Purification and biochemical characterisation of the *EcoR124* type I modification methylase

Ian Taylor, Jaynish Patel, Keith Firman and Geoff Kneale*

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

Received November 25, 1991; Accepted December 18, 1991

ABSTRACT

Large scale purification of the type I modification methylase *EcoR124* has been achieved from an over-expressing strain by a two step procedure using ion-exchange and heparin chromatography. Pure methylase is obtained at a yield of 30mg per gm of cell paste. Measurements of the molecular weight and subunit stoichiometry show that the enzyme is a trimeric complex of 162 kDa consisting of two subunits of HsdM (58 kDa) and one subunit of HsdS (46 kDa). The purified enzyme can methylate a DNA fragment bearing its cognate recognition sequence. Binding of the methylase to synthetic DNA fragments containing either the *EcoR124* recognition sequence GAAN₆RT-CG, or the recognition sequence GAAN₇RTCG of the related enzyme *EcoR124/3*, was followed by fluorescence competition assays and by gel retardation analysis. The results show that the methylase binds to its correct sequence with an affinity of the order 10^8 M^{-1} forming a 1:1 complex with the DNA. The affinity for the incorrect sequence, differing by an additional base pair in the non-specific spacer, is almost two orders of magnitude lower.

INTRODUCTION

Type I restriction and modification (R-M) systems are of great interest, not only for their physiological role in providing the host bacteria with 'immunity' to infection by foreign DNA, but also as complex multisubunit enzymes with unusual DNA recognition properties at the molecular level (for a review, see ref.1). Unlike type II R-M systems, a number of genes are required for restriction activity. Moreover, the sequence specificity of type I enzymes is quite different to that of the simpler type II enzymes, since the recognition sequences of the former consist of two asymmetric half sites separated by some 6–8 non-specific bases. In contrast the majority of type II enzymes recognise a sequence of 4–8 bases consisting of two adjacent and symmetrical half sites, consistent with the homodimeric structure that is typical of such enzymes.

A number of type I enzymes have now been characterised from *Escherichia coli* and *Salmonella typhimurium* and have been

classified into three distinct families (1) but all have similar features. Each system is comprised of 3 genes, *hsdS*, *hsdM* and *hsdR*. For methylase activity, only the HsdS and HsdM proteins are required, but all three gene products are required for endonuclease activity. S-adenosyl methionine (SAM) and Mg^{++} are required for both activities, but the endonuclease requires the additional presence of ATP. The transcription of *hsdM* and *hsdS* genes is tightly coupled, since they are under the control of the same promoter. In contrast, the *hsdR* gene is transcribed from an independent promoter.

There is convincing genetic evidence that the *hsdS* gene is the determinant of DNA specificity. Comparison of the DNA sequences of the *hsdS* genes from a number of type I systems have indicated the presence of two variable domains, separated by a short conserved domain, with an additional conserved domain at the distal end of the gene (2,3). Secondary structure analysis of the predicted protein sequence indicates that the conserved domains are largely alpha-helical, and it was suggested that the specificity determinants resided in these domains (4). The latter proposal, however, is almost certainly incorrect and in an elegant series of domain swapping experiments it was subsequently established that specificity is conferred by the two variable domains of the protein, each corresponding to one half of the bipartite DNA recognition sequence (5,6,7).

The central conserved region of the HsdS sequence is thought to be critical for correct spacing of the two recognition domains. Direct evidence in favour of this comes from a comparison of the HsdS subunits of the type I restriction enzyme *EcoR124* and the related system *EcoR124/3*, which differ only in the presence of the an additional four amino acid residues in the central conserved region of the protein (8). This is sufficient to change the specificity of the system through the requirement for an additional base in the non-specific component of the recognition sequence, leaving the specific bases unchanged (i.e. GAAN₇RT-CG instead of GAAN₆RTCG; ref. 9). Indeed it has recently been demonstrated by site-directed mutagenesis of the spacer region that recognition of both sequences can occur when spacers of intermediate length and of the appropriate sequence are constructed in the *hsdS* gene (10).

Although much has been learnt by analysis of type I restriction systems at the genetic level, studies of the proteins at the

* To whom correspondence should be addressed

biochemical level has lagged behind, due in large part to the relatively low level of expression from their natural promoters. Initial studies on the *EcoR124* and *EcoR124/3* restriction endonucleases have been reported on the enzymes expressed at low level from their natural promoters (11). The enzyme was shown to contain subunits of M_r 116,000, 55,000 and 43,000 corresponding to the products of the *hsdR*, *hsdM* and *hsdS* genes respectively, and to possess DNA dependent ATPase activity and site-specific DNA-methylation activity.

We have recently successfully co-expressed the *hsdM* and *hsdS* genes of *EcoR124* at high level under the control of two independent T7 promoters, each with the appropriate translational signals, in the plasmid expression vector pJS4M. The presence of this plasmid confers site-specific DNA methylation activity *in vivo* (12). We show below that the enzyme exists predominantly as a multi-subunit complex of defined stoichiometry which can be purified in large amounts. The enzyme is shown to be capable of DNA methylation activity *in vitro*, and binds tightly to its cognate DNA recognition sequence, and with a lower affinity to a non-cognate sequence, according to gel retardation and fluorescence competition assays.

MATERIALS AND METHODS

Bacterial strains and plasmids

E. coli JM109 (DE3) and plasmid pJS4M were used to produce *EcoR124* methylase protein (12). Strains were grown in 2×YT media supplemented with ampicillin (150µg/ml).

Growth of cells

A 1:100 dilution of a saturated culture of *E. coli* JM109 (DE3) [pJS4M] was grown in 5L of 2×YT supplemented with ampicillin at 37°C in a Braun-Biolab fermenter for 4 hours. The culture was induced with 1 mM IPTG and grown for a further 6 hours. The cells were harvested by centrifugation, washed in 10 mM Tris-HCl, 100mM NaCl, pH 8.2 and stored as a cell paste at -20°C.

Preparation of cell extract

All procedures were carried out at 4°C. Column fractions were monitored by absorption at 280 nm. Typically 30g of frozen cell paste was suspended in 75ml 50mM Tris-HCl (pH 8.2), 25% Sucrose, 5mM EDTA, 3mM dithiothreitol (DTT), 1mM benzamidine and 100µM PMSF. Cells were disrupted by sonication (15×10 sec bursts) with cooling between pulses. Insoluble debris was removed by centrifugation at 40,000g for 30 mins. The supernatant was made 250mM NaCl by the addition of concentrated salt solution and large molecular weight particles pelleted by centrifugation at 300,000g for 2 hours. Approximately 5g of polyethylimine (PEI) cellulose was added to the supernatant and the suspension stirred for 30 min. The bound nucleic acids were removed by low speed centrifugation (10,000g for 20 mins) and the supernatant adjusted to 70% ammonium sulphate using a saturated solution. This suspension was again stirred for 30 min and the proteins pelleted by low speed centrifugation (10,000g for 20 mins). The pellets were resuspended in approximately 60 ml of 10mM Tris-HCl (pH 8.2), 1mM EDTA, 50mM NaCl (Buffer A) and dialysed exhaustively against the same buffer.

DEAE Sephacel chromatography

The dialysed sample was applied at 1.67 ml/min to 10cm×5.0cm² column of DEAE Sephacel equilibrated with buffer A. The column was washed with buffer A until the absorbance of the eluate had reached baseline. The bound proteins were eluted from the column using a linear gradient of NaCl (50–800 mM, 400ml total volume) in buffer A. The *EcoR124* methylase eluted as a broad peak at around 230 mM NaCl, as judged by SDS-PAGE. These fractions were pooled and dialysed against 10mM Tris-HCl (pH 8.2), 1mM EDTA and 100mM NaCl (Buffer B).

Heparin affinity chromatography

The pooled sample from ion-exchange chromatography was applied to a heparin column (Biorad Econo-Pac cartridge, 5ml) equilibrated with buffer B at a flow rate of 2ml/min in 20 mg aliquots. The column was washed with buffer B until the absorbance at 280 nm reached baseline. The bound protein was eluted with a linear gradient of NaCl (0.1 M–1.0 M, 48 ml total volume) in buffer B. The methylase eluted in a sharp peak at 225 mM NaCl.

Concentration and storage

Purified methylase was concentrated to approximately 10 mg/ml using a Centricon ultrafiltration device (30,000 molecular weight cutoff) and then diluted 1:2 with glycerol. The solution was then made up to 10mM Tris-HCl, pH 8.2, 1mM EDTA, and 300 mM NaCl using concentrated stock solutions. Protein solutions of this type could be stored at -20°C for up to 9 months. Before further use, samples of the methylase were put into an appropriate buffer by gel filtration on Biorad 10DG desalting columns.

Protein analysis

Sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS PAGE) was carried out on 12.5% gels and were stained using Coomassie brilliant blue R250. Concentrations of purified proteins were determined from the UV absorption at 280 nm using extinction coefficients derived from their aromatic amino acid composition (HsdM: $E_{280} = 43,100 \text{ M}^{-1}\text{cm}^{-1}$; HsdS: $E_{280} = 71,600 \text{ M}^{-1}\text{cm}^{-1}$).

Size exclusion chromatography

A 190 ml bed volume column of Sephacryl S-300 was equilibrated with buffer B and calibrated with molecular mass markers (table 1). A 1ml sample (5 mg/ml) of purified methylase was applied to the column and allowed to run on to the bed. Buffer B was run on to the column at a flow rate of 0.2 ml/min and the accurate elution volume of the methylase determined.

Table 1. Molecular mass calibration markers

marker	molecular mass	K _{av}
Blue dextran	ca. 5,000,000	0.00
Apo ferritin	443,000	0.08
Alcohol dehydrogenase	150,000	0.31
Bovine Serum Albumin	66,000	0.38
Ovalbumin	46,000	0.41
Carbonic anhydrase	29,000	0.52
Ribonuclease A	13,700	0.65
Thymidine	242	1.00

Densitometry of gels

Coomassie blue stained gels were scanned using a Hoefer gel scanner equipped with an analogue to digital convertor, to facilitate data collection by computer. Peak areas of each protein band were determined by integration after baseline subtraction. Integrated peak areas of the HsdM and HsdS bands of the methylase were compared with those of a series of defined mixtures of the purified subunits. Extinction coefficients of the latter were calculated from their amino acid sequence.

Spectroscopic determination of Tryptophan content

Tryptophan content was determined following the method of Chrastil (13). A sample of methylase protein was prepared in 20% acetonitrile, 2mM Tris-HCl (pH 8.2), 0.2mM EDTA, 10% glycerol, 60mM NaCl. To 1ml of this sample was added 0.05mls of 0.75% formaldehyde followed by 0.5 mls of 85% sulphuric acid and the reaction mixed and allowed to stand overnight. The absorbance was read at 455nm and the molar concentration of tryptophan calculated from the expression $[\text{Trp}] = \text{OD}_{455} \times 1.55 / 2800$ (where 1.55 = sample dilution factor, 2800 = E_{455} of the coloured tryptophan formaldehyde adduct). Measurement of the OD_{280} of the original protein solution then permitted calculation of the molar extinction coefficient at this wavelength.

N-terminal Sequencing

Determination of the N-terminal sequence of the methylase and of the separately cloned HsdM and HsdS proteins was carried out by automated Edman degradation on an Applied Biosystems 477A pulsed liquid amino acid sequencer, by Dr M.Gore, University of Southampton. Details of the purification and characterisation of the individually cloned proteins will be presented elsewhere.

Preparation of oligonucleotide duplexes

Oligonucleotides containing the *Eco*R124 recognition sequence were purchased from Oswel DNA Services (University of Edinburgh). The oligonucleotides were mixed in equimolar proportions to generate the duplex (I) shown below. The recognition sequence including the non-specific spacer sequence (underlined) is shown in bold.

duplex I 5'-CCGTGCAGAATTCGAGGTCGACGGATCCGG-3'
3'-GGCACGTCTTAAGCTCCAGCTGCCTAGGCC-5'

Oligonucleotides containing the *Eco*R124/3 recognition sequence were synthesised on a Cruachem PS250 DNA synthesizer and purified on Nensorb-prep columns (Dupont) and were mixed in equimolar proportions to generate the duplex (II) shown below.

duplex II 5'-CCGTGCAGAATTCGACGGTCGACGATCCGG-3'
3'-GGCACGTCTTAAGCTGCCAGCTGCTAGGCC-5'

The 30 b.p. fragments differ by the addition of an extra G.C base pair in the non-specific spacer region of the sequence, and the deletion of a similar base pair from outside the recognition sequence. A molar extinction coefficient of $E_{260} = 396,000$ was used for both duplexes.

³²P end-labelling of oligonucleotides

In a typical case, approximately 10μg of oligonucleotide was used in a phosphorylation reaction as described in (14). The labelled oligonucleotide was purified from the unincorporated material

by size exclusion chromatography on a 1.5ml Sephadex CM-25 column. Oligonucleotide-containing fractions were pooled and precipitated with ethanol. The fragments were resuspended in 400μl of dH₂O. The concentration of each fragment was determined from their OD_{260} values. Specific activities were measured by Cerenkov counting in a liquid scintillation counter.

DNA Methylation Assay

Samples of *Eco*R124 methylase were incubated at 37°C in 300μl of 50mM Tris-HCl (pH 8.0), 5mM MgCl₂, 1mM DTT with 5μM oligonucleotide duplex I. The reaction was started with the addition of 3μM ³H S-adenosyl-methionine (Amersham: 81 Ci/mmol) and 25μl aliquots of the reaction mix were withdrawn at timed intervals. Samples were made up to 75% ethanol and 250mM NaCl and the DNA precipitated overnight at -20°C. Pellets were washed with 70% ethanol, dried and resuspended in 500μl of Optiphase 2 scintillation fluid. The amount of incorporated label was determined by scintillation counting the suspended pellets, and converted to picomoles of incorporated methyl groups.

ANS Fluorescence Assay

The fluorescent probe 1-anilinonaphthalene-8-sulphonic acid (ANS) was obtained in a highly purified form (Molecular Probes Inc.) since it is important to avoid the presence of the dimeric form bis-ANS (15). Fluorescence spectra were recorded on a Perkin-Elmer LS5B luminescence spectrophotometer. Buffer conditions were 10mM Tris-HCl (pH 8.2), 1mM EDTA, 100mM NaCl, 5mM MgCl₂, 2% (v/v) glycerol. The excitation wavelength was set to 370 nm, with excitation and emission slits set to 2.5 nm. The emission spectra were recorded between 400 and 600 nm with the detector set at 90° to the excitation beam. For titrations, 0.5μM aliquots of protein were added to a cuvette containing either 1μM oligonucleotide and 100μM ANS or just 100μM ANS alone. Titration curves were generated by subtraction of oligonucleotide plus ANS curves from ANS curves alone.

Gel retardation experiments

In a typical experiment, end labelled oligonucleotide was added to a series of tubes containing methylase over a range of

Figure 1. DEAE-Sepacel chromatography. SDS polyacrylamide gel of fractions from DEAE-Sepacel ion exchange chromatography of JM109(DE3)[pJS4M] soluble extract. Lanes 1-10 correspond to fractions of increasing NaCl concentrations from 100mM to 250mM. The HsdS and HsdM proteins fractionate together on the gradient.

concentrations in a binding buffer of 10% glycerol, 50mM Tris-HCl (pH 8.2), 5mM $MgCl_2$, 1mM DTT. These were incubated for 15 minutes at 4°C then loaded onto a 6% native polyacrylamide gel (40mM Tris-acetate, pH 7.4, 1mM EDTA) which was run at 100V and 4°C. After electrophoresis, gels were dried under vacuum and autoradiographed.

RESULTS

Purification of the *EcoR124* methylase from an over producing strain

EcoR124 methylase was purified from JM109(DE3) bearing the phagemid expression vector pJS4M, a high level expression system in which the *hsdM* and *hsdS* genes are co-expressed from two tandemly arranged T7 gene 10 promoters (12). We have devised a relatively simple two step chromatographic process to produce high yields of pure enzyme from this highly enriched soluble extract, after removal of cell debris and precipitation of nucleic acids by the use of polyethylimine-cellulose. The high capacity and good purification levels afforded by the use of DEAE Sephacel made it an excellent choice for a first step purification. The two subunits of the *EcoR124* methylase elute together as a broad peak around 230 mM NaCl (Figure 1). We estimate the purity to be around 90% after the first step. In order to remove the remaining contaminants heparin affinity chromatography was employed. Protein contaminants and also any remaining nucleic

acids passed straight through the column whilst the methylase bound and could be eluted within the salt gradient as a sharp peak at 225 mM NaCl (Figure 2). Analysis of the protein within this peak by SDS-PAGE and Coomassie blue staining revealed two bands of molecular mass 58,000 and 46,000 respectively. The results suggest that the *EcoR124* methylase exists as a multimeric complex containing the HsdM and HsdS proteins.

Solution molecular mass

In order to determine the subunit composition of the *EcoR124* methylase, the molecular mass of the protein in solution was measured by size exclusion chromatography on a calibrated S-300 column (Pharmacia). The elution profile and the column calibration curve are shown in Figure 3. Analysis of the major peak by SDS PAGE confirms both polypeptides are present. The minor peak which elutes closer to the void volume also contains both polypeptides and we suggest that at high concentrations a small proportion of the methylase exists as a dimer. The K_{av} value for the *EcoR124* methylase in the major peak is 0.24 which, from the calibration curve, indicates a solution molecular mass of 170,000 for the multimeric enzyme. The subunit composition closest to this value would be M_2S (calculated $M_r = 162$ kDa). However we cannot, from this data alone, unambiguously rule out the possibility of, for example, MS_3 (calculated $M_r = 196$ kDa). In order to determine the precise subunit composition it is necessary to examine the ratio of subunits in the intact methylase.

Figure 2. (a) Elution profile of the heparin affinity column following ion-exchange chromatography of the soluble cell extract. (b) track 2: SDS polyacrylamide gel analysis of the major peak in (a) and track 1: molecular weight standards (66, 46, 36, 29, 24, and 20.1 kDa) (c) Densitometer scan of track 2, showing the purity of the sample and the ratio of the two subunits, HsdM and HsdS.

Figure 3. (a) Elution profile of *EcoR124* methylase on a Sephacryl S-300 size exclusion column. (b) Calibration curve used to calculate the solution molecular mass. Elution volume is represented in terms of K_{av} , where $K_{av} = (V_e - V_0) / (V_t - V_0)$. V_e is the elution volume, V_0 the void volume (82 ml) and V_t the column bed volume (190 ml). The elution position of the methylase is marked with an arrow.

Stoichiometry and N-terminal Analysis

Coomassie blue is frequently used for staining protein gels following SDS PAGE and is also employed for the estimation of protein concentrations in solution (16). However, it is well known that there can be a considerable variation in the extinction coefficient of the coloured product from protein to protein (17) and for quantitative estimation it is essential to use calibration standards. For the purposes of our analysis, it was sufficient to obtain the relative efficiencies of staining of the two polypeptide components of the methylase, HsdM and HsdS. Three different loadings of a 2:1 molar ratio of the purified polypeptides were run on SDS PAGE, together with similar loadings of the purified methylase. Integrated peak areas of the 2:1 mixtures gave a ratio of 2.02 ± 0.08 (M:S), indicating that the efficiency of staining of the two components is essentially identical in this case. A similar analysis of tracks containing the pure methylase indicated a M:S ratio of 1.98 ± 0.08 (figure 2c). Taken together with the molecular mass determination of the methylase by gel filtration, the subunit composition is unambiguously established as M_2S . In such a complex, there will be 14 tryptophans according to the predicted protein sequences of the subunits, from which the molar extinction coefficient was determined experimentally ($E_{280} = 159,700$) by quantitative analysis of tryptophan. This value is in excellent agreement with the extinction coefficient predicted directly from the sum of the aromatic contributions to the spectrum ($E_{280} = 157,800$) and confirms the stoichiometry.

N-terminal analysis of the purified methylase and the purified polypeptide components (obtained from separate clones) was also carried out. The N-terminal sequences of the purified HsdS protein was established as SEMSYLEK in accordance with the sequence predicted from the open reading frame of the gene (8). The N-terminal sequence of the HsdM protein was established as (M)TSIQQRAEL (the N-terminal methionine in this sequence was observed to be only partially processed). In this case the sequence differs from that predicted in (8), since the first two or three amino acids are missing. It is possible that the sequence Met-Lys-(Met) has been processed post-translationally. A more

likely explanation, however, is that translation in fact starts from the second methionine codon, since this reading frame possesses the better Shine-Dalgarno consensus sequence. Analysis of the N-terminal sequencing data for the intact methylase was complicated by the observation that the N-terminal methionine of HsdM was only partially processed, such that there were approximately equal quantities of methionine, threonine and serine in the first cycle. Further cycles were also consistent with a stoichiometry of 2:1 (M:S) with approximately 50% of HsdM having the sequence MTSIQQRAEL.

Methylation Assay

Previous studies of the methylation activity of type I endonucleases have indicated considerable variation in rates of methylation within this family (18). Since these assays employed plasmid DNA as substrate, the role of non-specific DNA binding on the kinetics of the reaction could be significant. To circumvent these problems, we have investigated the methylation of synthetic DNA fragments bearing the recognition sequence for *EcoR*124 using a well characterised and highly purified enzyme preparation. *In vitro* methylation of a DNA fragment bearing the *EcoR*124 recognition sequence (duplex I) was monitored by the transfer of a tritiated methyl group from *S*-adenosyl methionine (SAM) to the DNA fragment. The incorporation of methyl groups as a function of time is shown in Figure 4, for protein concentrations of 140 nM and 420 nM. It is evident from the figure that the rate of incorporation is slow, and the reaction occurs over a period of hours despite the relatively high concentration of the enzyme. Taking into account the concentration of enzyme present, the initial velocity of the reaction can be estimated as $0.005 \mu\text{mol} \cdot \text{min}^{-1} (\mu\text{mol enzyme})^{-1}$ under the conditions tested. This rate of reaction is more typical of that observed for *EcoK* methylase (a type I(A)

Figure 4. *In vitro* methylation of a DNA fragment. Enzyme activity was monitored by measuring the transfer of a tritiated methyl group from *S*-adenosyl methionine ($3 \mu\text{M}$) to the DNA fragment (duplex I; $5 \mu\text{M}$). $25 \mu\text{l}$ samples were counted for levels of radioactivity at fixed times. The enzyme concentration was 140 nM (lower curve) or 420 nM (upper curve).

Figure 5. Fluorescence quenching of ANS. The fluorescence spectrum of $100 \mu\text{M}$ ANS was measured alone (lower curve), in the presence of $1 \mu\text{M}$ *EcoR*124 methylase (upper curve), and in the presence of a 1:1 complex of the methylase ($1 \mu\text{M}$) with DNA duplex (I) (middle curve).

enzyme) than that of the faster *EcoA* methylase (a type 1(B) enzyme) (18). Price et al. (11) have monitored methylation of plasmid DNA by the *EcoR124/3* endonuclease; however, it is not possible to compare the kinetic parameters quantitatively with our data on *EcoR124* methylase since the enzyme concentration was not defined.

Measurement of DNA binding by displacement of a fluorescent probe

Binding of oligonucleotides to proteins is frequently measured by monitoring the change in intrinsic fluorescence of either tyrosine or tryptophan in the polypeptide as a function of a DNA titration (see, for example, ref. 19). The *EcoR124* methylase, however, does not show any measurable change in the intrinsic fluorescence spectrum during such a titration (data not shown), suggesting that none of the 14 tryptophans in the methylase are directly involved at the DNA binding site of the enzyme, or that any such change is masked. We cannot rule out the possibility of tyrosine and/or phenylalanine involvement since their contribution to the fluorescence spectrum is small. A more indirect approach to monitor DNA binding was therefore adopted: the displacement of a bound fluorescent probe by a synthetic DNA fragment.

The fluorescent probe 1-anilinonaphthalene-8-sulphonic acid (ANS) and its derivatives have previously been used to study

protein structure (20) and more recently for the analysis of protein-nucleic acid interactions (21). The fluorescence of ANS is enhanced approximately 100 fold when transferred from an aqueous environment to less polar solvents such as methanol, and is accompanied by a 50 nm blue wavelength shift. ANS can show similar fluorescence enhancement and wavelength shifts, although less pronounced, when bound to proteins. Thus bound molecules of ANS fluoresce very strongly at a shorter wavelength than free molecules of ANS in an aqueous solvent. The above properties make ANS ideal for investigating hydrophobic patches and clefts on the surface of proteins. ANS has been used to monitor binding of co-repressor (tryptophan) to *trp* repressor, by fluorimetric monitoring of the displaced ANS (22). We have used a related approach, but in this case to measure binding of oligonucleotides to the *EcoR124* methylase.

Figure 5 shows the fluorescence spectral changes that result from the addition of the *EcoR124* methylase to ANS, indicating binding of the fluorescent probe to hydrophobic regions of the protein. The fluorescence is enhanced approximately twofold, and the emission maximum shifts from 520nm to 480nm. Competition with a DNA fragment (duplex I) produces a reversal of this effect on the spectrum, due to the displacement of a substantial fraction of bound ANS molecules from sites on the protein. Even though the fluorescent probe is in 100× molar excess over the DNA fragment, the effect of ANS displacement is appreciable, indicating that the affinity of the methylase for the DNA duplex is very much stronger than that for ANS, even at micromolar concentrations of DNA.

In three separate experiments, ANS was titrated with the methylase alone, and in the presence of either duplex I or duplex II, the latter sequence containing an additional base-pair in the non-specific spacer region of the DNA sequence (Figure 6). Both oligonucleotides compete effectively with the fluorescent probe and suppress the enhancement of fluorescence, showing that both sequences bind with affinities of at least 10^6 M^{-1} . However, binding of the specific sequence (duplex I) is more pronounced and has a more defined break point. The break point in the titration occurs at a molar ratio of approximately 1:1, indicating a stoichiometry of 1 duplex bound per methylase. It is clear from these results that, under the conditions of the experiment, *EcoR124* methylase binds strongly to DNA fragments containing either the cognate recognition sequence or a non-cognate

Figure 6. ANS displacement assay. (a) Fluorescence at 480nm of 100 μM ANS was measured as a function of added *EcoR124* methylase (upper curve). The experiment was repeated in the presence of DNA duplex I (1 μM; lower curve) or DNA duplex II (1 μM; middle curve). (b) Decrease in observed fluorescence due to competition of ANS with duplex I (upper curve) or duplex II (lower curve), obtained by subtraction of the appropriate titration curves in (a).

Figure 7. Oligonucleotide gel retardation assay. Increasing concentrations of purified methylase were added to (a) Duplex I, containing the *EcoR124* recognition sequence (30nM) and (b) Duplex II containing the *EcoR124/3* recognition sequence (64nM). Protein concentrations were as follows: (a) Tracks 1–6; 18, 27, 36, 45, 67.5, and 90 nM methylase; track 7, no protein (b) Tracks 1–8; 26, 38.5, 51, 64, 96, 128, 192, and 252 nM methylase; track 9, no protein.

sequence. In order to estimate the degree of sequence discrimination by *EcoR124* methylase a higher sensitivity technique was used.

Gel Retardation Analysis

The Interaction of *EcoR124* methylase with both oligonucleotide duplexes was investigated by gel retardation analysis. Figure 7(a) shows the mobility of duplex I (30nM) in the presence of increasing concentrations of protein. Binding is very close to stoichiometric under these conditions, since almost all the DNA is complexed to the protein above a protein:DNA ratio of 1. Appreciable DNA binding could be observed at much lower concentrations of DNA (3nM) and a K_d of the order of 10nM was estimated for this interaction. Figure 7(b) shows a similar titration with the non-cognate DNA fragment, duplex II (64nM) but in this case the binding is much weaker. Even at an oligonucleotide concentration of 160nM, the fraction of DNA bound was far from stoichiometric (data not shown). We estimate a K_d in the region of 500nM for the interaction of the methylase with duplex II.

In the experiment with duplex II higher molecular weight bands in addition to the major complex band were observed at higher protein concentrations, indicating the formation of complexes with two or more proteins bound. We suggest that the affinity of *EcoR124* methylase for the specific sequence GAAN₆RTCG is sufficiently high that it is unable to redistribute on the 30mer DNA fragment (duplex I) to accommodate additional protein molecules at high protein:DNA ratios. In the experiment with duplex II, at a 1:1 stoichiometry the methylase also appears to prefer a single site in (most probably the central pseudo-recognition sequence GAAN₇RTCG); however, at higher concentrations of protein it is able to slide (or dissociate) from this site rather more readily, and allow the formation of additional non-sequence specific binding with two or three copies of the protein bound per duplex. Such complexes are apparently unable to form on the duplex bearing the precise recognition sequence, since only a single complex band is observed under all conditions tested (up to a DNA concentration of 155 nM and a molar protein:DNA ratio of 2.0).

DISCUSSION

We have purified large quantities of the intact *EcoR124* DNA methylase from a high level expression system in order to characterise the solution properties of the enzyme and to analyse its DNA binding properties. The methylase is shown to exist as a trimeric complex of molecular mass 162 kDa, consisting of two copies of the HsdM subunit (58 kDa each) and a single copy of the HsdS subunit (46 kDa); there is also evidence for the existence of a small proportion of an aggregate consisting of two copies of the trimer. The stoichiometry of M₂S we have established for the *EcoR124* methylase is the same as that determined for *EcoK* methylase (D.Dryden and N.E.Murray, personal communication), showing further similarities between type IC enzymes and other type I methylases. The purified *EcoR124* methylase is capable of methylation of a DNA fragment containing its recognition sequence GAAN₆RCTG *in vitro*. Since a 30b.p. oligonucleotide fragment containing this sequence is capable of displacement of the fluorescent ligand ANS, which binds to hydrophobic sites on proteins, we suggest that the active site of the enzyme contains apolar residues which may contribute

to the binding energy, in addition to the polar interactions that will be necessary for sequence recognition.

Gel retardation analysis shows that the methylase binds tightly to a DNA fragment containing the specific *EcoR124* sequence with a 1:1 stoichiometry and an estimated binding constant of the order of 10^8 M^{-1} . Binding to this sequence is approximately two orders of magnitude higher than binding to a DNA fragment bearing the *EcoR124/3* site. This degree of specificity in binding is not unreasonable for the addition of an extra base pair in the non-specific spacer region of the DNA recognition sequence, and could be accounted for by a certain degree of flexibility between the two DNA binding domains of the HsdS subunit of the methylase. In general, the specificity of DNA binding need not necessarily reflect the specificity of enzyme activity, since additional discrimination can be shown in the catalytic step, as exemplified by the case of the type II restriction enzyme *EcoRV* (23). In the case of *EcoR124* methylase, however, the level of discrimination of the *EcoR124* and *EcoR124/3* recognition sequences we find by gel retardation is of a similar magnitude to that estimated for methylation activity *in vivo* and perhaps also *in vitro* (10). This suggests that there is little or no additional specificity provided by catalysis, at least in the ability to discriminate the length of the non-specific spacer. Further experiments are in progress in our laboratory to quantitate both the binding and catalytic specificities of *EcoR124* methylase for a range of point mutations within the specific bipartite recognition sequence.

ACKNOWLEDGEMENTS

This work has been supported by research grants from SERC and the Wellcome Trust. The Royal Society are gratefully acknowledged for the award of a Leverhulme Trust Senior Research Fellowship (to G.G.K.). We thank SERC for the award of a research studentship (to IT), Dr M.Gore (University of Southampton) for performing amino acid sequencing, and Dr D.Dryden (University of Edinburgh) for helpful discussion.

REFERENCES

1. Bickle, T.A. (1987) in *Escherichia coli and Salmonella typhimurium: Cellular and Molecular Biology*. American Society for Microbiology, Washington D.C.
2. Gough, J.A. and Murray, N.E. (1983) J. Mol. Biol. 166, 1–19
3. Kannon, P., Cowan, G.M., Daniel, A.S., Gann, A.A.F. and Murray, N.E. (1989) J. Mol. Biol. 209, 335–344
4. Argos, P. (1985) EMBO J. 4, 1351–1355
5. Fuller-Pace and Murray, N.E. (1986) Proc. Nat. Acad. Sci. 81, 6095–6099
6. Gann, A.A.F. et al., (1987) Mol. Microbiol. 1, 13–22
7. Cowan, G.M., Gann, A.A.F., and Murray, N.E. (1989) Cell 56, 103–109
8. Price, C., Lingner, J., Bickle, T.A., Firman, K. and Glover, S.W. (1989) J. Mol. Biol. 205, 115–125
9. Price, C., Shepherd, J.C.W. and Bickle, T.A. (1987a) EMBO J. 6, 1493–1497
10. Gubler, M. and Bickle, T.A. (1991) EMBO J. 10, 951–957
11. Price, C., Pripfl, T. and Bickle, T.A. (1987b) Eur. J. Biochem 167, 111–115
12. Patel, J., Taylor, I., Dutta, C., Kneale, G.G. and Firman, K. (1991) Gene, in the press
13. Chrastil, J. (1986) Anal. Biochem. 158, 443–446
14. Maniatis, T., Fritsch, E.F. and Sambrook, J. (1982) Molecular Cloning: A Laboratory Manual (Cold Spring Harbor University Press, Cold Spring Harbour)
15. York, S.S., Lawson R.C., and Worah, D.M. (1978) Biochemistry 21, 4480–4486
16. Bradford, M.M. (1976) Anal. Biochem. 72, 248–254

17. Tal, M., Weissman, I., Silberstein, A. (1990) *J. Biochem. Biophys. Methods* 21, 247–266
18. Suri, B. and Bickle, T.A. (1985) *J. Mol. Biol.* 186, 77–85
19. Kneale, G.G. and Wijnaendts, R. (1985) *Eur. J. Biochem.* 149, 85–93
20. Slavik, J. (1982) *Biochim. Biophys. Acta* 694, 1–25
21. Secnik, J., Wang, Q., Chang, C-M., and Jentoft, J.E. (1990) *Biochemistry* 29, 7991–7997
22. Chou W-Y, Bieber, C. and Mathews, K. (1989) *J. Biol. Chem.* 264, 18309–313
23. Taylor, J.D., Badcoe, I.G., Clarke, A.R. and Halford, S.E. (1991) *Biochemistry* 30, 8743–8753