.5 μg/ml. To measure the amount of linear plasmid reaction product, the fluorescent gel was digitalised and quantified using NIH Image 1.62 software. The percentage of total DNA that was linear DNA (linear, supercoiled, nicked) was calculated. Data were not corrected for differences in binding of EtBr to supercoiled and linear DNA.

Acknowledgements

We are very grateful to Dr Maria MacWilliams for fruitful discussions and for comments on the manuscript, and to Professor N.M. Murray for kindly supplying the plasmid pFFP30. We thank Ursula Sandmeier for technical assistance, Ariel Lustig for analytical ultracentrifugation measurement and Thierry Mini for N-terminal sequencing. This work was supported by the Swiss National Science Foundation.

References

Abadjieva, A., Patel, J., Webb, M., Zinkevich, V. & Firman, K. (1993). A deletion mutant of the type IC restriction endonuclease *Eco*R124I expressing a novel DNA specificity. *Nucl. Acids Res.* **21**, 4435–4443.

Bickle, T. A. & Krüger, D. H. (1993). The biology of DNA restriction. *Microbiol. Rev.* **57**, 434–450.

Burckhardt, J., Weisemann, J. & Yuan, R. (1981). Characterization of the DNA methylase activity of the restriction enzyme from *Escherichia coli* K. *J. Biol. Chem.* **256**, 4024–4032.

Chang, A. C. Y. & Cohen, S. N. (1978). Construction and characterization of amplifiable multicopy DNA cloning vehicles derived from the P15A cryptic miniplasmid. *J. Bacteriol.* **134**, 1141–1156.

Cowan, G. M., Gann, A. A. F. & Murray, N. E. (1989). Conservation of complex DNA recognition domains between families of restriction enzymes. *Cell*, **56**, 103–109.

Dreier, J., MacWilliams, M. P. & Bickle, T. A. (1996). DNA cleavage by the type IC restriction-modification enzyme *Eco*R124II. *J. Mol. Biol.* **264**, 722–733.

Dryden, D. T. F., Cooper, L. P. & Murray, N. E. (1993). Purification and characterization of the methyltransferase from the type I restriction and modification system of *Escherichia coli* K12. *J. Biol. Chem.* **268**, 13228–13236.

Dryden, D. T. F., Sturrock, S. S. & Winter, M. (1995). Structural modelling of a type I DNA methyltransferase. *Nature Struct. Biol.* **2**, 632–635.

Endlich, B. & Linn, S. (1985). The DNA restriction endonuclease of *Escherichia coli* B. I. Studies of the DNA translocation and the ATPase activities. *J. Biol. Chem.* **260**, 5720–5728.

Eskin, B. & Linn, S. (1972). The deoxyribonucleic acid modification and restriction enzymes of *Escherichia coli* B. III. Studies of the restriction adenosine triphosphatase. *J. Biol. Chem.* **247**, 6192–6196.

Fuller-Pace, F. V., Bullas, L. R., Delius, H. & Murray, N. E. (1984). Genetic recombination can generate altered restriction specificity. *Proc. Natl Acad. Sci. USA,* **81**, 6095–6099.

Fuller-Pace, F. V., Cowan, G. M. & Murray, N. E. (1985). *Eco*A and *Eco*E: alternatives to the *Eco*K family of type I restriction and modification systems of *Escherichia coli*. *J. Mol. Biol.* **186**, 65–75.

Goldenberg, D. P. & Creighton, T. E. (1983). Circular and circularly permuted forms of bovine pancreatic trypsin inhibitor. *J. Mol. Biol.* **165**, 407–413.

Gough, J. A. & Murray, N. E. (1983). Sequence diversity among related genes for recognition of specific targets in DNA molecules. *J. Mol. Biol.* **166**, 1–19.

Graf, R. & Schachman, H. K. (1996). Random circular permutation of genes and expressed polypeptide chains: application of the method to the catalytic chains of aspartate transcarbamoylase. *Proc. Natl Acad. Sci. USA,* **93**, 11591–11596.

Gubler, M. & Bickle, T. A. (1991). Increased protein flexibility leads to promiscuous protein-DNA interactions in type IC restriction-modification systems. *EMBO J.* **10**, 951–957.

Gubler, M., Braguglia, D., Meyer, J., Piekarowicz, A. & Bickle, T. A. (1992). Recombination of constant and variable modules alters DNA sequence recognition by type IC restriction-modification enzymes. *EMBO J.* **11**, 233–240.

Hahn, M., Piotukh, K., Borriss, R. & Heinemann, U. (1994). Native-like in vivo folding of a circularly permuted jellyroll protein shown by crystal structure analysis. *Proc. Natl Acad. Sci. USA,* **91**, 10417–10421.

Jacob, F. & Wollman, E. L. (1954). Etude génétique d'un bactériophage tempéré d'*Escherichia coli*. I. Le système génétique du bactériophage l. *Ann. Inst. Pasteur.* **87**, 653–673.

Janscak, P., Abadjieva, A. & Firman, K. (1996). The type I restriction endonuclease R. *Eco*R124I: over-production and biochemical properties. *J. Mol. Biol.* **257**, 977–991.

Johnson, J. L. & Raushel, F. M. (1996). Influence of primary sequence transpositions on the folding pathways of ribonuclease T1. *Biochemistry,* **35**, 10223–10233.

Kannan, P., Cowan, G. M., Daniel, A. S., Gann, A. A. F. & Murray, N. E. (1989). Conservation of organization in the specificity polypeptides of two families of type I restriction enzymes. *J. Mol. Biol.* **209**, 335–344.

Kneale, G. G. (1994). A symmetrical model for the domain structure of Type I DNA methyltransferases. *J. Mol. Biol.* **243**, 1–5.

Koebnik, R. & Kramer, L. (1995). Membrane assembly of circularly permuted variants of the *E. coli* outer membrane protein OmpA. *J. Mol. Biol.* **250**, 617–626.

Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature,* **227**, 680–685.

Luger, K., Hommel, U., Herold, M., Hofsteenge, J. & Kirschner, K. (1989). Correct folding of circularly permuted variants of a beta alpha barrel enzyme *in vivo*. *Science,* **243**, 206–210.

MacWilliams, M. P. & Bickle, T. A. (1996). Generation of new DNA binding specificity by truncation of the type IC *Eco*DXXI hsdS gene. *EMBO J.* **15**, 4775–4783.

Meister, J., MacWilliams, M., Hubner, P., Jutte, H., Skrzypek, E., Piekarowicz, A. & Bickle, T. A. (1993). Macroevolution by transposition: drastic modification of DNA recognition by the type I restriction enzyme following Tn5 transposition. *EMBO J.* **12**, 4585–4591.

Mernagh, D. R., Reynolds, L. A. & Kneale, G. G. (1997). DNA binding and subunit interactions in the type I methyltransferase M. EcoR124I. *Nucl. Acids Res.* **25**, 987–991.

Nagaraja, V., Shepherd, J. C. W. & Bickle, T. A. (1985). A hybrid recognition sequence in a recombinant restriction enzyme and the evolution of DNA sequence specificity. *Nature,* **316**, 371–372.

Patel, J., Taylor, I., Dutta, C. F., Kneale, G. & Firman, K. (1992). High-level expression of the cloned genes encoding the subunits of and intact DNA methyltransferase, M. EcoR124. *Gene,* **112**, 21–27.

Pieper, U., Hayakawa, K., Li, Z. & Herzberg, O. (1997). Circularly permuted beta-lactamase from *Staphylococcus aureus* PC1. *Biochemistry,* **36**, 8767–8774.

Powell, L. M., Dryden, D. T. F., Willcock, D. F., Pain, R. H. & Murray, N. E. (1993). DNA recognition by the EcoK methyltransferase. The influence of DNA methylation and the cofactor *S*-adenosyl-L-methionine. *J. Mol. Biol.* **234**, 60–71.

Redaschi, N. & Bickle, T. A. (1996). DNA restriction and modification systems. In *Escherichia coli and Salmonella:Cellular and Molecular Biology* (Neidhardt, F. C., Curtiss, R., III, Ingraham, J. L., Brooks, Low K. B., Magasanik, B., Reznikoff, W. S., Riley, R., Schaechter, M. & Umbarger, H. E., eds), 2nd edit., pp. 773–781, ASM Press, Washington, DC.

Studier, F. W. & Bandyopadhyay, P. K. (1988). Model for how type I restriction enzymes select cleavage sites in DNA. *Proc. Natl Acad. Sci. USA,* **85**, 4677–4681.

Studier, F. W., Rosenberg, A. H., Dunn, J. J. & Dubendorff, J. W. (1990). Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* **185**, 60–89.

Suri, B. & Bickle, T. A. (1985). *Eco*A: the first member of a new family of type I restriction modification systems. Gene organization and enzymatic activities. *J. Mol. Biol.* **186**, 77–85.

Suri, B., Shepherd, J. C. W. & Bickle, T. A. (1984). The *Eco*A restriction and modification system of *Escherichia coli* 15T-: enzyme structure and DNA recognition sequence. *EMBO J.* **3**, 575–579.

Szczelkun, M. D., Dillingham, M. S., Janscak, P., Firman, K. & Halford, S. E. (1996). Repercussions of DNA tracking by the type IC restriction endonuclease *Eco*R124I on linear, circular and catenated substrates. *EMBO J.* **15**, 6335–6347.

Taylor, I., Patel, J., Firman, K. & Kneale, G. (1992). Purification and biochemical characterisation of the *Eco*R124 type I modification methylase. *Nucl. Acids Res.* **20**, 179–186.

Taylor, I., Watts, D. & Kneale, G. (1993). Substrate recognition and selectivity in the type IC DNA modification methylase M. EcoR124I. *Nucl. Acids Res.* **21**, 4929–4935.

Thorpe, P. H., Ternent, D. & Murray, N. E. (1997). The specificity of StySKI, a type I restriction enzyme, implies a structure with rotational symmetry. *Nucl. Acids Res.* **25**, 1694–1700.

Tyndall, C., Meister, J. & Bickle, T. A. (1994). The *Escherichia coli prr* region encodes a functional type IC DNA restriction system closely integrated with an anticodon nuclease gene. *J. Mol. Biol.* **237**, 266–274.

Webb, M., Taylor, I. A., Firman, K. & Kneale, G. G. (1995). Probing the domain structure of the type IC DNA methyltransferase M. EcoR124I by limited proteolysis. *J. Mol. Biol.* **250**, 181–190.

Willcock, D. F., Dryden, D. T. F. & Murray, N. E. (1994). A mutational analysis of the two motifs common to adenine methyltransferases. *EMBO J.* **13**, 3902–3908.

Woodcock, D. M., Crowther, P. J., Doherty, J., Jefferson, S., DeCruz, E., Noyer-Weidner, M., Smith, S. S., Michael, M. Z. & Graham, M. W. (1989). Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. *Nucl. Acids Res.* **17**, 3469–3478.

Yang, Y. R. & Schachman, H. K. (1993). Aspartate transcarbamoylase containing circularly permuted catalytic polypeptide chains. *Proc. Natl Acad. Sci. USA,* **90**, 11980–11984.

Yuan, R., Heywood, J. & Meselson, M. (1972). ATP hydrolysis by restriction endonuclease from *E. coli* K. *Nature New Biol.* **240**, 42–43.

Yuan, R., Hamilton, D. L. & Burckhardt, J. (1980). DNA translocation by the restriction enzyme from *E. coli* K. *Cell,* **20**, 237–244.

Edited by J. Karn

(Received 27 July 1998; received in revised form 24 September 1998; accepted 24 September 1998)