

Probing the Domain Structure of the Type IC DNA Methyltransferase M.*EcoR124I* by Limited Proteolysis

M. Webb, I. A. Taylor, K. Firman and G. G. Kneale*

Biophysics Laboratories
School of Biological Sciences
University of Portsmouth
Portsmouth, PO1 2DT, UK

Limited proteolysis has been used to probe the domain structure of the type I DNA methyltransferase M.*EcoR124I*. Trypsin digestion of the methyltransferase generates two fragments derived from the HsdS subunit, a 28 kDa N-terminal domain and a 19 kDa C-terminal domain, leaving the HsdM subunit intact. Extensive digestion by chymotrypsin, however, removes 59 amino acid residues from the N terminus of the HsdM subunit to leave a 52 kDa C-terminal domain. Binding of the cofactor *S*-adenosyl methionine has no appreciable effect on the rate of cleavage, but binding of a 30 bp DNA duplex containing the cognate recognition sequence confers almost total protection. Following trypsin cleavage of the methyltransferase, a stable proteolytic product is produced which has been purified for biochemical characterisation. The trypsinised enzyme is shown to be a multimeric complex containing two intact HsdM subunits and both fragments of the HsdS subunit, consistent with the circular model proposed for the organisation of domains in the specificity subunit in type IC methyltransferases. Gel retardation studies show that the proteolysed enzyme still retains DNA binding activity, but its specificity for the DNA recognition sequence is dramatically reduced.

Keywords: methyltransferase; restriction-modification; proteolysis; domain organisation; DNA specificity

*Corresponding author

Introduction

Type I restriction-modification (R-M) systems consist of complex multisubunit enzymes which exhibit restriction endonuclease and modification methylase activities (Wilson & Murray, 1991; Bickle & Kruger, 1993; Heitman, 1993). Their DNA recognition sequences are asymmetric, consisting of two half sites 3 to 5 bp in length, separated by a non-specific "spacer" sequence 6 to 8 bp long. Restriction endonuclease activity results in cleavage of unmodified DNA some distance from the recognition site; methyltransferase activity, however, results in methylation at the N-6 of a specific adenine within each half-site of the DNA recognition sequence and prevents cutting by the corresponding endonuclease. Methylation activity is dependent upon the cofactor *S*-adenosyl methionine (AdoMet), which acts as a methyl group donor.

Type I R-M systems have been classified into three families (IA, IB and IC) using a variety of genetic and

biochemical criteria (Murray *et al.*, 1982; Fuller-Pace *et al.*, 1985; Suri & Bickle, 1985; Price *et al.*, 1987). The genes for all type I systems encode three subunits: HsdS, HsdM and HsdR (responsible for specificity, methylation, and restriction, respectively). For restriction endonuclease activity all three subunits are needed, but for methylation of the target sequence, HsdS and HsdM subunits alone are sufficient (Suri *et al.*, 1984).

The HsdS subunit consists of two regions of highly variable sequence plus a number of regions that are conserved within a given family (Fuller-Pace & Murray, 1986; Kannan *et al.*, 1989; Tyndall *et al.*, 1994). Specificity of DNA binding is conferred by the two variable regions of HsdS, each being responsible for binding a half-site in the bipartite DNA recognition sequence (Cowan *et al.*, 1989; Gubler *et al.*, 1992; Meister *et al.*, 1993; Abadjieva *et al.*, 1993). The HsdM subunits of type I R-M systems contain a motif that is common to all methyltransferases, and mutational analysis shows that it is essential for methyltransferase activity (Willcock *et al.*, 1994). The conserved domains of HsdS are believed to interact with the HsdM subunits (Cooper & Dryden, 1994; Abadjieva *et al.*, 1994). A model has recently been proposed in

Abbreviations used: R-M, restriction-modification; DFP, di-isofluoropropanol; HPLC, high pressure liquid chromatography.

which pseudo-dyad symmetry is imposed on the methyltransferase by the circular organisation of the domains of the HsdS subunit when complexed to two copies of the HsdM subunit (Kneale, 1994).

Of the many type I methyltransferases now identified, only *EcoR124I* and *EcoKI* have been over-expressed and purified in sufficient quantities for detailed biochemical and biophysical analysis. Both enzymes have been well characterised in terms of their subunit composition, DNA binding characteristics and enzyme activity (Taylor *et al.*, 1992, 1993; Dryden *et al.*, 1993; Powell *et al.*, 1993). The *EcoR124I* methyltransferase (*M.EcoR124I*) consists of two copies of the HsdM subunit (each 58 kDa) and one HsdS subunit (46 kDa), forming a trimeric enzyme (162 kDa) with a subunit stoichiometry of M_2S_1 . The HsdS subunit alone is insoluble, and the interaction with HsdM is essential to maintain its solubility (Patel *et al.*, 1992) and perhaps also its structural integrity (Kneale, 1994).

The binding affinity of *M.EcoR124I* for its cognate DNA recognition sequence is $10^8 M^{-1}$ (Taylor *et al.*, 1992), approximately four orders of magnitude greater than for non-specific DNA binding. Methylation of the target adenine at either of the half-sites in the DNA recognition sequence reduces the DNA binding affinity around 30-fold, but methylation at other positions in the recognition sequence has only a small effect on the affinity for DNA. Despite the reduction in binding affinity, the rate of reaction of the enzyme was shown to be increased by two orders of magnitude when the DNA substrate was hemi-methylated (Taylor *et al.*, 1993).

X-ray solution scattering has been used to determine the structural parameters of *M.EcoR124I* and its complex with a DNA duplex containing the specific recognition sequence. These experiments showed that there was a dramatic reduction in the overall dimensions of the enzyme when bound to the DNA that was independent of the cofactor AdoMet (Taylor *et al.*, 1994). It was proposed that this structural transition involved a large rotation of the HsdM subunits to clamp the DNA, mediated by hinge-bending regions close to the interface of the HsdS and HsdM subunits of the enzyme, and driven in part by non-sequence-specific interactions with the methyltransferase outside of the DNA recognition sequence. This change in quaternary structure of the enzyme is accompanied by a conformational change in the DNA, as evidenced by the large increase in the circular dichroism signal (Taylor *et al.*, 1994).

Limited proteolysis is a powerful technique for the investigation of structural domains within DNA binding proteins (Plyte & Kneale, 1994). Such domains can be identified by amino acid sequencing of stable proteolytic fragments. Changes in the fragmentation pattern that accompany DNA binding can give further information on the accessibility of sites in the DNA-protein complex. Protease-resistant domains have been identified, for example in the type II methyltransferases *M.EcoRI* and *M.EcoRII* (Reich *et al.*, 1991; Friedman *et al.*, 1991) and other

DNA binding proteins (see, for example, Plyte & Kneale, 1993).

Structural evidence for the existence of domains in the type IA methyltransferase *M.EcoKI* has recently been obtained (Cooper & Dryden, 1994), where it was shown that the HsdS and HsdM subunits of this enzyme were equally susceptible to proteolysis. Here, we report the results of limited proteolysis of the type IC methyltransferase *M.EcoR124I* and its complex with DNA. In contrast to *M.EcoKI*, our results show that (1) the susceptibility of the HsdS subunit to limited proteolysis is far higher than that of the HsdM subunit; (2) the susceptibility to proteolysis is greatly reduced when the methyltransferase is bound to DNA; (3) binding of the cofactor AdoMet has no appreciable effect on proteolysis. We also show that the proteolysed enzyme exists as a multisubunit complex containing both proteolytic domains of the HsdS subunit, together with two intact HsdM subunits. Analysis of the DNA binding properties of this complex, however, show that DNA sequence specificity is drastically reduced.

Results

Limited proteolysis of the *M.EcoR124I*

The *EcoR124I* methyltransferase was digested with trypsin and the products of the reaction analysed by SDS/polyacrylamide gel electrophoresis. Under limiting conditions (trypsin to methylase ratio 1:500, w/w) the time-dependent disappearance of the HsdS band was concomitant with the appearance of two dominant peptide fragments, T1 and T2 (Figure 1(a)). The HsdM subunit appears to be considerably more resistant to cleavage by trypsin than the HsdS subunit under the same conditions and remains so at higher concentrations of trypsin (1:100, w/w), even when the HsdS fragment is completely degraded (Figure 1(b)).

Limited digestion of the methyltransferase with chymotrypsin at a ratio of 1:500 also produces two bands derived from the HsdS subunit (C1 and C2), the lower band (C2) being the more persistent of the two (Figure 1(c)). In this case, however, the intensity of the smaller fragment increases as the larger fragment decreases in intensity, suggesting that C2 is produced by further digestion of C1 (Figure 2(c)). With higher concentrations of chymotrypsin (1:100, w/w), the HsdM subunit becomes susceptible cleavage, resulting in the fragment C3 (Figure 1(d)). After very prolonged digestion, a further 59 amino acid residues are lost from the N terminus of C3 yielding the fragment C4 (results not shown).

When limited proteolysis is carried out, either with trypsin or chymotrypsin, in the presence of the cofactor AdoMet or its analogues *S*-adenosyl homocysteine and sinefungin, all at saturating ligand concentrations and under the same conditions as protein alone, no change is observed in the rate of the proteolysis (see Table 1). However, in the presence of a 30 bp DNA duplex containing the *EcoR124I*

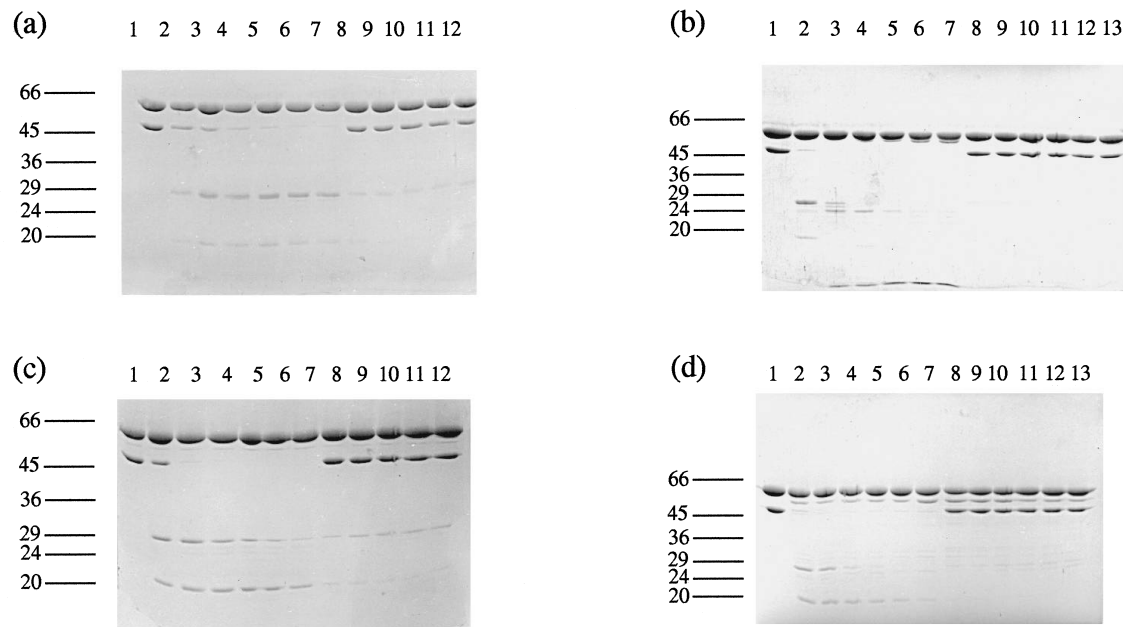


Figure 1. Time course of *M.EcoR124I* digestion by trypsin and chymotrypsin. (a) Trypsin digestion at a ratio of 1:500 (w/w). (b) Trypsin digestion at a ratio of 1:100 (w/w). (c) Chymotrypsin digestion at a ratio of 1:500 (w/w). (d) Chymotrypsin digestion at a ratio of 1:100 (w/w). Lanes 1 to 7 correspond to 0, 5, 10, 15, 20, 30 and 60 minute digests of the free methyltransferase. Lanes 8 to 12, ((a) and (c)) or 8 to 13 ((b) and (d)) show the result of digestion in the presence of duplex I, with digestion times of 10, 15, 20, 30 and 60 minutes ((a) and (c)) and 5, 10, 15, 20, 30 and 60 minutes ((b) and (d)). The positions of the molecular weight markers are indicated in kDa.

recognition sequence, the rate of proteolysis of the HsdS subunit by both trypsin and chymotrypsin is reduced by two orders of magnitude (Figure 2(a) and (b)). It is clear that binding of DNA to the methyltransferase confers very considerable protection from cleavage.

Peptide mapping of the proteolytic fragments of *M.EcoR124I*

The initial products produced by cleavage with trypsin are approximately 28 kDa (T1) and 19 kDa (T2), respectively, as judged by SDS/PAGE. N-

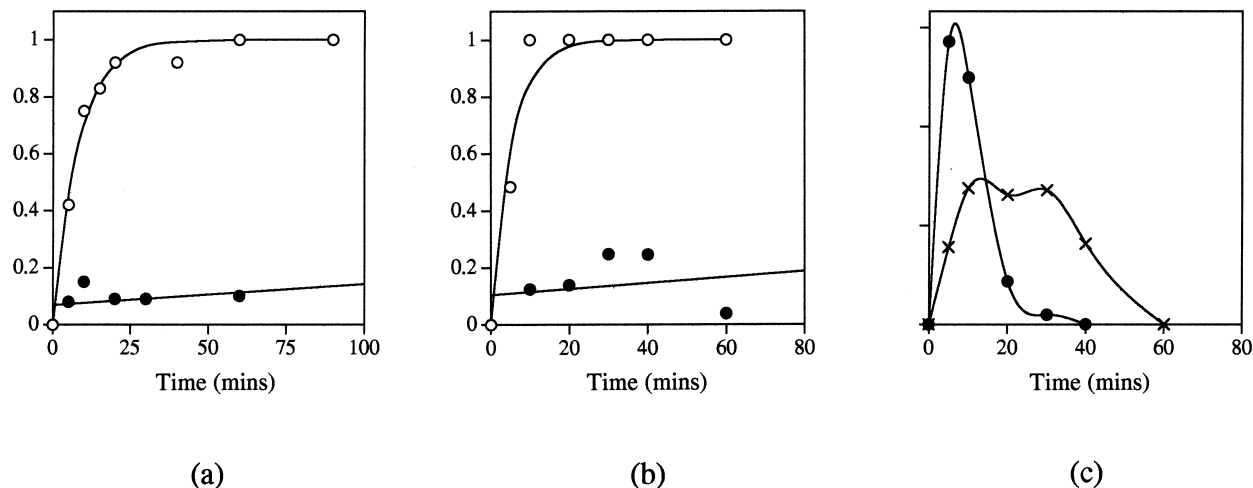


Figure 2. Kinetics of proteolysis of *M.EcoR124I* by (a) trypsin and (b) chymotrypsin in the absence of DNA (○). The degree of digestion was obtained by measuring the disappearance of the intact HsdS band as a function of time. The curves were fitted to a first-order rate equation to give the rate constants shown in Table 2. In the presence of DNA (●) there is a small but rapid rise in the extent of digestion which levels off very quickly. This corresponds to a small proportion (about 5 to 10%) of free methylase that is rapidly digested; the rate of digestion of the remaining complex was fitted as a linear function. (c) Time dependence of the appearance of chymotryptic fragments C1 (●) and C2 (×). In each case, the amount of each fragment was measured from the integrated intensity of the band after densitometry of the electrophoresis gel.

Table 1. Rates of digestion of the HsdS subunit of *M.EcoR124I*

Ligand	Protease	k (min ⁻¹)
—	Trypsin	0.128 (0.010)
—	Chymotrypsin	0.180 (0.040)
AdoMet	Trypsin	0.153 (0.013)
Sinefungin	Trypsin	0.111 (0.018)
DNA	Trypsin	0.001 (0.001)
DNA	Chymotrypsin	0.001 (0.002)

Rate constants (k) were obtained by fitting a first-order rate equation to data such as those presented in Figure 2, except for those in the presence of DNA which were fitted by a linear function. Estimated errors in k are given in parentheses. There was no detectable digestion of HsdM under the conditions used.

terminal sequencing of the peptides shows conclusively that they are derived from the HsdS subunit (Table 2). The two fragments produced by trypsin digestion arise from a single cut within the HsdS subunit at Arg240, T1 corresponding to the N-terminal fragment and T2 to the C-terminal fragment (Figure 3(a)). Upon further incubation, 37 amino acid residues are digested from the N terminus of T1 to produce T3, and the C-terminal fragment, T2, is degraded into small fragments.

Limited digestion with chymotrypsin also produces fragments of approximately 27 kDa and 19 kDa, both derived from the HsdS subunit. However, the cleavage pattern is quite different from that obtained with trypsin. The 27 kDa fragment C1 is produced by cleavage at Tyr137, and thus corresponds to the C-terminal two thirds of the HsdS subunit. The smaller fragment C2 arises from a region within this sequence and corresponds to the C-terminal third of the HsdS subunit (Figure 3(a)). This, together with the time dependence of the digestion pattern, suggests that C2 is derived by further digestion from the N terminus of C1.

The HsdM subunit is cleaved by chymotrypsin to produce a stable 52 kDa fragment, C3, in which the N-terminal 59 amino acid residues are removed (Figure 3(b)). Extended digestion times result in further cleavage from the N terminus to produce a 46 kDa fragment, C4, which is subsequently degraded fairly rapidly. (This fragment comigrates

with the intact HsdS subunit on PAGE, and its presence was only detected by amino acid sequencing of the band.)

Purification of the proteolysed methyltransferase

After mild trypsin digestion of the methyltransferase, the sample was applied directly to a heparin column. The bound fraction eluted at 250 mM NaCl and behaved identically in this respect to the native enzyme (Taylor *et al.*, 1992). Analysis of the peak by SDS/PAGE showed the presence of the two large fragments of the HsdS subunit, T1 and T2, together with the intact HsdM subunit (data not shown). Since HsdM will only bind to heparin when it is part of the intact methyltransferase enzyme (I. A. Taylor & G. G. Kneale, unpublished results), this indicates that the two large fragments of HsdS remain bound to HsdM in the form of a multimeric complex. The run-through peak contained non-complexed HsdM subunits and smaller peptides derived from HsdS, suggesting that further digestion of the HsdS domains abolishes the interaction with HsdM subunits.

Analytical gel filtration was employed to further characterise this complex (Figure 4). The trypsinised complex elutes as a single peak after 25.7 minutes, corresponding to a complex with a molecular mass of 162 kDa, indistinguishable from that of the native methyltransferase (Taylor *et al.*, 1993), providing clear evidence that the stoichiometry of the native enzyme is maintained. Native polyacrylamide gel electrophoresis further confirms this, since the trypsinised complex runs with an identical mobility to that of the intact methylase (Figure 5).

If the digestion with trypsin was allowed to progress somewhat further, a much more complex digestion pattern was obtained. Under these conditions, T2 was degraded to small fragments and T1 was degraded to T3. In this case, purification of the resulting complex by chromatography on heparin showed that T3 was able to bind to HsdM, thus indicating that deletion of about 40 amino acid residues from the N terminus of T1 does not abolish its ability to interact with HsdM (data not shown).

Table 2. Sequences of the proteolytic fragments

Fragment	N-terminal sequence	M_r (kDa)	Peptide sequence	Subunit
T1	SEMSY	28.2	S1-R240	S
T2	TLVDS	19.3	T241-N404	S
T3 (a)	AKDVHD	25.1	A36-R240	S
(b)	DVHDT	25.1	D38-R240	S
C1	SQKKIP	28.2	S138-R404	S
C2	ISPKM	19.3	I233-R404	S
C3	AKLDDS	52.0	A60-K517	M
C4	GYPSEA	48.0	G119-K517	M

Sequences and location of the fragments produced by limited proteolysis of *M.EcoR124I*, on the basis of N-terminal sequencing and estimated M_r . The numbering system for the amino acid residues is based on the N-terminal sequence SEMSY for HsdS and TSIQQ for HsdM (see Taylor *et al.*, 1992). Molecular weights were estimated by SDS/PAGE. Fragment T3 was shown by amino sequencing to consist of two components ((a) and (b)) differing by two residues at the N terminus.

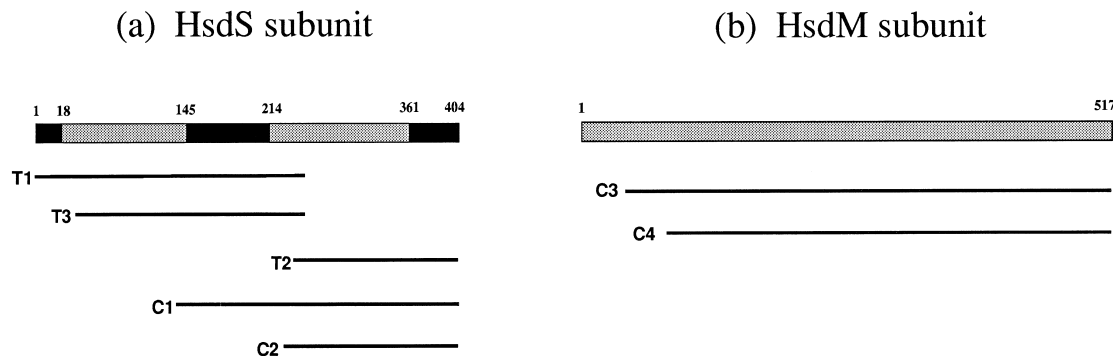


Figure 3. (a) Map of the HsdS subunit of *EcoR124I*, showing the variable (shaded) and conserved (dark) regions and the location of the sequenced proteolytic fragments. T1 and T2 denote the major tryptic fragments; T3 and T4 are produced by more extensive trypsin digestion. C1 and C2 are the major products of chymotryptic digestion. (b) Map of the HsdM subunit of *M.EcoR124I*, showing the sequence and location of the proteolytic products C3 and C4 obtained with chymotryptic digestion of the methyltransferase.

Further experiments were not carried out with this complex since it could not be obtained in sufficiently high yield and we were unable to determine the stoichiometry or molecular mass of the complex.

DNA binding studies on the proteolysed methyltransferase

Gel retardation assays were used to determine the extent to which the trypsinised methyltransferase was capable of binding to its cognate DNA recognition sequence, compared with the ability of the native enzyme to bind to the same sequence (Figure 6(a)). By adding increasing concentrations of protein to duplex I, the affinity of the proteolysed protein for the recognition sequence can be assessed. The native *M.EcoR124I* enzyme binds strongly to a 30 bp duplex containing its cognate DNA recognition sequence, and is fully bound to the DNA with a 1:1 stoichiometry (Taylor *et al.*, 1992). Since a two- to threefold excess of the proteolysed protein has to be added to achieve full retardation, it is clear that the affinity of the trypsinised enzyme for this sequence is greatly reduced. The presence of additional bands on the gel suggests that there are multiple binding sites for the proteolysed enzyme on the duplex. These complexes appear throughout the titration and indicate that non-specific DNA binding is occurring, i.e. that the capacity of the proteolysed enzyme to discriminate between the specific site and other sites on the duplex is greatly reduced in the proteolysed enzyme. In contrast, with the native enzyme multiple bands are weak and only occur once the specific site is fully occupied. Control experiments in which the methyltransferase was subjected to the same treatment, except for the addition of protease, confirmed that the loss of DNA binding specificity was due to cleavage of the HsdS subunit.

In order to directly compare the DNA sequence specificity of the proteolysed and native enzymes, competition assays were subsequently performed in which increasing concentrations of a non-specific duplex were used to dissociate the ^{32}P -labelled

specific duplex. In experiments with the proteolysed complex, the non-specific sequence (duplex II) was able to compete quite effectively with a DNA duplex containing the specific recognition sequence (duplex I). Substantial competition was observed with only a threefold excess of the non-specific duplex (Figure 6(b)). In parallel experiments with the native protein, even an 80-fold excess of competitor did not result in appreciable dissociation of the labelled complex (Figure 6(c)), consistent with the high specificity of the native enzyme previously observed for oligonucleotide duplexes containing the *EcoR124I* recognition sequence (Taylor *et al.*, 1992). It can be concluded that specificity for the *EcoR124I* DNA recognition sequence is almost completely abolished when the HsdS subunit is cleaved into two domains, even though these domains are associated with two HsdM subunits as in the native enzyme.

Discussion

Domain structure of HsdS and HsdM

A striking feature of the results obtained from limited proteolysis of *M.EcoR124I* with either trypsin or chymotrypsin is the simplicity of the resulting fragmentation pattern (compare, for example, with the results obtained with *M.EcoKI* by Cooper & Dryden, 1994). Moreover, the susceptibility of the HsdS subunit of *M.EcoR124I* to both trypsin and chymotrypsin is seen to be very much greater than that of HsdM; thus, under conditions where HsdS is fully digested, HsdM remains intact. We have shown that the HsdS subunit is initially cut by trypsin into two large folded domains, one consisting principally of the N-terminal variable region plus the central conserved region and the other containing the majority of the C-terminal variable region plus the C-terminal conserved region of the protein. Cleavage with chymotrypsin produces the same size fragments as with trypsin, but the cleavage sites map to different parts of the subunit. Together all these sites map close to the domain boundaries in the HsdS

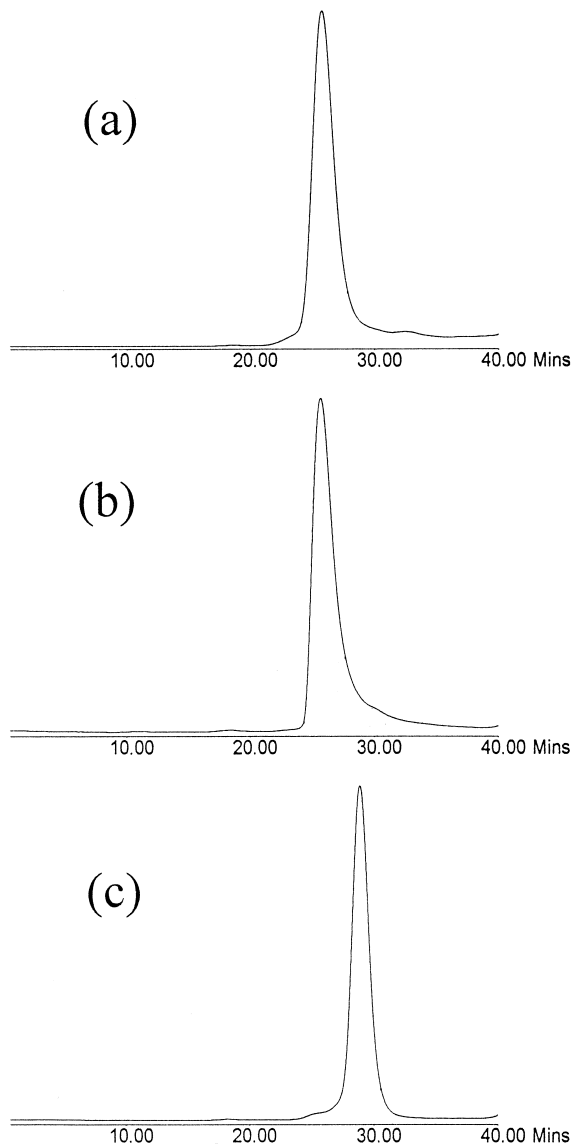


Figure 4. FPLC gel filtration profiles of (a) intact *M.EcoR124I*; (b) trypsinised *M.EcoR124I*; and (c) the purified HsdM subunit, with retention times of 25.7, 25.7 and 28.8 minutes, respectively. Peaks were monitored by their absorbance at 280 nm (arbitrary units). The trypsinised methyltransferase elutes with the same retention time as the intact enzyme.

subunit structure as deduced from sequence homologies and domain swapping experiments (see Figure 3(a)).

The major tryptic cleavage point within the HsdS subunit appears to lie just past the boundary of the central conserved region and the start of the C-terminal variable region. There is an additional cleavage point which is rather less susceptible to proteolysis just into the N-terminal variable domain. When these sites are considered in the context of the "circular" model proposed for the domain organisation in HsdS (Kneale, 1994) their locations are related by dyad symmetry. However, the rates of cleavage at these two sites are different

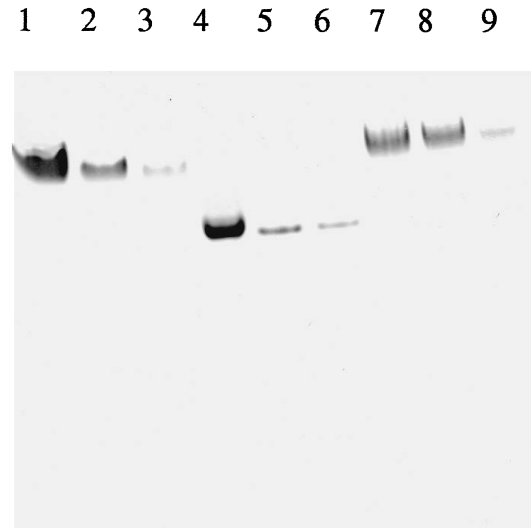


Figure 5. Analysis of trypsinised *M.EcoR124I* on a native polyacrylamide gel. Lanes 1 to 3: 10, 5 and 2 µg of the native *M.EcoR124I*. Lanes 4 to 6: 10, 5 and 2 µg of HsdM. Lanes 7 to 9: 10, 5 and 2 µg of trypsinised *M.EcoR124I*.

as the sequences surrounding the cutting sites are unrelated.

Limited proteolysis data on the HsdM subunit show that upon digestion with chymotrypsin, a stable 52 kDa domain is produced, resulting from removal of a 59 amino acid residue fragment from the N terminus of the HsdM subunit. The susceptibility of this region of the HsdM subunit to chymotrypsin suggests that this N-terminal sequence is a more flexible region of the protein. These results differ significantly from those obtained for *M.EcoKI* (Cooper & Dryden, 1994), where the HsdM subunit is cleaved into two large fragments, plus a small C-terminal "tail" under relatively mild reaction conditions. This difference could simply reflect a lack of proteolytic sites suitably placed on the surface of *M.EcoR124I*, but since there are potentially 49 tryptic sites and at least 55 potential chymotryptic sites distributed throughout the HsdM subunit, it would be surprising if neither protease were able to cleave at least once between putative domains. Alternatively, the domain structure of the HsdM subunits of type IA and type IC systems could be different.

Effect of DNA and cofactor binding

The binding of the specific DNA sequence to *M.EcoR124I* dramatically reduces the rate of cleavage of the HsdS subunit by either trypsin or chymotrypsin. We suggest that these sites are protected by protein-protein interactions due to a DNA-induced conformational change, together with some degree of additional protection arising directly from the interaction of HsdS with the DNA. From the results of small angle X-ray scattering (Taylor *et al.*, 1994) we have shown that there is a large conformational change in *M.EcoR124I* when DNA is bound, giving rise to a dramatic compaction in the DNA-protein

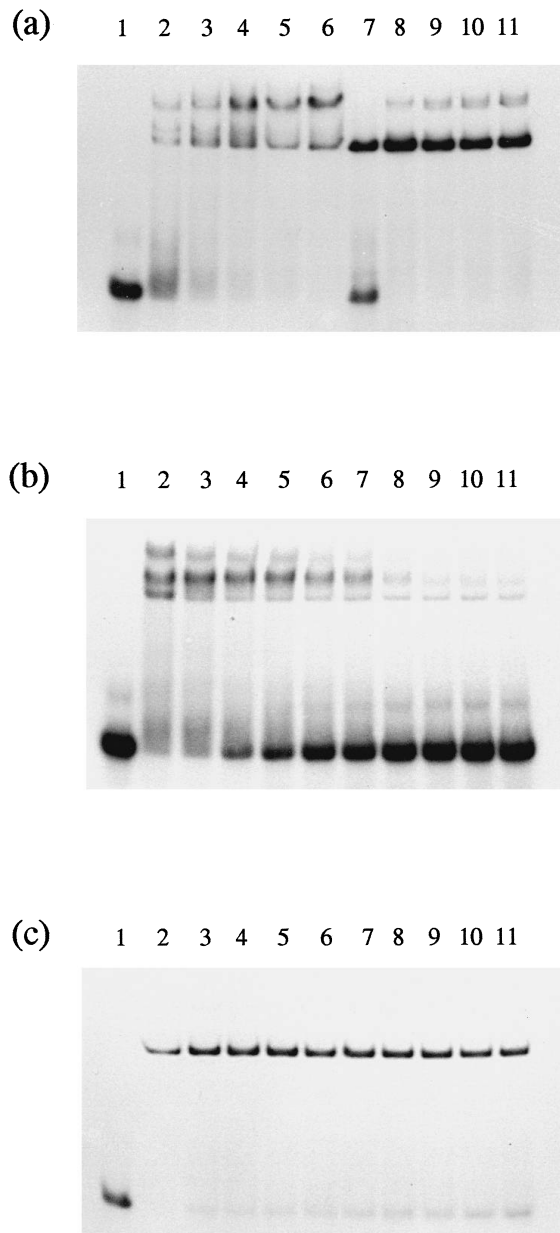


Figure 6. Gel retardation analysis. (a) Binding of trypsinised (lanes 2 to 6) and native (lanes 7 to 11) methyltransferase to a 30 bp duplex (280 nM) containing the *EcoR124I* recognition sequence (duplex I). Lane 1, no protein. Protein concentrations were 140 nM, 340 nM, 560 nM, 840 nM, and 1120 nM for lanes 2 to 6 and lanes 7 to 11. (b) Competition assay with proteolysed *M.EcoR124I*. Lane 1, duplex I alone; lane 2, duplex I (280 nM) and 840 nM protein. Lanes 3 to 11: as lane 2, but with increasing concentrations of unlabelled duplex II (0.33, 0.67, 0.9, 1.34, 1.8, 4.5, 6.7, 13.4, and 17.9 μ M). (c) Competition assay with native *M.EcoR124I*. Lane 1, duplex I alone; lane 2, duplex I (280 nM) with 280 nM protein. Lanes 3 to 11: as lane 2, but with increasing concentrations of unlabelled duplex II (0.33, 0.67, 0.9, 1.34, 1.8, 4.5, 6.7, 13.4, and 22.4 μ M).

complex which is mediated by movement of the HsdM subunits to enclose the DNA. It is very likely that the protection from cleavage that

accompanies DNA binding arises principally from this compaction.

The degree of protection conferred by DNA binding is very much greater than that observed in similar experiments with *M.EcoKI*, which was negligible (Cooper & Dryden, 1994). This suggests either (1) that any compaction induced by DNA binding is not as great for *M.EcoKI* as for *M.EcoR124I*, or (2) that the susceptible proteolytic sites in *M.EcoKI* are not sensitive to large conformational changes in the enzyme. The latter possibility seems unlikely, given the large number of cutting sites in *M.EcoKI*, at least one of which might be expected to reveal some significant differences in the rate of digestion if a large conformational change had taken place.

Binding of either AdoMet or its analogues appears to have no significant effect upon the digestion of the HsdS subunit. Limited proteolysis therefore provides no evidence for a conformational change in *M.EcoR124I* accompanying cofactor binding, unlike the situation found for *M.EcoKI* (Cooper & Dryden, 1994). This is in agreement with the small angle X-ray scattering data, which showed that neither AdoMet nor its analogues cause any change in the structure of *M.EcoR124I* (Taylor *et al.*, 1994). In contrast, *M.EcoKI* has been shown to undergo a significant structural change when AdoMet is bound (Powell *et al.*, 1993). Moreover, binding of AdoMet causes an increase in the affinity of *M.EcoKI* for its specific DNA recognition sequence (Powell *et al.*, 1993) but has no effect on the DNA binding affinity of *M.EcoR124I* (I. A. Taylor & G. G. Kneale, unpublished data). The results presented here therefore add to the evidence now accumulating that there are major differences between the two enzymes in the effects of cofactor binding.

Properties of the trypsinised methyltransferase

Cleavage of the HsdS subunit of the *EcoR124I* methyltransferase into two large fragments clearly does not impair its ability to form the appropriate intersubunit contacts with the HsdM subunits and to maintain the overall stoichiometry of the multisubunit enzyme. Since it has been argued that the complex is held together by interactions of the HsdM subunits with the conserved regions of HsdS and not by direct interactions between the two HsdM subunits themselves (Cooper & Dryden, 1994; Kneale, 1994) these results are, at first sight, rather surprising. If the two major conserved regions (the central and C-terminal regions) are each associated with one HsdM subunit, then it might be expected that cleavage of the HsdS subunit into two fragments (thereby separating the two major conserved regions) would result in the formation of two smaller complexes, each containing one HsdM subunit and one of the two HsdS fragments.

The fact that a larger complex than this is observed can be accounted for if there are non-covalent interactions, either direct or indirect, between the two halves of HsdS. Indirect interactions would arise if the small N-terminal conserved domain of HsdS

Table 3. Oligonucleotide extinction coefficients

Oligonucleotide	ϵ_{260} (calculated) ($M^{-1} \text{ cm}^{-1}$)	ϵ_{260} (corrected) ($M^{-1} \text{ cm}^{-1}$)	Hyperchromicity (%)
Duplex I (30-mer)	597,500	417,055	30.2
A-strand	310,300	240,242	22.5
B-strand	287,200	236,940	17.5
Duplex II (31-mer)	637,400	433,857	31.9
A-strand	321,700	248,192	22.8
B-strand	315,700	252,244	20.1

Calculated values were based on base composition using extinction coefficients ϵ_{260} ($M^{-1} \text{ cm}^{-1}$) for A, C, G and T of 14,700, 6100, 11,800, and 8700, respectively. The experimentally corrected extinction coefficients shown were used in the determination of all subsequent oligonucleotide concentrations.

bound to the same HsdM subunit as the C-terminal conserved domain, as proposed in the "circular" model for the domain structure of HsdS (Kneale, 1994). Thus, an HsdM subunit would essentially act as a "crosslink" between the two halves of the HsdS subunit, and the methyltransferase would remain as a 162 kDa complex following cleavage of the HsdS subunit into two fragments.

Although no fragments are released from the methyltransferase when the HsdS subunit is cleaved into two, gel retardation analysis shows quite clearly that some function is lost. The proteolysed complex can bind to DNA, but its ability to discriminate the specific *EcoR124I* recognition sequence from other sequences is drastically reduced. The precise spacing between the DNA recognition domains has been shown to be affected by the length of the central conserved region of HsdS (Price *et al.*, 1989). Moreover, it has been shown by genetic analysis that small deletions and insertions in the central conserved region linking the two variable (DNA recognition) domains of *EcoR124I* leads to promiscuous DNA recognition, together with reduced methyltransferase activity *in vivo* (Gubler & Bickle, 1991).

It has been suggested that this "elbow" region between the two conserved arms of the central conserved region of HsdS must have a certain degree of flexibility that is matched by a corresponding flexibility in certain regions of the HsdM subunit (Kneale, 1994). Our experiments show that cleavage of the HsdS subunit to sever the covalent connection between the two DNA recognition domains results in a dramatic reduction in the ability of *M.EcoR124I* to recognise its cognate sequence. Although the two variable domains remain bound to the two HsdM subunits, the increased flexibility of the system is

sufficient to perturb the spacing between the two DNA recognition domains, leading to an almost total loss of DNA binding specificity.

Materials and Methods

Protein purification

M.EcoR124I was overexpressed in *E. coli* JM109(DE3) from plasmid pJS4M as described (Patel *et al.*, 1992). The multisubunit enzyme was purified to homogeneity from crude cell extracts by ion exchange and heparin chromatography following published procedures (Taylor *et al.*, 1992). The integrity of the sample was checked by analytical gel filtration, SDS/PAGE and reverse phase high pressure liquid chromatography. Aliquots of the purified protein were concentrated using a pro-dicon vacuum filtration unit (30,000 molecular weight cut off), made up to 50% in glycerol and stored at -20°C . For subsequent experiments the methylase was buffer-exchanged by gel filtration using a BioRad 10DG desalting column. The concentration of the methylase was determined by its absorbance at 280 nm, using the extinction coefficient $\epsilon_{280} = 160,400 \text{ M}^{-1} \text{ cm}^{-1}$ (Taylor *et al.*, 1992).

Preparation of oligonucleotide duplexes

Oligonucleotides were purchased HPLC-purified from Oswel DNA services (University of Edinburgh). Extinction coefficients for the oligonucleotides and duplexes were calculated by digestion to completion with snake venom phosphodiesterase and summing the contributions from individual nucleotides, as described (Taylor *et al.*, 1994). The experimentally corrected extinction coefficients are shown in Table 3. Two duplexes were used in this investigation: a 30 bp duplex containing the *EcoR124I* recognition sequence, and a 31 bp duplex which lacks this site (see Figure 7). Duplexes were prepared by annealing equimolar

A

5' -CCGTGCAGAA**TTCGAGGTCG**ACGGATCCGG-3' DUPLEX I (specific)
3' -GGCACGT**CTT**AAGCTCC**AGC**TGCCTAGGCC-5'

B

5' -CAGGGATCCGAAGATATCGTTCAAGCTTCGC-3' DUPLEX II (non-specific)
3' -GTCCCTAGGCTTCTATAGCAAGTTCGAAGCG-5'

Figure 7. Sequences of synthetic oligonucleotide duplexes used in gel shift and proteolysis experiments. Duplex I is a 30 bp duplex containing the specific recognition sequence for *EcoR124I* (shown in bold). Duplex II is a 31 bp duplex lacking the recognition sequence, as used in competition binding assays.

concentrations of the two complementary strands, as described (Taylor *et al.*, 1992).

Proteolytic digestion

Samples for proteolytic digestion were prepared in 50 mM Tris-HCl (pH 7), 100 mM NaCl, 1 mM EDTA at a protein concentration of 1 mg ml⁻¹. Trypsin or chymotrypsin were added at w/w ratios of 1:500 or 1:100 and the reactions incubated at 37°C. Aliquots were withdrawn at various time points, the reaction terminated by addition of 1 mM di-isofluoropropanol (DFP) and the products separated by SDS/12.5% PAGE. Where appropriate, the cofactors *S*-adenosyl-L-methionine (AdoMet), *S*-adenosyl-L-homocysteine and sinefungin were included in the reaction mixture at a concentration of 400 µM. Digestion of DNA complexes was performed at a protein concentration of 1 mg/ml⁻¹, using a DNA to protein ratio of 2:1.

Protein sequencing

Proteolytic fragments which were to be N-terminally sequenced were separated on 12.5% polyacrylamide/SDS gels. The gel was then equilibrated in transfer buffer (10 mM Caps, 10% (v/v) methanol, pH 10) and electroblotted onto Problot membranes, (Applied Biosystems), prewetted with 100% methanol. When the transfer was complete the membrane was stained with amido black and destained with distilled water. Fragments were N-terminally sequenced using automated Edman degradation on an Applied Biosystems 477A pulsed liquid sequencer (University of Southampton).

Purification of trypsinised methyltransferase

Trypsin-treated methylase was produced by digestion of 5 mg methylase (w/w, ratio 1:500) at 37°C for ten minutes in 50 mM Tris-HCl (pH 7), 100 mM NaCl, 1 mM EDTA. The reaction was then inhibited with 1 mM DFP and applied directly to a 5 ml heparin cartridge (BioRad) equilibrated with buffer A (100 mM NaCl, 10 mM Tris-HCl (pH 7.5), 1 mM EDTA) at a flow rate of 2 ml min⁻¹. The column was washed with buffer A to remove non-heparin binding proteins. The bound protein was eluted with a linear gradient of NaCl from 0.1 M to 1 M in buffer A. Fractions were analysed on 12.5% polyacrylamide/SDS gels. Those fractions containing HsdM and trypsinised HsdS were pooled, concentrated and applied in 100 µl aliquots to a Superose 12 (10/30) FPLC column (Pharmacia) equilibrated in buffer A. The protein eluted as a single peak at 25 minutes. The fractions were pooled, concentrated to 2 mg ml⁻¹, made 50% in glycerol and stored at -20°C.

Gel electrophoresis

SDS/12.5% polyacrylamide gels (Laemmli, 1970) were routinely employed to separate proteins according to their molecular weight. Gels were run at a constant current of 45 mA until the bromophenol blue had reached the bottom of the gel. For the analysis of samples on native gels, samples were loaded in 10 mM DTT and 10% glycerol onto a 6% polyacrylamide/TAE gel (40 mM Tris-acetate (pH 7.4), 1 mM EDTA) and run at 100 V at 4°C until the bromophenol blue dye was two-thirds down the gel. All gels were stained in 50% (v/v) methanol, 10% (v/v) acetic acid and 0.1 (w/v) Coomassie blue, for one hour at 60°C and destained in 10% methanol and 10% acetic acid overnight.

Gel retardation assays

Oligonucleotide duplexes were 5'-end-labelled with ³²P as described (Taylor *et al.*, 1993). The concentration of labelled DNA was determined from the optical density at 260 nm. End-labelled duplexes of known concentration were incubated with *M.EcoR124I* (or the proteolysed enzyme) over a range of various concentrations and incubated at 4°C for 15 minutes in a binding buffer consisting of 10% glycerol, 50 mM Tris-HCl (pH 8.2), 5 mM MgCl₂, 1 mM DTT. For competition experiments, increasing concentrations of a non-labelled 31 bp oligonucleotide duplex were added to compete off a labelled 30 bp duplex containing the *EcoR124I* recognition sequence. Samples were subsequently loaded onto a 6% polyacrylamide native TAE gel. Gels were run at 100 V at 4°C until the bromophenol blue dye was two-thirds down the gel. After electrophoresis, gels were vacuum dried and autoradiographed.

Acknowledgements

We thank Professor N. E. Murray and Dr D. T. F. Dryden for communicating results prior to publication. Amino acid sequencing was carried out at the University of Southampton (Institute of Biomolecular Sciences). This work has been supported by a project grant and a research studentship from the Wellcome Trust.

References

- Abadjieva, A., Patel, J., Webb, M., Zinkevich, V. & Firman, K. (1993). A deletion mutant of the type IC restriction endonuclease *EcoR124I* expressing a novel DNA specificity. *Nucl. Acids Res.* **21**, 4435-4443.
- Abadjieva, A., Webb, M., Patel, J., Zinkevich, V. & Firman, K. (1994). Deletions within the DNA recognition subunit of *M.EcoR124I* that identify a region involved in protein-protein interactions between HsdS and HsdM. *J. Mol. Biol.* **241**, 35-43.
- Bickle, T. A. & Kruger, D. H. (1993). Biology of DNA restriction. *Microbiol. Rev.* **57**, 434-450.
- Cooper, L. P. & Dryden, D. T. F. (1994). The domains of a type I DNA methyltransferase. Interactions and role in recognition of DNA methylation. *J. Mol. Biol.* **236**, 1011-1021.
- Cowan, G. M., Gann, A. A. F. & Murray, N. E. (1989). Conservation of complex DNA recognition domains between families of restriction enzymes. *Cell*, **56**, 103-109.
- Dryden, D. T. F., Cooper, L. P. & Murray, N. E. (1993). Purification and characterisation of the methyltransferase from the type I restriction and modification system of *E. coli* K12. *J. Biol. Chem.* **268**, 13228-13236.
- Friedman, S., Som, S. & Yang, L.-F. (1991). The core element of the *EcoRII* methylase as defined by protease digestion and deletion analysis. *Nucl. Acids Res.* **19**, 5403-5408.
- Fuller-Pace, F. V. & Murray, N. E. (1986). Two DNA recognition domains of the specificity polypeptides of a family of type I restriction enzymes. *Proc. Natl Acad. Sci. USA*, **83**, 9368-9372.
- Fuller-Pace, F. V., Cowan, G. M. & Murray, N. E. (1985). *EcoA* and *EcoE*: alternatives to the *EcoK* family of type I restriction and modification systems of *Escherichia coli*. *J. Mol. Biol.* **186**, 65-75.

- Gubler, M. & Bickle, T. A. (1991). Increased protein flexibility leads to promiscuous protein-DNA interactions in type IC restriction-modification systems. *EMBO J.* **10**, 951–957.
- Gubler, M., Braguglia, D., Meyer, J., Piekarowitz, 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.
- Heitman, J. (1993). On the origins, structures and functions of restriction-modification enzymes. *Genetic Engineering* (Setlow, J. K., ed.), vol. 15, pp. 57–108, Plenum Press, New York.
- Kannan, P., Cowan, C. M., Daniel, A. S., Gann, A. A. F. & Murray, N. E. (1989). Conservation of organisation in the specificity polypeptides of two families of type I restriction enzymes. *J. Mol. Biol.* **209**, 335–344.
- Kneale, G. G. (1994). A symmetrical model for the domain structure of type I DNA methyltransferases. *J. Mol. Biol.* **243**, 1–5.
- Laemmli, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature*, **227**, 680–685.
- Meister, J., McWilliams, 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.
- Murray, N. E., Gough, J. A., Suri, B. & Bickle, T. A. (1982). Structural homologies among type I restriction and modification systems. *EMBO J.* **1**, 535–539.
- Patel, J., Taylor, I., Dutta, C., Kneale, G. G. & Firman, K. (1992). High-level expression of the cloned genes encoding the subunits of and the intact DNA methyltransferase, *M.EcoR124*. *Gene*, **112**, 21–27.
- Plyte, S. E. & Kneale, G. G. (1993). Identification of a compact DNA binding domain in the gene 5 protein of Pf1 bacteriophage. *Biochemistry*, **32**, 3623–3628.
- Plyte, S. E. & Kneale, G. G. (1994). Limited proteolysis of protein-nucleic acid complexes. In *Methods in Molecular Biology* (Kneale, G. G., ed.), vol. 30, pp. 151–160, Humana Press, Totowa, NJ.
- Powell, L. M., Dryden, D. T. F., Willcock, D. F., Pain, R. H. & Murray, N. E. (1993). DNA Recognition by the *EcoK* methyltransferase. *J. Mol. Biol.* **234**, 60–71.
- Price, C., Pripfl, T. & Bickle, T. A. (1987). *EcoR124* and *EcoR124/3*: 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 *EcoR124* and *EcoR124/3* type I DNA restriction and modification enzymes. *J. Mol. Biol.* **205**, 115–125.
- Reich, N. O., Maegley, K. A., Shoemaker, D. D. & Everett, E. (1991). Structural and functional analysis of *EcoRI* DNA methyltransferase by proteolysis. *Biochemistry*, **30**, 2940–2946.
- Suri, B. & Bickle, T. A. (1985). *EcoA*: the first member of a new family of type I restriction modification systems. Gene organization and enzymatic activities. *J. Mol. Biol.* **186**, 77–85.
- Suri, B., Nagaraja V. & Bickle, T. A. (1984). Bacterial DNA modification. *Current Top. Microbiol. Immunol.* **108**, 1–9.
- Taylor, I. A., Patel, J., Firman, K. & Kneale, G. G. (1992). Purification and biochemical characterisation of the *EcoR124* type I modification methylase. *Nucl. Acids Res.* **20**, 179–186.
- Taylor, I. A., Watts, D. & Kneale, G. G. (1993). Substrate recognition and selectivity in the type IC DNA modification methylase *M.EcoR124I*. *Nucl. Acids Res.* **21**, 4929–4935.
- Taylor, I. A., Davis, K. G., Watts, D. & Kneale, G. G. (1994). DNA binding induces a major structural transition in a type I methyltransferase. *EMBO J.* **13**, 5772–5778.
- Tyndall, C., Meister, J. & Bickle, T. A. (1994). The *E. coli prr* region encodes a functional type IC DNA restriction system closely integrated with an anticodon nuclease gene. *J. Mol. Biol.* **327**, 266–274.
- Willcock, D. F., Dryden, D. T. F. & Murray, N. E. (1994). A mutational analysis of the two motifs common to adenine methyltransferases. *EMBO J.* **13**, 3902–3908.
- Wilson, G. G. & Murray, N. (1991). Restriction and modification systems. *Annu. Rev. Genet.* **25**, 585–627.

Edited by J. Karn

(Received 28 February 1995; accepted 24 April 1995)