JOURNAL OF BACTERIOLOGY, Aug. 1991, p. 5220–5223
0021-9193/91/165220-04$02.00/0

Vol. 173, No. 16

NOTES

A Novel Activity in *Escherichia coli* K-12 That Directs Restriction of DNA Modified at CG Dinucleotides

J. E. KELLEHER AND E. A. RALEIGH*
New England Biolabs, 32 Tozer Road, Beverly, Massachusetts 01915

Received 22 January 1991/Accepted 25 March 1991

The restriction systems McrA and McrB of *Escherichia coli* K-12 are known to attack DNA containing modified cytosine. In strains lacking both activities, however, we observed that DNA methylated at CG dinucleotides (as is mammalian DNA) was still significantly restricted. We show that this substantial barrier to the acceptance of 5-methylcytosine-containing DNA is attributable to a hitherto unknown activity of the Mrr restriction system. Strikingly, the multiple systems used by this gut inhabitant to determine the fate of invading DNA will all limit genetic exchange with its mammalian host(s), reinforcing the idea that one role of DNA methylation is to serve as a "molecular passport" (E. A. Raleigh, R. Trimarchi, and H. Revel, Genetics 122:279–296, 1989).

At least four different site-specific restriction systems in the enteric bacterium *Escherichia coli* K-12 recognize foreign DNA: McrA, McrBC, Mrr, and *Eco*K. DNA sensitive to one or more of these systems is biologically inactivated upon entering the cell. Only the *Eco*K endonuclease degrades unmethylated DNA (2), while the others all restrict specifically methylated substrates. Mrr is active on N^6-methyladenine (mA)-containing DNA (8); McrA and McrBC recognize cytosine modifications, including 5-methylcytosine (^{5}mC), 5-hydroxymethylcytosine, and, for McrBC, N^4-methylcytosine (14, 15). Importantly, McrA and McrBC independently inactivate substrates that have a mammalian methylation pattern.

Some observations have suggested that yet another ^{5}mC-specific restriction system might be present in various K-12 strains. Since mutations that abolish the restriction activities of *E. coli* K-12 have important practical implications for the preparation of genomic libraries, for which high tolerance of foreign DNA is required, the McrA and McrBC phenotypes of laboratory derivatives of *E. coli* K-12 have been extensively surveyed (13). In strains lacking *Eco*K, *mcrA* and *mcrBC* mutations greatly increase the acceptance of plant and mammalian DNA and limit biased representation of differently methylated clones in genomic libraries (5–7, 13, 20, 22, 23). However, when McrA$^-$ strains carry a deletion spanning the cluster of restriction genes that is designated the "immigration control region" (ICR) (14) (Fig. 1A), they are significantly more permissive than McrA$^-$ strains that are individually mutant for the ICR loci *Eco*K and McrBC (5, 7, 23). These findings have at least three possible explanations: the *mcrB1* mutation found in many *mcrB* mutant strains might be leaky, resulting in residual Mcr activity (7, 23); an unidentified Mcr restriction locus within the ICR might be encoded by the deleted DNA (7); or the Mrr locus, also encoded by the ICR, might restrict DNA containing ^{5}mC (23). The individual mutations in *mcrB* that have been tested all behave in a similar manner; they include two small deletions and the *mcrB1* allele (23). In vitro transcription-translation of the latter does not detect synthesis of any *mcrB*-encoded polypeptides (11). Leaky expression of McrB thus does not explain the observations. We hoped to distinguish between the other models.

Each of the four restriction systems of *E. coli* K-12 can be assayed independently by monitoring restriction of specific "tester" phage. Completely unmodified λ (λ.O) is sensitive to *Eco*K. λ protected by K-specific methylation (λ.K) and not otherwise modified is not restricted by *E. coli* K-12 (2). If λ.K is additionally methylated by M.*Taq*I (TCG*A), the phage becomes a target for Mrr restriction alone (9; also see Table 1), while M.*Hae*III (GG*CC) or M.*Hpa*II (C*CGG) modification confers susceptibility only to McrBC or McrA, respectively (4, 15). When expressed from a plasmid, McrB can restrict M.*Msp*I (*CCGG)-modified phage in the absence of McrC (4) (Table 1).

Mutations and clones are available for all four systems (8, 14, 17). Three are encoded by the 14-kb ICR (Fig. 1A) of the *E. coli* K-12 chromosome. The *Eco*K endonuclease, specified by the *hsdRMS* genes, is inactivated by the *hsdR2* and *hsdS3* mutations (21). Upstream of *hsdR*, the *mrr* locus is inactivated by the *mrr-2*::Tn*5* insertion (8), and flanking *hsd* on the other side (14), the McrB and McrBC activities are inactivated by the *mcrB1* mutation (4). McrA is not encoded by the ICR but by a distant locus found on the excisable element e14 (14). Spontaneous induction and loss of this element have resulted in many e14$^-$ and consequently many McrA$^-$ strains (13). Clones (14, 17) and subclones of the *mcrCB-hsdRMS-mrr* genes are indicated in Fig. 1A. Strains carrying the (*mcrC-mrr*)*114*::IS*10* deletion have lost the entire DNA segment shown (14). The *mcrA* region is represented in Fig. 1B; plasmid pJK20 was derived from pER137 (14) by deleting e14 DNA between its *Bgl*II and *Bam*HI sites. DNA manipulations were performed as previously described (4); DNA restriction endonucleases were from New England Biolabs.

The DNA of mammals is modified at CG dinucleotides although in a variable fashion that is little understood (1, 18). A similar ^{5}mC modification pattern is conferred by the

* Corresponding author.

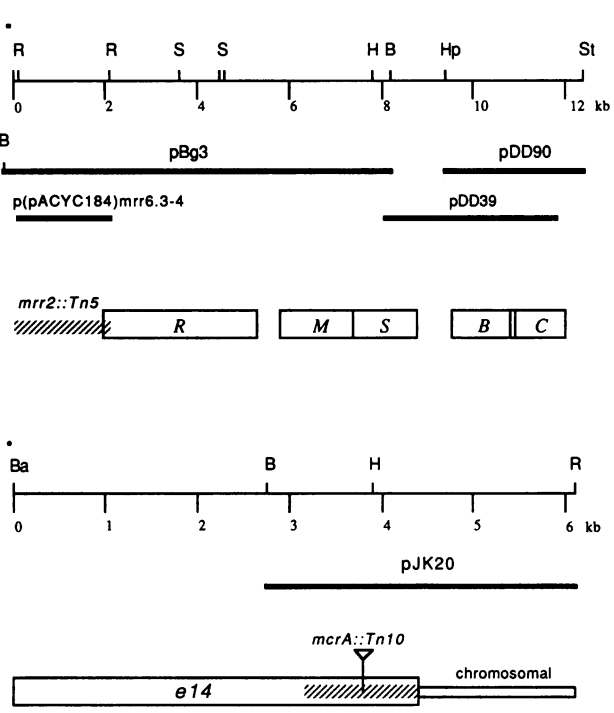

FIG. 1. (A) ICR of *E. coli* K-12. A restriction map of the ICR is shown at the top. Abbreviations for the enzymes indicated: R, *Eco*RI; S, *Sal*I; H, *Hin*dIII; B, *Bgl*II; Hp, *Hpa*I; St, *Stu*I. Directly beneath are ICR-derived clones, pBg3 (17), pDD90 (4), and pDD39 (4), and the pBg3-derived subclone p(pACYC184)mrr6.3-4 (19). Thick lines represent *E. coli* K-12 DNA contained in the plasmids. Boxes below the map show the positions of sequenced genes (4, 10, 17) indicated: *R*, *hsdR*; *M*, *hsdM*; *S*, *hsdS*; *B*, *mcrB*; *C*, *mcrC*. The hatched area marks the region within which the *mrr-2*::Tn*5* insertion is located (8). (B) McrA region of *E. coli* K-12. Representation is as described above, except that an additional restriction enzyme, *Bam*HI (Ba), is indicated. Beneath the map (3), boxes represent e14 and chromosomal DNA in the region. Hatched area marks the approximate position of *mcrA* (14) with the *mcrA*::Tn*10* insertion (3) indicated.

CG-specific bacterial methyltransferase M.*Sss*I (16); we used this methylase to mimic the expected methylation pattern of DNA of higher organisms. Unlike eukaryotic DNA, substrates methylated by M.*Sss*I are thought to be modified at all potential target sites (12). Because of the high level of modification, λ.K.*Sss*I is a particularly stringent probe for restriction directed against CG methylation.

We have used λ.K.*Sss*I to assay an isogenic series of strains carrying different combinations of mutations inactivating known *E. coli* K-12 restriction activities (Table 1). As described above, both unmethylated and site-specifically methylated phage were used to verify the phenotype of each strain (Table 1). The tester phage were λ*vir*, either unmodified (λ.O), modified in vivo (4) (λ.K and λ.K.*Msp*I), or made by in vitro methylation and subsequent in vitro packaging of λ*vir*.K DNA (λ.K.*Hpa*II, λ.K.*Hae*III, λ.K.*Taq*I, and λ.K.*Sss*I). DNA modification enzymes were obtained from New England Biolabs; in vitro packaging extracts were from Stratagene. The results shown in Table 1 each represent an average of three full-plate titers. Media, phage growth, and restriction assays were as described previously (14).

λ.K.*Sss*I is strongly restricted, plating 5,000-fold less efficiently on the restriction-proficient strain ER1370 (Table 1) than on the permissive host ER1793, which is defective for all known restriction activities. This saturating level of restriction is not significantly decreased by an *mcrA* mutation (Table 1, strain ER1644). Double mutants, retaining only one system, give intermediate levels (Table 1, strains ER1700, ER1810, ER1565, and ER1969). Since McrA restricts sites modified by M.*Hpa*II (C*CGG) (15) and all such sites are also M.*Sss*I sites, sensitivity of λ.K.*Sss*I to McrA was expected; similarly, susceptibility to McrBC was anticipated because M.*Sss*I methylation also provides some known targets for McrBC restriction, e.g., modification at *Hha*I sites (G*CGC) (15). Notably, however, strains with individual mutations in both *mcrA* and *mcrB* nevertheless retain significant mCG specific restriction (Table 1, strains ER1810 and ER1565). This is true of *mcrB1* strains, whether McrA activity is removed by e14 deletion (Table 1, strain ER1810) or by transposon insertion (Table 1, strain ER1565). Remarkably, the residual restriction of ^{5}mC-containing DNA is completely eliminated by the *mrr*-2::Tn*5* mutation, which, in accordance with previous data (8), also removes restriction of the mA-containing, Mrr-specific tester λ.K.*Taq*I (Table 1, strain ER1968). This *mrr*-dependent restriction of ^{5}mC DNA is large, reducing the plating efficiency of λ.K.*Sss*I by at least two orders of magnitude (Table 1, strains ER1810 and ER1565).

Subclones of individual restriction loci complement the restriction defects of the permissive host ER1793 in the expected manner (Table 1; plasmids shown in Fig. 1). Although the degree of restriction obtained with plasmid constructs is less than that shown by a chromosomal copy of any system, the qualitative pattern is in accordance with the data described above. Plasmids encoding McrA, McrB, McrBC, or Mrr each confer the ability to restrict λ.K.*Sss*I and restore restriction only of the substrate diagnostic for the particular system (Table 1). Even the smallest complementing subclone of the *mrr* locus (19) restores to the triple mutant strain ER1968 (*mcrA mcrB mrr*::Tn5) the capacity to restrict cytosine-methylated DNA [compare strains ER1810 and ER1968 with ER1968(p*mrr*) in Table 1]. No other gene in the ICR is required [compare strains ER1793(p*mrr*) and ER1968(p*mrr*)].

These findings prove genetically that the *mrr* locus mediates an Mcr (modified cytosine restriction) activity, hereafter named McrF, in addition to its effect on mA-containing DNA. It remains unclear how recognition of DNA containing such disparate residues as ^{5}mC and mA is achieved. McrF and Mrr activities could be properties of two distinct assemblies that share a subunit. Alternatively, the effects might be mediated by alternative recognition of a common structural feature in ^{5}mC- and mA-containing DNA; or perhaps a sequence containing both modified residues at the same time is recognized. Our results appear to argue against the last two possibilities. *mrr*-dependent restriction is not directed against all ^{5}mC or mA residues (endogenous ^{5}mC in Dcm sites and mA in Dam sites [8] are not recognized) and therefore is not simply directed against a structural feature of the modified base. McrF restriction also is not affected by either the presence or absence of Dam methylation of the substrate. λ.K.*Sss*I phage used in our experiments (Table 1) was made by in vitro methylation of Dam$^-$ Dcm$^-$ λ.K DNA and digested with *Dpn*I (G*ATC) and *Bst*UI (CGCG) endonucleases prior to packaging. In this way we ensured that the resulting phage were free of Dam methylation, since *Dpn*I cleaves Dam-modified substrates, but were modified at CG dinucleotides and thus protected against cleavage by *Bst*UI. Dam-methylated λ.K.*Sss*I behaved identically (data not

TABLE 1. Chromosomal loci, plasmid-borne genes, and plating efficiency of test phage for strains used in this study[a]

Strain	Chromosomal locus				Plasmid-borne gene(s)	Plating efficiency of test phage						
	mcrA	*mcrBC*	*hsd*	*mrr*		λ.K.*Hpa*II	λ.K.*Taq*I	λ.K.*Hae*III	λ.K.*Msp*I	λ.O	λ.K.*Sss*I	λ.K
ER1370	+	+	+	+	NA	**1×10^{-2}**	**5×10^{-2}**	**6×10^{-3}**	**7×10^{-3}**	**3×10^{-4}**	**2×10^{-4}**	1
ER1644	e14$^-$	+	+	+	NA	1	**4×10^{-2}**	**6×10^{-3}**	**7×10^{-3}**	**3×10^{-4}**	**3×10^{-4}**	1
ER1700	+	Δ	Δ	Δ	NA	**7×10^{-3}**	1	1	1	1	**4×10^{-3}**	1
ER1810	e14$^-$	*mcrB1*	*hsdS3*	+	NA	1	**3×10^{-2}**	1	1	1	**4×10^{-3}**	1
ER1565	Tn*10*	*mcrB1*	*hsdR2*	+	NA	1	**4×10^{-2}**	1	1	1	**5×10^{-3}**	1
ER1658	+	*mcrB1*	+	Tn*5*	NA	**8×10^{-3}**	1	1	0.9	**3×10^{-4}**	**5×10^{-3}**	1
ER1968	e14$^-$	*mcrB1*	+	Tn*5*	NA	1	1	1	1	**4×10^{-4}**	1	1
ER1969	e14$^-$	+	+	Tn*5*	NA	1	1	**1×10^{-2}**	**1×10^{-2}**	**4×10^{-4}**	**1×10^{-2}**	0.9
ER1793	e14$^-$	Δ	Δ	Δ	NA	(1)	(1)	(1)	(1)	(1)	(1)	(1)
ER1793(pDD90)	e14$^-$	Δ	Δ	Δ	*mcrBC*	1	1	**2×10^{-2}**	**2×10^{-2}**	0.9	**2×10^{-2}**	1
ER1793(pDD39)	e14$^-$	Δ	Δ	Δ	*mcrB*	1	1	0.9	**7×10^{-2}**	1	**9×10^{-2}**	1
ER1793(pJK20)	e14$^-$	Δ	Δ	Δ	*mcrA*	**2×10^{-3}**	1	0.9	1	1	**2×10^{-3}**	1
ER1793(pBg3)	e14$^-$	Δ	Δ	Δ	*mrr, hsdRM*	0.9	**0.3**	0.8	1	1	**2×10^{-3}**	1
ER1793(p*mrr*)	e14$^-$	Δ	Δ	Δ	*mrr*	0.9	**0.2**	0.9	1	1	**3×10^{-4}**	1
ER1968(p*mrr*)	e14$^-$	*mcrB1*	+	Tn*5*	*mrr*	0.9	**0.1**	1	0.9	**4×10^{-4}**	**2×10^{-4}**	1

[a] Plating efficiencies are calculated as the titer of phage on strain X divided by the titer of phage on permissive host ER1793. Seven different test phage were used, five diagnostic for known restriction activities: λ.K.*Hpa*II for McrA, λ.K.*Taq*I for Mrr, λ.K.*Hae*III for McrBC, λ.K.*Msp*I for McrB, and λ.O for *Eco*K. λ.K is the nonrestricted control, and λ.K.*Sss*I tests for ^{5}mC G-specific restriction. λ.K.*Hpa*II, λ.K.*Taq*I, and λ.K.*Hae*III were restricted by *mcrA*, *mrr*, and *mcrBC*, respectively. λ.K.*Msp*I was restricted by *mcrB* and *mcrBC*, and λ.O was restricted by *Eco*K. λ.K.*Sss*I was restricted by *mcrA*, *mcrB*, *mcrBC*, and *mrr*. Host strains are indicated in the first column; complete genotypes and details of constructions are available from the authors upon request. All strains were derived from ER1370 (4) differing only in their restriction loci as shown; ER1644, ER1810, ER1968, ER1969, and ER1793 are *arg*$^+$; ER1700, ER1810, ER1565, ER1658, and ER1793 are *ser*$^+$. Δ in the chromosomal locus columns indicates deletion of those loci. + indicates an active gene. Plasmid-borne genes are indicated where appropriate (NA, not applicable). Numbers in boldface represent restriction-positive combinations of plasmid, host strain, and phage. p*mrr* is p(pACYC184)mrr6.3-4 (19) (see Fig. 1A).

shown) to this Dam$^-$ tester (Table 1). This illustrates that mA and ^{5}mC recognized by Mrr and McrF activities, respectively, are unlikely to be each part of a single recognition site, since reducing the expected number of mA residues per molecule of λ.K.*Sss*I from 242 (Dam and *Eco*K modified) to 10 (*Eco*K modified) has no effect upon McrF restriction.

McrF activity is quantitatively the most significant *mrr*-mediated effect that we detect. *mrr*-dependent McrF restriction of λ.K.*Sss*I (Table 1, strains ER1810 and ER1565) is equivalent to that due to McrA (Table 1, strains ER1700 and ER1658) and McrBC (Table 1, strain ER1969), each of which markedly affects recovery of genomic clones of mammalian and plant DNA (5–7, 13, 20, 22, 23). That the triple mutant strain ER1968 (e14$^-$ *mcrB1 mrr-2*::Tn*5*) is as permissive as the e14$^-$, ICR-deleted host ER1793 further demonstrates that these three mutations effectively abolish all known restriction systems specific for modified DNA in this background.

We are grateful to Jack Benner and Phyllis Waite-Rees for kindly supplying the plasmid subclone p(pACYC184)mrr6.3-4. We thank Joe Heitman, Noreen Murray, and Ellen Sutherland for critical review of the manuscript.

REFERENCES

1. **Bestor, T. H.** 1990. DNA methylation: evolution of a bacterial immune function into a regulator of gene expression and genome structure in higher eukaryotes. Philos. Trans. R. Soc. Lond. B Biol. Sci. **326**:179–187.
2. **Bickle, T. A.** 1987. DNA restriction and modification systems, p. 692–696. *In* F. C. Neidhardt, J. L. Ingraham, K. B. Low, B. Magasanik, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology, vol. 1. American Society for Microbiology, Washington, D.C.
3. **Brody, H., A. Greener, and C. W. Hill.** 1985. Excision and reintegration of the *Escherichia coli* K-12 chromosomal element *e14*. J. Bacteriol. **161**:1112–1117.
4. **Dila, D., E. Sutherland, L. Moran, B. Slatko, and E. A. Raleigh.** 1990. Genetic and sequence organization of the *mcrBC* locus of *Escherichia coli* K-12. J. Bacteriol. **172**:4888–4900.
5. **Gossen, J. A., and J. Vijg.** 1988. *E. coli* C: a convenient host strain for rescue of highly methylated DNA. Nucleic Acids Res. **16**:9343.
6. **Graham, M. W., J. P. Doherty, and D. M. Woodcock.** 1990. Efficient construction of plant genomic libraries requires the use of *mcr*− host strains and packaging mixes. Plant Mol. Biol. Rep. **8**:33–42.
7. **Grant, S. G. N., J. Jessee, F. R. Bloom, and D. Hanahan.** 1990. Differential plasmid rescue from transgenic mouse DNAs into *Escherichia coli* methylation-restriction mutants. Proc. Natl. Acad. Sci. USA **87**:4645–4649.
8. **Heitman, J., and P. Model.** 1987. Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. J. Bacteriol. **169**:3243–3250.
9. **Kelleher, J. E., and E. A. Raleigh.** Unpublished data.
10. **Loenen, W. A. M., A. S. Daniel, H. D. Braymer, and N. E. Murray.** 1987. Organization and sequence of the *hsd* genes of *Escherichia coli* K-12. J. Mol. Biol. **198**:159–170.
11. **Noyer-Weidner, M.** Personal communication.
12. **Nur, I., M. Szyf, A. Razin, G. Glaser, S. Rottem, and S. Razin.** 1985. Procaryotic and eucaryotic traits of DNA methylation in spiroplasmas (mycoplasmas). J. Bacteriol. **164**:19–24.
13. **Raleigh, E. A., N. E. Murray, H. Revel, R. M. Blumenthal, D. Westaway, A. D. Reith, P. W. J. Rigby, J. Elhai, and D. Hanahan.** 1988. McrA and McrB restriction phenotypes of some *E. coli* strains and implications for gene cloning. Nucleic Acids Res. **16**:1563–1575.
14. **Raleigh, E. A., R. Trimarchi, and H. Revel.** 1989. Genetic and physical mapping of the *mcrA* (*rglA*) and *mcrB* (*rglB*) loci of *Escherichia coli* K-12. Genetics **122**:279–296.
15. **Raleigh, E. A., and G. Wilson.** 1986. *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. Proc. Natl. Acad. Sci. USA **83**:9070–9074.
16. **Renbaum, P., D. Abrahamove, A. Fainsod, G. G. Wilson, S. Rottem, and A. Razin.** 1990. Cloning, characterization, and expression in *Escherichia coli* of the gene coding for the CpG DNA methylase from *Spiroplasma sp.* strain MQ1(M.*Sss*I). Nucleic Acids Res. **18**:1145–1151.
17. **Sain, B., and N. E. Murray.** 1980. The *hsd* (host specificity) genes of *E. coli* K-12. Mol. Gen. Genet. **180**:35–46.

18. **Selker, E. U.** 1990. DNA methylation and chromatin structure: a view from below. Trends Biochem. Sci. **15**:103–107.
19. **Waite-Rees, P. A., C. J. Keating, L. S. Moran, B. E. Slatko, L. J. Hornstra, and J. S. Benner.** 1991. Characterization and expression of the *Escherichia coli* Mrr restriction system. J. Bacteriol. **173**:5207–5219.
20. **Whittaker, P. A., A. J. B. Campbell, E. M. Southern, and N. E. Murray.** 1988. Enhanced recovery and restriction mapping of DNA fragments cloned in a new λ vector. Nucleic Acids Res. **16**:6725–6736.
21. **Wood, W. B.** 1966. Host specificity of DNA produced by *E. coli*: bacterial mutations affecting the restriction and modification of DNA. J. Mol. Biol. **16**:118–133.
22. **Woodcock, D. M., P. J. Crowther, W. P. Diver, M. Graham, C. Bateman, D. J. Baker, and S. S. Smith.** 1988. RglB facilitated cloning of highly methylated eukaryotic DNA: the human L1 transposon, plant DNA, and DNA methylated *in vitro* with human DNA methyltransferase. Nucleic Acids Res. **16**:4465–4482.
23. **Woodcock, D. M., P. J. Crowther, J. Doherty, S. Jefferson, E. DeCruz, M. Noyer-Weidner, S. S. Smith, M. Z. Michael, and M. W. Graham.** 1989. Quantitative evaluation of *Escherichia coli* host strains for tolerance to cytosine methylation in plasmid and phage recombinants. Nucleic Acids Res. **17**:3469–3478.