The EMBO Journal vol.13 no.23 pp.5772–5778, 1994

DNA binding induces a major structural transition in a type I methyltransferase

I.A.Taylor, K.G.Davis, D.Watts and G.G.Kneale[1]

Biophysics Laboratories, University of Portsmouth, St Michael's Building, Portsmouth PO1 2DT, UK
[1]Corresponding author

Communicated by T.Bickle

The type IC DNA methyltransferase M.*Eco*R124I is a complex multisubunit enzyme that recognizes the non-palindromic DNA sequence GAAN$_6$RTCG. Small angle X-ray scattering has been used to investigate the solution structure of the methyltransferase and of complexes of the enzyme with unmethylated and hemi-methylated 30 bp DNA duplexes containing the specific recognition sequence. A major change in the quaternary structure of the enzyme is observed following DNA binding, based on a decrease in the radius of gyration from 56 to 40 Å and a reduction in the maximum dimension of the enzyme from 180 to 112 Å. The structural transition observed is independent of the methylation state of the DNA. CD shows that there is no change in the secondary structure of the protein subunits when DNA is bound. In contrast, there is a large increase in the CD signal arising from the DNA, suggesting considerable structural distortion which may allow access to the bases targeted for methylation. We propose that DNA binding induces a large rotation of the two HsdM subunits towards the DNA, mediated by hinge bending domains in the specificity subunit HsdS.

Key words: circular dichroism/DNA methylation/DNA–protein interactions/quaternary structure/restriction–modification/small angle X-ray scattering

Introduction

Type I restriction–modification (R–M) systems consist of complex multisubunit enzymes which can exhibit restriction endonuclease and modification methylase activity for a specific target recognition sequence (Wilson and Murray, 1991; Bickle and Kruger, 1993). The target recognition sequence for such systems is invariably asymmetric, consisting of two half-sites 3–5 bp in length, separated by a non-specific 'spacer' sequence 6–8 bp long. Restriction endonuclease activity results in cleavage of unmodified DNA some distance from the recognition site; methyltransferase activity is confined to the methylation at the N6 of specific adenines on opposite strands of the DNA recognition sequence. The normal biological substrate for modification is hemi-methylated DNA, in which only one strand of the recognition sequence is methylated following replication of the bacterial genome but *de novo* methylation must also be possible (Kellerher *et al.*, 1991). All type I systems consist of three genes which encode the 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 the HsdS and HsdM subunits alone are sufficient (Suri and Bickle, 1984).

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; Suri and Bickle, 1984; Fuller-Pace *et al.*, 1985; Price *et al.*, 1987). Analysis of the DNA sequences of the genes for all three families of type I enzymes has 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 (Fuller-Pace and Murray, 1986; Kannan *et al.*, 1989; Tyndall *et al.*, 1994). Domain swapping experiments, in which the variable regions from members of the same family have been exchanged, have established that the specificity of DNA binding is conferred by the two variable domains of HsdS, each variable domain being responsible for binding a 'half-site' in the bipartite DNA recognition sequence (Cowan *et al.*, 1989; Gubler *et al.*, 1992).

HsdS deletion mutants of two type IC systems can be generated which recognize a palindromic DNA sequence, probably by dimerization of two half copies of HsdS each containing one DNA recognition domain (Abadjieva *et al.*, 1993; Meister *et al.*, 1993). There is evidence that the conserved domains of HsdS interact with the HsdM subunits (Cooper and Dryden, 1994), the sequences of which are also conserved within a family (Sharp *et al.*, 1992). Based on analysis of repeated sequences in HsdS, we have proposed a symmetrical model for the organization of the conserved domains which takes account of these observations (Kneale, 1994).

Only two type I methyltransferases, *Eco*R124I and *Eco*KI, have been overexpressed and purified in sufficient quantities for detailed *in vitro* analysis. Both enzymes have been well characterized in terms of their subunit composition, DNA binding and enzyme activity (Taylor *et al.*, 1992, 1993; Dryden *et al.*, 1993; Powell *et al.*, 1993). The *Eco*R124I methyltransferase (M.*Eco*R124I) has been shown to consist 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. The binding affinity of M.*Eco*R124I for its cognate DNA recognition sequence (GAAN$_6$RTCG) is 10^8 M^{-1}; methylation of either strand of the DNA recognition sequence increases the catalytic activity of the enzyme at least 100-fold, but reduces the DNA binding affinity ~30-fold. On this basis, it was suggested that a conformational change might accompany

© Oxford University Press

methylation at one half-site, which could be communicated to the other half-site through protein–protein interactions (Taylor *et al.*, 1993).

The crystal structure of the type II methyltransferase M.*Hha*I has been reported recently and shows a number of novel features when bound to its target recognition site. In particular, the target base (in this case a cytosine) in the DNA recognition sequence is rotated out of the helix into a pocket in the methyltransferase; this is accompanied by a large movement of a segment of the protein situated in the catalytic domain of the enzyme (Cheng *et al.*, 1993; Kumar *et al.*, 1994). No structural studies have been reported to date on the more complex multisubunit type I methyltransferases. To investigate possible structural transitions of M.*Eco*R124I when bound to its DNA recognition sequence, we have performed small angle X-ray scattering experiments. In addition, we have investigated the secondary structure of the free and bound methyltransferase by CD spectroscopy. The results show that unexpectedly large global changes in the quaternary structure of the enzyme accompany DNA binding, and that changes in secondary structure are confined to the DNA component of the complex.

Results

The *Eco*R124I methyltransferase was overexpressed from an overproducing plasmid containing both the M and S genes of the methyltransferase (Patel *et al.*, 1992). The enzyme was purified to homogeneity and fully characterized as described elsewhere (Taylor *et al.*, 1992). Analytical ultracentrifugation and analytical gel filtration confirmed the molecular weight and stoichiometry (M_2S_1) of the multisubunit enzyme, and quantitative gel retardation experiments showed that the enzyme was fully active in DNA binding. The same preparation was subsequently used for small angle X-ray scattering and CD experiments.

Small angle X-ray scattering

Small angle X-ray scattering measurements were made on solutions of the purified *Eco*R124I methyltransferase at a variety of concentrations. Analysis of the scatter curves was initially performed by the use of Guinier plots [ln$I$(**Q**) versus $\mathbf{Q}^2$, where **Q** is the scattering vector and I is the intensity of the scattered X-rays] to obtain the radius of gyration, R_g (Figure 1). The slopes of the Guinier plots were unaffected by the concentration of the enzyme in the range 5–16 mg/ml (data not shown). Subsequent measurements were therefore made at a concentration of 10 mg/ml. The addition of the cofactor *S*-adenosyl methionine, or the substrate analogues *S*-adenosyl homocysteine and sinefungin, had no significant effect on the measured R_g (data not shown).

Identical experiments were carried out on the methyltransferase bound to a 30 bp DNA duplex containing the recognition sequence for the enzyme, under conditions known to allow formation of a specific DNA–protein complex of the correct stoichiometry (Taylor *et al.*, 1992). Guinier analysis of these data showed a large reduction in R_g, from 56 Å for the free enzyme to 40 Å for the enzyme–DNA complex (Table I). X-ray scattering data were also collected for complexes of the methyltransferase with hemi-methylated DNA duplexes in which either the

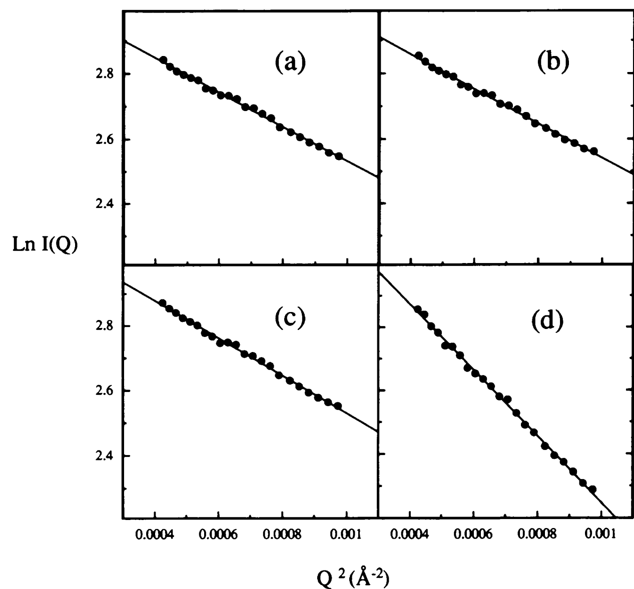

Fig. 1. Guinier plots for M.*Eco*R124I complexed to the 30 bp duplex AB (**a**), and the hemi-methylated duplexes AB13 (**b**) and A10B (**c**) at a 1:1 molar ratio, compared with the uncomplexed enzyme (**d**). All samples were 10 mg/ml in protein concentration. R_g were calculated from the slope of the Guinier plots, ln$I$(**Q**) versus $\mathbf{Q}^2$, where I is the intensity of the scattered X-rays (arbitrary units) and **Q** is the scattering vector (= 4πsinθ/λ).

Table I. Structural parameters are presented for M.*Eco*R124I and complexes with an unmethylated 30 bp duplex (AB) and hemi-methylated duplexes (A10B, AB13)

Sample	R_g (Å)	d_{max} (Å)	M_r (Da)
AB duplex	28	105	19 800
M.*Eco*R124I	56	180	162 000
M.*Eco*R124I/AB	40	112	181 800
M.*Eco*R124I/A10B	41	115	181 800
M.*Eco*R124I/AB13	40	112	181 800

Radii of gyration (R_g) were calculated from Guinier analysis of small angle scatter curves, with an estimated error of ±0.5 Å. The maximum dimension (d_{max}) was obtained from the relevant $p(r)$ plot, with an estimated error of ±3 Å. The molecular mass of each complex is calculated from the M_r and known stoichiometry of the components.

A or B strand was methylated at the target adenine in the recognition sequence (Taylor *et al.*, 1993). These complexes gave essentially identical scattering profiles to those of complexes with the unmethylated duplex. The addition of $MgCl_2$ (5 mM) had no effect on scattering from the DNA–protein complexes. In contrast, the methyltransferase alone showed significant aggregation in the presence of $MgCl_2$, as judged by severe non-linearity of the Guinier plots. However, the effect could be reversed by the addition of ethylene diamine tetraacetic acid (EDTA).

Complete X-ray scattering profiles are shown in Figure 2 for the methyltransferase and for the complex with the 30 bp DNA duplex. The shape of the curves confirms that a large structural change accompanies DNA binding, where the more pronounced shoulder (around **Q** = 0.12 $Å^{-1}$) is indicative of a more compact structure for the DNA–protein complex than for the protein alone, consistent with the observed decrease in R_g. Transformation of the scattering curves yields the distance distribution

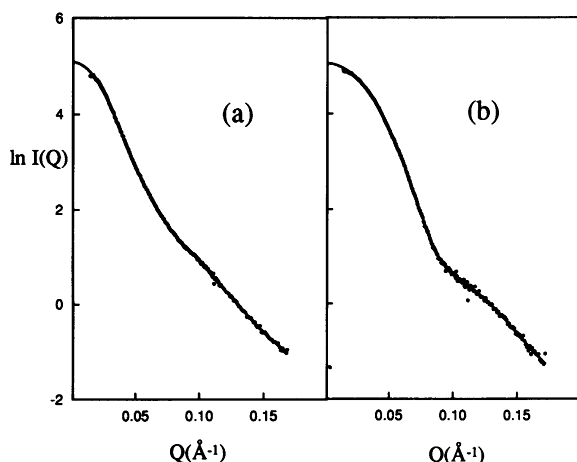

Fig. 2. Small angle X-ray scattering profiles ln*I*(Q) versus **Q** (as defined in Figure 1) for (**a**) M.*Eco*R124I and (**b**) the complex of M.*Eco*R124I with the 30 bp duplex AB at a 1:1 molar ratio. Experimental data points are shown. The solid line fitting the data points was obtained from a Fourier transformation of the distance distribution function, $p(r)$.

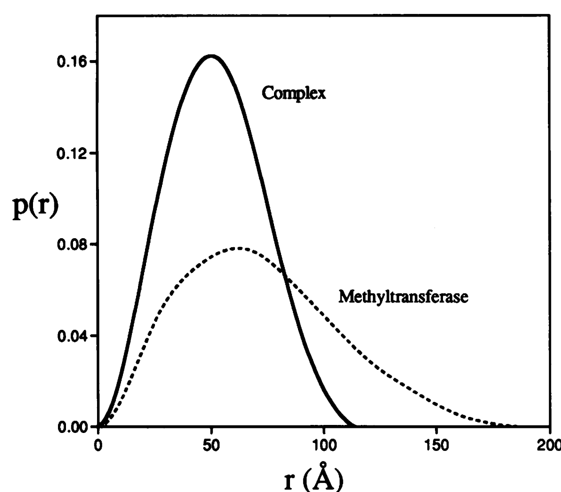

Fig. 3. Distance distribution functions $p(r)$ for M.*Eco*R124I (dashed line) and the complex of M.*Eco*R124I with duplex AB at a 1:1 molar ratio (solid line). Curves were obtained by transformation of the data shown in Figure 2. The point at which $p(r)$ drops to zero represents the maximum dimension of the particle. The area under each curve has been normalized to take account of the relative sizes of the protein and DNA–protein complex.

functions $p(r)$ for the methyltransferase and its complex with DNA (Figure 3). This function represents the distribution of vectors between all atoms in the particle; the distance at which the function drops to zero represents the maximum dimension of the macromolecular complex (for a review see Perkins, 1988). These curves show most clearly the extent of the structural change, since the maximum dimension changes from 180 Å in the free enzyme to 112 Å when DNA is bound.

CD spectroscopy

CD is highly sensitive to the conformation and environment of optically active chromophores and is a valuable probe of secondary structure in proteins and DNA. In view of the large structural change observed by X-ray scattering, we have employed CD spectroscopy to look for changes in the conformation of the protein and DNA components following complex formation between M.*Eco*R124I and the 30 bp DNA duplex containing the recognition sequence for the enzyme.

The protein component. The far UV CD spectra of the free enzyme and its complex with DNA are shown in Figure 4a. The overall features of the protein spectrum, in particular the double peak at 210–220 nm, suggests a substantial α-helical component in the multisubunit enzyme. Analysis of the CD spectrum by the method of Compton and Johnson (1986) leads to an estimate of the secondary structure content for the methyltransferase of 45% α-helix, with ~25% β-structure and 30% irregular structure. However, there is no significant change in the CD spectrum of the protein when DNA is bound, indicating that the secondary structure of the protein subunits in the protein–DNA complex remains unchanged. There is a small decrease (~5%) in the size of the negative CD signal at ~210–220 nm, but this can be entirely accounted for by the small positive peak in this region of the spectrum that arises from the DNA duplex itself.

The near UV CD spectrum of proteins is typically of a much smaller magnitude than that of the far UV, with bands arising from aromatic amino acid residues rather than the polypeptide backbone. Measurement of the near UV spectrum of the free methyltransferase shows that there is a peak of positive ellipticity at ~260 nm and a smaller peak at ~290 nm, but these are two orders of magnitude smaller than the far UV CD. These bands will be sensitive to changes in the tertiary structure of the protein, but any changes will be masked by the signal arising from the DNA itself and cannot be measured.

The DNA component. CD is a particularly useful probe of the conformation of the DNA component of nucleoprotein complexes since the near UV CD signal arises almost exclusively from the DNA bases (Carpenter and Kneale, 1994). Figure 4b shows the CD spectrum of the unmethylated 30 bp oligonucleotide duplex. The spectrum is typical of B-form DNA, having a positive peak centred at 272 nm and a negative peak of similar magnitude at 245 nm. As increasing quantities of M.*Eco*R124I are added to the DNA, a linear increase in the magnitude of the 272 nm peak is observed up to a stoichiometry of 1:1, beyond which no further change in CD occurs (Figure 4b and c). Changes in the negative CD peak at 245 nm have not been analysed, as in this region of the spectrum the contribution of the protein completely predominates.

The CD spectra of the hemi-methylated duplexes do not differ substantially from the CD spectrum of the unmethylated duplex. The addition of M.*Eco*R124I to each of the modified duplexes gave rise to similar increases in the 272 nm peak at saturation (data not shown), indicating that essentially identical conformational changes are occurring whether the enzyme binds to unmethylated, hemi-methylated or fully methylated DNA.

Discussion

Significant structural changes in DNA binding proteins can sometimes be induced when the protein binds to its DNA recognition sequence. Indeed, some of the smaller DNA binding proteins are partially unstructured in solution

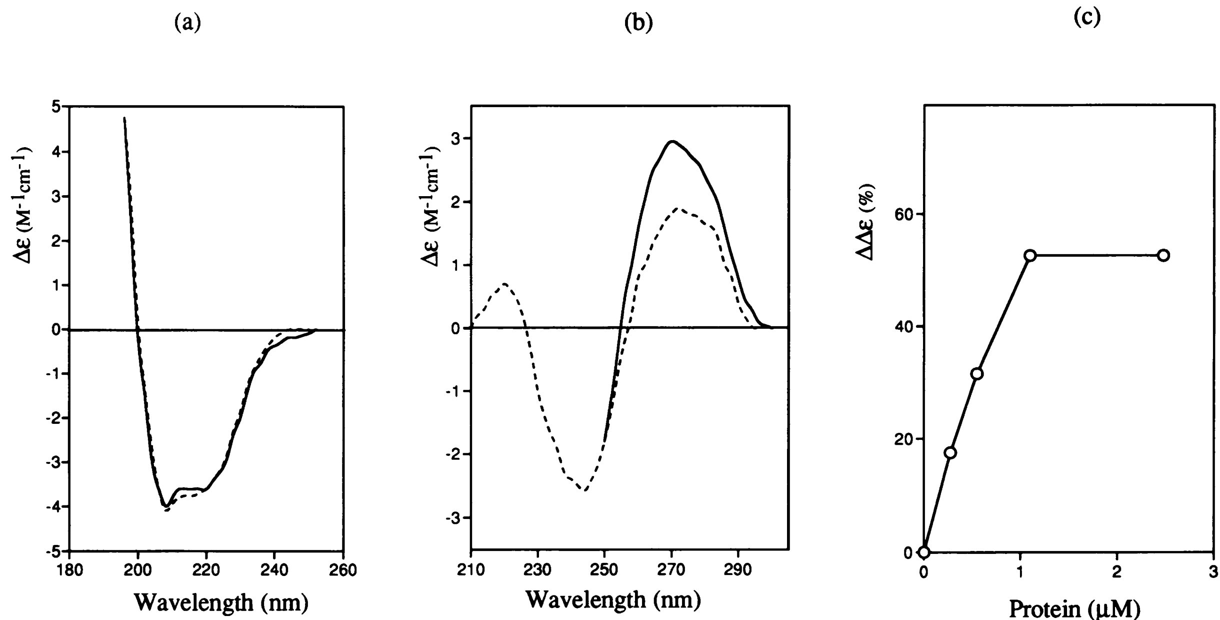

Fig. 4. CD spectroscopy. (**a**) Far UV CD spectra of M.*Eco*R124I (dashed line) and the complex with the 30 bp DNA duplex AB (solid line) at a 1:1 molar ratio. For both spectra, the protein concentration was 2.2 μM. Δε is expressed in terms of amino acid residue concentration (molar). (**b**) Near UV CD spectra of 1.1 μM DNA alone (dashed line) and complexed to M.*Eco*R124I at a 1:1 molar ratio (solid line). Δε is expressed in terms of nucleotide concentration (molar). The small contribution of the protein has been subtracted from the spectrum of the complex for wave lengths >250 nm, to give the CD spectrum corresponding to the bound DNA. Below 250 nm this subtraction is increasingly unreliable due to the dominance of the protein spectrum. (**c**) The change in Δε as a concentrated solution of M.*Eco*R124I (77 μM) is titrated into a fixed concentration of the 30 bp duplex (1.1 μM).

and their structure is imposed only when bound to the DNA, giving rise to a large change in their CD spectrum in the far UV (Weiss *et al.*, 1990). Recent crystallographic structures of *Hha*I methyltransferase and *Eco*RV endonuclease show that there are substantial structural changes in both enzymes when bound to their respective DNA recognition sequence (Winkler, 1992; Cheng *et al.*, 1993). Indeed, changes to the quaternary structure of proteins can also accompany DNA binding, as shown by a large rotation of the subunits of the λ-cro protein dimers when bound to the appropriate operator sequence (Brennan *et al.*, 1990).

Small angle X-ray solution scattering measurements presented here reveal that an unusually large compaction of the *Eco*R124I methyltransferase accompanies binding to an oligonucleotide duplex containing its cognate recognition site. The structural change is seen as a 16 Å decrease in the R_g of the enzyme on binding to the duplex, with a compaction of nearly 70 Å in the overall dimensions of the complex. The DNA contributes only 11% to the mass of the complex, and its direct contribution to the radius of gyration of the complex is almost negligible. The biggest reduction in R_g would be obtained if the DNA were at the centre of the complex, but this would only reduce the overall R_g by 2.4 Å (see Materials and methods). Thus, the large change in R_g must be attributed to a major structural change in the protein itself. The profound alteration in the shape of the distance distribution function when DNA is bound is further evidence of a significant global change in the structure of the enzyme. The magnitude of this structural change is greater than any reported previously, and represents a much larger change than, for example, the 25 Å movement of the catalytic loop of M.*Hha*I towards the DNA (Cheng *et al.*, 1993), which would have only a small effect on the radius of gyration.

On the basis of far UV CD spectroscopy, we conclude that there is no appreciable change in the secondary structure of the subunits of the methyltransferase when DNA is bound. Perturbation of the tertiary structure of the subunits of the methyltransferase (or of their domains) cannot be ruled out, but such changes could not account for the reduction of almost 70 Å in the dimensions of the enzyme when bound to DNA. The simplest explanation of the dramatic structural transition induced by DNA binding is that there is a quaternary structural change involving a major rearrangement of the subunits. This suggests that a hinge bending mechanism is involved in forming the DNA–protein complex, bringing the subunits closer together to form a much more compact structure than the free uncomplexed methyltransferase.

There is, however, a pronounced alteration in the conformation of the DNA in the complex with the methyltransferase, based on a large increase in the near UV peak in the CD spectrum. The increase we observe is in the same direction as that found when the bacterial repressors *Gal* and *Lac* bind to their cognate DNA sequences (Culard and Maurizot, 1981; Wartell and Adhya, 1988); however, for M.*Eco*R124I the magnitude is much larger (an increase of 52% in $\Delta\varepsilon_{272}$) and is comparable with the increase seen when the *Tet* repressor binds to its operator sequence (Altschmied and Hillen, 1984). The CD change is opposite in sign to that exhibited by *Cro* and *Cap* when they bind to their DNA recognition sites (Blazy *et al.*, 1987; Torigoe *et al.*, 1991). The large increase in CD induced by M.*Eco*R124I suggests a significant conformational change in the DNA, possibly involving local unwinding of the helix and/or an increase in propeller twist of the base pairs (Johnson *et al.*, 1981). This structural change is highly localized, and does not give rise to bending of the DNA, as judged by gel electrophoresis of complexes

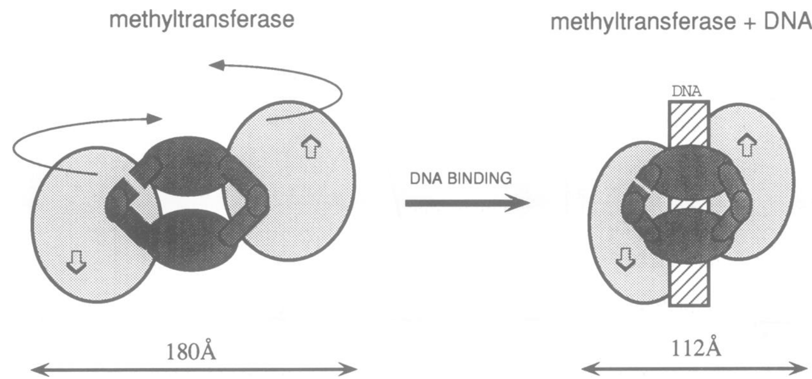

Fig. 5. Schematic model showing the open and closed forms of M.*Eco*R124I corresponding to the free enzyme and the complex with DNA. The two symmetrically arranged HsdM subunits (light grey) are attached to the single HsdS subunit (dark grey) via the conserved domains of HsdS. The 30 bp DNA duplex containing the recognition sequence is shown (hatched). The two variable domains of HsdS are separated by ~34 Å, corresponding to the spacing of the two half-sites of the recognition sequence to which they bind. Binding to the specific DNA sequence causes a rotation of the HsdM subunits around the DNA, making non-specific contacts with the back face of the DNA helix to generate a more compact structure.

formed with DNA fragments in which the binding site has been circularly permuted (D.Mernagh and G.G.Kneale, unpublished data).

It is possible that local unwinding of the DNA induced by M.*Eco*R124I may be required for the catalytic residues of the methyltransferase (and the methyl group donor, *S*-adenosyl methionine) to gain proper access to the N6 amino group of adenine that is the target for methylation by the enzyme. Indeed, a striking example is provided by the crystal structure of the type II methyltransferase M.*Hha*I, in which the base to be methylated (in this case a cytosine) is flipped completely out of the helix into the active site of the enzyme to allow the base to be modified. Whether such a drastic perturbation is required for adenine-specific methylases is still unclear (Suck, 1994), but the structural distortion of DNA by M.*Eco*R124I could serve a similar purpose.

We have proposed recently a schematic model for the structure of the specificity subunit (HsdS) of type I methyltransferases based on the symmetrical arrangement of the (conserved) repeated domains that are interspersed between the two (variable) DNA recognition domains. In this model, the N- and C-termini of the HsdS subunit are found in close proximity, giving rise to a 'circular' arrangement of domains which has the required pseudo-dyad symmetry to interact with the two identical HsdM subunits and the anti-parallel sugar-phosphate chains of the DNA (Kneale, 1994). The methyltransferase has a stoichiometry of M_2S_1, but HsdM itself is monomeric in solution (Taylor, 1992). This suggests that there are no contacts between the HsdM subunits in the trimeric enzyme, and that interactions are only made with the HsdS subunit. The HsdM subunits are believed to interact principally with the conserved repeated domains of HsdS, as suggested for a number of type I methyltransferases from genetic data (Gough and Murray, 1983) and more recently by limited proteolysis of the type IA methyltransferase M.*Eco*KI (Cooper and Dryden, 1994).

Figure 5 shows how it is possible to accommodate the transition from an extended ('open') structure in the free enzyme to a compact ('closed') structure in the DNA−protein complex in a way that could account for a reduction of ~70 Å in the maximum dimension of the complex. The variable domains of HsdS bind in the major groove of DNA to their cognate 'half-sites' separated by one helical turn. We propose that there is some flexibility, either within the conserved domains of HsdS or at the point of contact with HsdM subunits, which amplifies small structural changes in these regions into a large concerted movement of the two HsdM subunits. Thus, the specificity subunit, firmly attached to its recognition sequence on the DNA, can act as a pivot for the two methylation subunits. The resulting compact structure would be stabilized by contacts between the HsdM subunits and the DNA (which cannot form in the extended structure) and would serve to locate the catalytic site of each HsdM subunit near their target DNA sequence. These contacts are likely to include non-sequence-specific contacts extending well past the specific recognition site, as M.*Eco*R124I requires an oligonucleotide duplex of at least 18 bp to bind effectively (I.A.Taylor and G.G.Kneale, unpublished data; see also Powell *et al.*, 1993, for the case of *Eco*KI).

We have shown previously that the methylation state of the DNA has a profound effect on both binding to the methyltransferase and the rate of methylation by the enzyme (Taylor *et al.*, 1993); methylation on either strand reduces the binding affinity by a factor of 20–35, but increases enzyme activity by two orders of magnitude. These results suggested that there was communication between the two sites, probably mediated by a conformational change in the enzyme or in the DNA. However, the CD results demonstrate that the localized conformational changes in the DNA induced by binding of the methyltransferase are independent of the state of methylation of the DNA. Moreover, the X-ray scattering data show that the large structural transition that accompanies DNA binding is also unaffected by the methylation state of the DNA. Thus, if any conformational change in the subunits of the methyltransferase does occur following methylation of one DNA strand, it must be a small and localized effect, in contrast to the major structural transition seen in this study.

Materials and methods

Purification of M.EcoR124I

The *Eco*R124I methyltransferase was purified from an overexpressing strain of *Escherichia coli* JM109 (DE3) carrying the plasmid pJS4M in

which the *hsdM* and *HsdS* genes of EcoR124I are tandemly expressed from a T7 promoter (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). Purity and monodispersity of preparations were routinely analysed by SDS–PAGE, reverse phase HPLC and analytical gel filtration. The molar extinction coefficient ε_{280} = 160 400 $M^{-1}cm^{-1}$ was used to obtain the protein concentration by UV absorption spectroscopy (Taylor *et al.*, 1992). The ability of the purified protein to bind to its specific DNA recognition site was confirmed by gel retardation assays using the same duplex as used for the X-ray and CD experiments. The molecular masses of the enzyme and its complex with the 30 bp duplex were confirmed by analytical ultracentrifugation (manuscript in preparation).

Preparation of oligonucleotide duplexes

Synthetic oligonucleotides were purchased HPLC-purified from Oswel DNA Services (University of Edinburgh, UK) to make the 30 bp duplex containing the recognition sequences (shown in bold):

5′-CCGTGCA**GAA**TTCGAG**GTCG**ACGGATCCGG-3′ A strand
3′-GGCACGT**CTT**AAGCTC**CAGC**TGCCTAGGCC-5′ B strand

For the preparation of hemi-methylated duplexes, N6-methyl deoxyadenosine cyanoethyl phosphoramidite was used during the synthesis to incorporate N6-methyl adenine at each half-site of the recognition sequence. The molar extinction coefficient of each oligonucleotide was determined by digestion to completion with snake venom phosphodiesterase and summing the contributions from individual nucleotides. This value was adjusted to give the corrected value for the intact oligonucleotide, taking account of the hyperchromicity observed following digestion (O'Donohue, 1991). Final extinction coefficients for each strand were ε_{260} = 240 482 (A strand) and ε_{260} = 236 940 $M^{-1}cm^{-1}$ (B strand).

Duplexes were prepared by mixing oligonucleotides in equimolar proportions, heating to 85°C for 5 min and then cooling slowly to room temperature (Taylor *et al.*, 1993). Three duplexes were prepared, one unmodified (AB) and two hemi-methylated duplexes: one with N6-methyl adenine at position 10 in the A strand (A10B) and the other with the same modification but at position 13 in the B strand (AB13). Duplex formation was assayed by native gel electrophoresis and by analytical gel filtration on a Superose 12 HR 10/30 column (Pharmacia). An extinction coefficient for the duplex DNA was determined in the same way as for the single strands by adjustment of the summed extinction of the individual nucleotides, taking into account the hyperchromicity observed after digestion with snake venom phosphodiesterase I, to give ε_{260} = 412 275 $M^{-1}cm^{-1}$.

CD spectroscopy

Near UV (250–350 nm) and far UV (190–250 nm) CD spectra were recorded using a Jasco J600 spectropolarimeter purged with nitrogen gas. Near UV CD spectra were recorded in 1 cm pathlength cells and far UV spectra were recorded in 0.5 mm pathlength cells. All spectra were recorded in 50 mM Tris–HCl, 5 mM $MgCl_2$, 100 mM NaCl, pH 8.0 and corrected by subtraction of a buffer spectrum.

Far UV CD spectra were recorded for M.*Eco*R124I and for a 1:1 complex with a 30 bp oligonucleotide at a concentration of 2.2 μM. In this instance the small contribution from duplex DNA was not subtracted. In near UV titrations of oligonucleotide duplexes with M.*Eco*R124, the protein was added at molar ratios of 0.25, 0.5, 1.0 and 2.25 to a 1.1 μM solution of duplex. Spectra of M.*Eco*R124I were recorded at equivalent concentrations and subtracted to obtain the CD spectrum of the bound DNA. In other experiments, 1:1 complexes of methylase with the hemi-methylated duplexes were prepared at a concentration of 1.1 μM and near UV CD spectra recorded. Again the protein spectrum at the equivalent concentration was subtracted to obtain the CD spectrum of the bound DNA.

Collection of small angle X-ray data

Small angle X-ray scattering data were collected on Station 2.1 using the Synchrotron Radiation Source (SRS) at the SERC Daresbury Laboratory, the operation and configuration of which are detailed elsewhere (Towns-Andrews *et al.*, 1989). Using a camera length of 3.5 m, intensities were measured for values of the scattering vector **Q** extending from 0.009 to 0.230 $Å^{-1}$. Data were collected on a quadrant detector, calibrated using the diffraction pattern of wet rat tail collagen, and processed using the in-house software package OTOKO (Towns-Andrews *et al.*, 1989) to generate intensity against **Q** profiles, from which preliminary values of radii of gyration (R_g) were obtained. To guard against X-ray-induced sample damage, the data were collected in 10 s time frames and the R_g determined at intervals during the course of a scattering experiment, so that those frames in which the R_g had changed significantly could be excluded from subsequent analysis. However, the general experience was that there was no appreciable change in the R_g over collection times of up to 10 min duration. All samples were present in 20 mM Tris–HCl, pH 7.5, 3 mM DTT. Scattering from the buffer was measured for each sample and subtracted before analysis.

Data processing and analysis

Files processed using OTOKO were converted to ASCII format using the software package RECONV2. Further manipulations of intensity data were carried out using standard software packages on an Apple Macintosh computer. Guinier plots were constructed and the R_g derived from the slopes of the plots. Distance distribution functions were calculated using indirect Fourier transformation of intensities between **Q** = 0.014 and 0.170 $Å^{-1}$ using the software package GNOM (Semenyuk and Svergun, 1991). Calculated scatter curves in reciprocal space were obtained by Fourier transformation of the derived $p(r)$ curves, and gave a good fit to the experimental data points in all cases. R_g data were calculated directly from GNOM and these values were in excellent agreement with the results of the Guinier analysis.

An estimation of the contribution of the DNA duplex to the R_g of the complex was performed using the parallel axis theorem (see, for example, Lederer *et al.*, 1989):

$$R_g^2 = f_1R_1^2 + f_2R_2^2 + f_1f_2\Delta^2,$$

where f_1 and f_2 are the fraction of electrons contributed by the two components (protein and DNA), which approximates to their fractional contribution to the mass of the complex (0.89 and 0.11 respectively). R_1 and R_2 refer to the radii of gyration of the individual components (56 and 28 Å; see Table I). The smallest R_g possible on this basis is obtained when the distance between the centres of the two components (Δ) is zero, in which case an overall R_g of 53.6 Å is obtained.

Acknowledgements

We are grateful to Dr Alex Drake (Birkbeck College, University of London, UK) for the use of CD facilities, and to Drs John Harries, Geoff Mant and Gunter Grossmann (Daresbury Laboratory, UK) for generous advice and assistance with instrumentation and data processing. Both facilities are supported by SERC. We thank Dr D.L.Svergun for providing the GNOM data analysis package, and the Wellcome Trust and the Leverhulme Trust for financial support.

References

Abadjieva,A., Patel,J., Webb,M., Zinkevich,V. and Firman,K. (1993) *Nucleic Acids Res.*, **21**, 4435–4443.

Altschmied,L. and Hillen,W. (1984) *Nucleic Acids Res.*, **12**, 2171–2180.

Bickle,T.A. and Kruger,D.H. (1993) *Microbiol. Rev.*, **57**, 434–450.

Blazy,B., Culard,F. and Maurizot,J.C. (1987) *J. Mol. Biol.*, **195**, 175–183.

Brennan,R.G., Roderick,S.L., Takeda,Y. and Matthews,B.W. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 8165–8169.

Carpenter,M.L. and Kneale,G.G. (1994) *Methods Mol. Biol.*, **30**, 313–326.

Cheng,X., Kumar,S., Pflugrath,J.W. and Roberts,R.J. (1993) *Cell*, **74**, 299–307.

Compton,L.A. and Johnson,W.C.,Jr (1986) *Anal. Biochem.*, **155**, 155–161.

Cooper,L.P. and Dryden,D.T.F. (1994) *J. Mol. Biol.*, **236**, 1011–1021.

Cowan,G.M., Gann,A.A.F. and Murray,N.E. (1989) *Cell*, **56**, 103–109.

Culard,F. and Maurizot,J.C. (1981) *Nucleic Acids Res.*, **9**, 5175–5184.

Dryden,D.T.F., Cooper,L.P. and Murray,N.E. (1993) *J. Biol. Chem.*, **268**, 13228–13236.

Fuller-Pace,F.V and Murray,N.E. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 9368–9372.

Fuller-Pace,F.V., Cowan,G.M. and Murray,N.E. (1985) *J. Mol. Biol.*, **186**, 65–75.

Gough,J.A. and Murray,N.E. (1983) *J. Mol. Biol.*, **166**, 1–19.

Gubler,M., Braguglia,D., Meyer,J., Piekarowitz,A., and Bickle,T.A. (1992) *EMBO J.*, **11**, 233–240.

Johnson,B.B., Daki,K.S., Tinoco,I., Ivanov,V.I. and Zhurkin,V.B. (1981) *Biochemistry*, **20**, 73–78.

Kannan,P., Cowan,C.M., Daniel,A.S., Gann,A.A.F. and Murray,N.E. (1989) *J. Mol. Biol.*, **209**, 335–344.

Kellerher,J.E., Daniel,A.S. and Murray,N.E. (1991) *J. Mol. Biol.*, **221**, 431–440.
Kneale,G.G. (1994) *J. Mol. Biol.*, **243**, 1–5.
Kumar,S., Cheng,X., Klimasauskaus,S., Mi,S., Posfai,J., Roberts,R.J and Wilson,G.G. (1994) *Nucleic Acids Res.*, **22**, 1–10.
Lederer,H., Tovar,K., Baer,G., May,R.P., Hillen,W. and Heumann,H. (1989) *EMBO J.*, **8**, 1257–1263.
Meister,J., McWilliams,M., Hubner,P., Jutte,H., Skrzypek,E., Piekarowicz,A. and Bickle,T.A. (1993) *EMBO J.*, **12**, 4585–4591.
Murray,N.E., Gough,J.A., Suri,B. and Bickle,T.A. (1982) *EMBO J.*, **1**, 535–539.
O'Donohue,M.J. (1991) Ph.D. thesis, University of Portsmouth, UK.
Patel,J., Taylor,I., Dutta,C., Kneale,G.G. and Firman,K. (1992) *Gene*, **112**, 21–27.
Perkins,S.J. (1988) *Biochem. J.* **254**, 313–327.
Powell,L.M., Dryden,D.T.F., Willcock,D.F., Pain,R.H. and Murray,N.E. (1993) *J. Mol. Biol.*, **234**, 60–71.
Price,C., Pripfl,T. and Bickle,T.A. (1987) *Eur. J. Biochem.*, **167**, 111–115.
Semenyuk,A.V. and Svergun,D.I. (1991) *J. Appl. Crystallogr.*, **24**, 537–540.
Sharp,P.M., Kellerher,J.E., Daniel,A.S., Cowan,G.M. and Murray,N.E. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 9836–9840.
Suck,D. (1994) *Curr. Biol.*, **4**, 252–255.
Suri,B. and Bickle,T.A. (1984) *J. Mol. Biol.*, **186**, 77–85.
Taylor,I.A. (1992) PhD thesis, University of Portsmouth, UK.
Taylor,I.A., Patel,J., Firman,K. and Kneale,G.G. (1992) *Nucleic Acids Res.*, **20**, 179–186.
Taylor,I.A., Watts,D. and Kneale,G.G. (1993) *Nucleic Acids Res.*, **21**, 4929–4935.
Torigoe,C., Kidokoro,S., Takimoto,M., Kyoyoku,Y. and Wada,A. (1991) *J. Mol. Biol.*, **219**, 733–746.
Towns-Andrews,E., Berry,A., Bordas,J., Mant,G.R., Murray,P.K., Roberts,I., Worgan,J.S., Lewis,R. and Gabriel,A. (1989) *Rev. Sci. Instrum.*, **60**, 2346–2349.
Tyndall,C., Meister,J. and Bickle,T.A. (1994) *J. Mol. Biol.*, **327**, 266–274.
Wartell,R.M. and Adhya,S. (1988) *Nucleic Acids Res.*, **16**, 11531–11541.
Weiss,M.A., Ellenberger,T., Wobbe,C.R., Lee,J.P., Harrison,S.C. and Struhl,K. (1990) *Nature*, **347**, 575–578.
Wilson,G.G. and Murray,N.(1991) *Annu. Rev. Genet.*, **25**, 585–627.
Winkler,F.K. (1992) *Curr. Opin. Struct. Biol.*, **2**, 93–99.

Received on August 4, 1994; revised on September 12, 1994