**SGM SPECIAL LECTURE**

# Immigration control of DNA in bacteria: self versus non-self

Noreen E. Murray

Tel: +44 131 650 5374. Fax: +44 131 650 8650.
e-mail: noreen.murray@ed.ac.uk

Institute of Cell and Molecular Biology, Darwin Building, Mayfield Road, Edinburgh EH9 3JR, Scotland, UK

**2001 Fred Griffith Review Lecture**

(Delivered at the 148th Meeting of the Society for General Microbiology, 27 March 2001)

**Keywords:** Restriction and modification, control by proteolysis, DNA transfer, DNA modification, selfish genes

## Background and aims

Bacteria commonly endow their DNA with an identity mark. When DNA is transferred from one bacterium to another strain of the same species, DNA that lacks the identification mark of the recipient strain is recognized as 'foreign' rather than 'self'. Foreign DNA is commonly degraded. The first evidence for this discriminatory process was the demonstration of a barrier, albeit incomplete, to the productive infection of *Escherichia coli* strain K-12 by bacteriophage $\lambda$ previously propagated in either *E. coli* strain C or *E. coli* strain B (Bertani & Weigle, 1953). Much later it was proven that the growth of phages in *E. coli* K-12 can be 'restricted' by an endonuclease, a restriction enzyme (*Eco*KI), which attacks foreign DNA (Meselson & Yuan, 1968; Linn & Arber *et al.*, 1969). Occasionally phages escape restriction and they, like the resident bacterial chromosome, acquire a protective identification mark from a strain-specific modification enzyme that methylates defined bases within a specific target sequence (Arber & Dussoix, 1962; Smith *et al.*, 1972). This sequence-specific modification identifies the immediate provenance of bacterial, or phage, DNA (Fig. 1).

Classically, a restriction enzyme is accompanied by its cognate modification enzyme and together the two activities comprise a restriction and modification (R-M) system. There are, however, some restriction endonucleases, so-called modification-dependent restriction enzymes, which attack DNA only when specific nucleotide sequences in the DNA are methylated. The classical R-M systems and the modification-dependent restriction enzymes share the potential to attack DNA derived from different strains and thereby 'restrict' DNA transfer. While the modification activity of a classical R-M system is required to protect DNA from attack by the cognate restriction endonuclease, a modification enzyme specified by one strain may impart a signal that provokes the degradative activity of a modification-dependent restriction endonuclease found in a different strain.

It is often stated, though difficult to prove, that restriction systems exist to defend bacteria against invading phages. Recently, however, it has been argued that R-M systems are 'selfish' elements. This hypothesis emanates from the finding that bacterial cells die if they lose the genes that specify their R-M system (Naito *et al.*, 1995). It has been shown that the bacterial chromosome becomes susceptible to restriction as cell growth dilutes the modification enzyme (Handa *et al.*, 2000). However, while the loss of genes that specify some simple R-M systems leads to cell death, the loss of genes that specify other, more complex, R-M systems causes no detectable viability problem (O'Neill *et al.*, 1997; Kulik & Bickle, 1996; Makovets *et al.*, 1998). In this

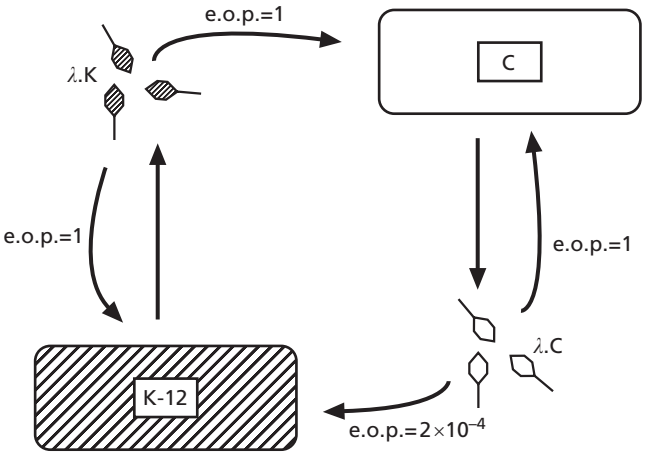


***Fig. 1.*** The phenomenon of restriction and modification. *E. coli* K-12 possesses, while *E. coli* C lacks, a type I R-M system. Phage $\lambda$ propagated in *E. coli* C ($\lambda$.C) is not protected from restriction by *Eco*KI and thus forms plaques with reduced efficiency (e.o.p. $2\times10^{-4}$) on *E. coli* K-12 as compared to *E. coli* C. The presence of modified DNA is indicated by hatching. Reproduced with permission from Barcus & Murray (1995).

review I wish to emphasize the different behaviour of *E. coli* strains dependent upon the nature, or type, of their R-M system (see also Murray, 2000). Some data will challenge our long-established belief that modification of DNA is essential to distinguish whether the DNA is 'self or foreign'. Experiments show that while modification of DNA is sufficient, it is not always essential to identify resident DNA as self (Makovets *et al.*, 1999; Doronina & Murray, 2001).

## Types of R-M systems

R-M systems have been subdivided according to the complexity and cofactor requirement of the enzymes, the nature of their target sequence, and the position of the site of DNA cleavage with respect to the target sequence. Three distinct, well-characterized types of classical R-M systems have been defined (types I, II and III; Fig. 2), although a few systems do not share all the characteristics of any of these three types (for general reviews see Wilson & Murray, 1991; Bickle & Kruger, 1993; Raleigh & Brooks, 1998). The first R-M systems identified in *E. coli* K-12 and *E. coli* B were designated type I, but the enzymes that serve as reagents in modern biology, type II R-M systems, are very much simpler. For this reason, they are described first.

A type II R-M system comprises two separate enzymes, a restriction endonuclease and a modification enzyme, or methyltransferase. The nuclease activity is dependent on $Mg^{2+}$ and the methyltransferase on *S*-adenosylmethionine (AdoMet) as the methyl donor. The restriction and modification enzymes recognize the same target sequence, usually a rotationally symmetrical sequence of 4–8 bp. Type II endonucleases are generally active as symmetrically arranged homodimers, an association that facilitates the co-ordinated cleavage of both strands of the DNA. The modification enzyme ensures that a specific base within the target sequence, one on each strand of the duplex, is methylated, but modification enzymes function as monomers, an organization consistent with their normal role in the methylation of newly replicated DNA (for reviews see Wilson & Murray, 1991; Roberts & Halford, 1993; Raleigh & Brooks, 1998; Pingoud & Jeltsch, 2001).

Genes encoding repressor-like proteins, referred to as C proteins for control, have been identified for some type II R-M systems (Ives *et al.*, 1992; Tao *et al.*, 1991; Tao & Blumenthal, 1992). The C-protein for the *Bam*HI system has been shown to activate efficient expression of the restriction gene (Ives *et al.*, 1992, 1995). Consequently, when R-M genes are transferred to a new environment in which there is no C protein, there will be preferential expression of the modification gene, and only after the production of C protein will transcription of the restriction gene be activated.

Type I R-M systems are heterooligomeric complexes that catalyse both restriction and modification (for reviews see Murray, 2000; Rao *et al.*, 2000; Dryden *et al.*, 2001). AdoMet is the methyl donor for modification but, importantly, endonuclease activity requires both AdoMet and ATP, in addition to $Mg^{2+}$. The restriction activity of type I R-M systems is associated with the hydrolysis of ATP, an activity that correlates with the bizarre characteristic of these enzymes, that of translocating DNA before they cut it at nonspecific sequences considerable distances from the target sequence (Davies *et al.*, 1999b). The nucleotide sequences recognized by type I enzymes are asymmetric and comprise two components, one of 3 or 4 bp and the other of 4 or 5 bp, separated by a non-specific spacer of 6–8 bp. The type I R-M enzyme binds to its target sequence in the presence of cofactors and the alternative activities of restriction or modification are determined by the methylation state of the target sequence. Hemimethylated target sequences are the substrate for modification but, if the target sequence is unmodified, the enzyme, while bound to its

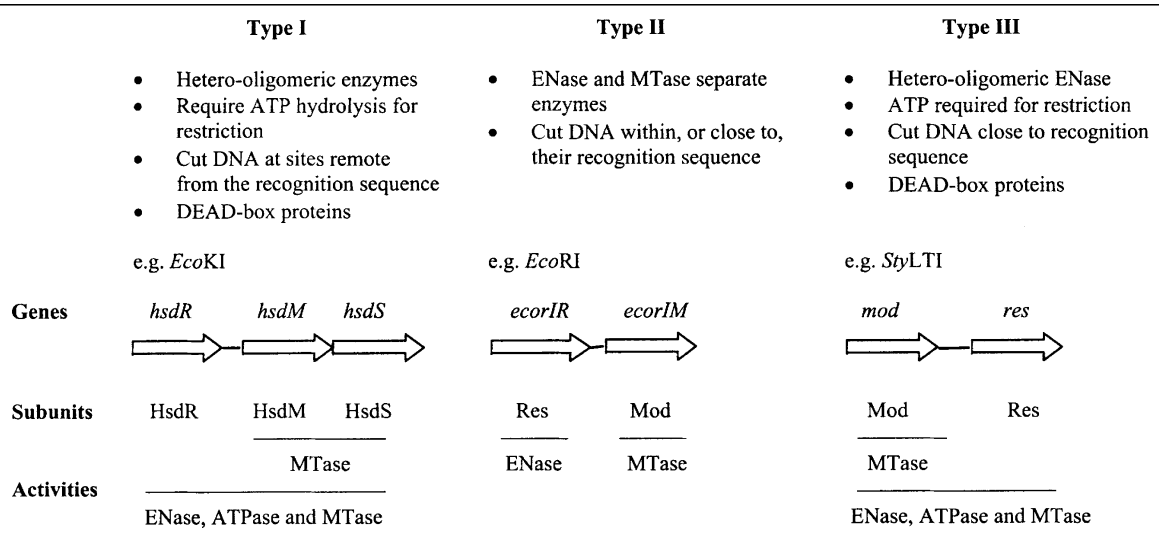


***Fig. 2.*** The characteristics and organization of the genetic determinants and subunits of different types of R-M systems. ENase, endonuclease activity; MTase, methyltransferase activity. Modified with permission from King & Murray (1994).

***Table 1.*** Number of potential restriction systems in microbial genomes based on computational analyses of DNA sequences

This table uses the information from Kong *et al.* (2000), derived from the database of Roberts & Macelis (2000).

| Organism | Genome size (Mb) | Type I | Type II | Type III | M-DRS* |
|---|---|---|---|---|---|
| *Aeropyrum pernix* | 1·67 | | 7 | | |
| *Aquifex aeolicus* | 1·55 | | | | 1 |
| *Archaeoglobus fulgidus* | 2·18 | 1 | 2 | 1 | |
| *Bacillus subtilis* | 4·21 | † | 2 | | 1 |
| *Borrelia burgdorferi* | 1·44 | | 2 | | |
| *Campylobacter jejuni* | 1·64 | 1 | 4 | | 1 |
| *Chlamydia muridarum* | 1·07 | | | | |
| *Chlamydia trachomatis* | 1·05 | | | | |
| *Chlamydia pneumoniae* AR39 | 1·23 | | | | |
| *Deinococcus radiodurans* | 2·65 | | 4 | | 3 |
| *Escherichia coli* K-12 | 4·60 | 1‡ | | | 3 |
| *Haemophilus influenzae* | 1·83 | 2 | 3 | 1 | |
| *Helicobacter pylori* 26695 | 1·66 | 3 | 14 | 2 | |
| *Helicobacter pylori* J99 | 1·64 | 3 | 16 | 2 | |
| *Methanobacterium thermoautotrophicum* | 1·75 | 1 | 1 | | 3 |
| *Methanococcus jannaschii* | 1·66 | 3 | 8 | | |
| *Mycobacterium tuberculosis* | 4·40 | 1 | 1 | | |
| *Mycoplasma genitalium* | 0·58 | 1 | | | |
| *Mycoplasma pneumoniae* | 0·81 | 1 | 1 | | |
| *Neisseria meningitidis* serotype A | 2·18 | 3 | 7 | 2 | |
| *Neisseria meningitidis* serotype B | 2·27 | 1 | 4 | 1 | |
| *Pyrococcus abyssi* | 1·77 | 1 | 4 | | |
| *Pyrococcus horikoshii* | 1·74 | | 3 | | |
| *Rickettsia prowazekii* | 1·10 | | | | |
| *Synechocystis* species | 3·57 | | 1 | | 1 |
| *Thermatoga maritima* | 1·80 | | 1 | | |
| *Treponema pallidum* | 1·16 | | | | |
| *Ureaplasma urealyticum* | 0·71 | 1 | 1 | | |

* Putative methylation-dependent restriction systems.

† Some strains of *B. subtilis* do have a type I R-M system.

‡ Many strains of *E. coli* have a chromosomally encoded type I R-M system; to date alleles conferring 11 different specificities have been identified (Barcus *et al.*, 1995).

target sequence, translocates the DNA from both sides towards itself in an ATP-dependent manner. DNA cleavage occurs when translocation is impeded (Studier & Bandyopadhyay, 1988; Janscak *et al.*, 1999a).

The three subunits of a type I R-M system are encoded by closely linked genes: *hsdR*, *hsdM* and *hsdS*. The acronym *hsd* denotes host specificity of DNA. *hsdM* and *hsdS* are transcribed from the same promoter; *hsdR* is from a separate one (Loenen *et al.*, 1987). The two subunits encoded by *hsdM* and *hsdS*, colloquially referred to as M and S, are both necessary and sufficient for methyltransferase activity. The third subunit, HsdR or R, is essential only for restriction. The specificity subunit, S, includes two target recognition domains (TRDs) that impart target-sequence specificity to the restriction and modification activities of the complex; the M subunits include the active site for DNA methylation and the R subunits that for nuclease activity. Two complexes are functional in bacterial cells: one comprises all three subunits ($R_2M_2S_1$) and is an R-M system, and the other lacks R ($M_2S_1$) and has only methyltransferase activity (Lautenberger & Linn, 1972; Suri & Bickle, 1985; Taylor *et al.*, 1992).

A separate promoter from which *hsdR* is transcribed suggests a means for regulating restriction activity, but experiments provide no evidence for the transcriptional regulation of any of those type I R-M systems for which data are available (Kulik & Bickle, 1996; Loenen *et al.*, 1987; Prakash-Cheng *et al.*, 1993). Evidence is accumulating for the role of post-translational regulation of restriction activity (see section on the mechanism by which restriction activity of *Eco*KI is controlled).

Type III R-M systems are less complex than type I systems but nevertheless share some similarities with them (see Rao *et al*., 2000). A single heterooligomeric complex catalyses both restriction and modification activities. Modification requires the cofactor AdoMet, and restriction requires $Mg^{2+}$ and ATP. Recent evidence indicates that type III restriction enzymes can translocate DNA. DNA cleavage is stimulated by collision of the translocating complexes and occurs close to, but on the 3′ side of, the target sequence (Meisel *et al*., 1995).

The foci of this review lecture are type I R-M systems, their extraordinary capacity for diversification and the acutely sensitive mechanisms for the control of their restriction activity: these mechanisms of control protect unmodified 'self' DNA from attack.

## Distribution of R-M systems

Commercial catalogues document the presence of type II R-M systems in a wide variety of bacterial strains. No 'market force' has driven searches for type I systems: nevertheless there is biological evidence for functional type I R-M systems in *Bacillus subtilis*, *Citrobacter freundii*, *Klebsiella pneumoniae*, *Lactococcus lactis*, *Mycoplasma pulmonis*, *Staphylococcus aureus* and many strains and species of *Salmonella* as well as those found in *E. coli* (see Murray, 2000). Computer-based analyses of the nucleotide sequences of bacterial genomes identify numerous putative R-M systems, of all types. Potential R-M genes within completed genomic sequences have recently been tabulated by Kong *et al*. (2000). Their survey indicates that >80% of the bacterial genomes for which completed sequences are available have at least one R-M system (see Table 1). Both type I and type II systems are prevalent throughout the Eubacteria and Archaea. It may be significant that strains for which screens of genomic sequences failed to identify putative R-M systems included those from very special environments, such as parasitic species in which bacterial growth may occur only within eukaryotic cells, e.g. *Chlamydia*, *Rickettsia* and *Treponema pallidum*, and *Aquifex aeolicus*, a thermophile that lives at extremely high temperatures.

## Diversification of sequence specificity

Type I R-M systems appear to be better suited to evolve new specificities than are the simpler type II systems. In summary, the following points seem relevant. First, the specificity of both the restriction and modification activities of a type I R-M complex is conferred by a single specificity subunit, S; therefore a change in specificity concomitantly affects restriction and modification. Second, those type I R-M systems that have been studied are sensitive to a sophisticated mechanism that controls their endonuclease activity, thereby protecting the resident chromosome from attack. Third, a specificity subunit that comprises two TRDs, each recognizing a different target sequence, offers more scope for diversification than a classical type II restriction endonuclease which, as a dimer of identical subunits, recognizes a symmetrical target sequence (Wilson & Murray, 1991).

Long repeated nucleotide sequences remain in the specificity genes of some type I R-M systems as evidence of gene duplication, providing an explanation for the origin of current specificity genes encoding two TRDs (Kannan *et al*., 1989). Early type I R-M systems with the subunit composition $R_2M_2S_2$ are likely to have recognized hyphenated symmetrical sequences, dictated by the symmetrical arrangement of two specificity subunits. Enzymes of this sort have been generated by deletions that truncate a specificity gene leading to an active enzyme comprising two symmetrically arranged truncated subunits (Abadjieva *et al*., 1993; Meister *et al*., 1993). Diversification of TRDs has led to the recognition of a variety of target sequences comprising 3–5 bp but always sequences within which an adenine residue is the substrate for methylation.

The evolution of a type I R-M system with a different specificity (see Fig. 3) was first witnessed by chance in the laboratory (Bullas *et al*., 1976) and later shown to be the result of recombination generating a hybrid S gene encoding a new combination of TRDs (Fuller-Pace *et al*., 1984). Similarly, a minor change in the length of the spacer sequence connecting the two TRDs was shown to alter the length of the spacer sequence separating the two components of the target sequence (Price *et al*., 1989). New combinations of TRDs can be generated experimentally quite readily, but attempts to generate new specificities as the result of changes within a TRD have been unsuccessful. The majority of many amino acid substitutions made within a TRD of *Eco*KI do not impair specificity (O'Neill *et al*., 2001). It seems likely that more than one amino acid substitution is necessary to change the specificity of a TRD. Even for type II R-M systems for which the structures of enzymes bound to their target sequences have been determined, it has not been possible to predict amino acid changes that lead to a new specificity. To date, *Bam*HI has been changed so that it prefers a methylated substrate (Dorner *et al*., 1999) and *Eco*RV has been engineered so that its preferred target sequence is 8 rather than 6 bp (Lanio *et al*., 2000).

## The immigration control region and the family concept

In *E. coli* K-12 the genes specifying *Eco*KI are flanked by genes that encode methylation-dependent restriction endonucleases (Mrr and McrBC). The segment of the genome that specifies these three endonucleases has been referred to as the immigration control region (Raleigh, 1992). Genetic analyses of other strains of *E. coli* and *Salmonella enterica* indicated considerable allelic diversity within, or close to, this region long before the era of genomic sequences (Boyer & Roulland-Dussoix, 1969; Bullas *et al*., 1980). *E. coli* strains K-12 and B, and *S. enterica* serovars *typhimurium* LT2 and *potsdam*, have alleles that specify type I R-M systems with

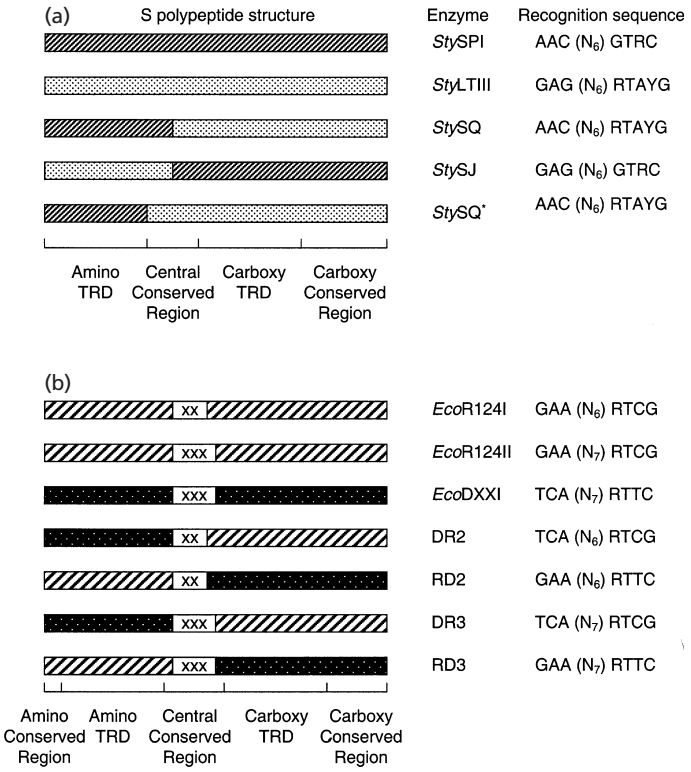


***Fig. 3.*** Evolution of type I R-M systems with new specificities. (a) Recombination between *hsdS* genes produces hybrid genes and chimeric S polypeptides. *Sty*SPI and *Sty*LTIII are naturally occurring type I R-M systems (see Table 2). *Sty*SQ and *Sty*SJ have hybrid *hsdS* genes (Fuller-Pace *et al.*, 1984; Gann *et al.*, 1987). The regions originating from *Sty*SPI are hatched and those originating from *Sty*LTIII are stippled. Reassortment of the TRDs accordingly gave rise to recombinant recognition sequences (Gann *et al.*, 1987; Nagaraja *et al.*, 1985). Site-directed mutagenesis of the central conserved region of the *Sty*SQ *hsdS* gene produced *Sty*SQ*, comprising only the amino-terminal variable region from *Sty*SPI and the remainder from *Sty*LTIII. The *Sty*SQ* target sequence confirms that the amino-terminal variable region is in fact a TRD responsible for recognition of the trinucleotide component of the sequence (Cowan *et al.*, 1989). (b) Sequence specificity may also be altered by changing the length of the nonspecific spacer of the target sequence. The S polypeptides of *Eco*RI24I and *Eco*RI24II differ only in the number of times a short amino acid motif (X = TAEL) is repeated within their central conserved regions (Price *et al.*, 1989), resulting in extension of the spacer in the target sequence from six nucleotides ($N_6$) for *Eco*RI24I to $N_7$ for *Eco*RI24II. The recognition sequence of *Eco*DXXI also contains a nonspecific spacer of 7 nt, corresponding to three TAEL repeats in its S polypeptide (Gubler *et al.*, 1992). Chimeric S polypeptides recognize the predicted target sequences (Gubler *et al.*, 1992). Modified with permission from a figure by Barcus & Murray (1995).

different specificities. The respective enzymes (*Eco*KI, *Eco*BI, *Sty*LTIII and *Sty*SPI) differ from each other in one or both of their TRDs (Gough & Murray, 1983; Fuller-Pace *et al.*, 1984). Of fundamental influence in our understanding of type I R-M systems has been the demonstration that these enzymes can be considered as members of a family within which the subunits of different enzymes are interchangeable. It came as a surprise, however, that alleles at this locus, in particular those specifying *Eco*AI in *E. coli* strain 15T$^-$, encode sufficiently dissimilar type I R-M systems to warrant their separation into a different family (Murray *et al.*, 1982). The initial evidence came from hybridization screens of bacterial DNAs and serological screens of bacterial extracts. As expected, the nucleotide sequences of *hsd* genes for *Eco*KI and *Eco*BI would hybridize to each other and antibodies raised against *Eco*KI reacted with *Eco*BI, but in contrast, DNA probes comprising the *Eco*KI genes failed to hybridize with those of *E. coli* 15T$^-$; similarly antibodies against *Eco*KI did not react with *Eco*AI. At least three families of type I R-M systems (IA, IB and ID) are encoded by alternative genes within the immigration control region of enteric bacteria (Fig. 4); currently these identify at least 16 specificities (Barcus *et al.*, 1995; Thorpe *et al.*, 1997; Titheradge *et al.*, 2001). The sequence of the genome of *E. coli* O157 (Perna *et al.*, 2001) identifies a type IB system.

An additional family (type IC), headed by *Eco*RI24I, was recognized initially via plasmid-encoded members (Glover *et al.*, 1983), but genes for a chromosomally encoded representative have been identified in an *E. coli* strain at a location distinct from the immigration control region (Tyndall *et al.*, 1994). While the only major difference between two enzymes within the same family is confined to their TRDs, the subunits of members of different families share only limited identity (15–35%) when their amino acid sequences are aligned. Sequence comparisons of the putative type I R-M systems predicted from genomic sequences suggest that family

***Table 2.*** Family-specific distance between target adenines

| Family | Enzyme | Recognition sequence* | Distance (bp)† | Reference |
|---|---|---|---|---|
| IA | *Eco*BI | TG**A**NNNNNNNN**T**GCT | 8 | Bickle (1987) |
| | *Eco*KI | A**A**CNNNNNNG**T**GC | 8 | Bickle (1987) |
| | *Eco*DI | TT**A**NNNNNNNG**T**CY | 8 | Bickle (1987) |
| | *Sty*LTIII | G**A**GNNNNNNR**T**AYG | 8 | Bickle (1987) |
| | *Sty*SPI | A**A**CNNNNNNG**T**RC | 8 | Bickle (1987) |
| IB | *Eco*AI | G**A**GNNNNNNNG**T**CA | 9 | Bickle (1987) |
| | *Eco*EI | G**A**GNNNNNNNA**T**GC | 9 | Cowan *et al.* (1989) |
| | *Cfr*AI | GC**A**NNNNNNNNG**T**GG | 9 | Kannan *et al.* (1989) |
| IC | *Eco*RI24I | GA**A**NNNNNNR**T**CG | 7 | Taylor *et al.* (1993); Price *et al.* (1987) |
| | *Eco*RI24IΔ | GA**A**NNNNNNN**T**TC | 7 | Abadjieva *et al.* (1993) |
| | *Eco*RI24II‡ | GA**A**NNNNNNNR**T**CG | 8 | Bickle (1987) |
| | *Eco*DXXI‡ | TC**A**NNNNNNNR**T**TC | 8 | Gubler *et al.* (1992) |
| | *Eco*prrI‡ | CC**A**NNNNNNNR**T**GC | 8 | Tyndall *et al.* (1994) |
| | *Eco*DXXIΔ‡ | TC**A**NNNNNNNN**T**GA | 8 | Meister *et al.* (1993) |
| ID | *Sty*SBLI | CG**A**NNNNNN**T**ACC | 6 | Titheradge *et al.* (2001) |

* N, any nucleotide; R, either purine; Y, either pyrimidine. Bold type identifies either the adenine that is the target for methylation or the thymine complementary to the target adenine. For *Eco*EI, *Cfr*AI and *Sty*SBLI the relevant adenine residues are not defined by experiments, but are the sole candidates within the target sequences.

† No. base pairs between target adenines.

‡ These type IC members have four more amino acids than *Eco*RI24I within the central conserved region, the region that links the TRDs.

affiliations extend across the Eubacterial kingdom (Titheradge *et al.*, 2001).

It seems likely that all type I R-M systems derive from a common ancestor (Sharp *et al.*, 1992), but systems allocated to different families are now so dissimilar that little evidence of homology remains at the level of gene sequences. One interesting exception is the 5′ part of the specificity genes of *Sty*LTIII (type IA) and *Eco*AI or *Eco*EI (type IB); these specify a TRD that recognizes the same trinucleotide target sequence (Table 2). An examination of the target sequences of the type I R-M systems (see Table 2) indicates that the evolution of different families of enzymes has enhanced the scope for diversification by varying the distance between those adenine residues within the target sequences that are the substrates for methylation. In the target sequences for members of the IB family, the adenine residues are separated by 9 bp, in the IA family by 8 bp, in the IC family by 7 or 8 bp and in type ID by only 6 bp; the variability in the IC family is dependent on whether a tetrapeptide sequence (TAEL) within the central conserved region is present in duplicate or in triplicate. The importance of the correct spacing between the adenine residues is illustrated by the target sequences for *Eco*RI24IΔ and *Eco*DXXIΔ (see Table 2). These are the systems that comprise symmetrically arranged truncated S polypeptides and their target sequences require an additional base pair in the spacer to maintain the distance between the adenine residues.

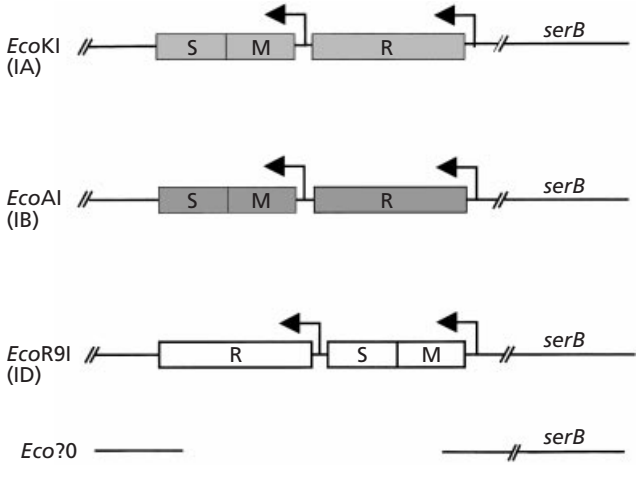


***Fig. 4.*** Alternatives at the *hsd* locus of *E. coli*. The diagrams identify the *hsd* genes within the immigration control region of *E. coli* K-12, *E. coli* 15T⁻, *E. coli* R9 and *E. coli* C.

In summary, diversity of specificity in type I systems, where two TRDs are present within the specificity subunit, not only depends on diversification of TRDs but is enhanced by different spacing between the TRDs and new combinations of TRDs.

## Is restriction an effective barrier to the acquisition of foreign DNA?

DNA in which the target sequences lack the correct identification mark is generally sensitive to restriction irrespective of whether the DNA enters the cell in single- or double-stranded form. Phage or plasmid DNA that enters in a single-stranded form becomes susceptible to restriction after the synthesis of the second strand. The fragmentation of foreign DNA reduces the efficiency of productive, or lysogenic, infection by phages, and the frequency of acquisition of conjugative plasmids. DNA fragments, particularly those that share sequence similarity with the resident chromosome, may be rescued by recombination. Early experiments in which gene transfer was monitored when unmodified donor DNA from an Hfr strain entered a restriction-proficient recipient showed the acquisition of early markers to be inefficient and linkage much reduced (Boyer, 1964; Pittard, 1964; Arber & Morse, 1965). However, many phage and conjugative plasmids, but not the well-known F factor, have the means of moderating their susceptibility to R-M systems. They may modify their DNA in unusual ways, or produce proteins that interfere with restriction, e.g. phage T7 or plasmid ColIb.

The modification of DNA by glucosylation, as in T-even phages, is effective against most restriction systems, while proteins that interfere with the activity of the enzyme may be specific to one enzyme, or one type of system. Phage T5 can inhibit the activity of *Eco*RI (Davison & Brunel, 1979) and a variety of host enzymes that modify DNA (see McCorquodale & Warner, 1988), but most of the anti-restriction functions currently identified are directed against type I systems. It seems unlikely that this bias towards functions that protect against type I systems simply reflects the fact that most work has been done with *E. coli* K-12 and *E. coli* B; *E. coli* strains specifying *Eco*RI have been in common laboratory use for 30 years. The bias could reflect the prevalence of type I systems in natural strains of *E. coli* enhanced by the fact that some feature common to type I R-M systems, or the conformation of their DNA substrates, permits the evolution of anti-restriction proteins that are able to combat all members of one family or even the members of different families of type I R-M systems.

The *0.3* gene products of phages T3 and T7 are the only anti-restriction functions available in significant quantities for detailed molecular analyses. These proteins, sometimes referred to as Ocr (overcoming classical restriction), bind to type I restriction complexes, both the R-M complex and the modification enzyme, and prevent them from binding to DNA (Atanasiu *et al*., 2001). The T3 product also destroys the cofactor AdoMet. It has been suggested that the *0.3* gene product, or Ocr, mimics the DNA substrate, thereby neutralizing the R-M complexes (Bandyopadhyay *et al*., 1985). Recent evidence based on the structure of Ocr supports this model: the protein is an elongated dimer that reflects both the size and shape of a bent DNA molecule (Atanasiu *et al*., 2001; M. Walkinshaw & D. Dryden, personal communication). An alternative proposal for the Ard (alleviation of restriction of DNA) proteins of conjugative plasmids, based on their acidic nature, is that an acidic surface mimics sequences of the specificity subunits of type I systems and the Ard proteins can displace the specificity subunit from the active R-M complex (Belogurov & Delver, 1995). Both the *0.3* gene product (C. Atanasiu & D. Dryden, personal communication) and ArdA (Read *et al*., 1992) are active against members of different families of type I R-M enzymes.

The efficacy of anti-restriction functions poses the critical question of how a protein specified by the unmodified DNA of a transmissible agent is able to act before the sequence that encodes it is attacked by the restriction enzyme. Bacteriophage P1 solves the problem by co-transfer of the protein with its DNA. In contrast, the *0.3* genes of T3 and T7 are transcribed early, prior to the internalization of the remainder of the genome. The *ard* genes of transmissible plasmids, like the *0.3* gene of T3 or T7, are located in the leading end of the DNA, but for conjugative plasmids it is single-stranded DNA that is transferred (5′ to 3′). Current evidence for IncI1 and ColIb supports a regulatory model in which the genes in the leading region of the DNA are transcribed from special promoters recognized within secondary structures of single-stranded DNA (Bates *et al*., 1999). This allows transcription of *ard* genes and the accumulation of anti-restriction protein before the transferred strand is converted into duplex DNA (see Fig. 5).

## Conserved sequences and active sites in type I R-M systems

While the S subunit confers sequence specificity to both the R-M ($R_2M_2S_1$) and modification ($M_2S_1$) complexes, the M subunit contributes the active site for modifi-

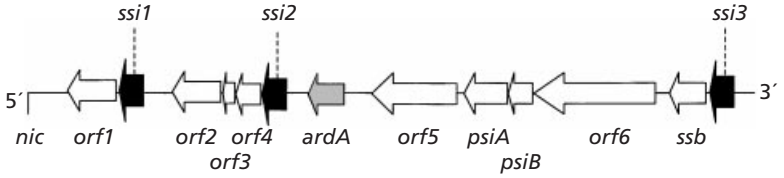


***Fig. 5.*** The ORFs in the leading region of ColIb. The direction of transfer from *nic* is from left to right. All ORFs (shown as arrows) are transcribed from right to left. *ard*A identifies the ORF specifying the anti-restriction protein. The regions identified as *ssi* are presumptive promoters for leftward transcription of the transferred strand of DNA. Reproduced with permission from Bates *et al*. (1999).

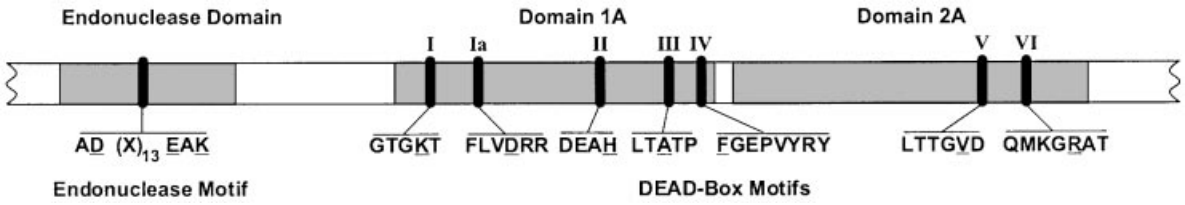


***Fig. 6.*** Domains, motifs and amino acid substitutions in the HsdR subunit of *Eco*KI. The N- and C-terminal regions are omitted. The two domains that include the DEAD-box motifs correlate with domains IA and 2A, as determined for structures of DNA helicases (see Davies *et al.*, 1999b). Substitutions for an underlined amino acid confer a restriction-deficient phenotype. These changes identify the restriction-deficient strains analysed for DNA translocation, ATPase and endonuclease activities. Reproduced with permission from Murray (2000).

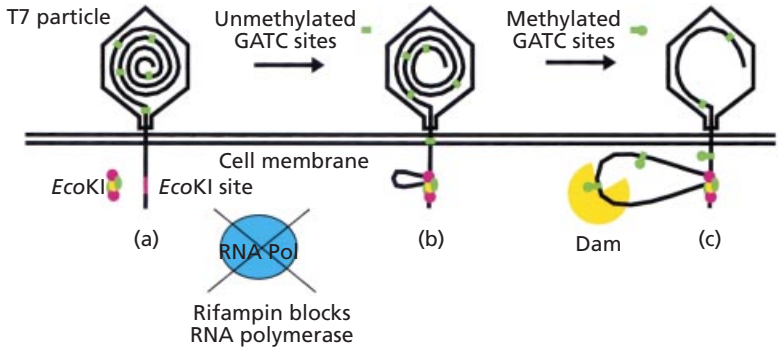


***Fig. 7.*** The experimental system in which DNA translocation by *Eco*KI was assayed by the transfer of phage T7 DNA into the bacterial cell (Garcia & Molineux, 1999; Davies *et al.*, 1999a). The methylation of the phage DNA by the Dam methylase of the recipient cell enables the identification of DNA within the cell by its susceptibility to *Dpn*I. The resulting fragments were identified on Southern transfers.

cation. Modification enzymes, whether type I, II or III, include motifs characteristic of methyltransferases. The type I systems transfer methyl groups to adenine residues and their M subunits include the sequence N/DPPF/Y/W as motif IV rather than the PC motif characteristic of cytosine methyltransferases (see Dryden, 1999). For *Eco*KI, amino acid substitutions within motif IV have been made that block the catalytic activity without impairing the binding of AdoMet, the methyl donor (Willcock *et al.*, 1994). In contrast, substitutions in motif I prevent binding of the methyl donor, a cofactor essential for restriction as well as modification.

The R subunits are essential for restriction but not modification. A type I R subunit includes motifs characteristic of ATP-binding proteins (Loenen *et al.*, 1987), consistent with the ATP-dependence of restriction. In addition, they include conserved sequences indicating the presence of motifs characteristic of ATP-dependent helicases (Gorbalenya & Koonin, 1991; Murray *et al.*, 1993; Titheradge *et al.*, 1996). It has been suggested that these motifs, the DEAD-box motifs, define an 'engine' that powers DNA translocation (Hall & Matson, 1999). Analyses of mutations in the *hsdR* gene of *E. coli* K-12 (Fig. 6) demonstrated that each of the seven DEAD-box motifs of *Eco*KI is essential for a restriction-proficient phenotype and for the DNA-dependent ATPase activity of the enzyme (Davies *et al.*, 1998, 1999a). Of special relevance was the finding that these restriction-deficient mutants lack DNA translocation activity (Davies *et al.*, 1999a). This activity was assayed by monitoring the *Eco*KI-dependent transfer of the T7 genome from the phage capsid to the bacterial cell (Fig. 7), an assay that relies on the inhibition of RNA polymerase activity, the normal means of DNA transfer, and the presence within the leading region of the T7 genome of a single target for *Eco*KI (Garcia & Molineux, 1999). The *Eco*KI complex bound to the unmodified *Eco*KI target can mobilize the 39 kb of T7 DNA at the rate of $\sim 100$ bp $s^{-1}$.

Additional conserved sequences in the N-terminal part of the R subunits of type I R-M systems (Titheradge *et al.*, 1996) show similarities with those motifs associated with DNA nicking in other nucleases (Davies *et al.*, 1999b). Site-directed mutagenesis proved the relevance of this motif to the endonuclease activity of *Eco*AI (Janscak *et al.*, 1999b) and *Eco*KI (Davies *et al.*, 1999a, b). Experiments *in vitro* for *Eco*AI (Janscak *et al.*, 1999b) and *in vivo* for *Eco*KI (Davies *et al.*, 1999a, b) showed that changes within the endonuclease motif do not block the ATPase and translocase activities of the R-M complex.

## Mechanism of action of type I restriction enzymes

Our current understanding of the mode of action of a type I restriction enzyme is essentially as outlined in the 'collision' model of Studier & Bandyopadhyay (1988). According to this model (Fig. 8), an enzyme binds to its target sequence and while remaining bound to this

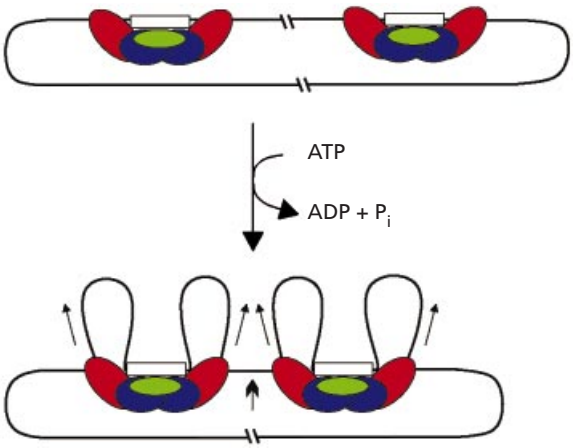


***Fig. 8.*** The model for DNA breakage described by Studier & Bandyopadhyay (1988). *Eco*KI bound to target sequences translocates DNA towards itself. Collision blocks translocation and stimulates the nicking of both DNA strands. Endonuclease activity may be stimulated when translocation is impeded by some other protein or structure (Janscak *et al.*, 1999a).

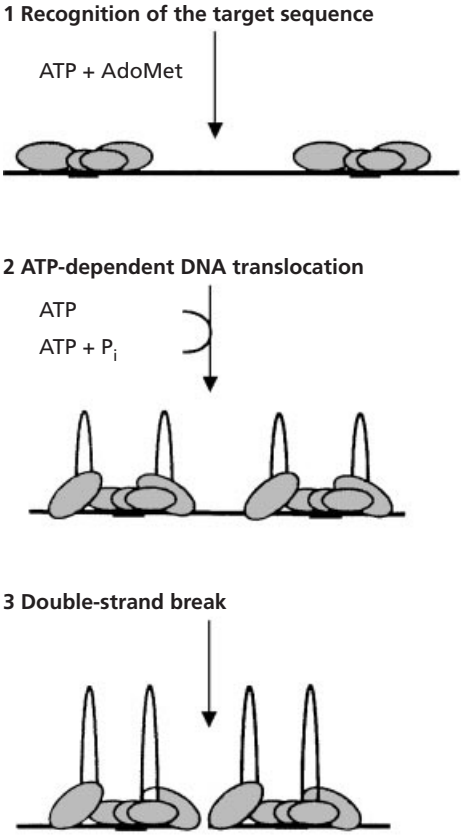


***Fig. 9.*** The restriction pathway. (1) The cofactors ATP and AdoMet are required for the specific binding of *Eco*KI to unmodified target sequences (Powell *et al.*, 1998). (2) ATP-dependent translocation is dependent on the DEAD-box motifs (Davies *et al.*, 1999a). Conservative substitutions in the endonuclease motif do not prevent ATP-dependent translocation (Davies *et al.*, 1999a). (3) Breakage of DNA is prevented by substitutions in the DEAD-box motifs as well as those in the endonuclease motifs. A mutation in *hsdM* that blocks methyltransferase activity but permits the binding of AdoMet does not block endonuclease activity (Doronina & Murray, 2001).

sequence it pulls in the DNA from both sides, simultaneously, in a process dependent upon the hydrolysis of ATP. When translocation is impeded, as for example by the collision of two translocating complexes, endonuclease activity is stimulated.

Representatives of three families of type I R-M systems have been studied *in vitro* (see Szcelkun, 2000). Each endonuclease is dependent upon AdoMet, ATP and $Mg^{2+}$, and all are believed to function in a similar way. For *Eco*KI, the addition of either ATP or a non-hydrolysable analogue in the presence of AdoMet allows tight binding of the enzyme to unmodified target sequences. DNA footprints demonstrate a conformational change that precedes the hydrolysis of ATP (Powell *et al.*, 1998). Enzymes with substitutions in DEAD-box motifs remain capable of the conformational change associated with target recognition, despite their failure to hydrolyse ATP and translocate DNA (Davies *et al.*, 1998, 1999a). Enzymes with conservative substitutions within the endonuclease motif retain their ability to translocate DNA, but these enzymes fail to hydrolyse phosphodiester bonds (Davies *et al.*, 1999a, b).

The *in vivo* and *in vitro* consequences of mutations in the *hsdR* gene of *E. coli* K-12 separate the restriction pathway into a series of steps in which AdoMet and ATP are required as cofactors for specific binding to the target sequence, while ATP hydrolysis is essential for the DNA translocation that precedes the eventual breakage of phosphodiester bonds in a $Mg^{2+}$-dependent reaction (Fig. 9). Known mutations in *hsdR* apparently fail to prevent the binding of ATP and they block either the second or the third step in the pathway. AdoMet binds to the M rather than the R subunit; a substitution in motif I of the M subunit of *Eco*KI, which prevents the binding of AdoMet (Willcock *et al.*, 1994), results in an enzyme incapable of either modification or restriction (Doronina & Murray, 2001). This defect is consistent with the predicted block in the first step of the restriction pathway. In contrast, a substitution in motif IV, which blocks methyltransferase activity but has little effect on the binding of AdoMet, leaves a complex able to translocate and break DNA. The expected consequence of this mutation *in vivo* would be fragmentation of the bacterial chromosome. However, recent experiments contradict this expectation (Makovets *et al.*, 1999; Doronina & Murray, 2001; Cromie & Leach, 2001). It would appear that when modification fails, the bacterial cell is endowed with the means of causing the restriction pathway to abort before the enzymes break the DNA. This effective control of the restriction activity of type I complexes is in stark contrast to the cell death that follows the concomitant loss of the genes that encode type II R-M system.

## Guarding the bacterial chromosome against DNA breakage

DNA modification marks and protects the chromosome of a restriction-proficient bacterium, but there are situations where unmodified targets could become exposed to a restriction enzyme and thereby jeopardise the integrity of the bacterial chromosome. An obvious example of this problem is encountered when a bacterium acquires genes that encode a different R-M system from any already present within the cell. One simple solution is to delay production of the restriction enzyme until the modification enzyme has had time to modify all the targets in the bacterial chromosome (Prakash-Cheng & Ryu, 1993). This process, however, takes many generations following the acquisition of the genes specifying *Eco*KI, because unmethylated DNA is a very poor substrate for modification (Makovets, 1999). For type II R-M systems, transcriptional control of gene expression is well documented (see Raleigh & Brooks, 1998), but transcriptional control has not been found to be relevant for any type I or type III system that has been investigated (Loenen *et al*., 1987; Prakash-Cheng *et al*., 1993; Kulik & Bickle, 1996; Redaschi & Bickle, 1996). The dependency of type II R-M systems on transcriptional regulation would explain why *E. coli* can cope with the acquisition of type II systems but is sensitive to their loss; following gene loss, transcriptional control is no longer possible and residual endonuclease will attack unmodified targets within the bacterial chromosome (Handa *et al*., 2000). The loss of genes encoding type I R-M systems is not associated with any loss of viability (O'Neill *et al*., 1997; Makovets *et al*., 1998). This may reflect loss of restriction activity by the dissociation of the R subunits of *Eco*KI to yield a complex ($M_2S_1$) with only modification activity.

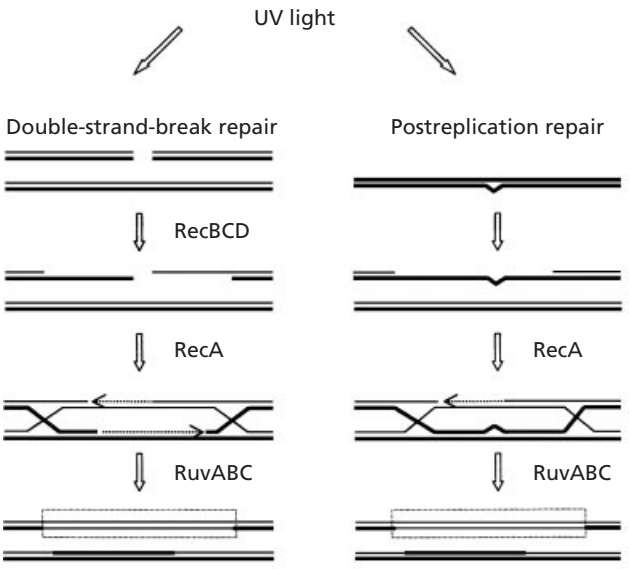


***Fig. 10.*** DNA damage can induce the alleviation of restriction. The diagram illustrates how unmodified target sequences could be generated following DNA damage. Methylated strands of DNA are shown as thick lines and unmethylated strands are shown as thin lines. Homologous recombination, involved in the repair of double-strand breaks or postreplicative repair, can generate regions of unmethylated double-stranded DNA via annealing of two unmethylated strands (regions within boxes). In addition, the SOS mutagenesis pathway leads to new (unmodified) target sequences as the result of base changes. 2-AP, a base analogue, is believed to create new target sequences as the result of base substitutions. Reproduced from Murray (2000) with permission.

The early experiments of Bertani & Weigle (1953) showed that the restriction proficiency of *E. coli* K-12 was alleviated following UV irradiation. Many experiments now document this phenomenon for type I systems, but not, so far, for any type II system. A similar response has been demonstrated for a variety of agents that damage DNA, including mutagens such as the base analogue 2-aminopurine (2-AP), and defects in some genes that affect DNA metabolism, e.g. *dam*, *topA* and *mutD* (*dnaQ*) (Efimova *et al*., 1988a, b; Thoms & Wackernagel, 1984; Makovets *et al*., 1999). DNA damage may generate unmodified target sequences as a consequence of the repair of double-strand breaks by homologous recombination (see Fig. 10), or directly by mutations that create target sequences. The original genetic evidence for the creation of vulnerable target sequences by mutation (Makovets *et al*., 1999) is now supported by the demonstration of breaks in the bacterial chromosome when *E. coli* K-12 is treated with 2-AP. The breaks are dependent on *Eco*KI and, as predicted if they arise by base substitutions, their generation requires two rounds of replication (Cromie & Leach, 2001).

Diversification of sequence specificity appears to be the hallmark of type I R-M systems and the control of restriction activity could facilitate the generation of new specificities. In *Mycoplasma pulmonis*, site-specific inversions of sequences within the specificity gene can 'switch' the sequence specificity of resident systems (Dybvig *et al*., 1998). This finding prompts the question of whether most cells that acquire an enzyme with a new specificity die, or whether the restriction potential of the new enzyme is controlled by a mechanism other than transcriptional regulation.

Finally, in the context of the generation of new specificities, it seems likely that the evolution of a TRD that recognizes a different nucleotide sequence will require a series of amino acid changes, some of which may initially impair the efficiency of modification. Our recent experiments show that even a modest drop in modification activity, one so small that the mutant strain still scores as modification proficient, elicits the modulation of restriction activity and this modulation is essential for the bacterium to survive (O'Neill *et al*., 2001). In this case a mutation in *hsdS* is associated with a restriction-deficient, modification-proficient phenotype!

## ClpX and ClpP are needed to modulate the restriction activity of some type I R-M systems

The efficient transmission of the genes encoding *Eco*KI requires some function specified by the recipient cell, if the recipient is modification deficient (Prakash-Cheng *et al*., 1993). Given the heterooligomeric nature of the R-M complex ($R_2M_2S_1$), an obvious way of alleviating the restriction activity within the recipient cells would be to destroy, or sequester, the R subunits of the complex. Energy-dependent proteases are known to play important regulatory roles in bacteria (see Gottesman, 1999), therefore mutants deficient in proteases were screened to check whether they might identify the unknown function. These experiments implicated the protease ClpXP (Makovets *et al*., 1998), which comprises two components, ClpX and ClpP. In the absence of either ClpX or ClpP, acquisition of *hsd* genes specifying either *Eco*KI (type IA) or *Eco*AI (type IB) led to the death of modification-deficient recipients (Makovets *et al*., 1998). Together, ClpX and ClpP form a large, but hollow, complex (see Gottesman, 1999); ClpX serves to recognize and unfold its substrate so that the polypeptide can be transported to the chamber within the complex where it becomes the target for degradation by ClpP. The alleviation of restriction in response to treatment with UV light, nalidixic acid or 2-AP, and to mutations in *dam*, *topA* or *mutD*, is dependent on ClpXP (Makovets *et al*., 1999). Similarly, survival of mutants in which methyltransferase activity is blocked (Makovets *et al*., 1999; Doronina & Murray, 2001), or even slightly impaired (O'Neill *et al*., 2001), requires ClpXP. The ClpXP protease provides a mechanism for controlling the restriction activity of type IA and IB systems, but it is not relevant to the control of all type I systems (see Murray, 2000).

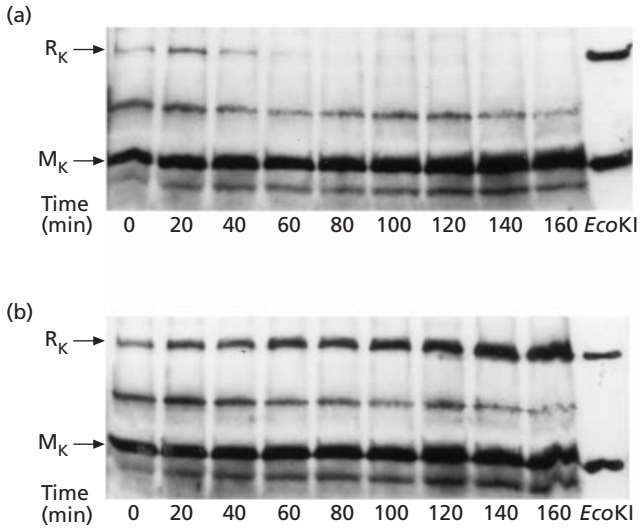


***Fig. 11.*** Treatment with 2-AP leads to a Clp-dependent deficiency of HsdR. The figure depicts a series of assays for HsdR and HsdM of *Eco*KI, following treatment with 2-AP. Panel (a) shows extracts from *clp*$^+$ bacteria, panel (b) from *clpX* bacteria. The polyclonal antibody used in the Western blots fails to detect HsdS, but detects some other *E. coli* proteins in addition to HsdR and HsdM. In the absence of 2-AP (data not shown), the assays for *clp*$^+$ and *clpX* bacteria were indistinguishable from those seen in (b). Taken with permission from Makovets *et al*. (1999).

## The mechanism by which the restriction activity of *Eco*KI is controlled

The fate of the subunits of *Eco*KI is readily monitored when restriction is alleviated in response to a DNA-damaging agent or because of a defect in modification activity. When *E. coli* was treated with 2-AP, a reduction in the concentration of the R polypeptide was observed if the cells were *clp*$^+$ but not if they were *clpX*$^-$ (see Fig. 11); the concentration of M appeared to be unaffected. In the absence of ClpX, the stability of the R polypeptide is enhanced in cells treated with 2-AP. These results are consistent with the activation of a control pathway in which R becomes susceptible to ClpXP-dependent proteolysis (Makovets *et al*., 1999). This susceptibility to proteolysis was found only when the R subunit was part of a functional restriction complex; neither a wild-type R subunit in the absence of M or S nor a defective R subunit in the presence of wild-type M and S was susceptible to proteolysis *in vivo*. These findings suggest that control of the restriction activity requires that the R-M complex can recognize its substrate and thereby embark on the restriction pathway. Such a mechanism would provide a remarkably specific control process that becomes effective only after the restriction pathway is initiated, but is able to act before damage is inflicted. An *Eco*KI complex with a substitution in motif IV of HsdM that blocks methyltransferase activity but not the endonuclease activity (Doronina & Murray, 2001), should initiate the restriction pathway on the resident DNA thereby making the R subunits vulnerable to ClpXP-dependent degradation. An examination of this mutant strain (*hsdM*F269G) revealed the predicted depletion of the wild-type R subunit, but depletion did not occur when the complex was impaired by a missense mutation in *hsdR* (Makovets *et al*., 1999). A modification-deficient *Eco*KI complex leaves a bacterial chromosome with around 600 unmodified target sequences. According to our model (Makovets *et al*., 1999), these targets will provide a powerful stimulus for the ClpXP-dependent alleviation of restriction by the degradation of R.

The available missense mutations in the *hsdR* gene of *E. coli* K-12 block either the ATP-dependent DNA translocation or the later step of DNA breakage. Both classes of mutants are defective in restriction. Are the R subunits of both classes of mutants refractory or susceptible to ClpXP-dependent proteolysis? A series of double mutants was made in which a mutation in *hsdR* was combined with the mutation (*hsdM*F269G) that pro-

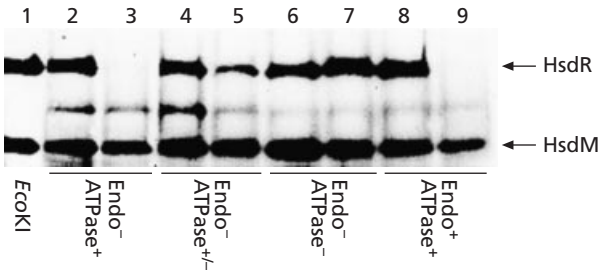


***Fig. 12.*** The effect of mutations in *hsdR* on the degradation of HsdR in response to a mutation in *hsdM* (substitution F269G) that blocks methyltransferase activity of the *Eco*KI complex. The bacteria in tracks 2, 4, 6 and 8 have no mutation in *hsdM*, and hence no stimulus to alleviate restriction, those in tracks 3, 5, 7 and 9 have the substitution F269G in HsdM. Degradation of HsdR correlates with the ATPase activity of the complex. Each of seven mutations that block ATPase activity, like that shown in track 7, prevents the degradation of HsdR; the two mutations that block endonuclease activity but have no effect on ATPase activity, like the one shown in track 3, had no effect on the degradation of HsdR (Doronina & Murray, 2001).

vokes degradation of HsdR. Each double mutant was monitored for the presence of the R subunit (Fig. 12). The R subunit of restriction-deficient mutants in which the ATP-dependent translocation activity was retained (Davies *et al*., 1999a) remained sensitive to proteolysis (as in track 3), but no depletion of R (see track 7) was observed in mutants where ATP-dependent translocation was blocked (Doronina & Murray, 2001).

The finding that ClpXP-dependent proteolysis protects the bacterial chromosome of *E. coli* K-12 from restriction in the complete absence of modification raises the following important question. Why do unmodified targets on the bacterial chromosome, but not those of infecting phage DNA, induce the alleviation of restriction? The classical view that modification is essential if a restriction system is to distinguish host DNA from foreign DNA is no longer tenable; apparently unmodified 'self' DNA is treated differently from unmodified foreign DNA. Host DNA may differ from invading DNA in its location, its association with other proteins and its topology. It seems improbable, however, that location alone will provide an adequate explanation. The DNA of both phage M13 and conjugative plasmids must enter the cytoplasm to be converted to a double-stranded form before it can be a substrate for restriction. M13 DNA and F factor DNA are recognized as foreign and restricted effectively in *clp*$^+$ and *clpX* cells (Doronina & Murray, 2001).

The role of ClpXP in the alleviation of restriction has been demonstrated for *Eco*AI (type IB) as well as for *Eco*KI (type IA). Members of the IC and ID families are also susceptible to restriction alleviation, but this may be dependent on an alternative mechanism (Makovets & Murray, unpublished observations).

## The effect of restriction on the acquisition of 'foreign' DNA

R-M systems in bacterial cultures are detected by their ability to restrict the acquisition of DNA from a different bacterial strain, or another bacterial species (Bertani & Weigle, 1953). It has been tempting to conclude that this biological phenomenon illustrates the role of R-M systems in nature, although attention has been drawn to the concept that DNA fragmentation by restriction endonucleases could potentiate recombination (S. Lederberg in Radding, 1973; Chang & Cohen, 1977; Price & Bickle, 1986; King & Murray, 1994; McKane & Milkman, 1995; Milkman *et al*., 1999; Kobayashi, 1998; Arber, 2000).

DNA molecules with ends are notoriously sensitive to degradation in *E. coli*; linear DNA fragments are degraded by a process dependent upon the ExoV activity of RecBCD (Simmon & Lederberg, 1972; see Telander-Muskavitch & Linn, 1981, for a review), the enzyme that catalyses an essential step in the major pathway of recombination in this bacterium. The degradation of $\lambda$ fragments by ExoV prevents detectable expression of those genes that are normally transcribed immediately after infection (Pilarski & Egan, 1973; Brammar *et al*., 1974). This implies an apparent conflict, or competition, between the alternative roles of RecBCD of either degrading the DNA fragments produced by a restriction system or rescuing them by recombination. Many experiments have shown that the DNA ends generated by cutting with *Eco*RI can serve to stimulate recombination, but these experiments often rely on recombination by alternative pathways under conditions in which DNA breakdown by the RecBCD nuclease is prevented (see, for example, Thaler *et al*., 1987; Eddy & Gold, 1992). Of more general relevance are experiments in which it was shown that DNA breakage by a type II restriction enzyme can stimulate RecBCD-mediated recombination in the presence of the nucleolytic activity of the wild-type enzyme (Stahl *et al*., 1983). In these experiments the recombination activity of RecBCD was assayed with a substrate that includes Chi, a specific nucleotide sequence of eight bases shown to be a hotspot for RecBCD-dependent recombination (see, for example, Kowalczykowski *et al*., 1994; Myers & Stahl, 1994; Smith *et al*., 1995; Kuzminov, 1999; Smith, 2001, for reviews). RecBCD enters a DNA molecule at an end. Genetic (Stahl *et al*., 1980) and biochemical (Taylor *et al*., 1985) evidence indicate that the Chi sequence must be oriented in the appropriate direction with respect to the approaching RecBCD enzyme, if it is to stimulate recombination. *In vitro*, the degradative behaviour of RecBCD prior to an encounter with a Chi sequence is dependent upon the relative concentrations of $Mg^{2+}$ and ATP (Ponticelli *et al*., 1985; Dixon & Kowalzcykowski, 1993). Hence, models for the mode of action of RecBCD *in vivo* differ; in one model both strands of DNA are degraded in the absence of a Chi sequence (see Myers & Stahl, 1994; Smith, 2001, for discussions of the models). It is, however, generally agreed that the presence of Chi

sites can impair the exonuclease activity of RecBCD and that this protective effect can be revealed *in trans* (Dabert *et al*., 1992; Kuzminov *et al*., 1994; Myers *et al*., 1995; Köppen *et al*., 1995; Taylor & Smith, 1999). A loss of ExoV activity in *E. coli* following the fragmentation of DNA that contains frequent Chi sites is consistent with the inactivation, or sequestration, of the RecD subunit (Köppen *et al*., 1995). *In vitro*, the RecBCD enzyme can disassemble into subunits following its encounter with a Chi sequence (Taylor & Smith, 1999).

In the chromosome of *E. coli* K-12, there is one Chi sequence per 4·6 kb (Blattner *et al*., 1997), roughly seven times more often than expected from a random association of nucleotides within the genome and 4- to 14-fold higher than in the DNA of seven non-enteric bacteria whose complete nucleotide sequences were analysed (Colbert *et al*., 1998). The frequency of Chi sequences is influenced by codon usage (Biaudet *et al*., 1998; Colbert *et al*., 1998). Chi sequences are predominantly within ORFs and predominantly oriented so that they will protect DNA from degradation should this proceed towards the origin of replication (Burland *et al*., 1993; Kuzminov *et al*., 1994; Blattner *et al*., 1997). The RecBCD enzyme of all enteric bacteria that have been tested uses the Chi sequence of *E. coli* K-12 (see Colbert *et al*., 1998). Other groups of related bacteria may have a functionally equivalent system in which the enzyme recognizes a different nucleotide sequence (Chedin *et al*., 2000). Chi, or an analogue, is likely to enhance the rescue of DNA from closely related bacteria. Even so, sequence divergence between members of close genera, e.g. *Escherichia* and *Salmonella*, can be sufficient to significantly limit genetic exchange (Matic *et al*., 1996). DNA fragments provoked by the R-M systems found within the same bacterial species are likely to be salvaged more efficiently than those generated within a bacterium from another genus, primarily because of sequence similarity, but aided perhaps by high frequencies of Chi sequences. These sequences, or their equivalents, should serve to stimulate recombination and, if present in abundance within the fragmented DNA, could convert a cell into an ExoV-deficient phenocopy that remains recombination proficient.

Some phages (e.g. T7 and P1) and many conjugative plasmids, as already mentioned, encode proteins that antagonize R-M systems. The F factor of *E. coli*, however, appears to lack an anti-restriction gene, and chromosomal DNA acquired by courtesy of an F factor is susceptible to restriction. DNA breakage reduces the linkage between markers transferred during conjugation (Pittard, 1964).

R-M systems seem likely to affect the flux of genetic information. DNA breakage followed by exonuclease activity may enhance the opportunity for the acquisition, and retention, of advantageous coding sequences in the absence of neighbouring deleterious ones (Milkman *et al*., 1999). This modulation of DNA transfer seems unlikely to provide the selective force for the allelic diversity detected for type I R-M systems in one species of enteric bacteria.

## Questions concerning the biological relevance of R-M systems

Most recently Kobayashi and colleagues have championed the case for R-M systems as 'selfish, mobile, genetic elements' (Kobayshi, 1998, 2001). The central theme for the premise that R-M systems are selfish elements rests on the finding that, under a variety of circumstances, the presence of an R-M system can lead to breakage of the bacterial chromosome and, consequently, to cell death. A particularly well-documented case of cell death follows the loss of genes specifying the type II R-M system *Eco*RI. The loss of R-M genes may be associated with the loss of a plasmid, or it may reflect the replacement of chromosomally encoded genes by recombination. Irrespective of the mechanism by which the genes specifying *Eco*RI are lost, cell growth leads to progeny that retain some active endonuclease at a time when they are no longer able to modify all the target sequences in their newly replicated DNA, hence the bacterial chromosome becomes the substrate for the residual endonuclease (Handa *et al*., 2000). No such susceptibility has been detected for strains specifying type I R-M systems (Kulik & Bickle, 1996; O'Neill *et al*., 1997); modulation of the restriction activity of type I R-M systems is extraordinarily effective in the protection of the bacterial chromosome. When the genes encoding *Eco*KI are deleted no viability problem is detected, even in the absence of the ClpXP protease (O'Neill *et al*., 1997; Makovets *et al*., 1998), and when the genes are replaced with those specifying another system, ClpXP alleviates restriction and permits survival. Furthermore, in contradiction to classical expectations, the presence of a mutation that destroys the modification activity of the *Eco*KI complex is not lethal: the restriction-proficient cells survive because ClpXP controls the endonuclease activity of the modification-deficient complex (Makovets *et al*., 1999; Doronina & Murray, 2001). Control by ClpXP was found to be essential for the survival of a cell in which only the balance between modification and restriction activities of the *Eco*KI complex was changed (O'Neill *et al*., 2001).

Two obvious questions arise about the mechanism and relevance of the alleviation of restriction by type I R-M systems. First, how are unmodified sequences in the resident bacterial chromosome distinguished from those in DNA that has recently entered the bacterial cell? Second, why do some, perhaps all, type I R-M systems have such elegant and sensitive mechanisms to control their activity and prevent cell death, while the genes for type II systems are apparently maintained by their failure to control endonuclease activity?

The translocation step in the complex restriction pathway of a type I system extends the opportunity for the bacterial proteins to counter-attack an R-M complex active on the resident chromosome. Perhaps the answer to the first question is simply that the translocation process on the resident chromosome is hindered by the nature of the bacterial nucleoid and this in turn increases

the opportunity for recognition by the ClpXP protease, or any alternative control system.

Experiments *in vivo*, using phage DNA substrates for restriction by *Eco*KI (Brammar *et al*., 1974; Garcia & Molineux, 1999), support the model (Studier & Bandyopadhyay, 1988) in which cutting occurs between two target sequences when the translocating complexes collide. It is not known whether any feature of the structure or organization of the nucleoid, or any process such as DNA replication, would either reduce the speed of DNA translocation, or alternatively halt translocation and stimulate endonuclease activity. *Eco*KI is very effective at displacing a repressor bound to its target sequence (Dreier *et al*., 1996). Therefore, collision with a protein that has a high affinity for its target sequence neither prevents translocation nor stimulates DNA breakage. *In vitro*, a fixed Holliday junction has been shown to stimulate cutting. Therefore one protein complex is sufficient to break the phosphodiester bonds in both strands of the duplex (Taylor & Smith 1990; Janscak *et al*., 1999a).

It is difficult to speculate about the fragmentation of the bacterial chromosome by type I R-M systems without knowing whether any events *in vivo*, other than the collision between translocating complexes, trigger DNA breakage. The spacing between unmodified target sequences is not obviously relevant. In some instances, as in response to treatment with 2-AP, ClpXP-dependent alleviation of restriction occurs when relatively few targets are unmodified while in others, such as the acquisition of R-M genes or the presence of a modification-deficient *Eco*KI complex, all or most of the genomic target sequences will be exposed. The only modification-deficient complex studied *in vitro* does, however, act more slowly than the wild-type enzyme (Doronina & Murray, 2001).

The behaviour of recipient bacteria following conjugation could be interpreted as support for a distinction between the nucleoid and other DNA. Unmodified DNA entering the cell by conjugation is recognized as foreign and attacked, but within 40 min of the time of entry restriction, assessed by infection with unmodified $\lambda$, is alleviated (Glover & Colson, 1965). This alleviation of restriction was found to be ClpXP-dependent, and was not detected in a *recA* recipient (Doronina & Murray, unpublished observations). These observations are explained if fragmented donor DNA must be incorporated into the resident chromosome by recombination before unmodified DNA is identified as 'self', and can evoke the ClpXP-dependent alleviation of restriction. Alternatively, RecA protein itself could be necessary for activation of the alleviation pathway. RecA is necessary for the alleviation of restriction in response to UV irradiation (Thoms & Wackernagel, 1984; Salaj-Šmic *et al*., 1997) but it is not necessary for the alleviation of restriction in response to treatment with 2-AP (Makovets *et al*., 1999).

The mechanism of distinction between unmodified 'self' and unmodified 'foreign' DNA should be susceptible to analysis. However, the biological relevance of the distinctly different behaviours of type I and type II R-M systems may be more difficult to determine: the differences caution against generalized speculations for the evolutionary strategies of R-M systems.

## Acknowledgements

The award of the Fred Griffith Review Lecture at the end of my career provides an appropriate opportunity for me to thank the many people (students, teachers, colleagues in Edinburgh and elsewhere) who have catalysed and sustained my interest in microbial genetics since 1956. Special thanks are due to former members of my group including, of course, those whose thoughts and experiments have contributed to the work reported in this review lecture, also to Alexander Gann, Kenneth Murray, Gerry Smith and Frank Stahl for their constructive comments on the manuscript. I am grateful to Alix Fraser for the preparation of the manuscript, and to the Medical Research Council for their support throughout my time in Edinburgh (1968–2001).

## References

**Abadjieva, A., Patel, J., Webb, M., Zinkevich, V. & Firman, K. (1993).** A deletion mutant of the type IC restriction endonuclease *Eco*R1241 expressing a novel DNA specificity. *Nucleic Acids Res* **21**, 4435–4443.

**Atanasiu, C., Byron, O., McMiken, H., Sturrock, S. S. & Dryden, D. T. F. (2001).** Characterisation of the structure of ocr, the gene 0.3 protein of bacteriophage T7. *Nucleic Acids Res* **29**, 3059–3068.

**Arber, W. (2000).** Genetic variation: molecular mechanisms and impact on microbial evolution. *FEMS Microbiol Rev* **24**, 1–7.

**Arber, W. & Dussoix, D. (1962).** Host specificity of DNA produced by *Escherichia coli*. I. Host controlled modification of bacteriophage lambda. *J Mol Biol* **5**, 18–36.

**Arber, W. & Morse, M. L. (1965).** Host specificity of DNA produced by *Escherichia coli*. VI. Effects on bacterial conjugation. *Genetics* **51**, 137–148.

**Bandyopadhyay, P. K., Studier, F. W., Hamilton, D. L. & Yuan, R. (1985).** Inhibition of the type I restriction-modification enzymes *Eco*B and *Eco*K by the gene 0.3 protein of bacteriophage T7. *J Mol Biol* **182**, 567–578.

**Barcus, V. A. & Murray, N. E. (1995).** Barriers to recombination: restriction. In *Population Genetics of Bacteria* (Society for General Microbiology symposium no. 52), pp. 31–58. Edited by S. Baumberg, J. P. W. Young, E. M. H. Wellington & J. R. Saunders. Cambridge: Cambridge University Press.

**Barcus, V. A., Titheradge, A. J. & Murray, N. E. (1995).** The diversity of alleles at the *hsd* locus in natural populations of *Escherichia coli*. *Genetics* **140**, 1187–1197.

**Bates, S., Roscoe, R. A., Althorpe, N. J., Brammar, W. J. & Wilkins, B. M. (1999).** Expression of leading region genes on IncI1 plasmid ColIb-P9: genetic evidence for single-stranded DNA transcription. *Microbiology* **145**, 2655–2662.

**Belogurov, A. A. & Delver, E. P. (1995).** A motif conserved among the type I restriction-modification enzymes and antirestriction proteins: a possible basis for mechanism of action of plasmid-encoded antirestriction functions. *Nucleic Acids Res* **23**, 785–787.

**Bertani, G. & Weigle, J. J. (1953).** Host controlled variation in bacterial viruses. *J Bacteriol* **65**, 113–121.

**Biaudet, V., El Karoui, M. & Gruss, A. (1998).** Codon usage can explain GT-rich islands surrounding Chi sites on the *Escherichia coli* genome. *Mol Microbiol* **29**, 666–669.

**Bickle, T. A. (1987).** Restriction and modification systems. In *Escherichia coli and Salmonella*, 1st edn, pp. 692–696. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.

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

**Blattner, F. R., Plunkett, G., III, Bloch, C. A & 14 other authors (1997).** The complete genome sequence of *Escherichia coli* K-12. *Science* **277**, 1453–1474.

**Boyer, H. (1964).** Genetic control of restriction and modification in *Escherichia coli*. *J Bacteriol* **88**, 1652–1660.

**Boyer, H. W. & Roulland-Dussoix, D. (1969).** A complementation analysis of the restriction and modification of DNA in *Escherichia coli*. *J Mol Biol* **41**, 459–472.

**Brammar, W. J., Murray, N. E. & Winton, S. (1974).** Restriction of *λtrp* bacteriophages by *Escherichia coli* K. *J Mol Biol* **90**, 633–647.

**Bullas, L. R., Colson, C. & Van Pel, A. (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.

**Bullas, L. R., Colson, C. & Neufeld, B. (1980).** Deoxyribonucleic acid restriction and modification systems in *Salmonella*: chromosomally located systems of different serotypes. *J Bacteriol* **141**, 275–292.

**Burland, V., Plunkett, G., III, Daniels, D. L. & Blattner, F. R. (1993).** DNA sequence and analysis of 136 kilobases of the *Escherichia coli* genome: organizational symmetry around the origin of replication. *Genomics* **16**, 551–561.

**Chang, A. C. Y. & Cohen, S. N. (1977).** *In vivo* site-specific genetic recombination promoted by the *Eco*RI restriction endonuclease. *Proc Natl Acad Sci U S A* **74**, 4811–4815.

**Chedin, F., Ehrlich, S. D. & Kowalczykowski, S. C. (2000).** The *Bacillus subtilis* AddAB helicase/nuclease is regulated by its cognate Chi sequence *in vitro*. *J Mol Biol* **298**, 7–20.

**Colbert, T., Taylor, A. F. & Smith, G. R. (1998).** Genomics, Chi sites and codons: islands of preferred DNA pairing are oceans of ORFs. *Trends Genet* **14**, 485–488.

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

**Cromie, G. A. & Leach, D. R. (2001).** Recombination repair of chromosomal DNA double-strand breaks generated by a restriction endonuclease. *Mol Microbiol* **41**, 873–884.

**Dabert, P., Ehrlich, S. D. & Gruss, A. (1992).** $\chi$ sequence protects against RecBCD degradation of DNA *in vivo*. *Proc Natl Acad Sci U S A* **89**, 12073–12077.

**Davies, G. P., Powell, L. M., Webb, J. L., Cooper, L. P. & Murray, N. E. (1998).** *Eco*KI with an amino acid substitution in any one of seven DEAD-box motifs has impaired ATPase and endonuclease activities. *Nucleic Acids Res* **26**, 4828–4836.

**Davies, G. P., Kemp, P., Molineux, I. J. & Murray, N. E. (1999a).** The DNA translocation and ATPase activities of restriction-deficient mutants of *Eco*KI. *J Mol Biol* **292**, 787–796.

**Davies, G. P., Martin, L., Sturrock, S. S., Cronshaw, A., Murray, N. E. & Dryden, D. T. (1999b).** On the structure and operation of type I DNA restriction enzymes. *J Mol Biol* **290**, 565–579.

**Davison, J. & Brunel, F. (1979).** Restriction insensitivity in bacteriophage T5. I. Genetic characterization of mutants sensitive to *Eco*RI restriction. *J Virol* **29**, 11–16.

**Dixon, D. A. & Kowalczykowski, S. C. (1993).** The recombination hotspot $\chi$ is a regulatory sequence that acts by attenuating the nuclease activity of the *E. coli* RecBCD enzyme. *Cell* **73**, 87–96.

**Dorner, L. F., Bitinaite, J., Whitaker, R. D. & Schildkraut, I. (1999).** Genetic analysis of the base-specific contacts of *Bam*HI restriction endonuclease. *J Mol Biol* **285**, 1515–1523.

**Doronina, V. A. & Murray, N. E. (2001).** The proteolytic control of restriction activity in *Escherichia coli* K-12. *Mol Microbiol* **39**, 416–428.

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

**Dryden, D. T. F. (1999).** Bacterial DNA methyltransferases. In *S-adenosylmethionine-Dependent Methyltransferases: Structure and Function*, pp. 283–340. Edited by X. Cheng & R. M. Blumenthal. River Edge, NJ: World Scientific Publishing.

**Dryden, D. T. F., Murray, N. E. & Rao, D. N. (2001).** Nucleoside triphosphate-dependent restriction enzymes. *Nucleic Acids Res* **29**, 3728–3741.

**Dybvig, K., Sitaraman, R. & French, C. T. (1998).** A family of phase-variable restriction enzymes with differing specificities generated by high-frequency gene rearrangements. *Proc Natl Acad Sci U S A* **95**, 13923–13928.

**Eddy, S. R. & Gold, L. (1992).** The DNA restriction endonuclease of *Escherichia coli* B. *J Biol Chem* **260**, 5729–5738.

**Efimova, E. P., Delver, E. P. & Belogurov, A. A. (1988a).** Alleviation of type I restriction in adenine methylase (*dam*) mutants of *Escherichia coli*. *Mol Gen Genet* **214**, 313–316.

**Efimova, E. P., Delver, E. P. & Belogurov, A. A. (1988b).** 2-Aminopurine and 5-bromouracil induce alleviation of type I restriction in *Escherichia coli*: mismatches function as inducing signals? *Mol Gen Genet* **214**, 317–320.

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

**Gann, A. A., Campbell, A. J., Collins, J. F., Coulson, A. F. & Murray, N. E. (1987).** Reassortment of DNA recognition domains and the evolution of new specificities. *Mol Microbiol* **1**, 13–22.

**Garcia, L. R. & Molineux, I. J. (1999).** Translocation and specific cleavage of bacteriophage T7 DNA *in vivo* by *Eco*KI. *Proc Natl Acad Sci U S A* **96**, 12430–12435.

**Glover, S. W. & Colson, S. (1965).** The breakdown of the restriction mechanism in zygotes of *Escherichia coli*. *Genet Res* **6**, 153–155.

**Glover, S. W., Firman, K., Watson, G. & Price, C. (1983).** The alternative expression of two restriction and modification systems. *Mol Gen Genetics* **190**, 65–69.

**Gorbalenya, A. E. & Koonin, E. V. (1991).** Endonuclease (R) subunits of type-I and type-III restriction-modification enzymes contain a helicase-like domain. *FEBS Lett* **291**, 277–281.

**Gottesman, S. (1999).** Regulation by proteolysis: developmental switches. *Curr Opin Microbiol* **2**, 142–147.

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

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

**Hall, M. C. & Matson, S. W. (1999).** Helicase motifs: the engine that powers DNA unwinding. *Mol Microbiol* **34**, 867–877.

**Handa, N., Ichige, A., Kusano, K. & Kobayashi, I. (2000).** Cellular responses to postsegrational killing by restriction-modification genes. *J Bacteriol* **182**, 2218–2229.

**Ives, C. L., Nathan, P. D. & Brooks, J. E. (1992).** Regulation of the *Bam*HI restriction-modification system by a small intergenic open reading frame, *bamHIC*, in both *Escherichia coli* and *Bacillus subtilis*. *J Bacteriol* **174**, 7194–7201.

**Ives, C. L., Sohail, A. & Brooks, J. E. (1995).** The regulatory C proteins from different restriction-modification systems can cross-complement. *J Bacteriol* **177**, 6313–6315.

**Janscak, P., MacWilliams, M. P., Sandmeier, U., Nagaraja, V. & Bickle, T. A. (1999a).** DNA translocation blockage, a general mechanism of cleavage site selection by type I restriction enzymes. *EMBO J* **18**, 2638–2647.

**Janscak, P., Sandmeier, U. & Bickle, T. A. (1999b).** Single amino acid substitutions in the HsdR subunit of the type IB restriction enzyme *Eco*AI uncouple the DNA translocation and DNA cleavage activities of the enzyme. *Nucleic Acids Res* **27**, 2638–2643.

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

**King, G. & Murray, N. E. (1994).** Restriction enzymes in cells, not Eppendorfs. *Trends Microbiol* **2**, 465–469.

**Kobayashi, I. (1998).** Selfishness and death: raison d'etre of restriction, recombination and mitochondria. *Trends Genet* **14**, 368–374.

**Kobayashi, I. (2001).** Behaviour of restriction-modification systems as selfish mobile elements and its relationship with genome evolution. *Nucleic Acids Res* **29**, 3742–3756.

**Kong, H., Lin, L.-F., Porter, N., Stickel, S., Byrd, D., Posfai, J. & Roberts, R. J. (2000).** Functional analysis of putative restriction-modification system genes in the *Helicobacter pylori* J99 genome. *Nucleic Acids Res* **28**, 3216–3223.

**Köppen, A., Krobitsch, S., Thoms, B. & Wackernagel, W. (1995).** Interaction with the recombination hot spot $\chi$ *in vivo* converts the RecBCD enzyme of *Escherichia coli* into a $\chi$-independent recombinase by inactivation of the RecD subunit. *Proc Natl Acad Sci U S A* **92**, 6249–6253.

**Kowalczykowski, S. C., Dixon, D. A., Eggleston, A. K., Lauder, S. D. & Rehrauer, W. M. (1994).** Biochemistry of homologous recombination in *Escherichia coli*. *Microbiol Rev* **58**, 401–465.

**Kulik, E. M. & Bickle, T. A. (1996).** Regulation of the activity of the type IC *Eco*R124I restriction enzyme. *J Mol Biol* **264**, 891–906.

**Kuzminov, A. (1999).** Recombinational repair of DNA damage in *Escherichia coli* and bacteriophage lambda. *Microbiol Mol Biol Rev* **63**, 751–813.

**Kuzminov, A., Schabtach, E. & Stahl, F. W. (1994).** $\chi$ sites in combination with RecA protein increase the survival of linear DNA in *Escherichia coli* by inactivating ExoV activity of RecBCD nuclease. *EMBO J* **13**, 2764–2776.

**Lanio, T., Jeltsch, A. & Pingoud, A. (2000).** On the possibilities and limitations of rational protein design to expand the specificity of restriction enzymes: a case study employing *Eco*RV as the target. *Protein Enz* **13**, 275–281.

**Lautenberger, J. A. & Linn, S. (1972).** The deoxyribonucleic acid modification and restriction enzymes of *Escherichia coli* B. I. Purification, subunit structure, and catalytic properties of the modification methylase. *J Biol Chem* **247**, 6176–6182.

**Linn, S. & Arber, W. (1968).** Host specificity of DNA produced by *Escherichia coli*. X. *In vitro* restriction of phage fd replication form. *Proc Natl Acad Sci U S A* **59**, 1300.

**Loenen, W. A., Daniel, A. S., Braymer, H. D. & Murray, N. E. (1987).** Organization and sequence of the *hsd* genes of *Escherichia coli* K-12. *J Mol Biol* **198**, 159–170.

**McCorquodale, J. D. & Warner, H. R. (1988).** Bacteriophage T5 and related phages. In *The Bacteriophages*, pp. 439–476. Edited by R. Calendar. New York: Plenum.

**McKane, M. & Milkman, R. (1995).** Transduction, restriction and recombination patterns in *Escherichia coli*. *Genetics* **139**, 35–43.

**Makovets, S., Titheradge, A. J. B. & Murray, N. E. (1998).** ClpX and ClpP are essential for the efficient acquisition of genes specifying type IA and IB restriction systems. *Mol Microbiol* **28**, 25–35.

**Makovets, S., Doronina, V. A. & Murray, N. E. (1999).** Regulation of endonuclease activity by proteolysis prevents breakage of unmodified bacterial chromosomes by type I restriction enzymes. *Proc Natl Acad Sci U S A* **96**, 9757–9762.

**Matic, I., Taddei, F. & Radman, M. (1996).** Genetic barriers among bacteria. *Trends Microbiol* **4**, 69–72.

**Meisel, A., Mackeldanz, P., Bickle, T. A., Krüger, D. V. & Schroeder, C. (1995).** Type III restriction endonucleases translocate DNA in a reaction driven by recognition site-specific ATP hydrolysis. *EMBO J* **14**, 2958–2966.

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

**Meselson, M. & Yuan, R. (1968).** DNA restriction enzyme from *E. coli*. *Nature* **217**, 1110–1114.

**Milkman, R., Raleigh, E. A., McKane, M., Cryderman, D., Bilodeau, P. & McWeeny, K. (1999).** Molecular evolution of the *Escherichia coli* chromosome. V. Recombination patterns among strains of diverse origin. *Genetics* **153**, 539–554.

**Murray, N. E. (2000).** Type I restriction systems: sophisticated molecular machines (a legacy of Bertani and Weigle). *Microbiol Mol Biol Rev* **64**, 412–434.

**Murray, N. E., Gough, J. A., Suri, B. & Bickle, T. A. (1982).** Structural homologies among type I restriction–modification systems. *EMBO J* **1**, 535–539.

**Murray, N. E., Daniel, A. S., Cowan, G. M. & Sharp, P. M. (1993).** Conservation of motifs within the unusually variable polypeptide sequences of type I restriction and modification enzymes. *Mol Microbiol* **9**, 133–143.

**Myers, R. S. & Stahl, F. W. (1994).** $\chi$ and the RecBCD enzyme of *Escherichia coli*. *Annu Rev Genet* **28**, 49–70.

**Myers, R. S., Kuzminov, A. & Stahl, F. W. (1995).** The recombination hot spot $\chi$ activates RecBCD recombination by converting *Escherichia coli* to a *recD* mutant phenocopy. *Proc Natl Acad Sci U S A* **92**, 6244–6248.

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

**Naito, T., Kusano, K. & Kobayashi, I. (1995).** Selfish behaviour of restriction-modification systems. *Science* **267**, 897–899.

**O'Neill, M., Chen, A. & Murray, N. E. (1997).** The restriction-modification genes of *Escherichia coli* K-12 may not be selfish: they do not resist loss and are readily replaced by alleles conferring different specificities. *Proc Natl Acad Sci U S A* **94**, 14596–14601.

**O'Neill, M., Powell, L. & Murray, N. E. (2001).** Target recognition by *Eco*KI: the recognition domain is robust and restriction-deficiency commonly results from the proteolytic control of enzyme activity. *J Mol Biol* **307**, 951–963.

**Perna, N. T., Plunkett, G., III, Burland, V. & 25 other authors (2001).** Genome sequence of enterohaemorrhagic *Escherichia coli* O157 : H7. *Nature* **409**, 529–533.

**Pilarski, L. M. & Egan, J. B. (1973).** Role of DNA topology in transcription of coliphage $\lambda$ *in vivo*: M DNA topology protects the template from exonuclease attack. *J Mol Biol* **76**, 257–266.

**Pingoud, A. & Jeltsch, A. (2001).** Structure and function of type II restriction endonucleases. *Nucleic Acids Res* **29**, 3705–3727.

**Pittard, J. (1964).** Effect of phage-controlled restriction on genetic linkage in bacterial crosses. *J Bacteriol* **87**, 1256–1257.

**Ponticelli, A. S., Schultz, D. W., Taylor, A. F. & Smith, G. R. (1985).** Chi-dependent DNA strand cleavage by RecBC enzyme. *Cell* **41**, 145–151.

**Powell, L. M., Dryden, D. T. & Murray, N. E. (1998).** Sequence-specific DNA binding by *Eco*KI, a type IA DNA restriction enzyme. *J Mol Biol* **283**, 963–976.

**Prakash-Cheng, A. & Ryu, J. (1993).** Delayed expression of in vivo restriction activity following conjugal transfer of *Escherichia coli hsdK* (restriction-modification) genes. *J Bacteriol* **175**, 4905–4906.

**Prakash-Cheng, A., Chung, S. S. & Ryu, J. (1993).** The expression and regulation of *hsdK* genes after conjugative transfer. *Mol Gen Genet* **241**, 491–496.

**Price, C. & Bickle, T. A. (1986).** A possible role for DNA restriction in bacterial evolution. *Microbiol Sci* **3**, 296–299.

**Price, C., Pripfl, T. & Bickle, T. A. (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.

**Price, C., Lingner, J., Bickle, T. A., Firman, K. & Glover, S. W. (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.

**Radding, C. M. (1973).** Molecular mechanisms in genetic recombination. *Annu Rev Genet* **7**, 87–111.

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

**Raleigh, E. A. & Brooks, J. E. (1998).** Restriction–modification systems; where they are and what they do. In *Bacterial Genomes: Physical Structure and Analysis*, pp. 78–92. Edited by F. J. de Bruijn, J. R. Lupski & G. M. Weinstock. New York: Chapman & Hall.

**Rao, D. N., Saga, S. & Krishnamurthy, V. (2000).** The ATP-dependent restriction enzymes. *Prog Nucleic Acids Res Mol Biol* **64**, 1–63.

**Read, T. D., Thomas, A. T. & Wilkins, B. M. (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.

**Redaschi, N. & Bickle, T. A. (1996).** DNA restriction and modification systems. In *Escherichia coli and Salmonella: Cellular and Molecular Biology*, 2nd edn, pp. 773–781. Edited by F. C. Neidhardt and others. Washington, DC: American Society for Microbiology.

**Roberts, R. J. & Halford, S. E. (1993).** Type II endonucleases. In *Nucleases*, pp. 35–88. Edited by S. M. Linn, R. S. Lloyd & R. J. Roberts. Cold Spring Harbor, NY: Cold Spring Harbor Press.

**Roberts, R. J. & Macelis, D. (2000).** REBASE – restriction enzymes and methylases. *Nucleic Acids Res* **28**, 306–307.

**Salaj-Šmic, E., Maršic, N., Trgovcevic, Z. & Lloyd, R. G. (1997).** Modulation of *Eco*KI restriction *in vivo*: role of the $\lambda$Gam protein and plasmic metabolism. *J Bacteriol* **179**, 1852–1856.

**Sharp, P. M., Kelleher, J. E., Daniel, A. S., Cowan, G. M. & Murray, N. E. (1992).** Roles of selection and recombination in the evolution of type I restriction-modification systems in enterobacteria. *Proc Natl Acad Sci U S A* **89**, 9836–9840.

**Simmon, V. F. & Lederberg, S. (1972).** Degradation of bacteriophage lambda deoxyribonucleic acid after restriction by *Escherichia coli* K-12. *J Bacteriol* **112**, 161–169.

**Smith, G. R. (2001).** Homologous recombination near and far from DNA breaks: alternative roles and contrasting views. *Annu Rev Genetics* **35**, 243–274.

**Smith, J. D., Arber, W. & Kuhnlein, U. (1972).** Host specificity of DNA produced by *Escherichia coli*. XIV. The role of nucleotide methylation in *in vivo* B-specific modification. *J Mol Biol* **63**, 1–8.

**Smith, G. R., Amundsen, S. K., Dabert, P. & Taylor, A. F. (1995).** The initiation and control of homologous recombination in *Escherichia coli*. *Philos Trans R Soc Lond B Biol Sci* **347**, 13–20.

**Stahl, F. W., Stahl, M. M., Malone, R. E. & Craseman, J. M. (1980).** Directionality and nonreciprocality of Chi-stimulated recombination in phage lambda. *Genetics* **94**, 235–248.

**Stahl, M. M., Kobayashi, I., Stahl, F. W. & Huntingdon, S. K. (1983).** Activation of Chi, a recombinator, by the action of endonuclease at a distant site. *Proc Natl Acad Sci U S A* **80**, 2310–2313.

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

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

**Szczelkun, M. D. (2000).** How do proteins move along DNA? Lessons from type I and type III restriction endonucleases. In *Essays in Biochemistry*, vol. 35. Edited by G. Banting & S. J. Higgins. London: Portland Press.

**Tao, T. & Blumenthal, R. M. (1992).** Sequence and characterization of *pvuIIR*, the *PvuII* endonuclease gene, and of *pvuIIC*, its regulatory gene. *J Bacteriol* **174**, 3395–3398.

**Tao, T., Bourne, J. C. & Blumenthal, R. M. (1991).** A family of regulatory genes associated with type II restriction-modification systems. *J Bacteriol* **173**, 1367–1375.

**Taylor, A. F. & Smith, G. R. (1990).** Action of RecBCD enzyme on cruciform DNA. *J Mol Biol* **211**, 117–134.

**Taylor, A. F. & Smith, G. R. (1999).** Regulation of homologous recombination: Chi inactivates RecBCD enzyme by disassembly of the three subunits. *Genes Dev* **13**, 890–900.

**Taylor, A. F., Schultz, D. W., Ponticelli, A. S. & Smith, G. R. (1985).** RecBC enzyme nicking at Chi sites during DNA unwinding: location and orientation dependence of cutting. *Cell* **41**, 153–163.

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

**Taylor, I., Watts, D. & Kneale, G. (1993).** Substrate recognition and selectivity in the type IC DNA modification methylase M *Eco*R124I. *Nucleic Acids Res* **21**, 4929–4935.

**Telander-Muskavitch, K. M. & Linn, S. (1981).** RecBC-like enzymes: exonuclease V deoxyribonucleases. *Enzymes* **14A**, 233–250.

**Thaler, D. S., Stahl, M. M. & Stahl, F. W. (1987).** Tests of the double-strand-break repair model for Red-mediated recombination of phage $\lambda$ and plasmid dv. *Genetics* **116**, 501–511.

**Thoms, B. & Wackernagel, W. (1984).** Genetic control of damage-inducible restriction alleviation in *Escherichia coli* K12: an SOS function not repressed by *lexA*. *Mol Gen Genet* **197**, 297–303.

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

**Titheradge, A. J. B., Ternent, D. & Murray, N. E. (1996).** A third family of allelic *hsd* genes in *Salmonella enterica*: sequence comparisons with related proteins identify conserved regions implicated in restriction of DNA. *Mol Microbiol* **22**, 437–447.

**Titheradge, A. J. B., King, J., Ryu, J. & Murray, N. E. (2001).** Families of restriction enzymes: an analysis prompted by molecular and genetic data for type ID restriction and modification systems. *Nucleic Acids Res* **29**, 4195–4205.

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

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

**Wilson, G. G. & Murray, N. E. (1991).** Restriction and modification systems. *Annu Rev Genet* **25**, 585–627.