Kurz, L. C., & Drysdale, G. R. (1987) *Biochemistry 26*, 2623–2627.
Layton, E. M., Kross, R. D., & Fassel, V. A. (1956) *J. Chem. Phys. 25*, 135–138.
MacClement, B. A. E., Carriere, R. G., Phelps, D. J., & Carey, P. R. (1981) *Biochemistry 20*, 3438–3447.
Pimentel, G. C., & McClellan, A. L. (1960) *The Hydrogen Bond*, Freeman, London.
Sans Cartier, L. R., Storer, A. C., & Carey, P. R. (1988) *J. Raman Spectrosc. 19*, 117–121.
Smolarsky, M. (1980) *Biochemistry 19*, 478–484.
Thijs, R., & Zeegers-Huyskens, T. (1984) *Spectrochim. Acta 40A*, 307–313.
Tonge, P. J., & Carey, P. R. (1989) *Biochemistry 28*, 6701–6709.
Wells, J. A., Cunningham, B. C., Graycar, T. P., & Estell, D. A. (1986) *Philos. Trans. R. Soc. London, A 317*, 415–423.
Willis, K. J., & Szabo, A. G. (1989) *Biochemistry 28*, 4902–4908.

Articles

Fidelity of DNA Recognition by the *Eco*RV Restriction/Modification System in Vivo†

John D. Taylor, Annette J. Goodall, Christian L. Vermote, and Stephen E. Halford*
Department of Biochemistry, Centre for Molecular Recognition, University of Bristol, Bristol BS8 1TD, U.K.
Received June 28, 1990; Revised Manuscript Received August 22, 1990

ABSTRACT: The *Eco*RV restriction/modification system consists of two enzymes that recognize the DNA sequence GATATC. The *Eco*RV restriction endonuclease cleaves DNA at this site, but the DNA of *Escherichia coli* carrying the *Eco*RV system is protected from this reaction by the *Eco*RV methyltransferase. However, in vitro, the *Eco*RV nuclease also cleaves DNA at most sites that differ from the recognition sequence by one base pair. Though the reaction of the nuclease at these sites is much slower than that at the cognate site, it still appears to be fast enough to cleave the chromosome of the cell into many fragments. The possibility that the *Eco*RV methyltransferase also protects the noncognate sites on the chromosome was examined. The modification enzyme methylated alternate sites in vivo, but these were not the same as the alternate sites for the nuclease. The excess methylation was found at GATC sequences, which are also the targets for the *dam* methyltransferase of *E. coli*, a protein that is homologous to the *Eco*RV methyltransferase. Methylation at these sites gave virtually no protection against the *Eco*RV nuclease: even when the *Eco*RV methyltransferase had been overproduced, the cellular DNA remained sensitive to the *Eco*RV nuclease at its noncognate sites. The viability of *E. coli* carrying the *Eco*RV restriction/modification system was found instead to depend on the activity of DNA ligase. Ligase appears to proofread the *Eco*RV R/M system in vivo: DNA, cut initially in one strand at a noncognate site for the nuclease, is presumably repaired by ligase before the scission of the second strand.

Restriction/modification (R/M)[1] systems possess two enzyme activities: a modification methyltransferase that recognizes a specific DNA sequence and catalyzes the transfer of a methyl group from AdoMet to a particular base within the recognition sequence and a restriction endonuclease that cleaves the DNA provided that neither strand has been methylated (Arber, 1979; Smith, 1979). These systems are widespread in prokaryotes, and their function is to maintain the integrity of the bacterial DNA. DNA that lacks the appropriate pattern of methylation is cleaved by the restriction enzyme, while the cellular DNA is protected by the methyltransferase (Arber, 1979). However, restriction enzymes can also cleave DNA at sequences other than their recognition sites (Bennett & Halford, 1989), often making the double strand break at these sequences by first cutting one strand of the DNA and then the second (Taylor & Halford, 1989; Thielking et al., 1990; Lesser et al., 1990). These reactions could be lethal to the cell unless the bacterium has a mechanism to protect the alternative sequences on its chromosome. One possibility is that the modification enzyme has a less stringent specificity for the recognition site than the nuclease, so that any alternative site for the nuclease is already methylated. This appears to be the case with the *Eco*RI R/M system (Woodbury et al., 1980a,b). An alternative is that another enzyme in the cell, perhaps DNA ligase, acts to proofread the specificity of the nuclease by selectively repairing DNA nicked at noncognate sites. A proofreading scheme for the *Eco*RV restriction enzyme involving DNA ligase has been modeled in vitro (Taylor & Halford, 1989), but it remains to be determined whether or not such a scheme operates in vivo.

*Eco*RI and *Eco*RV are both type II R/M systems (Smith, 1979). With the type II systems, in contrast to type I or III (Bickle, 1987), the restriction and modification activities are due to two separate enzymes [reviewed by Modrich and

†This work was funded by the Science and Engineering Research Council.

[1] Abbreviations: AdoMet, *S*-adenosylmethionine; Ap, ampicillin (with superscripts r and s to denote resistance and sensitivity); bp, base pair(s); BSA, bovine serum albumin; βME, β-mercaptoethanol; Cm, chloramphenicol; DMSO, dimethyl sulfoxide; kb, 1000 bp; Kn, kanamycin; R/M, restriction/modification; Sm, streptomycin; [ ], plasmid carrier state.

0006-2960/90/0429-10727$02.50/0

Roberts (1982) and Bennett and Halford (1989)]. The recognition sequence for the *Eco*RV R/M system is GATATC. In the presence of Mg^{2+}, the *Eco*RV restriction enzyme cleaves both strands at the center of this site (D'Arcy et al., 1985). In the presence of AdoMet, the *Eco*RV modification enzyme methylates the first adenine in this sequence (Nwosu et al., 1988). In all type II systems characterized to date, including *Eco*RV (Bougueleret et al., 1984), the recognition of the same DNA sequence by the restriction and the modification enzymes is not a consequence of homology between the two proteins: in each system, the two proteins have completely dissimilar amino acid sequences (Wilson, 1988). Moreover, with both *Eco*RI and *Eco*RV, the mechanism by which the methyltransferase recognizes the target sequence differs from that of the nuclease (Modrich & Rubin, 1977; Brennan et al., 1986; Newman et al., 1990). However, the modification enzymes from several different R/M systems are homologous to each other and to other DNA methyltransferases (Lauster et al., 1989; Posfai et al., 1989). For example, the *Eco*RV modification enzyme is homologous to both the *Dpn*II modification enzyme and the *dam* methyltransferases of *Escherichia coli* and phage T4, which recognize GATC (Lauster et al., 1987). The *dam* methyltransferase is associated with mismatch repair, gene expression, and DNA replication but not with R/M (Modrich, 1987; Barras & Marinus, 1989). The *Eco*RV and the *dam* methyltransferases have thus evolved not only to recognize different DNA sequences but also to fulfill different biological functions.

Experimental Procedures

Bacterial Strains. *E. coli* GM33 and K12ΔHIΔ*trp* have been described previously (Marinus & Morris, 1974; Remaut et al., 1983). The former is *dam* and the latter *dam*$^+$ λ*cI*857. Dam phenotypes were checked by using the cellular DNA as substrates for the *Mbo*I and *Dpn*I restriction enzymes (Geier & Modrich, 1979). *E. coli* strains N1624, N1626, and N2668 carry respectively *lig*$^+$, *lig*4, and *lig*ts7 but are otherwise isogenic: all three are Sm^r (Gottesman et al., 1973). All bacterial cultures were in L-broth or on LB agar except for the measurements of the viabilities of *lig* strains, where H-broth was used instead (Konrad et al., 1973). Media for strains carrying one or more of the plasmids listed below always contained the relevant antibiotics at 50 μg/mL, and the same concentration was used for Sm^r strains. Temperature inductions before either DNA or protein purifications were carried out by first growing the culture at 28 °C to A_{550} 0.4, then adding an equal volume of broth at 55 °C, continuing the growth at 42 °C, and subsequently harvesting the cells 2–4 h later. Transformations were by the $CaCl_2$/RbCl method as described in Maniatis et al. (1982).

Plasmids. The sources and the relevant characteristics of the plasmids used were as follows: pAT153 (Twigg & Sherratt, 1980), Ap^r; pLBM (Bougueleret et al., 1985), Cm^r, *ecoRVM* expressed from its natural promoter; pVIC1 (Nwosu et al., 1988), Ap^r, *ecoRVM* expressed from λP_L; pTZ115 (Bougueleret et al., 1985), Ap^r, *ecoRVR* expressed from λP_L; pEMA5/T96K (this laboratory), an Ap^r derivative of pMa5-8 (Stanssens et al., 1989) carrying *ecoRVR* expressed from λP_R but with a spontaneous mutation within *ecoRVR* that converts Thr96 to Lys; p*cI*857 (Remaut et al., 1983), Kn^r, λ*cI*857; pMetB (this laboratory), Kn^r, λ*cI*857, *ecoRVM* (constructed by inserting the 1.3-kb *Bam*HI–*Hin*dIII fragment from pLBM, which contains *ecoRVM*, at the *Bst*EII site on p*cI*857). [R/M genes are named as in Szybalski et al. (1988).] Except for pLBM, p*cI*857, and pMetB, all of the above carry the same origin of replication as pAT153. The three exceptions stem from pACYC184 and are thus compatible with the other plasmids.

DNA. Genomic DNA was isolated from *E. coli* as described by Hopwood et al. (1985). Plasmids were purified through two CsCl/ethidium bromide density gradients (Halford & Johnson, 1981). All other manipulations of DNA were essentially as in Maniatis et al. (1982).

Proteins. *Eco*RV methyltransferase was purified from either *E. coli* K12ΔHIΔ*trp* that had been transformed with pVIC1 or *E. coli* GM33 that had been transformed first with p*cI*857 and then with pVIC1. In both cases, cells were harvested 4 h after induction and the purification was that of Nwosu et al. (1988). Units of *Eco*RV methyltransferase were evaluated from the protection of phage λ DNA against the *Eco*RV nuclease (Nwosu et al., 1988). The *Eco*RV nuclease was purified as in Luke et al. (1987). Reactions of the nuclease was carried out either in buffer A or in buffer D (Taylor & Halford, 1989): buffer A is 100 mM NaCl, 50 mM Tris, 10 mM $MgCl_2$, 10 mM βME, and 100 μg/mL BSA, pH 7.5; buffer D is 50 mM Tris, 10 mM $MgCl_2$, 10 mM βME, 100 μg/mL BSA, and 10% (v/v) DMSO, pH 8.5. Protein concentrations were by the method of Bradford (1976), and the molarities given here are for the active forms of each enzyme: the dimer for the nuclease (D'Arcy et al., 1985) and the monomer for the methyltransferase (Garnett & Halford, 1988). *Bsp*AI (Mullings et al., 1986) was a gift from L. R. Evans (this department). All other enzymes were from Gibco-BRL and were used as advised by the supplier.

Results and Discussion

Protection at the Cognate Site. This study employed two plasmids that encode the *Eco*RV methyltransferase. On one, pLBM, *ecoRVM* is expressed constitutively from its natural promoter (Bougueleret et al., 1985). On the other, pVIC1, *ecoRVM* is linked to the λP_L promoter whose activity can be regulated by the temperature-sensitive *cI*857 repressor from phage λ (Nwosu et al., 1988): the latter was encoded by a compatible plasmid, p*cI*857. When cells transformed with pVIC1 are grown at 28 °C, expression of *ecoRVM* is inhibited by the repressor but, after derepression at 42 °C, the cells produce nearly 20% of their protein as *Eco*RV methyltransferase.

Genomic DNA was isolated from *E. coli* strains carrying these plasmids and subsequently digested with the *Eco*RV nuclease (Figure 1). The concentration of the *Eco*RV nuclease used here was sufficient to cleave all unmodified recognition sites on these DNA samples, but it was insufficient for reactions at any other DNA sequences (Luke et al., 1987; Taylor & Halford, 1989). The genomic DNA obtained from the *E. coli* strain in the absence of either plasmid was cleaved as expected by the *Eco*RV nuclease to a series of smaller fragments (Figure 1). Likewise, the DNA from the cells containing pVIC1 grown at 28 °C was also cleaved by the *Eco*RV restriction enzyme. In contrast, the DNA from cells carrying either pLBM or pVIC1 (the latter induced at 42 °C) was unaffected by the *Eco*RV nuclease (Figure 1). Hence, these two samples of DNA must have been fully protected in vivo at all *Eco*RV recognition sites.

In order to determine the amount of the *Eco*RV methyltransferase in these strains, cultures of *E. coli* GM33 [pLBM], grown at 37 °C, and of GM33 [p*cI*857, pVIC1], grown initially at 28 °C and then transferred at 42 °C for 4 h, were harvested by centrifugation. The cell pellets were resuspended and disrupted by sonication and the resultant extracts assayed for *Eco*RV methyltransferase activity [all as in Nwosu et al. (1988)]. The strain carrying pLBM yielded 2×10^3 units of

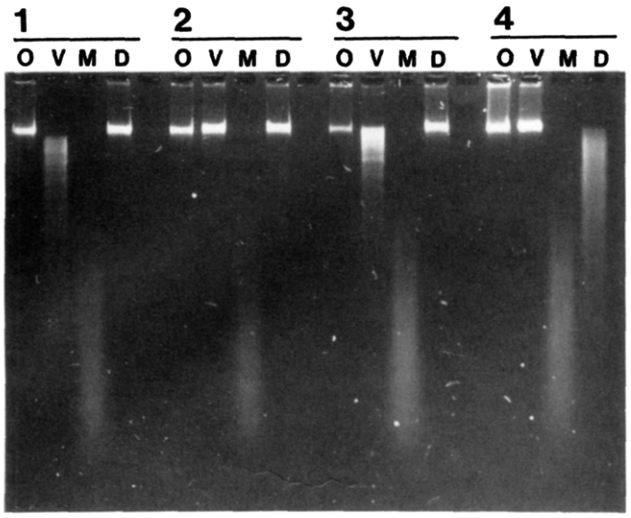

FIGURE 1: Methylation of cellular DNA. Genomic DNA was extracted from the following strains of *E. coli*, numbered in sets as indicated above the gel: 1, GM33 grown at 37 °C; 2, GM33 [pLBM] grown at 37 °C; 3, GM33 [p*cI*857, pVIC1] grown at 28 °C; 4, GM33 [p*cI*857, pVIC1] grown initially at 28 °C, then transferred to 42 °C, and harvested 2 h later. Samples (2 μg) from each DNA preparation were analyzed by electrophoresis through 1% agarose without any prior restriction digest (lane 0 in each set) or were first subjected to 1-h reactions with either the *Eco*RV (lane V), *Mbo*I (lane M), or *Dpn*I (lane D) restriction enzymes. The *Eco*RV reactions were with 0.5 nM enzyme in 20 μL of buffer A at 37 °C. The *Mbo*I and *Dpn*I reactions were with 7 and 12 units, respectively.

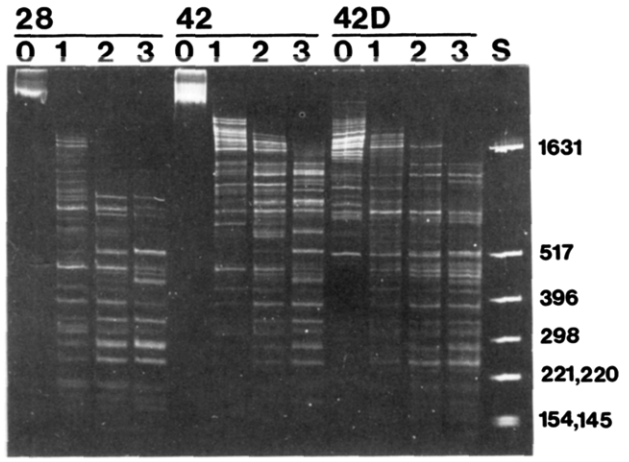

FIGURE 2: Lack of protection at noncognate sites. The plasmids were purified from *E. coli* GM33 [p*cI*857, pVIC1] that had been either grown at 28 °C (in the set labeled 28 above the gel) or grown initially at 28 °C, then transferred to 42 °C, and harvested 2 h later (in the sets labeled 42 and 42D). Samples from each preparation of the plasmid DNA (5 μg in 100 μL of buffer D) were reacted at 37 °C with 600 nM *Eco*RV restriction enzyme. Aliquots (20 μL) were removed from these reactions at timed intervals, mixed immediately with phenol, washed with diethyl ether, and subsequently analyzed by electrophoresis through 6% polyacrylamide: lane 0 (in each set), before the addition of *Eco*RV; lane 1, 5 min after *Eco*RV; lane 2, 15 min; lane 3, 60 min. For the set labeled 42D, the plasmids from the cells grown at 42 °C were digested with *Dpn*I prior to the *Eco*RV reaction. Lane S is a size marker generated by *Hin*fI on pAT153 (fragment sizes, in bp, given on the right of the gel).

methyltransferase/g (wet weight) of cells while that with pVIC1 had 5×10^6 units/g of cells. From the specific activities of these two extracts, we calculate that the in vivo concentration of *Eco*RV methyltransferase expressed from pLBM is about 500 nM and that after induction of pVIC1 is about 1 mM. Though pLBM is a multicopy plasmid, the former level of methyltransferase is probably a fair reflection of the situation in the native strain, for the original isolate of *E. coli* carried the *Eco*RV R/M system on a multicopy colE1-like plasmid (Bougueleret et al., 1984).

Protection at the Noncognate Site. The *Eco*RV restriction enzyme cleaves DNA not only at its cognate recognition site but also at a number of alternative sequences, the latter being described as either primary or secondary noncognate sites (Halford et al., 1986). The primary noncognate sites comprise all sequences that differ from GATATC by 1 bp except for those where the guanine is replaced by a pyrimidine or the symmetric equivalent, a purine in place of the cytosine (Halford et al., 1986). In buffer A at 20 °C, the values of k_{cat}/K_m for double-strand breaks at either the cognate site or one particular noncognate site (GTTATC at position 1734 on pAT153) are respectively 3×10^7 M^{-1} s^{-1} and 36 M^{-1} s^{-1} (Taylor & Halford, 1989). The latter value leads to the prediction that, in vivo, the reactions of the *Eco*RV nuclease at its noncognate sites would cleave the chromosome of the cell into about 100 fragments within each cell division cycle of 30 min (Taylor & Halford, 1989).

Whether or not this prediction is correct depends on whether the activity of the nuclease at noncognate sites in vivo corresponds to that measured in vitro. The DNA within *E. coli* will be complexed with both polyanions such as spermine and basic proteins such as HU (Drlica, 1987), both of which might inhibit the *Eco*RV restriction enzyme: this has been observed with *Eco*RI (Pingoud et al., 1984). However, two factors suggest that the activity in vivo may be higher than that in vitro. First, the kinetics of the *Eco*RV nuclease at the noncognate site on pAT153 were measured by Taylor and Halford (1989) at 20 °C rather than at 37 °C, and this reaction is faster at the higher temperature (data not shown). Second, the experiments were carried out in buffers that contained chloride ions while the predominant anion in the cytoplasm of *E. coli* is glutamate (Richey et al., 1987). The *Eco*RV restriction enzyme is more active in the presence of glutamate (Leirmo et al., 1987). We have confirmed this observation by measuring steady-state velocities for the reaction of the *Eco*RV nuclease at its single cognate site on pAT153 [as in Halford and Goodall (1988): data not shown]. In 0.1 M sodium glutamate, the reaction velocity was only 1.5 times higher than that in 0.1 M NaCl [1.4–0.9 M/(M·min)]. But in 0.3 M sodium glutamate [the level in vivo: Richey et al. (1987)], it was 60 times faster than that in 0.3 M NaCl [0.6–0.01 M/(M·min)]. The reaction rates at the noncognate sites on pAT153 were also faster in glutamate than in chloride (data not shown). Hence, as with other systems (Richey et al., 1987; Leirmo et al., 1987; Bracco et al., 1989), the in vitro measurements of this DNA–protein interaction may underestimate the efficiency of the interaction in vivo. We describe below a further example of this: a reaction of the *Eco*RV methyltransferase that occurs in vivo but not in vitro.

It seems likely that an *E. coli* cell carrying the *Eco*RV R/M system needs to protect its chromosome, not only at the cognate sites for the nuclease but also at noncognate sites. One mechanism for this could be that the modification enzyme discriminates against noncognate sites less efficiently than the restriction enzyme, so that these sites are also methylated. To test this, plasmid DNA was isolated from *E. coli* GM33 [p*cI*857, pVIC1] that had been either cultured at 28 °C or grown initially at 28 °C and then at 42 °C.[2] Both prepa-

[2] Both plasmids are about 4.0 kb, and their relative yields were determined from the amount of each DNA after *Hin*dIII digests (Figure 3). At 28 °C, the preparation was about 75% pVIC1 and 25% p*cI*857 while, at 42 °C, the ratio was reversed. As pVIC1 stems from pAT153, it normally has the higher copy number (Twigg & Sherratt, 1980), but presumably the transcription from λP_L at 42 °C inhibits its replication.

rations were then digested with the *Eco*RV nuclease under conditions where the enzyme cleaves noncognate sites, and the appearance of products was monitored with time (Figure 2). Since neither pc*I*857 nor pVIC1 possesses the cognate sequence for *Eco*RV, any cutting of these DNAs must be at noncognate sites.[3] At the concentration of the *Eco*RV nuclease used in Figure 2, in reaction buffer D, all of the primary noncognate sites would have been cleaved within 1 h (Halford et al., 1986; Luke et al., 1987).

The plasmids from the 28 °C culture were, as expected, cleaved by the *Eco*RV nuclease into a large number of DNA fragments, nearly all of which were <500 bp (Figure 2). The same plasmids, from cells transferred to 42 °C, were again cleaved by the *Eco*RV nuclease to a large number of fragments, many of which matched those from the 28 °C DNA, but this reaction also yielded several other fragments of >500 bp (Figure 2). Reactions with higher concentrations of the nuclease, or for longer times, failed to convert these fragments of >500 bp into smaller fragments (data not shown). The DNA in the cells induced at 42 °C would have been exposed in vivo to concentrations of the *Eco*RV methyltransferase that are about 2000 times higher than that in the native strain (determined above). Yet even this level failed to confer full protection of the cellular DNA against the *Eco*RV nuclease at its noncognate sites. A small fraction of the noncognate sites for the nuclease are protected upon overexpression of the methyltransferase in vivo, but the majority of these sites remain susceptible.

These results with *Eco*RV differ markedly from previous data with the *Eco*RI R/M system. In vitro, the *Eco*RI modification enzyme can introduce many more methyl groups onto DNA than can be accounted for by the number of *Eco*RI recognition sites (Woodbury et al., 1980a). This overmethylation prevents the *Eco*RI restriction enzyme from cleaving noncognate sites that otherwise would have been susceptible (Woodbury et al., 1980b). Moreover, DNA isolated from an *E. coli* strain carrying the *Eco*RI R/M system cannot be further methylated in vitro, so the excess methylation may also occur in vivo (Woodbury et al., 1980a). Hence, it appears that the *Eco*RI methyltransferase confers protection from the *Eco*RI nuclease at both cognate and noncognate sites.

Methylation at dam Sites. Since the *Eco*RV modification enzyme failed to methylate most DNA sequences that differ from its recognition site by 1 bp, and given the homology between the *Eco*RV and the *dam* methyltransferases (Lauster et al., 1987), it may be that *dam* sites are the alternative sites for the *Eco*RV enzyme. This was examined by using two restriction enzymes that cleave DNA at the *dam* site, GATC: *Mbo*I cleaves here only when the adenine is not methylated while *Dpn*I requires the adenine to be methylated (Geier & Modrich, 1979). A third enzyme, *Bsp*AI, was used as a control as this cleaves at GATC regardless of the state of adenine methylation (Mullings et al., 1986).

The genomic DNA for the *dam* strain, GM33, was cleaved to a heterogeneous series of small fragments by *Mbo*I, but no reactions by *Dpn*I were observed (Figure 1). The same results were obtained with the DNA from cells of this strain carrying either pLBM or pVIC1 at 28 °C (Figure 1). Hence, in these three samples, none of the *dam* sites were methylated. In control experiments, DNA isolated from a *dam*$^+$ strain underwent no reaction with *Mbo*I but was cleaved by *Dpn*I (data not shown). In contrast to either of these patterns, the DNA obtained from *dam* cells carrying pVIC1, with induction at 42 °C, was cleaved by both *Mbo*I and *Dpn*I, though *Dpn*I produced much larger fragments than *Mbo*I (Figure 1). In cells that had overproduced the *Eco*RV modification enzyme, some DNA methylation at GATC sequences must have occurred, though with either a low yield of methylation at each site or, alternatively, complete methylation at a subset of these sites.

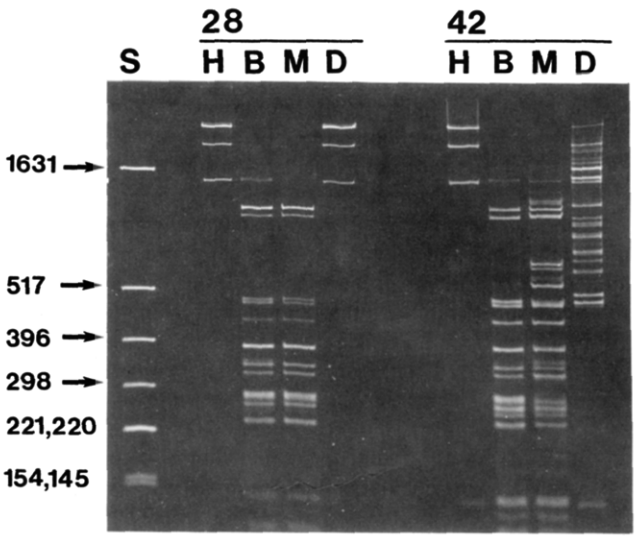

FIGURE 3: Methylation at *dam* sites. The plasmids were purified from *E. coli* GM33 [pc*I*857, pVIC1] that had been either grown at 28 °C (in the set labeled 28 above the gel) or grown initially at 28 °C, then transferred to 42 °C, and harvested 2 h later (in the set labeled 42). Samples (1 μg) of each preparation were digested for 1 h with *Hind*III alone (lane H in both sets) or with *Hin*dIII and one of the following: lane B, *Bsp*AI; lane M, *Mbo*I; lane D, *Dpn*I. (*Hin*dIII cleaves pc*I857* to three fragments of 2.6, 1.3, and 0.14 kb and pVIC1 once to a linear DNA of 4.0 kb.) The samples were subsequently analyzed by electrophoresis through 6% polyacrylamide. Lane S is a size marker generated by *Hin*fI on pAT153 (fragment sizes, in bp, given on the left of the gel).

DNA that had been exposed to high levels of the *Eco*RV methyltransferase in vivo was further characterized, by carrying out *Bsp*AI, *Mbo*I, and *Dpn*I digests on the plasmids purified from *E. coli* GM33 [pc*I*857, pVIC1] after growth at either 28 or 42 °C (Figure 3). This showed that, at 28 °C, none of the *dam* sites had been methylated: the *Mbo*I digest yielded the same DNA fragments as *Bsp*AI, and no reaction at all was detected with *Dpn*I (Figure 3). After exposure to 42 °C, the two plasmids gave the same *Bsp*AI fragments as before (though the yield from each plasmid differed),[2] while both the *Mbo*I and *Dpn*I digests failed to go to completion (though the pattern generated by *Mbo*I was closer to that with *Bsp*AI than was the case with *Dpn*I). These reactions were repeated with increasing amounts of either *Mbo*I of *Dpn*I (up to 50 units) and for increasing times (up to 3 h), but no alteration from the patterns shown in Figure 3 was observed. The products of the *Dpn*I digest were a large number of DNA fragments of 2.5–0.5 kb, which must be formed in partial yield as the sum of the sizes of these fragments is larger than that of the starting material. Hence, the overproduction of the *Eco*RV methyltransferase resulted in partial methylation of a large number of *dam* sites in vivo. If only a subset of the *dam* sites had been methylated, *Dpn*I would have generated a small number of DNA products, each in stoichiometric yield.

[3] The DNA sequence from PVIC1 was assembled from its component parts and analyzed with UWGCG programs (Devereaux et al., 1984). It lacks the recognition sequence for *Eco*RV, but primary noncognate sequences are present at 21 locations. We do not have the complete sequence for pc*I*857 but found experimentally that it lacked the cognate site, and we would expect, on statistical grounds, that it has about 20 noncognate sites.

Two of the noncognate sites for the *Eco*RV nuclease, GATCTC and GAGATC, contain the *dam* sequence, GATC, but the number of *Dpn*I cleavages on the two plasmids used here far exceeds the number of such sequences on this DNA: pVIC1 contains only two of these sites.[3] However, methylation at these two sequences might account for why the *Eco*RV nuclease failed to cleave all of its noncognate sites on the 42 °C preparations of the plasmids: it left products of >500 bp while almost all of the products from the 28 °C preparation were <500 bp (Figure 2). This was examined by digesting the 42 °C preparation with *Dpn*I before the *Eco*RV reaction at its noncognate sites: several (but not all) of the larger products from the *Eco*RV reaction were no longer detected (Figure 2). Hence, at least part of the low level of protection at noncognate sites for the nuclease is due to methylation at these two sequences.

We have attempted to reproduce in vitro the methylation of *dam* sites by the *Eco*RV modification enzyme. Plasmid pAT153, obtained from the *dam* strain GM33, was incubated with the purified *Eco*RV methyltransferase under a variety of reaction conditions and the DNA subsequently tested with either the *Eco*RV or *Dpn*I nucleases for methylation at *Eco*RV and *dam* sites, respectively (data not shown). The concentrations of *Eco*RV methyltransferase were up to 2000 times the miminum needed for methylation of *Eco*RV sites, and the reaction conditions covered a wide range of pH values (from 6 to 10), different NaCl concentrations (0–0.2 M), and different concentrations of AdoMet (0.1–10 mM). Reactions were also carried out in the presence of DMSO, which enhances the activity of the *Eco*RV nuclease at its noncognate sites (Halford et al., 1986). When purified from *E. coli* K12ΔHIΔ*trp* [pVIC1], as described by Nwosu et al. (1988), the resultant preparations methylated *dam* sites: this needed 500 times more protein than that for the methylation of *Eco*RV sites. However, the *E. coli* strain used by Nwosu et al. (1988) is *dam*$^+$, and when purified from GM33 [p*cI*857, pVIC1], the preparations had no detectable activity at *dam* sites under all conditions tested. The former preparations of the *Eco*RV methyltransferase are likely to have contained trace amounts of the *dam* enzyme, and no methylation of *dam* sites can be ascribed to the *Eco*RV modification enzyme in vitro.

The methylation of GATC sequences was observed only after the overproduction of the *Eco*RV modification enzyme in vivo, but this still may be significant in respect to the homology between the *Eco*RV and the *dam* methyltransferases. These two proteins possess 27% amino acid identity, but most of the identities are located in four discrete regions of the polypeptide chain (Lauster et al., 1987). One explanation for the segmental homology is that the conserved regions are responsible for functions that are common to both methyltransferases (perhaps AdoMet binding or catalysis), while regions that lack homology are responsible for the recognition of the different DNA sequences (GATATC for *Eco*RV, GATC for *dam*). However, the observation that the *Eco*RV enzyme can methylate *dam* sites raises the possibility that DNA recognition may also be due to one or more of the conserved regions, with the difference between the two target sequences being caused by a very small number of amino acid substitutions. The recognition of DNA sequences that are shorter than the canonical site for the *Eco*RV modification enzyme, shown here in vivo, is fully consistent with data in vitro from Newman et al. (1990). Oligonucleotides containing nucleoside analogues in place of either the first adenine within GATATC or the first thymine were methylated slowly (or not at all), but equivalent substitutions at either the second adenine or the second thymine had only minor effects on the rate of methylation (Newman et al., 1990).

Proofreading in Vivo. In this study, we have confirmed the current view on the operation of a R/M system in vivo (Arber, 1979; Smith, 1979), in that the DNA of *E. coli* carrying the *Eco*RV system is fully protected by the methyltransferase at all the cognate sites for the nuclease. However, the extrapolation of the activity of the *Eco*RV nuclease at its noncognate sites measured in vitro, to conditions in vivo, indicated that this could destroy the chromosome of the cell. In vivo, the *Eco*RV methyltransferase failed to protect the majority of these noncognate sites, even when it had been overproduced in the cell. Thus the cell may possess some other system, either to protect its DNA from cleavage at noncognate sites or to repair the DNA after such reactions.

At its cognate site, each reaction of the *Eco*RV restriction enzyme normally introduces a double-strand break to the DNA: the two strands are cleaved by the dimeric protein in a coupled reaction with no detectable delay between cutting the first and the second strands (Halford & Goodall, 1988). In contrast, at a noncognate site, the *Eco*RV nuclease cleaves duplex DNA in two sequential reactions: first cutting one strand, and then freely dissociating from the nicked DNA and only subsequently cutting the second strand in a separate reaction (Taylor & Halford, 1989). Potentially, any phosphodiester bond hydrolyzed by the *Eco*RV enzyme can be resynthesized by DNA ligase, but ligase has a much higher activity at repairing nicks in DNA duplexes than at joining together two separate duplexes (Lehman, 1974). Blunt-ended DNA fragments of the type formed by *Eco*RV are joined by the ligase from *E. coli* at extremely low rates (Zimmerman & Pheiffer, 1983). Hence, by repairing the initial nicks introduced by *Eco*RV at its noncognate sites, without concomitant repair of double-strand breaks, DNA ligase could effectively proofread the activities of the *Eco*RV restriction enzyme (Taylor & Halford, 1989). This has been modeled in vitro: the addition of the ligase from *E. coli* to reactions of the *Eco*RV nuclease on pAT153 made no difference to the rate at which the product from cleaving the cognate site was formed, but the products from reactions at noncognate sites were no longer detected (Taylor & Halford, 1989). Similar schemes have been suggested for *Eco*RI (Lesser et al., 1990; Thielking et al., 1990).

In order to determine whether ligase proofreads *Eco*RV in vivo, experiments were conducted with three isogenic strains of *E. coli* that carried respectively the *lig*$^+$, *lig*4, and *lig*ts7 alleles, the latter two being temperature-sensitive mutants of DNA ligase (Gottesman et al., 1973; Konrad et al., 1973). When assayed at 30 °C, the ligase encoded by *lig*4 has 35% of the end-joining activity of the wild type, but this falls to <1% at 42 °C, though its activity measured by enzyme adenylation remains as wild type (Gottesman et al., 1973). However, the *lig*4 strain is viable at 42 °C, so its ligase activity at this temperature must be adequate for the cell (Lehman, 1974). In contrast to *lig*4, the *lig*ts7 strain is not viable at 42 °C and its DNA ligase has a more severe defect: even at 25 °C, its end-joining activity is <5% of wild type and its adenylation activity is also depressed (Konrad et al., 1973). The three strains were used in conjunction with two plasmids, pMetB and pTZ115. The former encodes the *cI*857 repressor from phage λ and also the *Eco*RV methyltransferase expressed from its natural promoter (as in pLBM): DNA in cells containing pMetB is methylated at all cognate sites for *Eco*RV. The latter expresses the *Eco*RV nuclease from the λP_L promoter, and this can be regulated by the repressor from pMetB.

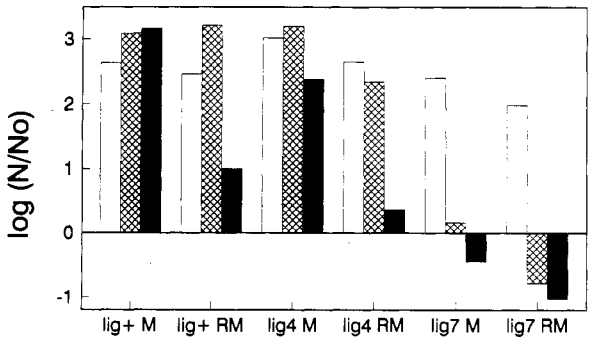

FIGURE 4: Proofreading in vivo. *E. coli* strains N1624 (noted above as *lig*$^+$), N1626 (*lig*4), and N2668 (*lig*7) were transformed either with pMetB alone (noted with each strain as M) or with pMetB and pTZ115 (noted with each strain as RM). Transformants were cultured in H-broth at 28 °C to A_{595} 0.4 and then diluted 1:200 in H-broth at either 37 °C (white columns), 39.5 °C (hatched columns), or 42 °C (black columns). After 4 h at these temperatures, viable cell titers were determined by plating triplicate aliquots on LB agar, which were then incubated at 30 °C. Viable titers were also determined at the start of 4-h periods at elevated temperatures. The values given on the ordinate are the log of the ratio of that after the 4-h incubation (N) to that before (N_0) and are the means from 2–5 repeat experiments. The standard errors about the means were ±10%.

In cells containing both pMetB and pTZ115, virtually no expression of the gene for the nuclease was seen at 37 °C and that at 39.5 °C was about 20% of the level found at 42 °C (data not shown).

The *lig*$^+$, *lig*4, and *lig*ts7 strains were each transformed with either pMetB alone or first with pMetB and then with pTZ115, and the effect of the latter plasmid on the viability of these strains was then measured at three different temperatures. (By itself, pMetB caused no change to cell viability.) At 37 °C, pTZ115 had no significant effect on the viability of any of these strains (Figure 4). At 39.5 °C, the *lig*ts7 strain transformed with pMetB alone was almost stationary, but the same strain with pTZ115 showed a net loss of viable cells (Figure 4). At 42 °C, pTZ115 caused a reduction in the number of viable cells in all three strains, but more *lig*$^+$ cells survived than *lig*4, which in turn yielded more survivors than *lig*ts7: in all three strains, fewer cells survived with both pTZ115 and pMetB than with pMetB alone (Figure 4). The variations in cell viability shown in Figure 4 were confirmed by measuring the rate at which the number of viable cells in each culture changed as a function of time, following the transfer from 28 °C to the higher temperatures. In all cases tested, the respective increases or decreases in log (N/N_0), as in Figure 4, were approximately linear with time over the 4-h period examined (data not shown). Viability experiments were also carried out in L-broth instead of H-broth, and though these yielded less cell death [see Konrad et al. (1973)], the trends between 37, 39.5, and 42 °C and between *lig*$^+$, *lig*4, and *lig*ts7 were again as in Figure 4 (data not shown).

The losses in cell viability shown in Figure 4 could have been due to either the activity of the *Eco*RV endonuclease, upon induction of pTZ115 at temperatures >37 °C, or simply the overproduction of the protein per se. The overload of the genetic machinery might be more deleterious in the *lig* mutants than in *lig*$^+$. To distinguish between these possibilities, the experiments in Figure 4 were repeated but with, instead of pTZ115, a plasmid (pEMA5/T96K) that encodes a mutant of the *Eco*RV nuclease. Upon temperature induction, this yielded a similar amount of protein to that from pTZ115, but it has 1% of the specific activity of the wild type. When tested as in Figure 4, both *lig*$^+$ and *lig*4, transformed with pMetB and pEMA5/T96K, yielded essentially the same number of viable cells as these strains transformed with pMetB alone (data not shown). Hence, by itself, protein overproduction was not lethal.

The reduced viabilities were caused by the catalytic activity of the endonuclease. Previously, loss of viability of the *lig*4 strain had been observed upon transformation with a plasmid encoding a mutant of the *Eco*RI restriction enzyme (Heitman et al., 1989), but that study had been carried out in the absence of the companion methyltransferase, so the cellular DNA could then have been cleaved at the cognate sites. However, in our experiments, the cells contained sufficient *Eco*RV methyltransferase for the full protection of all *Eco*RV recognition sites, so the reactions of the *Eco*RV nuclease could only have occurred at noncognate sites. Hence, the proofreading scheme for the *Eco*RV restriction enzyme, by DNA ligase (Taylor & Halford, 1989), operates in vivo. The effects on cell viability were not large and were observed only after the overproduction of the endonuclease. A plasmid on which *ecoRVR* is expressed from its natural promoter caused no increase in the mortality of the *lig*ts7 strain at 42 °C (data not shown). This is probably due to the fact that *E. coli* cells normally make much more ligase than is needed to maintain essential functions (Gottesman et al., 1973). *Lig*$^+$ *E. coli* cells contain enough DNA ligase for about 7500 ligations/min, yet chromosomal replication is thought to require only 200 ligations/min (Modrich et al., 1973; Lehman, 1974). We estimate that proofreading the *Eco*RV system normally requires about 3 ligations/min.

ACKNOWLEDGMENTS

We thank Bernard Connolly and Fritz Winkler for materials and extensive advice and discussion, Martin Gellert for the *lig* strains, Linda Jen-Jacobson and Alfred Pingoud for preprints, and John Grinsted, Tony Clarke, and Barry Vipond for comments on the manuscript.

Registry No. *Eco*RV restriction endonuclease, 83589-02-0; *Eco*RV methyltransferase, 91448-94-1; DNA ligase, 9015-85-4.

REFERENCES

Arber, W. (1979) *Science (Washington, D.C.) 205*, 361–365.

Barras, F., & Marinus, M. G. (1989) *Trends Genet. 5*, 139–143.

Bennett, S. P., & Halford, S. E. (1989) *Curr. Top. Cell. Regul. 30*, 57–104.

Bickle, T. A. (1987) in *Escherichia coli and Salmonella typhimurium, Cellular and Molecular Biology* (Neidhardt, F. C., Ed.) pp 692–696, American Society of Microbiology, Washington, DC.

Bougueleret, L., Schwartzstein, M., Tsugita, A., & Zabeau, M. (1984) *Nucleic Acids Res. 12*, 3659–3676.

Bougueleret, L., Tenchini, M. L., Botterman, J., & Zabeau, M. (1985) *Nucleic Acids Res. 13*, 3823–3839.

Bracco, L., Kotlarz, D., Kolb, A., Diekmann, S., & Buc, H. (1989) *EMBO J. 8*, 4289–4296.

Bradford, M. M. (1976) *Anal. Biochem. 72*, 248–254.

Brennan, C. A., Van Cleve, M. D., & Gumport, R. I. (1986) *J. Biol. Chem. 252*, 7273–7278.

D'Arcy, A., Brown, R. S., Zabeau, M., Van Resandt, R. W., & Winkler, F. K. (1985) *J. Biol. Chem. 260*, 1987–1990.

Devereaux, J., Haeberli, P., & Smithies, O. (1984) *Nucleic Acids Res. 12*, 387–395.

Drlica, K. (1987) in *Escherichia coli and Salmonella typhimurium, Cellular and Molecular Biology* (Neidhardt, F. C., Ed.) pp 91–103, American Society for Microbiology, Washington, DC.

Garnett, J., & Halford, S. E. (1988) *Gene 74*, 73–76.
Geier, G. E., & Modrich, P. (1979) *J. Biol. Chem. 254*, 1408–1413.
Gottesman, M. M., Hicks, M. L., & Gellert, M. (1973) *J. Mol. Biol. 77*, 531–547.
Halford, S. E., & Johnson, N. P. (1981) *Biochem. J. 199*, 767–777.
Halford, S. E., & Goodall, A. J. (1988) *Biochemistry 27*, 1771–1777.
Halford, S. E., Lovelady, B. M., & McCallum, S. A. (1986) *Gene 41*, 173–181.
Heitman, J., Zinder, N. D., & Model, P. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 2281–2285.
Hopwood, D. A., Bibb, M. J., Chater, K. F., Bruton, C. J., Kieser, T., Keiser, H. M., Lydiate, D. J., Smith, C. P., Ward, J. M., & Schrempf, H. (1985) *Genetic Manipulation of Streptomyces: Laboratory Manual*, John Innes Foundation, Norwich.
Konrad, E. B., Modrich, P., & Lehman, I. R. (1973) *J. Mol. Biol. 77*, 519–529.
Lauster, R., Kriebardis, A., & Guschlbauer, W. (1987) *FEBS Lett. 220*, 167–176.
Lauster, R., Trautner, T. A., & Noyer-Weidner, M. (1989) *J. Mol. Biol. 206*, 305–312.
Lehman, I. R. (1974) *Science (Washington, D.C.) 186*, 790–797.
Leirmo, S., Harrison, C., Cayley, D. S., Burgess, R. R., & Record, M. T., Jr. (1987) *Biochemistry 26*, 2095–2101.
Lesser, D. R., Kurpiewski, M. R., & Jen-Jacobson, L. (1990) *Science (Washington, D.C.)* (in press).
Luke, P. A., McCallum, S. A., & Halford, S. E. (1987) *Gene Amplif. Anal. 5*, 183–205.
Maniatis, T., Fritsch, E. E., & Sambrook, J. (1982) *Molecular Cloning, a Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.
Marinus, M. G., & Morris, N. R. (1974) *J. Mol. Biol. 85*, 309–322.
Modrich, P. (1987) *Annu. Rev. Biochem. 56*, 435–466.
Modrich, P., & Rubin, R. A. (1977) *J. Biol. Chem. 252*, 7273–7278.
Modrich, P., & Roberts, R. J. (1982) in *Nucleases* (Linn, S. M., & Roberts, R. J., Eds.) pp 109–154, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York.
Modrich, P., Anraku, Y., & Lehman, I. R. (1973) *J. Biol. Chem. 248*, 7495–7502.
Mullings, R., Evans, L. R., & Brown, N. L. (1986) *FEMS Microbiol. Lett. 37*, 237–240.
Newman, P. C., Nwosu, V. U., Williams, D. M., Cosstick, R., Seela, F., & Connolly, B. A. (1990) *Biochemistry 29*, 9891–9901.
Nwosu, V., Connolly, B. A., Halford, S. E., & Garnett, J. (1988) *Nucleic Acids Res. 16*, 3705–3720.
Pingoud, A., Urbanke, C., Alves, J., Ehbrecht, H.-J., Zabeau, M., & Gualerzi, C. (1984) *Biochemistry 23*, 5697–5703.
Posfai, J., Bhagwat, A. S., Posfai, G., & Roberts, R. J. (1989) *Nucleic Acids Res. 17*, 2421–2435.
Remaut, E., Tsao, H., & Fiers, W. (1983) *Gene 22*, 103–113.
Richey, B., Cayley, D. S., Mossing, M. C., Kolka, C., Anderson, C. F., Farrar, T. C., & Record, M. T., Jr. (1987) *J. Biol. Chem. 262*, 7157–7164.
Smith, H. O. (1979) *Science (Washington, D.C.) 205*, 455–462.
Stanssens, P., Opsomer, C., McKeown, Y. M., Kramer, W., Zabeau, M., & Fritz, H.-J. (1989) *Nucleic Acids Res. 17*, 4441–4454.
Szybalski, W., Blumenthal, R. M., Brooks, J. E., Hattman, S., & Raleigh, E. A. (1988) *Gene 74*, 279–280.
Taylor, J. D., & Halford, S. E. (1989) *Biochemistry 28*, 6198–6207.
Thielking, V., Alves, J., Fleiss, A., Maass, G., & Pingoud, A. (1990) *Biochemistry 29*, 4682–4691.
Twigg, A. J., & Sherratt, D. J. (1980) *Nature (London) 283*, 216–218.
Wilson, G. G. (1988) *Trends Genet. 4*, 314–318.
Woodbury, C. P., Downey, R. L., & von Hippel, P. H. (1980a) *J. Biol. Chem. 255*, 11526–11533.
Woodbury, C. P., Hagenbuchle, O., & von Hippel, P. H. (1980b) *J. Biol. Chem. 255*, 11534–11546.
Zimmerman, S. B., & Pheiffer, B. H. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 5852–5856.