Nucleic Acids Research, 1997, Vol. 25, No. 5 **987–991**

DNA binding and subunit interactions in the type I methyltransferase M.*Eco*R124I

D. R. Mernagh, L. A. Reynolds and G. G. Kneale*

Biophysics Laboratories, School of Biological Sciences, University of Portsmouth, St Michael's Building, Portsmouth PO1 2DT, UK

Received November 25, 1996; Revised and Accepted January 12, 1997

ABSTRACT

The type I DNA methyltransferase M.*Eco*R124I consists of two methylation subunits (HsdM) and one DNA recognition subunit (HsdS). When expressed independently, HsdS is insoluble, but this subunit can be obtained in soluble form as a GST fusion protein. We show that the HsdS subunit, even as a fusion protein, is unable to form a discrete complex with its DNA recognition sequence. When HsdM is added to the HsdS fusion protein, discrete complexes are formed but these are unable to methylate DNA. The two complexes formed correspond to species with one or two copies of the HsdM subunit, indicating that blocking the N-terminus of HsdS affects one of the HsdM binding sites. However, removal of the GST moiety from such complexes results in tight and specific DNA binding and restores full methylation activity. The results clearly demonstrate the importance of the HsdM subunit for DNA binding, in addition to its catalytic role in the methyltransferase reaction.

INTRODUCTION

Type I DNA methyltransferases are complex multisubunit enzymes which bind to a target recognition sequence with high affinity and methylate a specific base within this sequence. The host DNA is thereby protected from restriction by the corresponding endonuclease (1,2). The target recognition sequence for type I systems is asymmetrical, consisting of two half-sites 3–5 bp in length, separated by a non-specific 'spacer' sequence of 6–8 bp. Methylation occurs at the N6 of two specific adenines on opposite strands, one within each half-site of the DNA recognition sequence. Type I restriction–modification (R–M) systems are encoded by three genes, encoding the subunits HsdS, HsdM and HsdR (responsible for specificity, methylation and restriction respectively). For methylation of the target sequence, only the HsdS and HsdM subunits are required (3).

Sequence comparisons of type I R–M systems have shown that the HsdS subunit consists of two highly variable domains of 150–180 amino acid residues and two or more regions that are well conserved within a given family (4–6). The two variable regions of the HsdS subunit form independent target recognition domains (TRDs), each being responsible for recognition of one half of the bipartite DNA recognition sequence (7,8). Based on analysis of repeated sequences in HsdS, we have proposed a circular model for the domain organization in HsdS, driven by the interaction with HsdM, which thus locates the DNA binding domains of the specificity subunit (9).

The *Eco*R124I methyltransferase (M.*Eco*R124I) is amongst the best studied of all type I enzymes at the molecular level. It consists of two copies of the HsdM subunit (each 58 kDa) and one HsdS subunit (46 kDa), to form a trimeric enzyme (162 kDa) with a subunit stoichiometry of M_2S_1 (10). We have shown for M.*Eco*R124I that DNA binding confers considerable protection from proteolysis (11). Likewise, chemical modification experiments have shown that a large fraction of the lysine residues on the surface of the protein are inaccessible in the DNA–protein complex, including a number of sites in regions implicated in intersubunit contacts, as well as likely DNA binding residues (12). These results all indicate that the methyltransferase undergoes a significant conformational change when it binds to DNA.

X-Ray solution scattering and circular dichroism have been used to determine the structural parameters of M.*Eco*R124I and its complex with DNA (13). A dramatic reduction is observed in the overall dimensions of the enzyme following DNA binding, which was proposed to involve a large rotation of the HsdM subunits to clamp the DNA, resulting in additional non-sequence-specific interactions outside the DNA recognition sequence. The circular dichroism spectrum shows that this structural transition in the enzyme is accompanied by considerable distortion of the DNA structure in the complex.

The DNA recognition subunit (HsdS) of M.*Eco*R124I is completely insoluble when expressed independently from a variety of expression vectors, in contrast to the methylation subunit (HsdM), which is highly soluble even at very high levels of expression (14). Studies to date have therefore been conducted entirely with the intact methyltransferase. Expression as fusion proteins can frequently improve the solubility of insoluble proteins (15,16). In this paper we adopt such an approach to allow us to investigate the DNA binding properties of the individual subunits of M.*Eco*R124I.

Purification of GST–HsdS fusion protein

JM109 cells containing the plasmid pHSDS124A (17) were grown at 25°C until an OD_{600} of 0.6 was obtained. The cells were then induced with 0.1 mM IPTG and grown for 4 h. The cells were harvested by centrifugation at 2000 *g* at 4°C for 15 min. The

*To whom correspondence should be addressed. Tel: +44 1705 842678; Fax: +44 1705 842053; Email: geoff.kneale@port.ac.uk

supernatant was removed and the cells dried and frozen at –20°C until required.

The cell pellets were resuspended in 0.3 M TNE supplemented with 0.1 mM PMSF, 1 mM benzamidine, 2 mM DTT and 1 mg/ml lysozyme. The resuspended cells were frozen in liquid nitrogen and allowed to thaw slowly. Once thawed, the samples were disrupted by sonication at 0°C. The insoluble debris was removed by centrifugation at 40 000 *g*. PEI–cellulose (5 g) was added to the supernatant and the suspension stirred slowly for 1 h at 4°C. The bound nucleic acids were removed by centrifugation at 5000 *g* in a swing out rotor at 4°C. The supernatant was removed and stored awaiting additional purification.

GST fusion extracts were purified using a 10 ml column containing glutathione covalently attached to a Sepharose 6B matrix (Pharmacia). The crude preparation was loaded onto the column in buffer A (0.15 M TNE, 2 mM DTT). Subsequently, the bound protein was eluted by the addition of buffer B (0.15 M TNE, 10 mM glutathione, 2 mM DTT). Fractions were analysed by SDS–PAGE, those containing GST–HsdS being pooled for further purification.

The pooled fractions were further purified using a 5 ml Econopac heparin cartridge (BioRad). The sample was loaded in buffer A (0.15 M TNE, 2 mM DTT) and the bound protein subsequently eluted with a linear gradient of buffer B (1 M TNE, 2 mM DTT). The GST–HsdS was passed down a Q ion exchange column for final purification and eluted in the same manner as described for the heparin chromatography. The samples were then concentrated and dialysed into a buffer containing 50% glycerol, 150 mM NaCl, 50 mM Tris–HCl, pH 8.2, 5 mM DTT. Protein concentrations were estimated by UV spectroscopy.

End-labelling of DNA

Aliquots of 1 nmol 30 bp DNA were mixed with 50 µCi [γ-^{32}P]ATP (6000 Ci/mmol) and 50 U T4 polynucleotide kinase (New England Biolabs) in 25 µl kinase buffer (50 mM Tris–HCl, pH 7.6, 10 mM $MgCl_2$, 5 mM DTT, 10 mM spermidine) and incubated for 30 min at 37°C. The solution was made up to 400 µM with non-labelled ATP and incubated for an additional 30 min at 37°C. The enzyme was heat inactivated at 68°C and the solution made up to 0.15 M NaCl. The unincorporated ATP was separated from the labelled oligonucleotides using a Nuctrap push column (Stratagene).

In vitro methylation assay for reconstituted *Eco*R124I methyltransferase

The reconstituted methyltransferase was incubated overnight at 37°C with the radiolabelled 30 bp fragment (100 nM DNA, 12.5 nM protein) in binding buffer (50 mM Tris–HCl, pH 7.8, 10 mM $MgCl_2$, 1 mM DTT) supplemented with 1 mM AdoMet. Identical experiments were carried out with the native methyltransferase and the GST–HsdS–HsdM complex and, as controls, with GST–HsdS, HsdM and cleaved GST–HsdS respectively, under the same conditions. Samples were heat inactivated, made up in *Eco*RI reaction buffer (50 mM NaCl, 100 mM Tris–HCl, 10 mM $MgCl_2$, 0.025% Triton X-100, pH 7.5, at 25°C) and digested with *Eco*RI. After this time, the reactions were again heat inactivated. Finally, formamide loading dye was added to the samples, which were denatured by heating, before loading onto a 24% sequencing gel. The gels were fixed in 10% acetic acid, 50% methanol, then dried and the bands visualized by autoradiography.

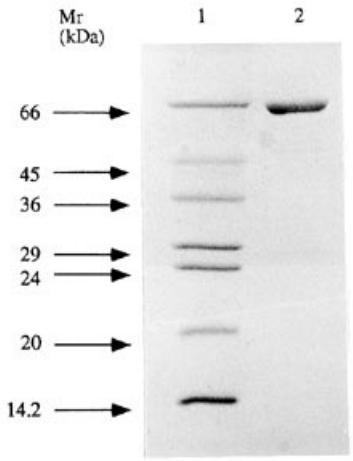

Figure 1. SDS gel electrophoresis of purified GST–HsdS. Lane 1, molecular weight marker; lane 2, purified GST–HsdS.

Gel retardation analysis

Gel retardation experiments were performed using non-denaturing gel electrophoresis. Proteins were incubated at various concentrations with the radiolabelled 30 bp DNA duplex containing the M.*Eco*R124I recognition sequence and the complexes were allowed to form at 4°C for a period of 20 min. At this point the samples were loaded onto a 6% non-denaturing acrylamide gel running at 4°C in 1× TAE buffer (40 mM Tris-acetate, pH 7.4, 1 mM EDTA). After an appropriate time the gels were dried under vacuum and visualized by autoradiography.

RESULTS

Over-expression and purification of GST–HsdS

HsdS was expressed as a fusion with glutathione *S*-transferase (GST) from the plasmid pHSDS124A (17). In order to achieve acceptable levels of expression and solubility of the fusion protein, it was necessary to grow the bacterial cells at 25°C. At this temperature, despite relatively low levels of expression, the majority of the fusion protein was found in the soluble fraction of the bacterial lysate.

Purification of the fusion protein was achieved initially by single step chromatography on a glutathione–Sepharose affinity column. However, it was necessary to follow this procedure with additional purification using both Q-Sepharose and heparin in order to remove nucleic acid and contaminating proteins that co-purify with the GST–HsdS fusion protein. The purified protein could be seen as essentially a single ~65 kDa band by SDS–PAGE (Fig. 1).

DNA binding characteristics of the GST–HsdS fusion

In previous experiments (10) we have shown that M.*Eco*R124I binds specifically to a 30 bp oligonucleotide containing its cognate DNA recognition sequence with high affinity ($K_d \approx 10^{-8}$ M). The same oligonucleotide duplex was used to compare the binding affinity of GST–HsdS before and after

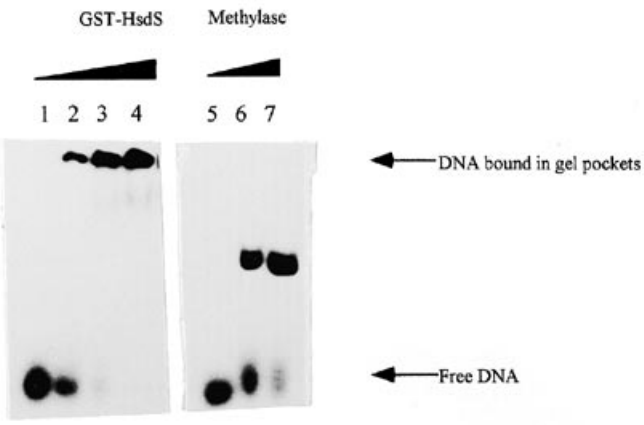

Figure 2. Gel retardation assay of GST–HsdS. Lane 1, free DNA; lanes 2–4, GST–HsdS:DNA ratios of 4:1, 12:1 and 20:1 respectively; lanes 5–7, a control experiment with native methylase at protein:DNA ratios of 0, 0.3 and 0.6. The 30 bp oligonucleotide probe containing the M.*Eco*R124I recognition sequence was kept at constant concentration (335 nM) throughout.

thrombin cleavage by gel retardation analysis. Initially, binding reactions with GST–HsdS were performed at concentrations similar to those used for the experiments with the native methylase, but these resulted in no observable shift in the DNA. At higher concentrations of both protein and DNA, binding was observed, but this was in the form of aggregated species that did not enter the gel. At very high protein concentrations (a 20-fold excess of protein), there was a faint band observable that migrated into the gel but this represented only a few per cent of the total (Fig. 2). Similar experiments were performed where the GST was first cleaved from the HsdS moiety by thrombin, before the addition of DNA, and in this case not even the aggregated species was observed (data not shown). Cleavage of GST–HsdS *in situ* in the DNA binding reaction gave an identical result.

Both HsdM and HsdS are required for efficient DNA binding

Since HsdS is soluble when co-expressed with HsdM during the preparation of wild-type M.*Eco*R124I (14), we decided to investigate the DNA binding properties of both GST–HsdS and HsdS in the presence of HsdM. The GST–HsdS fusion protein and the 30 bp DNA fragment were first combined at a ratio of 4:1 and the effect of increasing amounts of HsdM on DNA binding was assessed by gel retardation analysis. Three shifted complexes were observed (Fig. 3). One of these complexes remained at the top of the gel and represented an aggregated species. In addition, two species migrated as discrete complexes in the gel, the slower migrating species being predominant at the higher concentrations of HsdM. This suggests that these two bands represent nucleoprotein complexes with different numbers of HsdM subunits bound (see Discussion).

Parallel experiments were carried out with HsdS that had been cleaved from its fusion partner. Under these conditions, titration of HsdS with HsdM revealed a single shifted species that migrated on a non-denaturing gel with the same mobility as the native methylase (Fig. 4). In this case, the binding affinity of this protein complex was similar to that observed for the native enzyme, with almost complete binding at the stoichiometric point. Control reactions where HsdS was omitted from the reaction were performed, but even at very high concentrations of

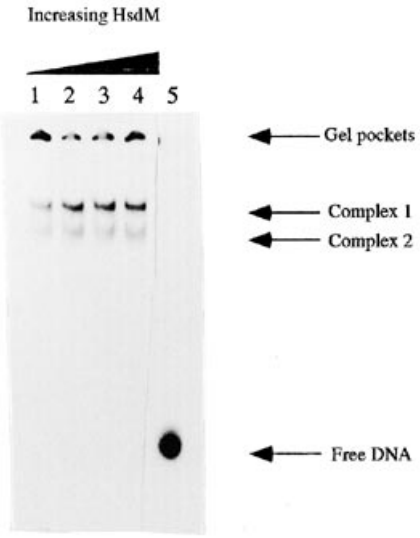

Figure 3. Gel retardation assay of GST–HsdS–HsdM complexes. Lanes 1–4, 1.67 μM GST–HsdS titrated with HsdM at concentrations of 1.67, 3.34, 6.68 and 13.36 μM respectively (corresponding to HsdM:GST–HsdS ratios of 1:1, 2:1, 4:1 and 8:1); lane 5, free DNA. The 30 bp oligonucleotide probe containing the M.*Eco*R124I recognition sequence was kept at constant concentration (335 nM) throughout.

Figure 4. Gel retardation assay of reconstituted M.*Eco*R124I. Lanes 1–2, HsdM was added to GST–HsdS at a 2:1 molar ratio, cleaved with thrombin and varying amounts of the reconstituted enzyme (molar ratios 1.0 and 0.5 respectively) added to the DNA. Lanes 3–5, control experiment with native methylase at protein:DNA ratios of 0, 0.5 and 1.0. The 30 bp oligonucleotide containing the M.*Eco*R124I recognition sequence was at constant concentration (335 nM) throughout.

HsdM, no shifted species were observed (data not shown). Additionally, the order in which the assay was performed was shown to be irrelevant. Prior cleavage of the GST–HsdS fusion or *in situ* cleavage in the binding reaction gave identical results.

The effect of GST fusion on enzymatic activity

Having shown that full DNA binding activity could be obtained from the HsdS–HsdM complex, we subsequently investigated the ability of this complex to methylate its DNA recognition sequence. A synthetic oligonucleotide substrate was constructed containing overlapping recognition sites for *Eco*RI and M.*Eco*R124I. The target adenine for methylation on one strand of the

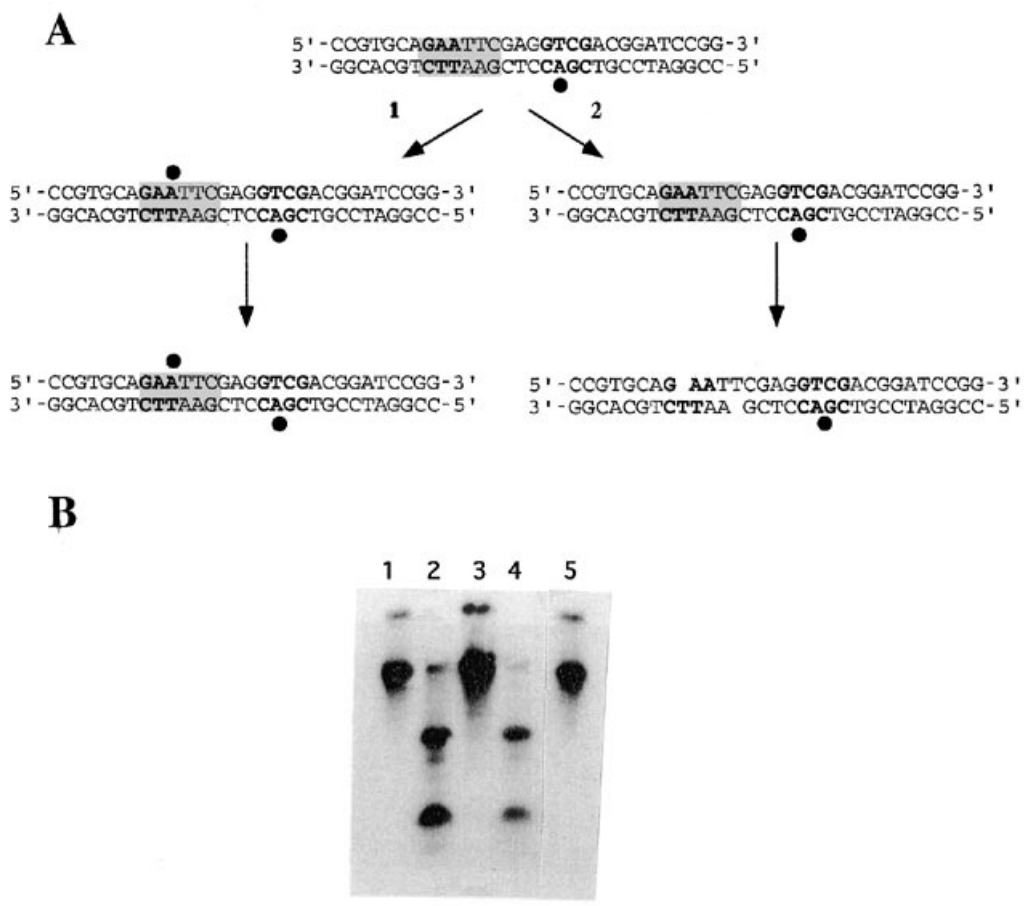

Figure 5. Methylation assay. (**A**) The M.*Eco*R124I recognition site in the DNA substrate overlaps with the GAATTC recognition sequence of *Eco*RI (on the upper strand M.*Eco*RI and M.*Eco*R124I modify DNA at the same adenine). Methylation at the M.*Eco*R124I site (path **1**) prevents cleavage by the *Eco*RI endonuclease. Otherwise *Eco*RI will digest the DNA (path **2**) to produce four fragments, two of which will be visible by autoradiography if the 5′-end of the DNA is labelled. The M.*Eco*R124I site is shown in bold and methylated bases are indicated by ●. (**B**) The radiolabelled DNA substrate was incubated with various protein samples at a 1:8 protein:DNA ratio and then challenged with *Eco*RI endonuclease. The DNA products were separated on a 24% sequencing gel and the dried gel visualized by autoradiography. Lane 1, native *Eco*R124I methylase; lane 2, GST–HsdS–HsdM complex; lane 3, HsdS–HsdM reconstituted complex; lane 4, oligonucleotide cut with *Eco*RI as a control; lane 5, undigested oligonucleotide as a control.

DNA duplex is identical to that modified by the *Eco*RI methylase. Thus methylation of this base by an active M.*Eco*R124I methyltransferase will result in protection from cleavage by *Eco*RI, which can readily be assayed on a denaturing gel using an end-labelled substrate (see Fig. 5).

Figure 5 clearly shows that the HsdS–HsdM complex prevents cleavage by the endonuclease and is thus active as a site-specific DNA methyltransferase. Since the substrate is in 8-fold excess over the reconstituted enzyme and the substrate is fully protected, the methyltransferase must be capable of turnover. However, if the GST–HsdS fusion protein was not cleaved with thrombin, then the complex with HsdM was unable to protect the substrate from *Eco*RI cleavage and this complex is therefore unable to methylate the DNA recognition sequence. The presence of the GST moiety at the N-terminus of HsdS thus abolishes enzyme activity.

DISCUSSION

Gel retardation analysis has shown that the GST–HsdS fusion protein is unable to bind to DNA to form a specific complex; rather, it forms a high molecular weight complex which is unable to enter the gel. However, this interaction is weak and requires a large excess of protein (at stoichiometric ratios the DNA is almost entirely unbound). Presumably, this is due to aggregation of the protein and/or multiple (non-specific) interactions with the DNA. When the GST moiety is cleaved from HsdS, no binding at all can be observed, reflecting the total insolubility of the native subunit.

However, cleavage of the GST–HsdS fusion protein prior to (or concomitant with) the addition of stoichiometric quantities of HsdM gives rise to a multisubunit protein complex with native DNA binding characteristics. High affinity binding to the DNA recognition sequence is then observed and the mobility of the resulting DNA–protein complex on non-denaturing gels is identical to that formed by the native methyltransferase.

The results differ from those of Kusiak *et al.* (17), who reported that DNA binding activity was possible with the HsdS subunit in the absence of HsdM. However, the significance of these findings is unclear. Although the concentrations of DNA and protein were not defined, the shifted species represented only a very small fraction of the total DNA; thus the binding appeared to be exceptionally weak, even though the gel was greatly overexposed. Moreover, the mobility of the HsdS complex was identical to that of intact methylase complex, despite enormous differences in molecular weight (46 kDa compared with 162 kDa) and nor did the addition of HsdM to the HsdS complex give rise to the expected ‘supershift’.

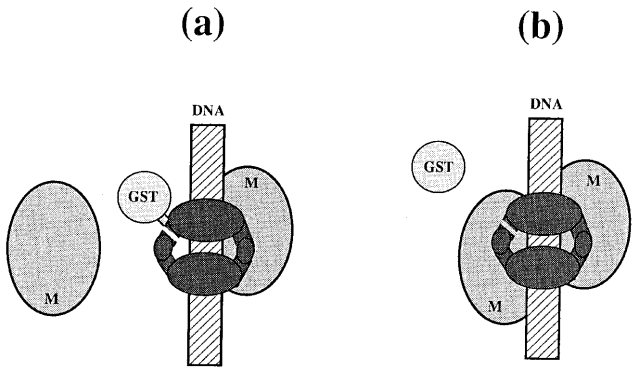

Figure 6. Model for the binding of GST–HsdS–HsdM to DNA. (**A**) A schematic representation of GST–HsdS with one HsdM subunit attached in the protein–DNA complex. The presence of GST at the N-terminus of HsdS prevents the second HsdM subunit from binding correctly to the HsdS subunit, thus destabilizing the organization of HsdS domains. (**B**) Removal of GST from HsdS results in correct binding to its cognate DNA recognition sequence and restores full methylation activity.

Although it is known from genetic data that HsdS is the determinant of specificity, our results demonstrate that this subunit is unable to bind DNA independently of HsdM. Both HsdM and HsdS are required for effective DNA binding and thus HsdM must play a structural role, in addition to its catalytic role in the methyl transfer reaction. It is already clear that HsdM is required to maintain the solubility of HsdS (14), presumably through hydrophobic interactions between sites on the surfaces of both subunits which would otherwise be exposed. The conserved regions of HsdS have been implicated in intersubunit contacts (11,18) and since there are two such regions in HsdS, each one interacting with a single HsdM subunit (9), their exposure to solvent in the free subunits is likely to be more critical for HsdS than for HsdM, which may explain the insolubility of the former. Thus, the role of HsdM in promoting DNA binding could simply be one of maintaining solubility. However, although the fusion of HsdS with GST also enhances its solubility, the fusion protein alone is still unable to properly bind its DNA recognition sequence, which argues for an additional and more specific role for HsdM.

It has been proposed that HsdM is necessary to maintain the circular organization of the domains of HsdS, in particular the spatial arrangement of the TRDs of this subunit, so that they can correctly locate the two half-sites in the DNA recognition sequence (9). This could explain, at least in part, why HsdM is required to enable HsdS to bind to its DNA recognition sequence. Additionally, the HsdM subunit could itself make a direct contribution to DNA–protein interactions in the complex. Non-sequence-specific interactions between HsdM and DNA have been suggested from the results of chemical modification experiments on M.*Eco*R124I, where lysine residues in HsdM as well as HsdS were implicated in contacts with the DNA (12). Experimental evidence from X-ray scattering (13) and DNA footprinting (19) indicate that the methyltransferase completely encloses the DNA when it binds. Stabilizing interactions between the HsdM subunit and DNA would be almost certain to occur, in order to drive the formation of such a complex.

Finally, how can we account for the presence of two well-defined species in the gel retardation experiments with GST–HsdS and HsdM? These bands most probably represent complexes with one HsdS subunit together with one or two copies of the HsdM subunit, since the faster and slower bands remain constant above HsdM:HsdS ratios of 1 and 2 respectively. According to our earlier model (9), the N-terminal region of HsdS is in close proximity to the C-terminal region, with the two held in place by interactions with one of the HsdM subunits. Thus, the presence of the bulky GST domain at the N-terminus of HsdS would interfere with one of the HsdM binding sites (Fig. 6). If the affinity of HsdM at this site were significantly reduced (but not completely abolished) then a population of species with one and two bound HsdM subunits would result. Even though both these complexes exhibit significant DNA binding affinity (although much weaker than that of the native enzyme), the methylation assays show that the presence of GST completely inhibits their ability to perform the more subtle tasks associated with the catalysis of DNA methylation.

The experiments we report refer to one member of the type IC family of methyltransferases, M.*Eco*R124I. Given the extensive homology between the conserved sequences within this family, one would expect that the HsdS subunit of other type IC systems (*Eco*DXXI and *Eco*PrrI) would be unable to bind their recognition sequences effectively in the absence of the HsdM subunit. Whether the same is true of the type IA and IB systems is less certain, but if the circular model for the structure of the HsdS domains applies more generally, as proposed (9), then one would predict that for these systems too HsdM subunit interactions would be required to correctly orient the target recognition domains of the HsdS subunit to allow effective DNA binding.

ACKNOWLEDGEMENTS

We are grateful to Dr David Hornby for freely providing the HsdS expression plasmid (pHSDS124A) used in these experiments and to Damian Watts for technical assistance.

REFERENCES

1 Bickle,T.A. and Kruger,D.H. (1993) *Microbiol. Rev.*, **57**, 434–450.
2 Wilson,G.G. and Murray,N.E. (1991) *Annu. Rev. Genet.*, **25**, 585–627.
3 Suri,B., Nagaraja,V. and Bickle,T.A. (1984) *Curr. Topics Microbiol. Immunol.*, **108**, 1–10.
4 Fuller-Pace,F.V. and Murray,N.E. (1986) *Proc. Natl. Acad. Sci. USA*, **83**, 9368–9372.
5 Kannan,P., Cowan,G.M., Daniel,A.S., Gann,A.A.F. and Murray,N.E. (1989) *J. Mol. Biol.*, **209**, 335–344.
6 Tyndall,C., Meister,J. and Bickle,T.A. (1994) *J. Mol. Biol.*, **237**, 266–274.
7 Cowan,G.M., Gann,A.A.F. and Murray,N.E. (1989) *Cell*, **56**, 103–109.
8 Gubler,M., Braguglia,D., Meyer,J., Piekarowicz,A. and Bickle,T.A. (1992) *EMBO J.*, **11**, 233–240.
9 Kneale,G.G. (1994) *J. Mol. Biol.*, **243**, 1–5.
10 Taylor,I., Patel,J., Firman,K. and Kneale,G. (1992) *Nucleic Acids Res.*, **20**, 179–186.
11 Webb,M., Taylor,I.A., Firman,K. and Kneale,G.G. (1995) *J. Mol. Biol.*, **250**, 181–190.
12 Taylor,I.A., Webb,M. and Kneale,G.G. (1996) *J. Mol. Biol.*, **258**, 62–73.
13 Taylor,I.A., Davis,K.G., Watts,D. and Kneale,G.G. (1994) *EMBO J.*, **13**, 5772–5778.
14 Patel,J., Taylor,I., Dutta,C.F., Kneale,G.G. and Firman,K. (1992) *Gene*, **112**, 21–27.
15 Fox,D.G. and Kneale,G.G. (1994) In Kneale,G.G. (ed.), *Methods in Molecular Biology.* Humana Press, Totowa, NJ, Vol. 30, pp. 169–184.
16 Ford,K.G., Whitmarsh,A.J. and Hornby,D.P. (1994) In Kneale,G.G. (ed.), *Methods in Molecular Biology.* Humana Press, Totowa, NJ, Vol. 30, pp. 185–198.
17 Kusiak,M., Price,C., Rice,D. and Hornby,D.P. (1992) *Mol Microbiol.*, **6**, 3251–3256.
18 Cooper,L.P. and Dryden,D.T.F. (1994) *J. Mol. Biol.*, **236**, 1011–1021.
19 Mernagh,D.R and Kneale,G.G. (1996) *Nucleic Acids Res.*, **24**, 4853–4859.