MICROBIOLOGICAL REVIEWS, June 1993, p. 434–450
0146-0749/93/020434-17$02.00/0
Copyright © 1993, American Society for Microbiology
Vol. 57, No. 2

Biology of DNA Restriction

THOMAS A. BICKLE[1]* AND DETLEV H. KRÜGER[2]

Department of Microbiology, Biozentrum, Basel University, Klingelbergstrasse 70, CH-4056 Basel, Switzerland,[1] and Institute of Virology, Charité School of Medicine, Humboldt University, D-O-1040 Berlin, Germany[2]

INTRODUCTION 434
TYPE I R-M SYSTEMS 434
Type I Systems Form Families of Related Enzymes 435
Structure of *hsd* Genes 435
Evolution of DNA Sequence Recognition by Recombination between *hsdS* Genes 436
Mutations Affecting Modification Activity 437
TYPE II R-M SYSTEMS 437
Evolutionary Aspects 437
Control of Expression of Type II RM Genes 438
Cytosine Can Be Methylated on Either C-5 Or N^4: Consequences for Mutagenesis 438
Type II Restriction Endonucleases That Require Two Recognition Sites for Cleavage 439
What Is the Function of Type IIS Enzymes? 440
TYPE III R-M SYSTEMS 440
Genetics of Type III Systems 440
DNA Recognition by Type III Enzymes: Different Sequence Requirements for Restriction and Modification 441
*Sty*LTI System 442
RESTRICTION SYSTEMS SPECIFIC FOR MODIFIED DNA 442
*Dpn*I and *Dpn*II 442
Rediscovery of Methyl-Directed Restriction in *E. coli* 442
McrBC 442
McrA 443
Mrr 443
PHAGE ANTIRESTRICTION 443
Recent Developments 443
"Nature Red in Tooth and Claw": The Case of Phage T4 444
CONCLUDING REMARKS 444
ACKNOWLEDGMENTS 445
REFERENCES 445

INTRODUCTION

The phenomenon of restriction and modification (R-M) of bacterial viruses was first described some 40 years ago (15, 125) and received a molecular explanation 10 years later: restriction was due to an endonucleolytic cleavage of foreign DNA (2, 3, 51). We now know that there are two different types of restriction systems: those in which cellular DNA is protected from restriction by modification methylation of adenosyl or cytosyl residues within the sequences recognized by the restriction enzymes (the classical R-M systems) and those which cleave only DNA that carries specific modifications. A large body of knowledge concerning DNA restriction has accumulated and has been the subject of many excellent reviews (5, 18, 19, 24, 57, 100, 102, 120, 124, 131–133, 157, 168, 169, 214, 219).

The purpose of this review is to describe the exciting developments in our understanding of the biology of DNA restriction during the last 10 years. We will begin by discussing recent findings with the classical type I, II, and III R-M systems. We will continue with the rediscovery of a group of restriction systems that specifically cleave DNA carrying certain modifications; host cell DNA in this case avoids restriction because it is not modified. Finally, we will discuss the ways in which bacteriophages avoid the effects of the R-M systems of their hosts. The major characteristics of the different kinds of R-M systems are set out in Table 1.

TYPE I R-M SYSTEMS

The type I R-M systems are the most complex so far discovered. The main enzyme is a three-subunit protein containing the products of the *hsdS*, *hsdM*, and *hsdR* genes (*hsd* for host specificity for DNA). This enzyme is multifunctional, catalyzing both restriction and modification reactions, and is also a restriction-dependent ATPase and a DNA topoisomerase (reviewed in reference 19). A second enzyme containing the products of the *hsdS* and *hsdM* genes can also be isolated. This enzyme is a monofunctional modification methylase, which may have no physiological role under normal conditions since the three-subunit enzyme is also an efficient modification methylase. Type I systems have so far only been found in members of the family *Enterobacteriaceae*, with the exception of an early report of a member in *Haemophilus influenzae* (65), which may actually be a type

* Corresponding author. Electronic mail address: bickle@urz.unibas.ch.

TABLE 1. Distinguishing features of different types of restriction enzymes

Type	Structural features	Cofactors	DNA cleavage
I	Multifunctional, multisubunit, three structural genes	Mg^{2+}, ATP, AdoMet	Cut randomly, far from asymmetrical recognition sites
II	Simple, separate enzymes for restriction and modification, two structural genes	Mg^{2+}	Cut at fixed positions within symmetrical recognition sites
IIS	Like type II (three structural genes in one case)	Mg^{2+}	Cut at fixed distance from asymmetrical recognition sites
III	Multisubunit, multifunctional, two structural genes	Mg^{2+}, ATP	As type IIS
Mcr/Mrr	Several subunits and genes, no modification enzymes	Mg^{2+}, GTP (McrBC)	Cut DNA with modified C and A residues

III system (the distinction between types I and III was first made in 1978 [88]).

Type I Systems Form Families of Related Enzymes

The major finding of the last few years has been that the enterobacterial type I systems are grouped in genetically related families, a fact that has important consequences for the understanding of the biology of these systems. Early studies had showed that many of the chromosomally coded type I R-M systems could genetically complement each other; indeed, complementation was the tool that allowed the number and function of the *hsd* genes to be defined (25). By chance, all of the R-M systems from *Escherichia coli* and *Salmonella* spp. that were investigated in the early days were from the same family and thus gave genetic complementation. The first hint that different families might exist came from a comparison of genes for different systems by DNA hybridization and from an immunochemical comparison of different enzymes by using hybridization probes and antisera specific for the *Eco*KI enzyme of *E. coli* K-12 (135). It was found that for all except one of the systems investigated there was considerable homology among the structural genes and that two of the three subunits of the enzymes were immunologically cross-reactive (no antisera were available for the third subunit). It was proposed that *Eco*AI, the exceptional nonhomologous enzyme, should be considered the first member of a new family of type I R-M systems (135).

What are the reasons for considering *Eco*AI to be a type I system? First, it is allelic to the other type I systems in the sense that the structural genes have the same chromosomal location (6, 40); second, the enzymes have a typical type I subunit structure, genetic organization, cofactor requirements, reaction mechanism, and recognition sequence (59, 190, 192). Finally, *Eco*AI is the first member of a new family, since two other members were identified later: *Eco*EI and *Cfr*II (40, 59).

A third family of type I systems that has no homology with the other two families and which will not complement them has since been found (63, 150, 155, 179). The first member of this family to be investigated was *Sty*R124I, coded by a *Salmonella* conjugative plasmid. Most members of the family are plasmid coded (for a possibly trivial exception, see reference 119), and thus they cannot be said to be allelic to the other two families. However, in all other features they are typical type I systems. It has been proposed that the three type I families be called type IA with *Eco*KI as the prototype, type IB (the *Eco*AI family), and type IC (*Sty*R124I and relatives) (19). All of the characterized type I R-M systems, together with their recognition sequences, are shown in Table 2.

Structure of *hsd* Genes

The structural genes for all three families of type I systems have now been characterized at the molecular level by DNA sequencing or DNA hybridization analysis (38, 40, 59, 60, 62, 64, 67, 86, 119, 122, 135, 154, 174). Several generalities emerge from a comparison of these sequences. First, all of the genes are arranged into two contiguous transcription units, with *hsdM* and *hsdS* forming an operon and *hsdR* being transcribed from its own promoter; the order of the two transcriptional units is different for different families (Fig. 1). Second, with one important exception (see below), there is no sequence homology between different families, apart from a few short amino acid sequence motifs common to DNA adenine methylases and to ATP-binding proteins. Third, within a given family the *hsdM* and *hsdR* genes are largely homologous throughout their length. Fourth, the *hsdS* genes contain two regions of high homology (three for the type IB family [86]) separating two extensive regions of nonhomology. There is good evidence (see the next section) for all three families that the regions of nonhomology code two protein domains, each of which recognizes one-half of

TABLE 2. Specificities of type I restriction enzymes

Enzyme	Recognition sequence[a]	Reference(s)
Type IA		
*Eco*BI	TGA(N_8)TGCT	115, 164, 183
*Eco*KI	AAC(N_6)GTGC	85
*Eco*DI	TTA(N_7)GTCY	138
*Sty*SBI	GAG(N_6)RTAYG	137
*Sty*SPI	AAC(N_6)GTRC	137
*Sty*SQI	AAC(N_6)RTAYG	136
*Sty*SJI[b]	GAG(N_6)GTRC	62
Type IB		
*Eco*AI	GAG(N_7)GTCA	97, 192
*Eco*EI	GAG(N_7)ATGC	39
*Cfr*AI	GCA(N_8)GTGG	86
Type IC		
*Eco*DXXI	TCA(N_7)RTTC	67, 149
*Sty*R124I	GAA(N_6)RTCG	156
*Sty*R124/3I	GAA(N_7)RTCG	156
*Eco*RD2[b]	GAA(N_6)RTTC	67
*Eco*RD3[b]	GAA(N_7)RTTC	67
*Eco*DR2[b]	TCA(N_6)RTCG	67
*Eco*DR3[b]	TCA(N_7)RTCG	67

[a] N, any nucleotide; R, either purine; Y, either pyrimidine. The subscript number within the parentheses indicates the number of residues in the nonspecific spacer.

[b] These R-M systems are artificial hybrids generated in the laboratory.

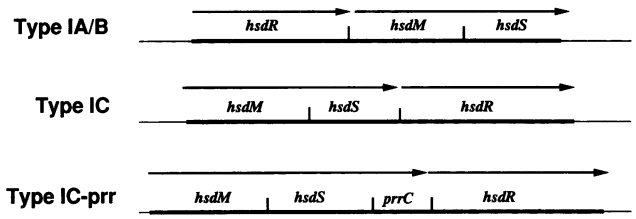

FIG. 1. Genetic organization of the *hsd* region of the different families of type I R-M systems. The arrows delineate the transcriptional units. The genes *hsdS*, *hsdM*, and *hsdR* encode the three subunits of type I restriction enzymes. The *prrC* gene encodes an RNA-based restriction system (see text). Not drawn to scale.

the split recognition sequences that are characteristic of type I R-M systems (Table 2). The conserved regions of the *hsdS* genes are thought to be responsible for protein-protein interactions with the *hsdM* and *hsdR* gene products, although there is no hard evidence for this. For the type IC family (see below), there is evidence that the central conserved region of the *hsdS* gene products also determines, in part, the DNA recognition properties of the enzymes.

Evolution of DNA Sequence Recognition by Recombination between *hsdS* Genes

Type I R-M systems are the only examples known to date of proteins that recognize specific DNA sequences and that can change their sequence specificities by well-understood natural processes. The first of these to be discovered was in the type IA family, when an attempt to transduce the structural genes for *Sty*SPI into cells containing the *Sty*SBI genes resulted in a new specificity, which was called *Sty*SQI (30). It was then shown, first by electron-microscopic heteroduplex analysis and later by DNA sequencing, that *Sty*SQI was the result of a recombination event within the central conserved region of the *hsdS* genes of *Sty*SPI and *Sty*SBI (58, 60). The DNA sequences recognized by the two parents and by the recombinant are shown in Table 2 (136, 137). It is clear that the *Sty*SQI recognition sequence is a hybrid between those of the two parents: the 5′ end of the sequence is the same as the 5′ half of the *Sty*SPI sequence, and the 3′ end of the sequence is identical to that of the 3′ half of the *Sty*SBI sequence.

The picture that emerges from these results is that the recombination event that created the *Sty*SQI *hsdS* gene reassorted two protein domains, each of which recognizes one-half of the split recognition sequence. This idea is strengthened by the following results. If the model is correct, the reciprocal recombinant, that is, with 5′ sequences from the *Sty*SBI *hsdS* gene and 3′ sequences from the *Sty*SPI *hsdS* gene, should have a predictable hybrid recognition sequence. This recombinant was made and proved to have the predicted recognition sequence (*Sty*SJI in Table 2) (62). Strong support for the hypothesis also comes from studies with the type IB enzyme *Eco*EI (Table 2). This enzyme recognizes GAG as the 5′ part of its recognition sequence, as do *Eco*AI itself and the type IA enzyme *Sty*SBI (Table 2). The region of the *hsdS* genes corresponding to the 5′ variable domains of *Eco*EI and *Sty*SBI shows considerable homology—about 50% on the amino acid sequence level—even though they belong to different families (38). The conclusion is inescapable that a conserved DNA-binding domain that recognizes GAG is present in both families. In principle, there are three possible explanations for this result: (i) the sequences were present in an ancestor common to both families; (ii) they represent the result of convergent evolution; or (iii) the sequence has passed from one family to the other by horizontal gene transfer.

The second example of spontaneous change in DNA specificity concerns the two type IC enzymes *Sty*R124I and *Sty*R124/3I. *Sty*R124I is coded by an IncFIV conjugative drug resistance plasmid (74). At a low frequency, the *Sty*R124I specificity is lost and is replaced by a new specificity, called *Sty*R124/3I. Cells expressing *Sty*R124/3I can switch back to expressing *Sty*R124I, with a concomitant loss of the *Sty*R124/3I specificity. The system thus switches back and forth between two alternative specificities (63).

The DNA sequences recognized by the two systems are shown in Table 2. They differ only in the length of the nonspecific spacer separating the two half-sites of the sequence: *Sty*R124I requires a 6-bp spacer, whereas *Sty*R124/3I requires a 7-bp spacer. In terms of enzymatic recognition of DNA, this is not a trivial difference. The extra base pair in the *Sty*R124/3I spacer moves the two half-sites 3.4 Å (0.34 nm) further apart and rotates them by 36° with respect to the *Sty*R124I sequence. The difference between the two enzymes was determined by sequencing the structural genes. The two sequences were identical except at one position in the middle of the *hsdS* genes, where a 12-bp in-frame sequence was directly repeated twice in the *Sty*R124I gene and three times in the *Sty*R124/3I gene (154). The mechanism of the switch in specificity was immediately clear: unequal crossing over between the two repeats of the *Sty*R124I *hsdS* gene would generate one copy of the gene with three repeats and a second copy with one repeat. A similar mechanism could reduce the number of repeats from three back to two again. This result led to the hypothesis that the region of the protein that contains the repeated sequences forms a spacer separating two DNA-binding domains, each of which binds to one-half of the recognition sequence. The extra four amino acids in the spacer of the *Sty*R124/3I protein would increase the physical distance between the DNA-binding domains sufficiently (6 Å [0.6 nm] if α-helical) that they could no longer bind a recognition sequence with 6 bp separating the two halves but could make productive interactions with a sequence containing a 7-bp spacer.

This hypothesis gained support from results of mutagenesis experiments examining the region of the repeats (66). Single-amino-acid replacements within the repeats had no effect on either activity or specificity, indicating that the exact amino acid sequence in this region of the protein is not crucial for its function. Changes in the length of the repeated region had profound effects on both activity and specificity. For example, variants with one or four copies of the repeated sequences were very inefficient restriction enzymes, 10^{-5} to 10^{-6} times less effective than the wild-type enzymes, which explains why these variants were not found naturally. They were, on the other hand, efficient modification enzymes: the variant with one repeat methylated the *Sty*R124I sequence, and the enzyme with four repeats methylated both *Sty*R124I and *Sty*R124/3I sequences, as though its long protein spacer was flexible enough to be floppy. These differences between restriction and modification are interesting and may indicate that the region of the spacer serves a double function, determining the length of the nonspecific spacer in the recognition sequence on the one hand and mediating protein-protein interactions between HsdS and HsdR on the other.

Another enzyme in the type IC family is *Eco*DXXI, which

recognizes a DNA sequence completely different from that recognized by *Sty*R124I or *Sty*R124/3I (Table 2). DNA heteroduplex analysis of the structural genes of *Eco*DXXI and *Sty*R124I showed that the overall structure was the same as in the type IA family: homology throughout the region except within the *hsdS* genes, where two long blocks of nonhomology were found. A limited DNA sequence analysis of the central region of the *hsdS* gene showed that *Eco*DXXI was very similar to *Sty*R124/3I; in particular, the three copies of the 12-bp repeated sequence were identical (67). All possible hybrids between the *Sty*R124I and *Eco*DXXI *hsdS* genes were constructed with one-half of each hybrid derived from one system, one-half from the other system, and either two or three copies of the 12-bp repeat in the middle. All of these hybrids were active in both restriction and modification, and the sequences that they recognized were determined. The results clearly demonstrate the correlation between protein structure and target site specificity. As with the type IA systems, the amino terminus of the protein recognizes the trinucleotide half of the recognition sequence, the central part of the protein determines the number of nucleotides in the nonspecific spacer, and the carboxyl terminus recognizes the other half of the recognition sequence (Table 2) (67).

Mutations Affecting Modification Activity

A crucial difference between eukaryotic and all prokaryotic DNA methyltransferases but one concerns the DNA substrate. The mammalian enzymes methylate the cytosine in mainly CG sequences to 5-methylcytosine (5-meC), but they do it efficiently only if the cytosine in the opposite strand already bears a methyl residue. The result is that CG sequences that are methylated perpetuate their methylated state following DNA replication whereas nonmethylated CG sequences normally remain nonmethylated. Methylation is correlated with gene inactivation, and it is thought that DNA methylation is one of the primary control processes in development: as yet undetected enzymes would inactivate a gene by de novo methylation of CG sequences, and the maintenance methylases, which require a hemimethylated substrate, would then ensure that the gene remained inactivated during further cell divisions (reviewed in reference 170). Most prokaryotic DNA methylases behave quite differently. They show no marked preference for hemimethylated over nonmethylated DNA and therefore have the properties expected for the postulated de novo eukaryotic enzyme. The type IA enzymes are an exception to this rule in that both the modification enzymes and the multifunctional restriction enzymes are bona fide maintenance methylases, methylating hemimethylated substrates much more efficiently than they methylate nonmethylated ones (31, 191, 205). Mutations in the type IA system *Eco*K that give it the properties expected of a de novo methylase have now been selected (90). The mutations are all loosely clustered within the *hsdM* gene, indicating that it is the HsdM polypeptide that senses the methylation state of the recognition site. Thus, elements of both HsdS and HsdM must interact directly with the recognition site, and it will be interesting to determine whether these interactions are simultaneous or sequential. Some of these mutations are also defective in restriction, but the molecular basis for the deficiency is not yet known (90).

TYPE II R-M SYSTEMS

The type II systems are the simplest of the R-M systems, with separate restriction and modification enzymes and no special cofactor requirements other than Mg^{2+} (Table 1). The fact that the restriction enzymes recognize and cleave simple sequences in DNA at fixed positions has made them indispensable in many if not most branches of molecular biology. These useful properties of type II enzymes have stimulated many laboratories to screen thousands of taxonomically diverse bacteria for enzymes with new specificities. As a result, we know of hundreds of enzymes, most of which, apart from their DNA recognition properties, have not been further characterized.

Evolutionary Aspects

So far, the structural genes from over 50 type II R-M systems have been sequenced (213). Surprisingly, no significant amino acid sequence homologies could be found between cognate pairs of endonucleases and methyltransferases. Although they recognize the same DNA sequence, even their target recognition domains (TRDs) are different (35, 114). Therefore, for type II enzymes, the endonucleases and methyltransferases are the products of independent evolution. One can envisage a scenario in which the methyltransferases first evolved primarily as enzymes, like the *dam* methylase of *E. coli* (reviewed in reference 12), with important functions in the regulation of gene expression, DNA repair, or replication, and then provided a genetic background with a methylated genome in which restriction endonucleases could evolve.

Not only do endonucleases have no homology with their cognate methyltransferases, but also there are no (with a few exceptions) homologies between different type II restriction endonucleases (35, 114). Even isoschizomers, that is, endonucleases from different bacterial species that recognize the same sequence and cut it at the same position, such as *Bsu*RI and *Ngo*PII, are very different (92, 189). It can be concluded that endonucleases have not evolved from a single primitive precursor and have acquired different sequence specificities by mutating the TRD. They must be of different evolutionary origin. The few cases in which more or less marked homology between isoschizomeric enzymes has been described are the enzyme pairs *Eco*RI and *Rsr*I (185), *Bsu*BI and *Pst*I (217), *Bsu*FI and *Msp*I (87), *Tth*HB8I and *Taq*I (11), *Fnu*DI and *Ngo*PII, and *Cfr*9I and *Xma*I (214). However, it should be pointed out that *Eco*RI is coded by a conjugative plasmid (218) and that the *Eco*RI-*Rsr*I pair may therefore be the result of horizontal gene transfer.

In contrast to the endonucleases, methyltransferases share many amino acid sequence similarities. This is especially true for the family of 5-meC transferases, which includes one eukaryotic enzyme (16, 112, 114, 152). Several groups have described a common building plan for these enzymes, with a pattern of highly conserved amino acid sequence motifs alternating with variable sequence elements (94, 114, 152). The TRD is part of the largest variable region, which is always located close to the center of the primary sequence (94, 114, 207, 211).

Studies of the 5-meC methyltransferases coded by the *Bacillus* phages SPβ, SPR, ϕ3T, $\rho 11_s$, and H2, which methylate several different DNA sequences, have contributed much to our understanding of the functional domains in these enzymes. Each DNA sequence is recognized by a different TRD. These TRDs have been identified, and the

feasibility of engineering new combinations to create enzymes with novel specificities has been demonstrated (9, 13, 111, 201, 202, 211). In addition to functional TDRs, the phages H2 and $\rho 11_s$ have silent TRDs in their methyltransferases which can be experimentally activated by mutation (110). It is not understood how such silent domains are maintained, how multispecific methyltransferases originated evolutionarily, or whether modules, perhaps even complete methyltransferase genes, can be exchanged by horizontal gene transfer between phage and cellular genomes (exchange of complete genes between phages has been demonstrated [143, 198]). It is thought that it is the presence of restriction systems in their hosts that selects for the methyltransferase genes and drives their evolution (110), even though not all of the required restriction specificities have so far been demonstrated in *Bacillus* strains.

Homologous segments are also found in the family of the N^6meA methyltransferases, and consensus motifs have been defined. In general, the homologies are less extensive than for the 5-meC methylases and the degree of conservation is not as high (35, 68, 112, 113, 139, 182). The N^4meC methyltransferases are (with the possible exception of *Mva*I) more similar to the N^6meA methyltransferases than to the 5-meC methyltransferases (27, 95, 197). This is perhaps not surprising since methylation of an exocyclic amino group (the N^4 of cytosine or the N^6 of adenosine) is mechanistically very different from the methylation of a ring carbon to 5-meC. Lauster (112) has developed a model for the evolution of type II methyltransferases by gene duplication and subsequent divergent evolution of the duplicated segments.

Control of Expression of Type II RM Genes

The expression of genes coding for methyltransferases and endonucleases must be tightly regulated. It is essential that the methyltransferase completely modify the cellular genome at all times to protect it from the lethal action of the endonuclease. This regulation is especially important when R-M genes first enter a cell, on conjugative plasmids for example, but it can also be important during changes in the physiology of the cell (e.g., entry into the stationary phase, starvation associated with a deficiency of methyl donors, and other stresses).

For at least some type II R-M systems, this regulation seems to be achieved by the product of an additional open reading frame (ORF) tightly linked to the R-M gene cluster. Tao et al. (195, 196) have sequenced this ORF in the *Pvu*II gene cluster and shown that its mutation leads to a restrictionless phenotype even though the structural gene for the endonuclease is intact. The defect can be complemented by a wild-type copy of the ORF in *trans*. They have proposed that the proteins coded by these ORFs should be known as C proteins (e.g., *Pvu*IIC), where C stands for controller. Brooks and coworkers (27, 140) have similarly shown that an ORF in the *Bam*HI system has regulatory functions. Disruption of the *Bam*HIC gene leads to both a decrease in restriction and an increase in modification methylase activity (82). The predicted amino acid sequences of these C proteins have some homology with DNA-binding proteins of the helix-turn-helix class (72), indicating that they may function as transcriptional repressors or activators.

Of the type II R-M systems for which the gene organization and sequence are known, nine systems have an ORF in addition to the R and M structural genes (213). Five of these systems were compared with each other, and for four of them (*Bam*HI, *Eco*RV, *Pvu*II, and *Sma*I) the ORFs were remarkably similar, especially in the region of the putative helix-turn-helix motif (195, 196). Since there is no homology between the corresponding four endonucleases and the homology between the four methyltransferases (all being either N^4meC or N^6meA methyltransferases) is also lower than for the C proteins, Tao et al. (196) concluded that the evolution of the C gene family could have been independent of the R and M genes. It is interesting that the *Bsu*RI ORF that has no homology with the other four is from an R-M system that specifies 5-meC (196). Other possible regulatory mechanisms for the expression of type II R-M systems have been proposed (75, 180, 199, 212).

One problem with studies on regulation is that they are almost always done with cloned genes examined in heterologous hosts, in which regulation need not necessarily be the same as in the original strain. Although this was probably not critical for the study with the *Pvu*II gene cluster in *E. coli*, because *Pvu*II originally comes from *Proteus vulgaris*, which is also a member of the *Enterobacteriacae*, it was definitely a factor in the work with the *Bam*HI genes. *Bam*HI comes from the gram-positive sporeformer *Bacillus amyloliquifaciens*, a close relative of *B. subtilis*. It was found that the regulation in the gram-negative *E. coli* was distinctly different from that in *B. subtilis* (27).

Cytosine Can Be Methylated on Either C-5 Or N^4: Consequences for Mutagenesis

All the early examples of DNA modification that had been studied were due either to methylation of the extracyclic amino group of adenosine to N^6meA or to methylation of a ring carbon of cytosine to 5-meC. In 1983 Janulaitis et al. discovered that the *Bcn*I methylase from *Bacillus centrosporus* methylated the extracyclic amino group of cytosine to N^4meC (84). N^4meC has since been shown to be the product of many modification methylases (32, 33, 54, 56).

It was proposed that N^4meC rather than 5-meC was an adaptation to thermophilia, which would avoid the hypermutability associated with 5-meC (54). Deamination of 5-meC is a highly temperature-dependent reaction which produces thymine, resulting in G · C-to-A · T transition mutations on DNA replication. N^4meC (as well as C) also undergoes spontaneous deamination, but the product is uridine, which is subject to correction by uridine-specific DNA repair pathways (49): 5-meC has long been known as a hot spot for spontaneous mutation, even at mesophilic temperatures (37, 49, 208). An initial screening of thermophilic bacteria showed that about half of the strains tested contained N^4meC in their DNA (54). However, by using improved methods the same authors could also demonstrate the presence of N^4meC in the DNA of many mesophilic species (56), as has also been noted by others (32).

The recognition sequences for common 5-meC methyltransferases in the genomes of prokaryotes and eukaryotes and some of their viruses (for example, CCWGG in *E. coli* or CpG in mammalian cells) are often statistically underrepresented (55, 104, 175). This is generally explained by the fact that 5-meC is a hot spot for mutation, leading, when the mutations are phenotypically neutral, to loss of the methylatable sequences during evolution. However, it is also possible that some methylatable sequences were counterselected to avoid the consequences of their potential methylation on regulatory processes (104); in some phages the recognition sites for adenine methylases are also underrepresented, even though N^6meA is not thought to be mutagenic (127, 175).

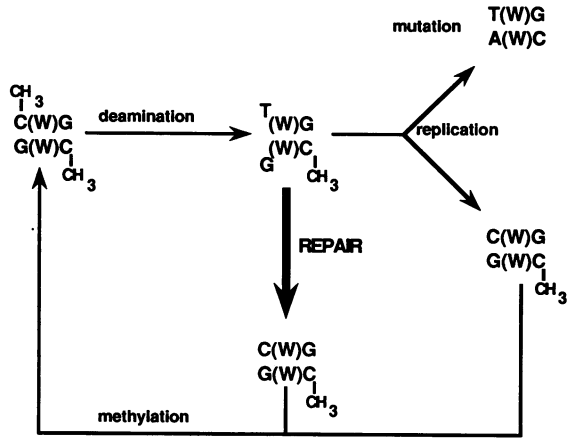

FIG. 2. Deamination of 5-meC in DNA is mutagenic if the lesion is not repaired. The specific example shown is part of the *E. coli* Dcm methylase site. Dcm methylates the inner C in the sequence CCWGG, where the symbol W denotes either A or T.

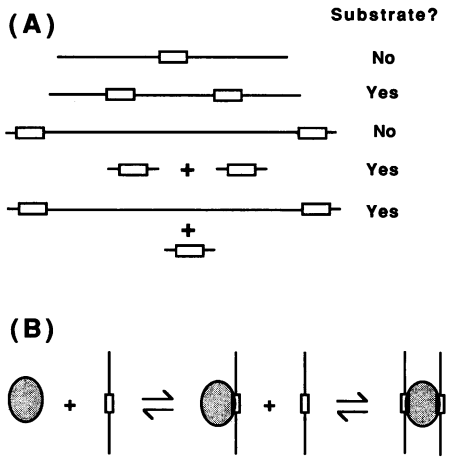

FIG. 3. DNA cleavage by *Eco*RII. Solid lines represent DNA molecules, open boxes represent *Eco*RII sites, and stippled ovals represent *Eco*RII enzyme molecules. (A) Single *Eco*RII sites in long DNA molecules or sites separated by large distances are not cleaved. Cleavage of such molecules can be stimulated by the addition of high concentrations of sites on small DNA molecules. (B) A possible mechanism for cleavage by *Eco*RII. See the text for details.

The magnitude of the mutation burden due to the deamination of 5-meC will depend on the efficiency with which the T · G mispair in DNA is detected and repaired by using the G-containing strand as the template (Fig. 2). Repair systems with precisely this specificity have been found in both prokaryotic (77, 117, 118, 182a, 220) and eukaryotic (28, 210) cells. In *E. coli* the repair pathway is known as very short patch repair and is due to the product of the *vsp* gene, which is located just downstream of the *dcm* methylase gene (182a). Recent results from the laboratory of G. Wilson (quoted in reference 143a) indicate that ORFs with homology to *vsp* are found in the vicinity of the genes encoding the *Alu*I, *Bsu*RI, and *Hpa*II 5-meC transferases, suggesting that very short patch-type repair may be widespread.

Type II Restriction Endonucleases That Require Two Recognition Sites for Cleavage

*Eco*RII was the first type II restriction endonuclease that was shown to require at least two recognition sites for endonucleolytic cleavage (99). The enzyme must bind to an activator site (which does not necessarily have to be cleaved or even be cleavable) before it will cut a cleavage site. The maximum distance between two sites in a DNA molecule which allows efficient cleavage by *Eco*RII is about 1 kb (146).

Figure 3A shows the cleavage susceptibility of *Eco*RII sites located on DNA molecules of different lengths and separated by different distances. Figure 3B shows an interpretation of these results in terms of a kinetic model. The enzyme first binds reversibly to a single site on a DNA molecule. Cleavage can occur only if a second specific DNA-binding site on the enzyme is occupied before the enzyme dissociates from the first DNA molecule. The probability of having both specific binding sites simultaneously occupied will depend both on the concentration of *Eco*RII sites and on the size of the DNA molecules carrying those sites: rapidly diffusing oligonucleotides will be more effective at eliciting cleavage than will slowly diffusing longer DNA molecules. Very long DNA molecules diffuse so slowly that the only possibility for cleavage is for the enzyme to simultaneously bind to two sites on the same DNA molecule. The kinetics of this reaction will depend on the distance between the sites: sites that are too far apart are effectively unlinked, whereas sites that are too close together should not allow cleavage, because the DNA between them cannot form a loop. It was recently shown that *Nae*I can also bind simultaneously to several sites on a DNA molecule to form loops visible by electron microscopy (200).

Several authors have pointed out that the requirement of *Eco*RII and *Nae*I for simultaneous binding to two recognition sites before they will act is similar to that of enzymes responsible for site-specific recombination and for transposition (17, 102, 200). Future investigations will show whether these restriction endonucleases execute other biological functions besides restriction of foreign DNA. For instance, a requirement for two unmodified sites for cleavage makes it less likely that the enzyme will suicidally restrict the rare unmodified sites in chromosomal DNA that may arise by DNA repair or by incomplete methylation.

Most probably, the group of type II enzymes requiring two recognition sites for cleavage is not limited to *Eco*RII and *Nae*I. For *Eco*RII and *Nae*I the cleavage of resistant recognition sites is stimulated by adding oligonucleotides containing recognition sites or a second DNA species with a high density of recognition sites (36, 99, 145). Similar stimulation has also been found for the restriction enzymes *Atu*BI, *Bsp*MI, *Cfr*9I, *Eco*57I, *Hpa*II, *Ksp*632I, *Nar*I, *Sac*II, and *Sau*BMKI (144, 167, 167a), which may, therefore, also require two sites for cleavage. However, not all enzymes that cleave certain sites inefficiently have the same mechanism. The enzymes *Eae*I and *Ecl*XI, for instance, which also cleave certain substrates poorly, are actually inhibited by addition of site-containing oligonucleotides (166).

The fact that some restriction sites in natural DNA molecules are inefficiently cleaved can have certain practical consequences. For example, the methylation status of DNA is often assessed by digestion with pairs of enzymes that recognize the same DNA sequence but have different sensitivities to methylation: one member of the pair will cleave the methylated sequence, whereas the other will not (141). In these experiments, if the methylation-sensitive enzyme has

resistant sites, these may be interpreted as methylated when in fact they are not. In these cases it may be helpful to use oligonucleotide stimulation of cleavage to ensure complete digestion of nonmethylated sites (167). In some experiments, in which ligation follows cleavage, for example, the presence of the cleaved stimulatory oligonucleotides may be a hindrance. For *Eco*RII (146) and *Nae*I (36a) this problem can be circumvented by the use of chemically modified oligonucleotides which stimulate cleavage of DNA but are themselves resistant to digestion.

What Is the Function of Type IIS Enzymes?

Some class II R-M systems recognize asymmetrical 4- to 7-bp DNA sequences and cut the DNA at a fixed distance (1 to 20 bp) outside this recognition sequence. These enzymes have been called type IIS to distinguish them from the majority of type II enzymes, which recognize symmetrical DNA sequences (for a review, see reference 194). All of the type IIS enzymes known were discovered by screening cell extracts for endonuclease activity; it is therefore not clear whether their biological function is the same as that of the more classical enzymes. The separation of the domains for DNA sequence recognition and DNA cleavage makes IIS enzymes interesting objects for studies of DNA-protein recognition (22). The genes coding several of these systems have been cloned and sequenced (213).

For classical type II R-M systems, the methylatable nucleotides are on both strands of the DNA, symmetrically disposed about the central axis of symmetry of the recognition site. The question of how the nonsymmetrical recognition sequences of type IIS systems are modified following DNA replication, which is necessary to avoid suicidal restriction of the cellular DNA, is of great interest. For most of the type IIS systems, this point has not been yet been investigated. The best-understood example is *Fok*I, in which the methyltransferase contains two independent N^6meA transferase domains, each specific for one of the two strands of the DNA (93, 109, 123, 187). Modification in the *Hga*I system is accomplished by two independent 5-meC methylases, each specific for one strand of the DNA (188). It has even been reported that some type IIS systems may methylate cytosine in one strand and adenosine in the other (J. Bitinaite, quoted in reference 194). However, for *Mbo*II it has been shown that only one strand becomes methylated (22, 128), and at least one other type IIS enzyme (*Rle*AI) has neither A nor C residues in one strand of its recognition sequence and can therefore be methylated only in one strand (204). One of several ways in which this last class of systems might survive without lethal restriction of host cell DNA would be by a mechanism similar to that of the type III enzymes, which also modify only one strand of their recognition sequence (see below).

TYPE III R-M SYSTEMS

Type III R-M systems constitute the smallest group of restriction systems, with only four known members, *Eco*P1, *Eco*P15, *Hin*fIII, and *Sty*LTI. *Eco*P1 and *Eco*P15 are coded by the P1 prophage and the related p15B plasmid of *E. coli*, respectively. *Hin*fIII is produced by *Haemophilus influenzae* Rf (with an isoschizomer in strain Re [147]) and is the only type III system so far found outside the family *Enterobacteriaceae* (88). *Sty*LTI is coded by the chromosome of most *Salmonella* strains (29) and has only recently been shown to be a type III system (42–44).

The following properties are shared by all type III R-M systems. (i) The restriction enzymes contain two subunits, one of which, the *mod* gene product, can function alone as a modification methylase. The second subunit, the *res* gene product, has no enzymatic activity when not complexed with Mod (not shown for *Hin*fIII). (ii) Restriction requires ATP, which is not hydrolyzed, and the reaction is stimulated by *S*-adenosylmethionine (AdoMet) (not shown for *Sty*LTI). (iii) the recognition sequences are asymmetrical and, when modified, have a single N^6meA in one strand of the DNA only, whereas the other strand has no modifications (8, 42, 70, 130, 148).

Genetics of Type III Systems

The only type III systems to be well characterized genetically are *Eco*P1 and *Eco*P15. These two systems complement each other: mutational defects in one system can be complemented by wild-type alleles from the other (6). Transposon mutagenesis, studies with the purified enzymes, and DNA heteroduplex and sequence analysis showed that, contrary to some earlier reports, there are only two structural genes in these systems, with a structure reminiscent of that found within a family of type I R-M systems; that is, the *res* gene is highly conserved over its whole length whereas the *mod* gene is a mosaic of conserved and nonconserved regions (69, 81). The *res* gene is necessary only for restriction, whereas the *mod* gene product provides DNA recognition for both restriction and modification as well as catalyzing modification. It is therefore a functional analog of the *hsdS* and *hsdM* gene products of type I R-M systems.

The *mod* genes can be divided into three roughly equal parts, the amino- and carboxy-terminal thirds of the gene, which are highly conserved between *Eco*P1 and *Eco*P15, and the central third, which is completely different in the two genes (80). Several lines of evidence support the idea that the central region of the gene codes the domain of the protein that recognizes DNA sequences. First, there does not seem to be enough amino acid sequence variation in the conserved regions to allow them to recognize different DNA sequences. Second, replacing a conserved region from one gene with the equivalent region from the other did not change the specificity (52). The conserved regions of the Mod proteins are most probably involved in protein-protein interactions with the Res subunit. A similar function has also been suggested for the conserved regions of the type I *hsdS* gene products (58, 64).

The structures of some unusual *Eco*P1 *mod* mutants have been elucidated recently. These are the clear-plaque *c*2 mutants of phage P1 (176, 177). The phenotype of these mutant phages is Mod^- Res^+. They cannot establish themselves as lysogens, because restriction in the absence of modification destroys the host cell chromosome. Since all infected cells are killed, regardless of whether they give rise to a productive infection, the plaques obtained are clear (171). The reason that the lytic cycle can succeed is that restriction activity is normally expressed only late after infection, when a lytically infected cell has already produced mature phage particles (7). Two independent *c*2 alleles were sequenced, and both mutations were found to be in the *mod* gene and changed two different threonine codons within the variable region to isoleucine codons (80). Both *c*2 mutant proteins were purified. As expected from the in vivo phenotype, they had no methylase activity and their primary defect was that they failed to bind the methyl donor, AdoMet (162). Other mutations in the nonconserved regions of the *mod*

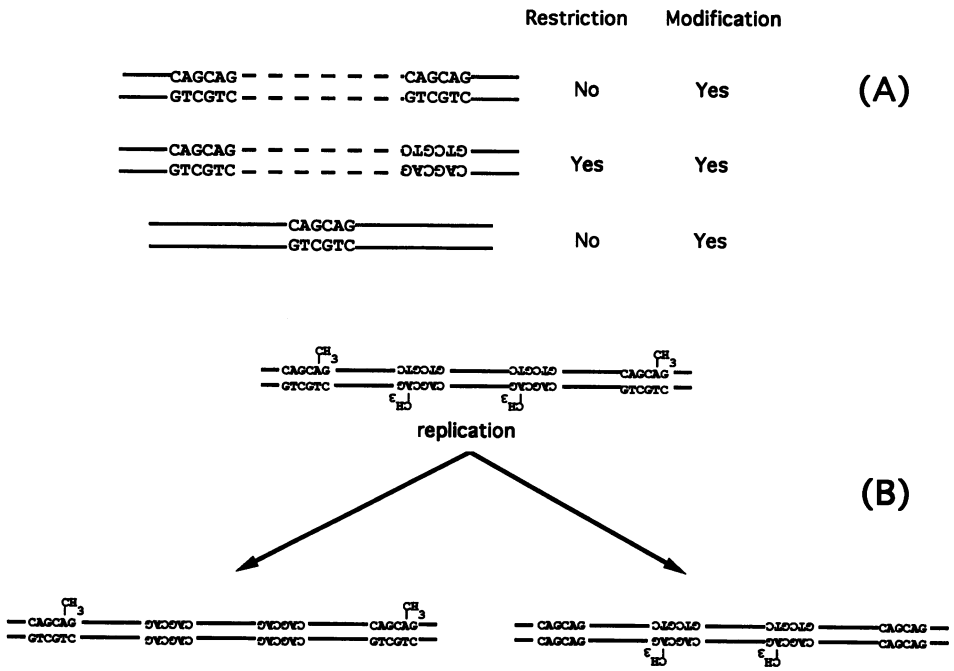

FIG. 4. Substrate requirements for the type III system *Eco*P15. (A) Only pairs of *Eco*P15 sites in inverse orientation are substrates for restriction. Both orientations of pairs of sites, as well as single sites, are substrates for modification methylation. (B) Replication of *Eco*P15-modified DNA leaves all unmodified sites in the daughter DNA molecules in the same orientation, where they are therefore not restricted.

gene have also been constructed, and they encode mutants with phenotypes similar to those of the *c*2 mutants (162). One of these mutants also had the interesting property that, unlike the wild-type enzyme, it was no longer subject to inhibition by high concentrations of AdoMet (163). All of these results are consistent with the idea that the variable region of the *mod* gene contains a DNA-binding domain and is also at least partially responsible for binding AdoMet (162).

DNA Recognition by Type III Enzymes: Different Sequence Requirements for Restriction and Modification

All of the type III enzymes recognize asymmetrical DNA sequences, and modified DNA bears methyl groups in one strand of the DNA only (8, 42, 70, 130, 148). For example, *Eco*P15 recognizes CAGCAG, the second adenosyl is methylated in modified DNA (130), and no methyl groups are found in the complementary strand (70). It was early recognized that the asymmetrical methylation of modified DNA posed a conceptual problem with no immediately obvious solution (8): when a modified recognition site is replicated, the single methyl group is inherited by one daughter DNA molecule, which remains completely modified, whereas the same site in the other daughter molecule is completely unmodified and ought to be subject to restriction, which would be lethal for the cell (most other R-M systems methylate residues in both strands of the DNA, and each daughter molecule inherits one of the parental methyl groups; hemimethylated recognition sites of this type are resistant to restriction).

A solution to this problem followed the fortuitous observation that phage T3 DNA is restricted by *Eco*P15 whereas the DNA of its close relative, T7, is not. This was not due to a lack of *Eco*P15 sites in the T7 genome: phage T7 DNA has 36 sites for *Eco*P15 (50), which, strikingly, all have the same orientation along the DNA; that is, all of the CAGCAG sequences are in one strand of the DNA and all of the CTGCTG complementary sequences are in the other strand (175). The T3 genome has not been entirely sequenced, but the sequenced portions contain both orientations of the *Eco*P15 site (101). These observations led to the hypothesis that restriction requires two recognition sites that have to be in inverse orientation; sites in direct orientation would not be subject to restriction but could still be modified (Fig. 4A). This prediction was tested by using phage M13 constructs with different constellations of *Eco*P15 sites in both direct and inverse orientations. It was shown that any combination of sites (or a single site) could be modified, whereas only sites in inverse orientation were restricted (129). In other words, the *Eco*P15 endonuclease recognition site is symmetrical and is interrupted at the center of symmetry by a nonspecific spacer of variable length.

Figure 4B shows the biological consequences of this model: all of the unmodified sites in freshly replicated DNA molecules are necessarily in direct orientation and thus are not subject to restriction. The model explains many early observations concerning type III restriction that at the time of publication were unexplainable. For example, density-labeled P1-modified phage λ DNA was grown for one cycle on a non-P1-modifying host and phages of intermediate density containing DNA with one P1-modified parental DNA strand and one unmodified newly synthesized DNA strand were isolated. These phages were totally resistant to *Eco*P1 restriction (4). A similar in vivo experiment with heteroduplex λ DNA containing one *Eco*P1 modified and one unmodified strand gave the same result (73).

*Sty*LTI System

The chromosomally encoded *Sty*LTI system was difficult to characterize, largely because it proved impossible to transfer the structural genes from one strain to another by conjugation or by cloning chromosomal DNA. Conjugation resulted in extensive DNA breakdown and cell death in the recipients as a result of uncontrolled expression of restriction before the chromosome of the recipients could be modified (43). The structural genes were finally cloned in a two-step procedure, involving first cloning of the modification functions and then cloning of the restriction functions into these modifying hosts on a second, compatible vector (44). This is unusual behavior for an R-M system; the structural genes for many (but not all) restriction systems can be transferred between strains without notable killing, indicating that the expression of restriction is closely controlled.

Most recently, the *Sty*LTI structural genes have been sequenced (41). The amino acid sequences of the *Sty*LTI Res and Mod proteins show a good deal of homology with the *Eco*P1 and *Eco*P15 proteins. In the Mod gene products, the homology is highest in the amino- and carboxy-terminal thirds of the protein and very weak in the central third (where several large gaps must be introduced to find any homology at all). This mosaic pattern of homologies strengthens the ideas about the functions of the various domains of the protein that were presented above: the central section most probably codes DNA recognition functions while the amino and carboxy termini would mediate protein-protein interactions. Homologies between the *Sty*LTI and *Eco*P1 Res proteins are generally weaker than those found in the Mod proteins (the *Eco*P15 *res* gene has not been sequenced). An exception is a region of about 50 amino acids in the middle of the proteins, which is very highly conserved between the two proteins and may well be involved in interactions with the Mod subunit.

RESTRICTION SYSTEMS SPECIFIC FOR MODIFIED DNA

*Dpn*I and *Dpn*II

The first restriction enzyme that was shown to require a methylated substrate was *Dpn*I from *Streptococcus pneumoniae*, which cleaves the sequence GATC only if the A is methylated. Other strains of *S. pneumoniae* have the complementary specificity and produce *Dpn*II, which recognizes and cleaves the same sequence only if it is not methylated (106, 107). The *Dpn*I and *Dpn*II pair of enzymes (together with the methyltransferases associated with *Dpn*II) present one of the clearest arguments in support of the idea that the main function of R-M systems is to protect cells from viral infection. Phages produced in *S. pneumoniae* strains that produce *Dpn*II have methylated genomes and are efficiently restricted in strains that produce *Dpn*I. Conversely, phages produced in the latter strains have nonmethylated genomes and are efficiently restricted in *Dpn*II producers. However, *S. pneumoniae* is naturally transformable with high efficiency, and homologous transforming DNA is only weakly restricted in either type of strain independently of its methylation status (108). This is because during transformation of *S. pneumoniae* the donor DNA is taken up in single-stranded form, and neither *Dpn*I nor *Dpn*II is active against single-stranded DNA. Moreover, integration of the donor DNA into the chromosome involves the transient formation of a heteroduplex between the donor DNA and the recipient DNA; neither enzyme will cleave a hemimethylated recognition site. *Dpn*II-expressing cells also favor transformation by expressing two modification methylases. One of these is a typical modification methylase, specific for double-stranded DNA, but the second enzyme is unusual in that it is very active on single-stranded DNA (34). Since transformation is thought to be important in the population biology of highly transformable organisms—*S. pneumoniae* can switch its specificity between *Dpn*I and *Dpn*II by transformation, for example (107)—it is intriguing that the R-M systems of *S. pneumoniae* have evolved to have a maximum effect on phage infection with virtually no effect on homologous, transforming DNA. Methyl-dependent restriction has also been demonstrated in other bacterial species, for instance, in *Streptomyces* species (126).

Rediscovery of Methyl-Directed Restriction in *E. coli*

Methylation-dependent restriction enzymes have been rediscovered in *E. coli* in recent years. The classical restriction enzymes operate by recognizing and destroying DNA that does not carry strain-specific modification. The new enzymes recognize and cut DNA carrying the DNA modification signature of other strains and, in consequence, are not accompanied by the equivalent of a classical modification enzyme. Three such systems have now been found in *E. coli*: McrA, McrBC, and Mrr (76, 142, 161, 206). The McrBC system is the most thoroughly studied of these.

Ironically, McrBC was the first restriction system to be described in *E. coli*, as early as 1952 (125). This is because most phage research in those days involved the T-even phages and McrBC is active against variants of T-even phages that have nonglycosylated hydroxymethylcytosine (hmC) in their DNA. It was later shown that two independent restriction systems were active against these phage variants, and they were named RglA and RglB (for restricts glucoseless DNA [reviewed in reference 168]). Because T-even phages were the only source of hmC-containing DNA likely to be found in *E. coli*, the Rgl systems were considered to be a defense mechanism specific for T-even phages (100, 168). The Rgl systems were rediscovered in the late 1980s in the course of cloning experiments, when it became apparent that 5meC-containing DNA from many sources, including higher plants and animals, was restricted in *E. coli* and that the Rgl systems were responsible for the restriction (20, 142, 159, 161, 209, 215, 216). It has been formally proposed that the Mcr nomenclature (for methylcytosine restricting) is more appropriate for general use but that Rgl is still useful for discussions of T-even phage biology (158; this reference also contains a brief review of the Mcr and Mrr systems). RglA corresponds to McrA, and RglB corresponds to McrBC.

McrBC

The *mcrBC* (*rglB*) locus in *E. coli* K-12 is adjacent to the *hsd* genes that code the classical type I *Eco*K restriction system. Interestingly, the gene coding the Mrr system (another methyl-dependent restriction enzyme [see below]) is just on the other side of the *hsd* genes (76), so that genes for three independent restriction systems are clustered in this small region of the genome, which has, in consequence, been dubbed the immigration control region (89, 160). There are conflicting reports in the literature concerning the organization of the *mcrBC* locus. However, a consensus seems to

have arisen that it contains two genes, *mcrB* and *mcrC*, organized as an operon and encoding at least three polypeptides. The products of both genes are necessary for most restriction reactions (47, 48, 96, 105, 151, 172, 173, 221, 222).

McrBC restricts DNA containing 5-meC, N^4meC, and hmC in specific sequence contexts (21, 142, 161, 168). The recognition sequences have not yet been fully defined, but it seems clear that they all include R-mC, where "m" is one of the three cytosine modifications and R is either purine (161). Investigations of the enzymology of McrBC have begun, and some fascinating results have emerged. The purified enzyme contains two of the polypeptides encoded by the operon and has an absolute requirement for GTP, the only known DNase with this property. ATP, a cofactor for many nucleases, is a competitive inhibitor of the reaction (193). Work with the purified enzyme has shed further light on the DNA sequence requirements of the enzyme, but much work remains to be done. The enzyme seems to require pairs of R-mC sequences separated by 40 to 80 bp of DNA, and it is necessary only that a single strand of the DNA be methylated (193).

McrA

McrA (RglA) is coded by the *e14* genetic element of many strains of *E. coli* K-12 (160). The *e14* element is most probably a defective prophage: it can excise from the chromosome as a nonreplicating circle following UV irradiation and is then eliminated from the cells by dilution during cell division (26). The *mcrA* locus has been mapped and sequenced (79, 165) and appears to code for a single polypeptide. The sequence specificity of McrA appears to be more restricted than that of McrBC, in that only DNAs methylated by the *Hpa*II and *Sss*I methylases (both cytosine methylases) have so far been shown to be substrates for the enzyme (79, 151, 161).

Mrr

The Mrr system was discovered as an activity that restricted DNA methylated by the *Hha*II or *Pst*I modification enzyme (76). Both of these enzymes methylate adenosyl residues to N^6meA, but it was later shown that some sequences containing 5-meC were also restricted by Mrr, and this has been referred to as McrF restriction (89, 206). The *mrr* locus maps next to the *hsd* locus on the other side from *mcrBC* (174). The locus has been sequenced and shown to contain a single ORF (206).

In summary, the existence of the methyl-dependent restriction systems reinforces the view that the purpose of restriction is to prevent successful phage infection. *E. coli* K-12, for example, expends considerable metabolic energy maintaining at least four independent restriction systems containing at least eight polypeptides in readiness for a situation that most cells never encounter: invasion by foreign DNA. Despite the apparent uselessness of these restriction systems for most cells, they have not been eliminated during evolution, indicating that the selective advantage of restriction in natural populations may be much higher than one would have predicted from laboratory experiments.

PHAGE ANTIRESTRICTION

Recent Developments

In this section we will describe results published since our last review on phage antirestriction mechanisms (100). A survey of the most important of these mechanisms is shown in Table 3. A large number of papers dealing with the influence of unusual bases in natural or artificial substrates on the reaction mechanism of endonucleases and methyltransferases have appeared. Studies concerning the influence of the naturally modified bases N^6meA, 5-meC, and N^4meC in different positions of restriction enzyme recognition sequences have been recently reviewed (91).

TABLE 3. Examples of phage antirestriction mechanisms

Mechanism	Example[a]
Virus-coded inhibitors of restriction enzymes	Ocr protein of phages T3 and T7, *Bacillus* phages ϕN2rH and ϕ1rH
Virus-coded DNA-modifying enzymes	Methylases with the same specificity as host R-M systems in some *Bacillus* phages, phage-specific modification [N^6-(1-acetimido)adenine for phage Mu]
Incorporation of unusual nucleotides into DNA	Several unusual nucleotides in various *Bacillus* phages
Coinjection of DNA and inhibitory internal head proteins	DarA and DarB proteins of phage P1
Counterselection of relevant restriction sites in phage genomes	In coliphage T7 and in the *Bacillus* phages ϕ1, ϕ29, and SPO1
Strand-biased asymmetrical recognition sequences	*Eco*P15 sites in the T7 genome
Hydrolysis of restriction enzyme cofactors	AdoMet hydrolase of phage T3

[a] For references, see text and reference 100.

The significance of counterselection of restriction enzyme recognition sites in phage genomes has been shown by further studies. For instance, two broad-host-range phages with very large genomes (pseudomonas phage ϕST1 and streptomycete phage FP22) have been shown to have no sites for some of the restriction enzymes found in their hosts and very few sites for other enzymes (61, 71).

In the genome of coliphage T7, the recognition sites for the chromosomally encoded *E. coli* methylases Dam (GATC) and Dcm (CCWGG) appear only six and two times, respectively (50). Schroeder et al. (175) calculated the expected frequencies of the tetra- and pentanucleotides by Markov chain analysis on the basis of the observed occurrences of the subsequences GAT, ATC, and AT (for the Dam sequence) and CCWG, CWGG, and CWG (for the Dcm sequence) in the T7 genome. From these calculations, GATC and CCWGG are expected to occur 114 and 56 times, respectively, in the absence of counterselection. This shows that selection was directed against these particular sequences (and not against subsets of the sequences which could be involved in different selection processes), indicating that their presence was disadvantageous for the reproduction of the virus (175). In independent studies, selection against Dam sites and restriction enzyme recognition sites in the genomes of other phages of members of the *Enterobacteriaceae* has been demonstrated (127, 178).

In some cases, a few recognition sites for certain restriction enzymes can be tolerated in phage genomes because they do not lead to restriction. The very few *Eco*RII sites in the T7 and T3 genomes are not restricted, because the enzyme requires at least two sites in relatively close proximity for restriction and the distance between the *Eco*RII sites in the T3 and T7 genomes is too great for the enzyme to function (99). For enzymes that recognize asymmetrical sequences, the distance between sites may not be the only criterion for avoiding restriction. As discussed above for the type III enzymes, some enzymes can be sensitive to the orientation of cleavage sites. The 36 *Eco*P15 sites in the phage T7 genome all have the same orientation, making the DNA refractory to *Eco*P15 cleavage (129, 175). On the other hand, restriction sites may be overrepresented in phage genomes. An example is the sequence AGACC, recognized by *Eco*P1, which occurs 126 times in the phage T7 genome (distributed on both DNA strands). The *Eco*P1 recognition sequence fortuitously overlaps the site recognized by the T7 DNA primase, an enzyme essential for T7 DNA replication. This overlap is most probably the reason for the high frequency of *Eco*P1 sites in the T7 genome and explains why T7 DNA replication is severely depressed in P1-lysogenic cells (175).

An efficient mechanism for antirestriction has been developed by phages T3 and T7, whose 0.3 gene codes a specific inhibition protein, Ocr (100). The Ocr protein binds to type I restriction enzymes and blocks both the endonuclease and methylase activities (10, 23, 103, 184). It is inactive against type II enzymes, both endonucleases and methylases (10). In early studies, an inhibition of the type III *Eco*P1 restriction was demonstrated in vivo but could not be shown in vitro with the purified restriction enzyme (100). More recent studies involving a different experimental approach failed to confirm the in vivo *Eco*P1 inhibition (134). The T3 (but not T7) 0.3 gene product has, in addition to its Ocr activity, a second enzymatic activity: the protein is an AdoMet hydrolase (186). As a result, the T3 protein not only blocks type I restriction enzymes but also has markedly inhibitory effects on type III restriction enzymes, which are stimulated by AdoMet (98).

Several phages of *B. subtilis*, such as SPβ, SPR, φ3T and ρ11, protect themselves from host restriction systems by phage-coded DNA methyltransferases (100). These enzymes are often multispecific; that is, they recognize several unrelated DNA sequences. The modular (and sometimes cryptic) structure of these enzymes has been worked out over the last few years by the group of Trautner and Noyer-Weidner (13, 110, 111, 202, 203, 211). These results are discussed in more detail in the section on type II cytosine methyltransferases.

Recently, a DNA methyltransferase gene has been demonstrated on the genome of the *Lactococcus lactis* phage φ50 (78). Unlike the *Bacillus* phage methyltransferase genes, which are unrelated to host cell genes, the φ50 gene has clearly been picked up recently from a plasmid of its host: some 1.3 kb of DNA sequence with homologies to other methyltransferase genes is identical between the phage and the plasmid (78).

Phage λ has an antirestriction mechanism called Ral that is specific for type IA R-M systems. Ral reduces restriction, but its primary effect is to stimulate the modification of completely unmodified DNA (100), a reaction that for type IA R-M systems is normally very inefficient. The *ral* gene has now been cloned and the antirestriction mechanism has been examined in more detail (121, 223). These studies confirm that the main action of Ral is on modification and that modification by both the three-subunit restriction enzyme and the two-subunit modification enzyme are stimulated; the results are consistent with the formation of a Ral-enzyme complex with altered methylation properties. In fact, the properties of the putative complex are very similar to those of the *hsdM* mutants described above, which convert *Eco*K from a maintenance to a de novo methylase (90), and it is tempting to speculate that Ral interacts with HsdM. Previous studies had indicated that Ral might interact with the Rho transcription termination factor, and it was suggested that Rho might participate directly in restriction (45), but the present results fail to confirm this notion (121).

Finally, it should be noted that some *E. coli* conjugative plasmids in incompatibility groups I and N have also been shown to code antirestriction functions. These functions are active only against type I R-M systems and are coded by genes (called *ard*) located close to the origin of conjugative transfer, where they are expressed in the recipient bacteria very early in the conjugation process (14, 14a, 46, 53, 165a).

"Nature Red in Tooth and Claw": The Case of Phage T4

The coevolution of bacterial defenses and phage attack strategies is perhaps most clearly seen with *E. coli* and the T-even family of virulent phages, which includes T2, T4, and T6. A plausible case can be made for the evolutionary scenario outlined in Fig. 5. Originally, the T-even phages would have had a genome composed of normal DNA and would have been subject to restriction by the classical R-M systems of their hosts. The evolution of a hmC-containing phage genome resistant to classical restriction can be seen as a response to the presence of these classical R-M systems. The evolutionary pressure that led to the development of the Mcr (Rgl) restriction systems specific for DNA with modified cytosines was surely challenge by phages such as the T-even phages with modified cytosines in their genomes. The most recent stage in this scenario is the acquisition by the phages of the ability to glycosylate their DNA, rendering them resistant to all known restriction systems of *E. coli*.

The struggle between host and phage is not confined to DNA restriction. Certain strains of *E. coli* have a restriction system specific for T-even phages that operates on the RNA level. This enzyme, the *prrC* gene product, is present but has no activity in uninfected cells. The activity is unmasked on infection with a T-even phage and proceeds to make a specific cleavage in the anticodon of the host-encoded $tRNA^{Lys}$. This inactivates protein synthesis and thus prevents phage multiplication. Phage T4 encodes two enzymes, polynucleotide kinase and RNA ligase, whose only known function in the T4 life cycle is to repair the damage caused by the anticodon nuclease (83, 116). A surprising finding was that the *prrC* gene is inserted into an *hsd* operon encoding an active type I R-M system (Fig. 1) (119). Moreover, the PrrC protein can be immunoprecipitated with anti-Hsd serum, indicating that it is physically complexed with the restriction enzyme. Both DNA- and RNA-based restriction systems are thus tightly integrated at both the genetic and the protein levels (1).

CONCLUDING REMARKS

It seems clear that the main role of restriction is to protect against phage infection; the wide variety of mechanisms developed by phages to counteract restriction attests to the importance of restriction for phages. This is particularly clear for the R-M systems of *S. pneumoniae*, which restrict

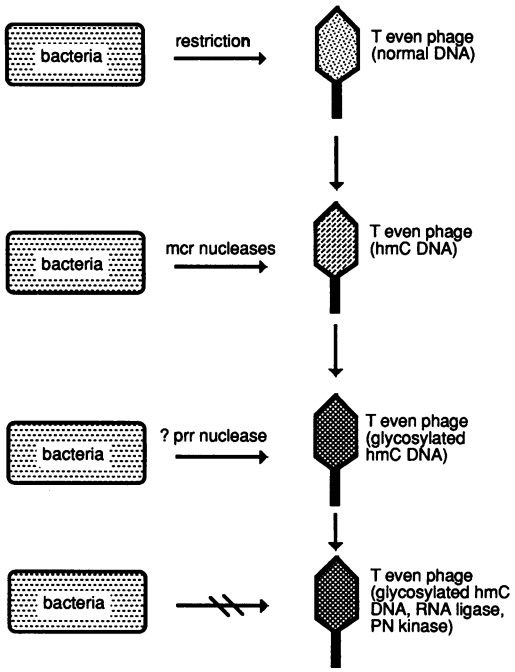

FIG. 5. Evolutionary interactions between T-even phages and their hosts. The idea here is that the phage evolves DNA modifications to avoid host restriction. The host then responds by evolving new enzymatic systems that recognize and restrict the modified DNA. See the text for details.

phages severely, whereas beneficial transformation is hardly affected. Other roles for restriction in the cellular economy are more difficult to document. For example, restriction may facilitate recombination by generating recombinogenic double-stranded breaks in homologous incoming DNA (181), and it has been suggested that type I R-M systems would be particularly effective in this role (153).

It should not be forgotten that a small fraction of invading DNA molecules evade restriction and become modified. Thus, although restriction provides a barrier to genetic mobility, the barrier is leaky, allowing the possibility of a low rate of evolution by horizontal gene transfer.

ACKNOWLEDGMENTS

We thank the many members of the nuclease community who sent us manuscripts prior to publication and Mario Noyer-Weidner for a critical appraisal of the manuscript of this review. Special thanks are due to Rich Roberts, who maintains a comprehensive data base of the restriction literature, which he makes freely available to the scientific community. This review would have been much more difficult to prepare without access to this data base.

Work from Basel has been supported by grants from the Swiss National Science Foundation. Recent work from Berlin was supported by the Deutsche Forschungsgemeinschaft (Kr 1293/1-1) and the Federal Minister for Research and Technology (grant 0319854 A).

REFERENCES

1. **Amitsur, M., I. Morad, D. Chapman-Shimshoni, and G. Kaufmann.** 1992. HSD restriction-modification proteins partake in latent anticodon nuclease. EMBO J. **11:**3129–3134.
2. **Arber, W.** 1965. Host specificity of DNA produced by *Escherichia coli*. V. The role of methionine in the production of host specificity. J. Mol. Biol. **11:**247–256.
3. **Arber, W., and D. Dussoix.** 1962. Host specificity of DNA produced by *Escherichia coli*. I. Host controlled modification of bacteriophage lambda. J. Mol. Biol. **5:**18–36.
4. **Arber, W., S. Hattman, and D. Dussoix.** 1963. On the host-controlled modification of bacteriophage lambda. Virology **21:**30–35.
5. **Arber, W., and S. Linn.** 1969. DNA modification and restriction. Annu. Rev. Biochem. **38:**467–500.
6. **Arber, W., and D. Wauters-Willems.** 1970. Host specificity of DNA produced by *Escherichia coli*. XII. The two restriction and modification systems of strain 15T$^-$. Mol. Gen. Genet. **108:**203–217.
7. **Arber, W., R. Yuan, and T. A. Bickle.** 1974. Strain-specific modification and restriction of DNA in bacteria, p. 3–22. *In* F. Antoni and A. Farago (ed.), Post synthetic modification of macromolecules. Elsevier Science Publishing, Inc., New York.
8. **Bächi, B., J. Reiser, and V. Pirrotta.** 1979. Methylation and cleavage sequences of the *Eco*P1 restriction-modification enzyme. J. Mol. Biol. **128:**143–163.
9. **Balganesh, T. S., L. Reiners, R. Lauster, M. Noyer-Weidner, K. Wilke, and T. A. Trautner.** 1987. Construction and use of chimeric SPR/ϕ3T DNA methyltransferases in the definition of sequence recognizing enzyme regions. EMBO J. **6:**3543–3549.
10. **Bandyopadhyay, P. K., F. W. Studier, D. L. Hamilton, and R. Yuan.** 1985. Inhibition of the type I restriction-modification enzymes *Eco*B and *Eco*K by the gene *O.3* protein of bacteriophage T7. J. Mol. Biol. **182:**567–578.
11. **Barany, F., M. Danzitz, J. Zebala, and A. Mayer.** 1992. Cloning and sequencing of genes encoding the *Tth*HB8I restriction and modification enzymes: comparison with the isoschizomeric *Taq*I enzymes. Gene **112:**3–12.
12. **Barras, F., and M. G. Marinus.** 1989. The great GATC: DNA methylation in *E. coli*. Trends Genet. **5:**139–143.
13. **Behrens, B., M. Noyer-Weidner, B. Pawlek, R. Lauster, T. S. Balganesh, and T. A. Trautner.** 1987. Organization of multispecific DNA methyltransferases encoded by temperate *Bacillus subtilis* phages. EMBO J. **6:**1137–1142.
14. **Belogurov, A. A., T. N. Yussifov, V. U. Kotova, and G. B. Zavilgelsky.** 1985. The novel gene(s) ARD of plasmid pKM101: alleviation of *Eco*K restriction. Mol. Gen. Genet. **198:**509–513.

14a. **Belogurov, A. A., E. P. Delver, and O. V. Rodzevich.** 1992. IncN plasmid pKM101 and IncI1 plasmid ColIb-P9 encode homologous antirestriction proteins in their leading regions. J. Bacteriol. **174:**5079–5085.

15. **Bertani, G., and J. J. Weigle.** 1953. Host controlled variation in bacterial viruses. J. Bacteriol. **65:**113–121.
16. **Bestor, T., A. Laudano, R. Mattaliano, and V. Ingram.** 1988. Cloning and sequencing of a cDNA encoding DNA methyltransferase of mouse cells. J. Mol. Biol. **203:**971–983.
17. **Bhagwat, A. S., B. Johnson, K. Weule, and R. J. Roberts.** 1990. Primary sequence of the *Eco*RII endonuclease and properties of its fusions with β-galactosidase. J. Biol. Chem. **265:**767–773.
18. **Bickle, T. A.** 1982. The ATP-dependent restriction endonucleases, p. 85–108. *In* S. M. Linn and R. J. Roberts (ed.), Nucleases. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
19. **Bickle, T. A.** 1987. DNA restriction and modification systems, p. 692–696. *In* J. L. Ingraham, K. B. Low, B. Magasanik, F. C. Neidhardt, M. Schaechter, and H. E. Umbarger (ed.), *Escherichia coli* and *Salmonella typhimurium*: cellular and molecular biology. American Society for Microbiology, Washington, D.C.
20. **Blumenthal, R. M.** 1987. The *Pvu*II restriction-modification system: cloning, characterization and use in revealing in *E. coli* barrier to certain methylases or methylated DNAs, p. 227–245. *In* J. G. Chirikjian (ed.), Gene amplification and analysis. Elsevier/North-Holland Publishing Co., Amsterdam.
21. **Blumenthal, R. M., S. A. Gregory, and J. S. Cooperider.** 1985. Cloning of a restriction-modification system from *Proteus vulgaris* and its use in analyzing a methylase-sensitive phenotype in *Escherichia coli*. J. Bacteriol. **164:**501–509.
22. **Bocklage, H., K. Heeger, and B. Müller-Hill.** 1991. Cloning and characterization of the *Mbo*II restriction-modification system.

Nucleic Acids Res. **19**:1007–1013.

23. **Bogdarina, I. G., M. Reuter, D. H. Krüger, Y. I. Buryanov, and A. A. Baev.** 1983. Methylation of DNA of phages T3 and T7 by various types of DNA-adenine methylases and inhibition of the *Eco*K methylase by the Ocr protein. Dokl. Akad. Nauk SSSR **273**:234–237.
24. **Boyer, H. W.** 1971. DNA restriction and modification mechanisms in bacteria. Annu. Rev. Microbiol. **25**:153–176.
25. **Boyer, H. W., and D. Roulland-Dussoix.** 1969. A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. J. Mol. Biol. **41**:459–472.
26. **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.
27. **Brooks, J. E., P. D. Nathan, D. Landry, L. A. Sznyter, P. Waite-Rees, C. L. Ives, L. S. Moran, B. E. Slatko, and J. S. Benner.** 1991. Characterization of the cloned *Bam*HI restriction modification system: its nucleotide sequence, properties of the methylase, and expression in heterologous hosts. Nucleic Acids Res. **19**:841–850.
28. **Brown, T. C., and J. Jiricny.** 1987. A specific mismatch repair event protects mammalian cells from loss of 5-methylcytosine. Cell **50**:945–950.
29. **Bullas, L. R., C. Colson, and B. Neufeld.** 1980. Deoxyribonucleic acid restriction and modification systems in *Salmonella*: chromosomally located systems of different serotypes. J. Bacteriol. **141**:275–292.
30. **Bullas, L. R., C. Colson, and A. van Pel.** 1976. DNA restriction and modification systems in *Salmonella*. SQ, a new system derived by recombination between the SB system of *Salmonella typhimurium* and the SP system of *Salmonella potsdam*. J. Gen. Microbiol. **95**:166–172.
31. **Burckhardt, J., J. Weisemann, and R. Yuan.** 1981. Characterization of the DNA methylase activity of the restriction enzyme from *Escherichia coli* K. J. Biol. Chem. **256**:4024–4032.
32. **Butkus, V., S. Klimasauskas, D. Kersulyte, D. Vaitkevicius, A. Lebionka, and A. A. Janulaitis.** 1985. Investigation of restriction-modification enzymes from *M. varians* RFL19 with a new type of specificity toward modification of substrate. Nucleic Acids Res. **13**:5727–5746.
33. **Butkus, V., S. Klimasauskas, L. Petrauskiene, Z. Maneliene, A. Lebionka, and A. A. Janulaitis.** 1987. Interaction of *Alu*I, *Cfr*6I and *Pvu*II restriction-modification enzymes with substrates containing either N4-methylcytosine or 5-methylcytosine. Biochim. Biophys. Acta **909**:201–207.
34. **Cerritelli, S., S. S. Springhorn, and S. A. Lacks.** 1989. *Dpn*A, a methylase for single-strand DNA in the *Dpn*II restriction system, and its biological function. Proc. Natl. Acad. Sci. USA **86**:9223–9227.
35. **Chandrasegaran, S., and H. O. Smith.** 1988. Amino acid sequence homologies among twenty-five restriction endonucleases and methylases, p. 149–156. *In* R. H. Sarma and M. H. Sarma (ed.), Structure and expression. From proteins to ribosomes. Adenine Press, Inc., New York.
36. **Conrad, M., and M. D. Topal.** 1989. DNA and spermidine provide a switch mechanism to regulate the activity of restriction enzyme *Nae*I. Proc. Natl. Acad. Sci. USA **86**:9707–9711.

36a. **Conrad, M., and M. D. Topal.** 1992. Modified DNA fragments activate *Nae*I cleavage of refractory sites. Nucleic Acids Res. **20**:5127–5130.

37. **Coulondre, C., J. H. Miller, P. J. Farabaugh, and W. Gilbert.** 1978. Molecular basis of base substitution hot spots in *Escherichia coli*. Nature (London) **274**:775–780.
38. **Cowan, G. M., A. A. F. Gann, and N. E. Murray.** 1989. Conservation of complex DNA recognition domains between families of restriction enzymes. Cell **56**:103–109.
39. **Cowan, G. M., A. A. F. Gann, and N. E. Murray.** 1989. Conservation of complex DNA recognition domains between families of restriction enzymes. Cell **56**:103–109.
40. **Daniel, A. S., F. V. Fuller-Pace, D. M. Legge, and N. E. Murray.** 1988. Distribution and diversity of *hsd* genes in *Escherichia coli* and other enteric bacteria. J. Bacteriol. **170**:1775–1782.
41. **Dartois, V., O. De Backer, and C. Colson.** Sequence of the *Salmonella typhimurium Sty*LTI restriction-modification genes. Homologies with *Eco*P1 and *Eco*P15 type III R-M systems and presence of helicase domains. Submitted for publication.
42. **De Backer, O., and C. Colson.** 1991. Identification of the recognition sequence for the M.*Sty*LTI methyltransferase of *Salmonella typhimurium* LT7: an asymmetric site typical of type-III enzymes. Gene **97**:103–107.
43. **De Backer, O., and C. Colson.** 1991. Transfer of the genes for the *Sty*LTI restriction-modification system of *Salmonella typhimurium* to strains lacking modification ability results in death of the recipient cells and degradation of their DNA. J. Bacteriol. **173**:1328–1330.
44. **De Backer, O., and C. Colson.** 1991. Two-step cloning and expression in *Escherichia coli* of the DNA restriction-modification system *Sty*LTI of *Salmonella typhimurium*. J. Bacteriol. **173**:1321–1327.
45. **Debrouwere, L., M. van Montague, and J. Schell.** 1980. The *ral* gene of phage λ. III. Interference with *E. coli* ATP dependent functions. Mol. Gen. Genet. **179**:81–88.
46. **Delver, E. P., V. U. Kotova, G. B. Zavilgelsky, and A. A. Belogurov.** 1991. Nucleotide sequence of the gene (*ard*) encoding the antirestriction protein of plasmid ColIb-P9. J. Bacteriol. **173**:5887–5892.
47. **Dila, D., and E. A. Raleigh.** 1988. Genetic dissection of the methylcytosine-specific restriction system *mcrB* of *Escherichia coli* K-12. Gene **74**:23–24.
48. **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.
49. **Duncan, B. K., and J. H. Miller.** 1980. Mutagenic deamination of cytosine residues in DNA. Nature (London) **287**:560–561.
50. **Dunn, J. J., and F. W. Studier.** 1983. The complete nucleotide sequence of bacteriophage T7 DNA, and the locations of T7 genetic elements. J. Mol. Biol. **166**:477–535.
51. **Dussoix, D., and W. Arber.** 1962. Host specificity of DNA produced by *Escherichia coli*. II. Control over acceptance of DNA from infecting phage lambda. J. Mol. Biol. **5**:37–49.
52. **Eberle, H. E., N. Redaschi, and T. A. Bickle.** Unpublished results.
53. **Efimova, E. P., E. P. Delver, and A. A. Belogurov.** 1988. Alleviation of type I restriction in adenine methylase (*dam*) mutants of *Escherichia coli*. Mol. Gen. Genet. **214**:313–316.
54. **Ehrlich, M., M. A. Gama-Sosa, L. H. Carreira, L. G. Ljungdahl, K. C. Kuo, and C. W. Gehrke.** 1985. DNA methylation in thermophilic bacteria: N4-methylcytosine, 5-methylcytosine, and N6-methyladenine. Nucleic Acids Res. **13**:1399–1412.
55. **Ehrlich, M., and R. Y.-H. Wang.** 1981. 5-Methylcytosine in eukaryotic DNA. Science **212**:1350–1357.
56. **Ehrlich, M., G. G. Wilson, C. K. Kenneth, and C. W. Gehrke.** 1987. *N*4-Methylcytosine as a minor base in bacterial DNA. J. Bacteriol. **169**:939–943.
57. **Endlich, B., and S. Linn.** 1981. Type I restriction enzymes, p. 137–156. *In* P. D. Boyer (ed.), The enzymes. Academic Press, Inc., New York.
58. **Fuller-Pace, F. V., L. R. Bullas, H. Delius, and N. E. Murray.** 1984. Genetic recombination can generate altered restriction specificity. Proc. Natl. Acad. Sci. USA **81**:6095–6099.
59. **Fuller-Pace, F. V., G. M. Cowan, and N. E. Murray.** 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.
60. **Fuller-Pace, F. V., and N. E. Murray.** 1986. Two DNA recognition domains of the specificity polypeptides of a family of type I restriction enzymes. Proc. Natl. Acad. Sci. USA **83**:9368–9372.
61. **Gachechiladze, K. K., N. S. Balardshishvili, R. S. Adamia, T. G. Chanishvili, and D. H. Krüger.** 1991. Host-controlled modification and restriction as a criterion of evaluating the therapeutical potential of *Pseudomonas* phage. J. Basic Microbiol. **31**:101–106.

62. **Gann, A. A. F., A. J. B. Campbell, J. F. Collins, A. F. W. Coulson, and N. E. Murray.** 1987. Reassortment of DNA recognition domains and the evolution of new specificities. Mol. Microbiol. **1:**13–22.
63. **Glover, S. W., K. Firman, G. Watson, C. Price, and S. Donaldson.** 1983. The alternate expression of two restriction and modification systems. Mol. Gen. Genet. **190:**65–69.
64. **Gough, J. A., and N. E. Murray.** 1983. Sequence diversity among related genes for recognition of specific targets in DNA molecules. J. Mol. Biol. **166:**1–19.
65. **Gromkova, R., and S. H. Goodgal.** 1976. Biological properties of a *Haemophilus influenzae* restriction enzyme, *Hin*dI. J. Bacteriol. **127:**848–854.
66. **Gubler, M., and T. A. Bickle.** 1991. Increased protein flexibility leads to promiscuous protein-DNA interactions in type IC restriction-modification systems. EMBO J. **10:**951–957.
67. **Gubler, M., D. Braguglia, J. Meyer, A. Piekarowicz, and T. A. Bickle.** 1992. Recombination of constant and variable modules alters DNA sequence recognition by type IC restriction-modification enzymes. EMBO J. **11:**233–240.
68. **Guschlbauer, W.** 1988. The DNA and S-adenosylmethionine-binding regions of *Eco*Dam and related methyltransferases. Gene **74:**211–214.
69. **Hadi, S. M., B. Bächi, S. Iida, and T. A. Bickle.** 1983. DNA restriction-modification enzymes of phage P1 and plasmid p15B. J. Mol. Biol. **165:**19–34.
70. **Hadi, S. M., B. Bächi, J. C. W. Shepherd, R. Yuan, K. Ineichen, and T. A. Bickle.** 1979. DNA recognition and cleavage by the *Eco*P15 restriction endonuclease. J. Mol. Biol. **134:**655–666.
71. **Hahn, D. R., M. A. McHenney, and R. H. Baltz.** 1990. Characterization of FP22, a large streptomycete bacteriophage with DNA insensitive to cleavage by many restriction enzymes. J. Gen. Microbiol. **136:**2395–2404.
72. **Harrison, H. C.** 1991. A structural taxonomy of DNA-binding domains. Nature (London) **353:**715–719.
73. **Hattman, S., J. E. Brooks, and M. Masurekar.** 1978. Sequence specificity of the P1 modification methylase (M.*Eco*P1) and the DNA methylase (M.*Eco*dam) controlled by the *Escherichia coli dam* gene. J. Mol. Biol. **126:**367–380.
74. **Hedges, R. W., and N. Datta.** 1972. R124, an fi+ R factor of a new compatibility class. J. Gen. Microbiol. **71:**403–405.
75. **Heidmann, S., W. Seifert, C. Kessler, and H. Domdey.** 1989. Cloning, characterization and heterologous expression of the *Sma*I restriction-modification system. Nucleic Acids Res. **17:** 9783–9796.
76. **Heitman, J., and P. Model.** 1987. Site-specific methylases induce the SOS DNA repair response in *Escherichia coli*. J. Bacteriol. **169:**3243–3250.
77. **Hennecke, F., H. Kolmar, K. Bründl, and H. J. Fritz.** 1991. The *vsr* gene product of *E. coli* K-12 is a strand- and sequence-specific DNA mismatch endonuclease. Nature (London) **353:** 776–778.
78. **Hill, C., L. A. Miller, and T. R. Klaenhammer.** 1991. In vivo genetic exchange of a functional domain from a type II a methylase between lactococcal plasmid pTR2030 and a virulent bacteriophage. J. Bacteriol. **173:**4363–4370.
79. **Hiom, K., and S. G. Sedgwick.** 1991. Cloning and structural characterization of the *mcrA* locus of *Escherichia coli*. J. Bacteriol. **173:**7368–7373.
80. **Hümbelin, M., B. Suri, D. N. Rao, D. P. Hornby, H. Eberle, T. Pripfl, S. Kenel, and T. A. Bickle.** 1988. Type III DNA restriction and modification systems *Eco*P1 and *Eco*P15. Nucleotide sequence of the *Eco*P1 operon, the *Eco*P15 *mod* gene and some *Eco*P1 *mod* mutants. J. Mol. Biol. **200:**23–29.
81. **Iida, S., J. Meyer, B. Bächi, M. Stålhammar-Carlemalm, S. Schrickel, T. A. Bickle, and W. Arber.** 1983. DNA restriction-modification genes of phage P1 and plasmid p15B. J. Mol. Biol. **165:**1–18.
82. **Ives, C. L., P. D. Nathan, and J. E. Brooks.** 1992. Regulation of the *Bam*HI restriction modification system by a small intergenic ORF, *bamHIC*, in both *Escherichia coli* and *Bacillus subtilis*. J. Bacteriol. **174:**7194–7201.
83. **Jabbar, M. A., and L. Snyder.** 1984. Genetic and physiological studies of an *Escherichia coli* locus that restricts polynucleotide kinase- and RNA ligase-deficient mutants of bacteriophage T4. J. Virol. **51:**522–529.
84. **Janulaitis, A. A., S. Klimasauskas, M. Petrusyte, and V. Butkus.** 1983. Cytosine modification in DNA by *Bcn*I methylase yields N4-methylcytosine. FEBS Lett. **161:**131–134.
85. **Kan, N. C., J. A. Lautenberger, M. H. Edgell, and C. A. Hutchison III.** 1979. The nucleotide sequence recognized by the *Escherichia coli* K12 restriction and modification enzymes. J. Mol. Biol. **130:**191–209.
86. **Kannan, P., G. M. Cowan, A. S. Daniel, A. A. F. Gann, and N. E. Murray.** 1989. Conservation of organization in the specificity polypeptides of two families of type I restriction enzymes. J. Mol. Biol. **209:**335–344.
87. **Kapfer, W., J. Walter, and T. A. Trautner.** 1991. Cloning, characterization and evolution of the *Bsu*FI restriction endonuclease gene of *Bacillus subtilis* and purification of the enzyme. Nucleic Acids Res. **19:**6457–6463.
88. **Kauc, L., and A. Piekarowicz.** 1978. Purification and properties of a new restriction endonuclease from *Haemophilus influenzae* Rf. Eur. J. Biochem. **92:**417–426.
89. **Kelleher, J. C., and E. A. Raleigh.** 1991. A novel activity in *Escherichia coli* K12 that directs restriction of DNA modified at CG dinucleotides. J. Bacteriol. **173:**5220–5223.
90. **Kelleher, J. E., A. S. Daniel, and N. E. Murray.** 1991. Mutations that confer *de novo* activity upon a maintenance methyltransferase. J. Mol. Biol. **221:**431–440.
91. **Kessler, C., and V. Manta.** 1990. Specificity of restriction endonucleases and DNA modification methyltransferases—a review (edition 3). Gene **92:**1–248.
92. **Kiss, A., G. Posfai, C. C. Keller, P. Venetianer, and R. J. Roberts.** 1985. Nucleotide sequence of the *Bsu*RI restriction-modification system. Nucleic Acids Res. **13:**6403–6420.
93. **Kita, K., H. Kotani, H. Sugisaki, and M. Takanami.** 1989. The *Fok*I restriction-modification system. I. Organization and nucleotide sequences of the restriction and modification genes. J. Biol. Chem. **264:**5751–5756.
94. **Klimasauskas, S., J. L. Nelson, and R. J. Robert.** 1991. The sequence specificity domain of cytosine-C5 methylases. Nucleic Acids Res. **19:**6183–6190.
95. **Klimasauskas, S., A. Timinskas, S. Menkevicius, D. Butkiene, V. Butkus, and A. Janulaitis.** 1989. Sequence motifs characteristic of DNA[cytosine-N4]methylases: similarity to adenine and cytosine-C5 DNA-methylases. Nucleic Acids Res. **17:** 9823–9832.
96. **Kretz, P. L., S. W. Kohler, and J. M. Short.** 1991. Identification and characterization of a gene responsible for inhibiting propagation of methylated DNA sequences in *mcrA mcrB1 Escherichia coli* strains. J. Bacteriol. **173:**4707–4716.
97. **Kröger, M., and G. Hobom.** 1984. The nucleotide sequence recognized by the *Escherichia coli* A restriction and modification enzyme. Nucleic Acids Res. **12:**887–899.
98. **Krüger, D. H.** Unpublished data.
99. **Krüger, D. H., G. J. Barcak, M. Reuter, and H. O. Smith.** 1988. *Eco*RII can be activated to cleave refractory DNA recognition sites. Nucleic Acids Res. **16:**3997–4008.
100. **Krüger, D. H., and T. A. Bickle.** 1983. Bacteriophage survival: multiple mechanisms for avoiding the DNA restriction systems of their hosts. Microbiol. Rev. **47:**345–360.
101. **Krüger, D. H., T. A. Bickle, M. Reuter, C. D. Pein, and C. Schroeder.** 1990. DNA methylation and restriction processes in *Escherichia coli*: insights by use of bacterial viruses T3 and T7, p. 113–124. *In* G. A. Clawson, D. B. Willis, A. Weissbach, and P. A. Jones (ed.), Nucleic acid methylation. Alan R. Liss, Inc., New York.
102. **Krüger, D. H., M. Reuter, L. S. Chernin, K. K. Gachechiladze, T. G. Chanishvili, and C. Schroeder.** 1990. Biological functions of restriction endonucleases. Biol. Zentralbl. **109:**257–266.
103. **Krüger, D. H., C. Schroeder, M. Reuter, I. G. Bogdarina, Y. I. Buryanov, and T. A. Bickle.** 1985. DNA methylation of bacterial viruses T3 and T7 by different DNA methylases in *Escherichia coli* K12 cells. Eur. J. Biochem. **150:**323–330.

104. **Krüger, D. H., C. Schroeder, M. Santibanez-Koref, and M. Reuter.** 1989. Avoidance of DNA methylation: a virus-encoded methylase inhibitor and evidence for counterselection of methylase recognition sites in viral genomes. Cell Biophys. **15:**87–95.
105. **Krüger, T., C. Grund, C. Wild, and M. Noyer-Weidner.** 1992. Characterization of the *mcrBC* region of *Escherichia coli* K12 wild-type and mutant strains. Gene **114:**1–12.
106. **Lacks, S., and B. Greenberg.** 1977. Complementary specificity of restriction endonucleases of *Diplococcus pneumoniae* with respect to DNA methylation. J. Mol. Biol. **114:**153–168.
107. **Lacks, S. A., B. M. Mannarelli, S. S. Springhorn, and B. Greenberg.** 1986. Genetic basis of the complementary *Dpn*I and *Dpn*II restriction systems of *S. pneumoniae*: an intercellular cassette mechanism. Cell **46:**993–1000.
108. **Lacks, S. A., and S. S. Springhorn.** 1984. Transfer of recombinant plasmids containing the gene for *Dpn*II DNA methylase into strains of *Streptococcus pneumoniae* that produce *Dpn*I or *Dpn*II restriction endonucleases. J. Bacteriol. **158:**905–909.
109. **Landry, D., M. C. Looney, G. R. Feehery, B. E. Slatko, W. E. Jack, I. Schildkraut, and G. G. Wilson.** 1989. M.*Fok*I methylates adenine in both strands of its asymmetric recognition sequence. Gene **77:**1–10.
110. **Lange, C., A. Jugel, J. Walter, M. Noyer-Weidner, and T. A. Trautner.** 1991. 'Pseudo' domains in phage-encoded DNA methyltransferases. Nature (London) **352:**645–648.
111. **Lange, C., M. Noyer-Weidner, T. A. Trautner, M. Weiner, and S. A. Zahler.** 1991. M.H2I, a multispecific 5C-DNA methyltransferase encoded by *Bacillus amyloliquefaciens* phage H2. Gene **100:**213–218.
112. **Lauster, R.** 1989. Evolution of type II DNA methyltransferases: a gene duplication model. J. Mol. Biol. **206:**313–321.
113. **Lauster, R., A. Kriebardis, and W. Guschlbauer.** 1987. The GATATC-modification enzyme *Eco*RV is closely related to the GATC-recognizing methyltransferases *Dpn*II and *dam* from *E. coli* and phage T4. FEBS Lett. **220:**167–176.
114. **Lauster, R., T. A. Trautner, and M. Noyer-Weidner.** 1989. Cytosine-specific type II DNA methyltransferases: a conserved enzyme core with variable target-recognizing domains. J. Mol. Biol. **206:**305–312.
115. **Lautenberger, J. A., N. C. Kan, D. Lackey, S. Linn, M. H. Edgell, and C. A. I. Hutchison.** 1978. Recognition site of *Escherichia coli* B restriction enzyme on ϕXsB1 and simian virus 40 DNAs: an interrupted sequence. Proc. Natl. Acad. Sci. USA **75:**2271–2275.
116. **Levitz, R., D. Chapman, M. Amitsur, R. Green, L. Snyder, and G. Kaufmann.** 1990. The optional *E. coli prr* locus encodes a latent form of phage T4-induced anticodon nuclease. EMBO J. **9:**1383–1389.
117. **Lieb, M.** 1983. Specific mismatch correction in bacteriophage lambda crosses by very short patch repair. Mol. Gen. Genet. **191:**118–125.
118. **Lieb, M.** 1991. Spontaneous mutation at a 5-methylcytosine hotspot is prevented by very short patch (VSP) mismatch repair. Genetics **128:**23–27.
119. **Linder, P., R. Doelz, M. Gubler, and T. A. Bickle.** 1990. An anticodon nuclease gene inserted into a *hsd* region encoding a type I DNA restriction system. Nucleic Acids Res. **18:**7170.
120. **Linn, S., and W. Arber.** 1968. Host specificity of DNA produced by *Escherichia coli*. X. In vitro restriction of phage fd replicative form. Proc. Natl. Acad. Sci. USA **59:**1300–1306.
121. **Loenen, W., and N. Murray.** 1986. Modification enhancement by restriction alleviation protein (Ral) of bacteriophage lambda. J. Mol. Biol. **190:**11–22.
122. **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* K12. J. Mol. Biol. **198:**159–170.
123. **Looney, M. C., L. S. Moran, W. E. Jack, G. R. Fehery, J. S. Benner, B. E. Slatko, and G. G. Wilson.** 1989. Nucleotide sequence of the *Fok*I restriction-modification system: separate strand-specificity domains in the methyltransferase. Gene **80:**193–208.
124. **Luria, S. E.** 1970. The recognition of DNA in bacteria. Sci. Am. **222:**88–102.
125. **Luria, S. E., and M. L. Human.** 1952. A nonhereditary, host-induced variation of bacterial viruses. J. Bacteriol. **64:**557–569.
126. **MacNeil, D. J.** 1988. Characterization of a unique methyl-specific restriction system in *Streptomyces avermitilis*. J. Bacteriol. **170:**5607–5612.
127. **McClelland, M.** 1985. Selection against *dam* methylation sites in the genomes of DNA of enterobacteriophages. J. Mol. Evol. **21:**317–322.
128. **McClelland, M., M. Nelson, and C. R. Cantor.** 1985. Purification of *Mbo*II methylase (GAAGmA) from *Moraxella bovis*: site specific cleavage of DNA at nine and ten base pair sequences. Nucleic Acids Res. **13:**7171–7182.
129. **Meisel, A., T. A. Bickle, D. H. Krüger, and C. Schroeder.** 1992. Type III restriction enzymes need two inversely oriented recognition sites for DNA cleavage. Nature (London) **355:**467–469.
130. **Meisel, A., D. H. Krüger, and T. A. Bickle.** 1991. M.*Eco*P15I methylates the second adenine in its recognition sequence. Nucleic Acids Res. **19:**3997.
131. **Meselson, M., R. Yuan, and J. Heywood.** 1972. Restriction and modification of DNA. Annu. Rev. Biochem. **41:**447–466.
132. **Modrich, P.** 1979. Structures and mechanisms of DNA restriction and modification enzymes. Q. Rev. Biophys. **12:**315–369.
133. **Modrich, P., and R. J. Roberts.** 1982. Type-II restriction and modification enzymes, p. 109–154. *In* S. Linn and R. J. Roberts (ed.), Nucleases. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
134. **Moffatt, B. A., and F. W. Studier.** 1988. Entry of bacteriophage T7 DNA into the cell and escape from host restriction. J. Bacteriol. **170:**2095–2105.
135. **Murray, N. E., J. A. Gough, B. Suri, and T. A. Bickle.** 1982. Structural homologies among type I restriction-modification systems. EMBO J. **1:**535–539.
136. **Nagaraja, V., J. C. W. Shepherd, and T. A. Bickle.** 1985. A hybrid recognition sequence in a recombinant restriction enzyme and the evolution of DNA sequence specificity. Nature (London) **316:**371–372.
137. **Nagaraja, V., J. C. W. Shepherd, T. Pripfl, and T. A. Bickle.** 1985. Two type I restriction enzymes from *Salmonella* species. Purification and DNA recognition sequences. J. Mol. Biol. **182:**579–587.
138. **Nagaraja, V., M. Stieger, C. Nager, S. M. Hadi, and T. A. Bickle.** 1985. The nucleotide sequence recognized by the *Escherichia coli* D type I restriction and modification enzyme. Nucleic Acids Res. **13:**389–399.
139. **Narva, K. E., J. L. Van Etten, B. E. Slatko, and J. S. Benner.** 1988. The amino acid sequence of the eukaryotic DNA [N6-adenine]methyltransferase, M.*Cvi*BIII, has regions of similarity with the prokaryotic isoschizomer M.*Taq*I and other DNA [N6-adenine] methyltransferases. Gene **74:**253–259.
140. **Nathan, P. D., and J. E. Brooks.** 1988. Characterization of clones of the *Bam*HI methyltransferase gene. Gene **74:**35–36.
141. **Nelson, M., and M. McClelland.** 1991. Site-specific methylation: effect on DNA modification methyltransferases and restriction endonucleases. Nucleic Acids Res. **19:**2045–2071.
142. **Noyer-Weidner, M., R. Diaz, and L. Reiners.** 1986. Cytosine-specific DNA modification interferes with plasmid establishment in *Escherichia coli* K12: involvement of *rglB*. Mol. Gen. Genet. **205:**469–475.
143. **Noyer-Weidner, M., S. Jentsch, B. Pawlek, U. Gunthert, and T. A. Trautner.** 1983. Restriction and modification in *Bacillus subtilis*: DNA methylation potential of the related bacteriophages Z, SPR, SPβ, ϕ3T, and ρ11. J. Virol. **46:**446–453.
143a. **Noyer-Weidner, M., and T. A. Trautner.** 1993. Methylation of DNA in prokaryotes, p. 39–108. *In* J. P. Jost and H. P. Saluz (ed.), DNA methylation: molecular biology and biological significance. Birkhäuser, Basel.
144. **Oller, A. R., W. Vanden Broek, M. Conrad, and M. D. Topal.** 1991. Ability of DNA and spermidine to affect the activity of restriction endonucleases from several bacterial species. Biochemistry **30:**2543–2549.

145. **Pein, C.-D., M. Reuter, D. Cech, and D. H. Krüger.** 1989. Oligonucleotide duplexes containing CC(A/T)GG stimulate cleavage of refractory DNA by restriction endonuclease *Eco*RII. FEBS Lett. **245:**141–144.

146. **Pein, C.-D., M. Reuter, A. Meisel, D. Cech, and D. H. Krüger.** 1991. Activation of restriction enzyme *Eco*RII does not depend on the cleavage of stimulator DNA. Nucleic Acids Res. **19:**5139–5142.

147. **Piekarowicz, A.** 1982. *Hin*eI is an isoschizomer of *Hin*fIII restriction endonuclease. J. Mol. Biol. **157:**373–381.

148. **Piekarowicz, A., T. A. Bickle, J. C. W. Shepherd, and K. Ineichen.** 1981. The DNA sequence recognised by the *Hin*fIII restriction endonuclease. J. Mol. Biol. **146:**167–172.

149. **Piekarowicz, A., and J. D. Goguen.** 1986. The DNA sequence recognized by the *Eco*DXXI restriction endonuclease. Eur. J. Biochem. **154:**295–298.

150. **Piekarowicz, A., J. D. Goguen, and E. Skrzypek.** 1985. The *Eco*DXXI restriction and modification system of *Escherichia coli* ET7. Purification, subunit structure and properties of the restriction endonuclease. Eur. J. Biochem. **152:**387–393.

151. **Piekarowicz, A., R. Yuan, and D. C. Stein.** 1991. Isolation of temperature-sensitive McrA and McrB mutations and complementation analysis of the *mcrBC* region of *Escherichia coli* K-12. J. Bacteriol. **173:**150–155.

152. **Posfai, J., A. S. Bhagwat, G. Posfai, and R. J. Roberts.** 1989. Predictive motifs derived from cytosine methyltransferases. Nucleic Acids Res. **17:**2421–2435.

153. **Price, C., and T. A. Bickle.** 1986. A possible role for DNA restriction in bacterial evolution. Microbiol. Sci. **3:**296–299.

154. **Price, C., J. Lingner, T. A. Bickle, K. Firman, and S. W. Glover.** 1989. Basis for changes in DNA recognition by the *Eco*R124 and *Eco*R124/3 Type I DNA restriction and modification enzymes. J. Mol. Biol. **205:**115–125.

155. **Price, C., T. Pripfl, and T. A. Bickle.** 1987. *Eco*R124 and *Eco*R124/3: the first members of a new family of type I restriction and modification systems. Eur. J. Biochem. **167:** 111–115.

156. **Price, C., J. C. W. Shepherd, and T. A. Bickle.** 1987. DNA recognition by a new family of type I restriction enzymes: a unique relationship between two different DNA specificities. EMBO J. **6:**1493–1497.

157. **Raleigh, E. A.** 1992. Organization and function of the *mcrBC* genes of *Escherichia coli* K-12. Mol. Microbiol. **6:**1079–1086.

158. **Raleigh, E. A., J. Benner, F. Bloom, H. D. Braymer, E. DeCruz, K. Dharmalingam, J. Heitman, M. Noyer-Weidner, A. Piekarowicz, P. L. Kretz, J. M. Short, and D. Woodcock.** 1991. Nomenclature relating to restriction of modified DNA in *Escherichia coli*. J. Bacteriol. **173:**2707–2709.

159. **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. **15:**1563–1575.

160. **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.

161. **Raleigh, E. A., and G. Wilson.** 1986. *Escherichia coli* K-12 restricts DNA containing 5-methylcytosine. Proc. Natl. Acad. Sci. USA **83:**9070–9074.

162. **Rao, D. N., H. Eberle, and T. A. Bickle.** 1989. Characterization of mutations of the bacteriophage P1 *mod* gene encoding the recognition subunit of the *Eco*P1 restriction and modification system. J. Bacteriol. **171:**2347–2352.

163. **Rao, D. N., M. G. P. Page, and T. A. Bickle.** 1989. Cloning, over-expression and the catalytic properties of the *Eco*P15 modification methylase from *Escherichia coli*. J. Mol. Biol. **209:**599–606.

164. **Ravetch, J. V., K. Horiuchi, and N. D. Zinder.** 1978. Nucleotide sequence of the recognition site for the restriction-modification enzyme of *Escherichia coli* B. Proc. Natl. Acad. Sci. USA **75:**2266–2270.

165. **Ravi, R. S., S. Sozhamannan, and K. Dharmalingam.** 1985. Transposon mutagenesis and genetic mapping of the *rglA* and *rglB* loci of *Escherichia coli*. Mol. Gen. Genet. **198:**390–392.

165a. **Read, T. D., A. T. Thomas, and B. M. Wilkins.** 1992. Evasion of type I and type II DNA restriction systems by IncI1 plasmid ColIb-P9 during transfer by bacterial conjugation. Mol. Microbiol. **6:**1933–1941.

166. **Reuter, M., D. Kupper, C.-D. Pein, M. Petrusyte, V. Siksnys, B. Frey, and D. H. Krüger.** 1993. Use of specific oligonucleotide duplexes to stimulate cleavage of refractory sites by restriction endonucleases. Anal. Biochem. **209:**232–237.

167. **Reuter, M., C.-D. Pein, V. Butkus, and D. H. Krüger.** 1990. An improved method for the detection of *Dcm* methylation in DNA molecules. Gene **95:**161–162.

167a. **Reuter, M., C.-D. Pein, D. Kupper, D. H. Krüger, B. Frey, G. Schmitz, M. Petrusyte, V. Siksnys, and V. Butkus.** 1990. A group of restriction enzymes which can be activated to cleave insensitive DNA sites, p. 122. Abstracts of the Second New England Biolabs Workshop on Biological DNA Modification, Berlin.

168. **Revel, H. R.** 1983. DNA modification: glucosylation, p. 156–165. *In* C. K. Mathews, E. M. Kutter, G. Mosig, and P. Berget (ed.), Bacteriophage T4. American Society for Microbiology, Washington, D.C.

169. **Revel, H. R., and S. E. Luria.** 1970. DNA-glucosylation in T-even phage: genetic determination and role in phage-host interaction. Annu. Rev. Genet. **4:**177–192.

170. **Riggs, A. D.** 1989. DNA methylation and cell memory. Cell Biophys. **15:**1–13.

171. **Rosner, J. L.** 1973. Modification-deficient mutants of bacteriophage P1. I. Restriction by P1 cryptic lysogens. Virology **52:**213–222.

172. **Ross, T. K., E. C. Achberger, and H. D. Braymer.** 1989. Identification of a second polypeptide required for McrB restriction of 5-methylcytosine-containing DNA in *Escherichia coli* K12. Mol. Gen. Genet. **216:**42–407.

173. **Ross, T. K., E. C. Achberger, and H. D. Braymer.** 1989. Nucleotide sequence of the McrB region of *Escherichia coli* K-12 and evidence for two independent translational initiation sites at the *mcrB* locus. J. Bacteriol. **171:**1974–1981.

174. **Sain, B., and N. E. Murray.** 1980. The *hsd* (host specificity) genes of *E. coli* K12. Mol. Gen. Genet. **180:**35–46.

175. **Schroeder, C., H. Jurkschat, A. Meisel, J. G. Reich, and D. Krüger.** 1986. Unusual occurrence of *Eco*P1 and *Eco*P15 recognition sites and counterselection of type II methylation and restriction sequences in bacteriophage T7 DNA. Gene **45:**77–86.

176. **Scott, J. R.** 1968. Genetic studies on bacteriophage P1. Virology **36:**564–574.

177. **Scott, J. R.** 1970. Clear plaque mutants of bacteriophage P1. Virology **41:**66–71.

178. **Sharp, P. M.** 1986. Molecular evolution of bacteriophages: evidence of selection against the recognition sites of host restriction enzymes. Mol. Biol. Evol. **3:**75–83.

179. **Skrzypek, E., and A. Piekarowicz.** 1989. The *Eco*DXXI restriction and modification system: cloning the genes and homology to type I restriction and modification systems. Plasmid **21:**195–204.

180. **Slatko, B. E., J. S. Benner, T. Jager-Quinton, L. S. Moran, T. G. Simcox, E. M. VanCott, and G. G. Wilson.** 1987. Cloning, sequencing and expression of the *Taq*I restriction-modification system. Nucleic Acids Res. **15:**9781–9796.

181. **Smith, G. R.** 1987. Mechanism and control of homologous recombination in *Escherichia coli*. Annu. Rev. Genet. **21:**179–201.

182. **Smith, H. O., T. M. Annau, and S. Chandrasegaran.** 1990. Finding sequence motifs in groups of functionally related proteins. Proc. Natl. Acad. Sci. USA **87:**826–830.

182a. **Sohail, A., M. Lieb, M. Dar, and A. S. Bhagwat.** 1990. A gene required for very short patch repair in *Escherichia coli* is adjacent to the DNA cytosine methylase gene. J. Bacteriol. **172:**4214–4221.

183. **Sommer, R., and H. Schaller.** 1979. Nucleotide sequence of the

recognition site of the B-specific restriction modification system in *Escherichia coli*. Mol. Gen. Genet. **168**:331–335.

184. **Spoerel, N., P. Herrlich, and T. A. Bickle.** 1979. A novel bacteriophage defence mechanism: the anti-restriction protein. Nature (London) **278**:30–34.
185. **Stephenson, F. H., B. T. Ballard, H. W. Boyer, J. M. Rosenberg, and P. J. Greene.** 1989. Comparison of the nucleotide and amino acid sequences of the *Rsr*I and *Eco*RI restriction endonucleases. Gene **85**:1–13.
186. **Studier, F. W., and N. R. Movva.** 1976. SAMase gene of bacteriophage T3 is responsible for overcoming host restriction. J. Virol. **19**:136–145.
187. **Sugisaki, H., K. Kita, and M. Takanami.** 1989. The *Fok*I restriction-modification system. II. Presence of two domains in *Fok*I methylase responsible for modification of different DNA strands. J. Biol. Chem. **264**:5757–5761.
188. **Sugisaki, H., K. Yamamoto, and M. Takanami.** 1991. The *Hga*I restriction and modification system contains two cytosine methylase genes responsible for modification of different DNA strands. J. Biol. Chem. **266**:13952–13957.
189. **Sullivan, K. M., and J. R. Saunders.** 1989. Nucleotide sequence and genetic organization of the *Ngo*PII restriction-modification system of *Neisseria gonorrhoeae*. Mol. Gen. Genet. **216**:380–387.
190. **Suri, B., and T. A. Bickle.** 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.
191. **Suri, B., V. Nagaraja, and T. A. Bickle.** 1984. Bacterial DNA modification. Curr. Top. Microbiol. Immunol. **108**:1–9.
192. **Suri, B., J. C. W. Shepherd, and T. A. Bickle.** 1984. The *Eco*A restriction and modification system of *Escherichia coli* 15T−: enzyme structure and DNA recognition sequence. EMBO J. **3**:575–579.
193. **Sutherland, E., L. Coe, and E. A. Raleigh.** 1992. McrBC: a multisubunit GTP-dependent restriction endonuclease. J. Mol. Biol. **225**:327–348.
194. **Szybalski, W., S. C. Kim, N. Hasan, and A. J. Podhajska.** 1991. Class-IIS restriction enzymes—a review. Gene **100**:13–26.
195. **Tao, T., and R. M. Blumenthal.** 1992. Sequence and characterization of *pvuIIR*, the *Pvu*II endonuclease gene, and of *pvuIIC*, its regulatory gene. J. Bacteriol. **174**:3395–3398.
196. **Tao, T., J. C. Bourne, and R. M. Blumenthal.** 1991. A family of regulatory genes associated with type II restriction-modification systems. J. Bacteriol. **173**:1367–1375.
197. **Tao, T., J. Walter, K. J. Brennan, M. M. Cotterman, and R. M. Blumenthal.** 1989. Sequence, internal homology and high-level expression of the gene for a DNA-(cytosine N4)-methyltransferase, M.*Pvu*II. Nucleic Acids Res. **17**:4161–4175.
198. **Terschuren, P.-A., M. Noyer-Weidner, and T. A. Trautner.** 1987. Recombinant derivatives of *Bacillus subtilis* phage Z containing the DNA methyltransferase genes of related methylation-proficient phages. J. Gen. Microbiol. **198**:945–952.
199. **Theriault, G., P. H. Roy, K. A. Howard, J. S. Benner, J. S. Brooks, A. F. Waters, and T. R. Gingeras.** 1985. Nucleotide sequence of the *Pae*R7 restriction/modification system and partial characterization of its protein products. Nucleic Acids Res. **13**:8441–8461.
200. **Topal, M. D., R. J. Thresher, M. Conrad, and J. Griffith.** 1991. *Nae*I endonuclease binding to pBR322 DNA induces looping. Biochemistry **30**:2006–2010.
201. **Tran-Betcke, A., B. Behrens, M. Noyer-Weidner, and T. A. Trautner.** 1986. DNA methyltransferase genes of *Bacillus subtilis* phages: comparison of their nucleotide sequences. Gene **42**:89–96.
202. **Trautner, T. A., T. Balganesh, K. Wilke, M. Noyer-Weidner, E. Rauhut, R. Lauster, B. Behrens, and B. Pawlek.** 1988. Organization of target-recognizing domains in the multispecific DNA (cytosine-5) methyltransferases of *Bacillus subtilis* phages SPR and ϕ3T. Gene **74**:267.
203. **Trautner, T. A., T. S. Balganesh, and B. Pawlek.** 1988. Chimeric multispecific DNA methyltransferases with novel combinations of target recognition. Nucleic Acids Res. **16**:6649–6658.
204. **Vesely, Z., A. Muller, G. G. Schmitz, K. Kaluza, M. Jarsch, and C. Kessler.** 1990. *Rle*AI: a novel class-IIS restriction endonuclease from *Rhizobium leguminosarum* recognizing CCCACA. Gene **95**:129–131.
205. **Vovis, G. F., K. Horiuchi, and N. D. Zinder.** 1974. Kinetics of methylation of DNA by a restriction endonuclease from *Escherichia coli* B. Proc. Natl. Acad. Sci. USA **71**:3810–3813.
206. **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.
207. **Walter, J., M. Noyer-Weidner, and T. A. Trautner.** 1990. The amino acid sequence of the CCGG recognizing DNA methyltransferase M.*Bsu*FI: implications for the analysis of sequence recognition by cytosine DNA methyltransferases. EMBO J. **9**:1007–1013.
208. **Wang, R. Y. H., K. C. Kuo, C. W. Gehrke, L. H. Huang, and M. Ehrlich.** 1982. Heat- and alkali-induced deamination of 5-methylcytosine and cytosine residues in DNA. Biochim. Biophys. Acta **697**:371–377.
209. **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.
210. **Wiebauer, K., and J. Jiricny.** 1991. *In vitro* correction of G/T mispairs to G/C pairs in nuclear extracts from human cells. Nature (London) **339**:234–236.
211. **Wilke, K., E. Rauhut, M. Noyer-Weidner, R. Lauster, B. Pawlek, B. Behrens, and T. A. Trautner.** 1988. Sequential order of target-recognizing domains in multispecific DNA-methyltransferases. EMBO J. **7**:2601–2609.
212. **Wilson, G. G.** 1988. Type II restriction-modification systems. Trends Genet. **4**:314–318.
213. **Wilson, G. G.** 1991. Organization of restriction-modification systems. Nucleic Acids Res. **19**:2539–2566.
214. **Wilson, G. G., and N. E. Murray.** 1991. Restriction and modification systems. Annu. Rev. Genet. **25**:585–627.
215. **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.
216. **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* strains for tolerance to cytosine methylation in plasmid and phage recombinants. Nucleic Acids Res. **17**:3469–3478.
217. **Xu, G., W. Kapfer, J. Walter, and T. A. Trautner.** Submitted for publication.
218. **Yoshimori, R., D. Roulland-Dussoix, and H. W. Boyer.** 1972. R factor-controlled restriction and modification of deoxyribonucleic acids: restriction mutants. J. Bacteriol. **112**:1275–1279.
219. **Yuan, R.** 1981. Structure and mechanism of multifunctional restriction endonucleases. Annu. Rev. Biochem. **50**:285–315.
220. **Zell, R., and H. J. Fritz.** 1987. DNA mismatch repair in *Escherichia coli* counteracting the hydrolytic deamination of 5-methylcytosine residues. EMBO J. **6**:1809–1815.
221. **Zheng, L., and H. D. Braymer.** 1991. Overproduction and purification of McrC protein from *Escherichia coli* K-12. J. Bacteriol. **173**:3918–3920.
222. **Zheng, L., X. Wang, and D. H. Braymer.** 1992. Purification and N-terminal amino acid sequences of two polypeptides encoded by the *mcrB* gene from *Escherichia coli*. Gene **112**:97–100.
223. **Zinkevich, V. E., A. M. Alekseev, and V. I. Tanyashin.** 1986. Influence of bacteriophage λ *ral* gene on the level of synthesis of the restriction endonuclease *Eco*K β-subunit. Mol. Biol. (Moscow) **20**:1638–1644. (In Russian.)