Determination of the order of substrate addition to *MspI* DNA methyltransferase using a novel mechanism-based inhibitor

Christine TAYLOR, Kevin FORD, Bernard A. CONNOLLY* and David P. HORNBY†

Krebs Institute for Biomolecular Research, Department of Molecular Biology, Sheffield University, Sheffield S10 2TN, U.K.,

and *Department of Biochemistry and Genetics, The University of Newcastle, Newcastle-upon-Tyne NE2 4HH, U.K.

The cloning and overexpression of the *MspI* DNA methyltransferase as a functional fusion with glutathione *S*-transferase is described. The fusion enzyme retains full biological activity and has been used to investigate the interaction of substrates and inhibitors with *MspI* DNA methyltransferase. The fusion enzyme has been purified to homogeneity in a single step on GSH-agarose and is free from contaminating exonuclease activity. The enzyme can be photolabelled with *S*-adenosyl-L-methionine and the level of incorporation of label is enhanced by the presence of a non-specific DNA duplex. In the presence of a cognate oligodeoxynucleotide, no photolabelling was observed since methyl transfer occurs instead. The inclusion of a mechanism-based inhibitor of C-5 deoxycytidine DNA methylation (an oligodeoxynucleotide containing the base 2-pyrimidinone-1- β -D-2'-deoxyribofuranoside in the position of the deoxycytidine to which methyl addition occurs), which is thought to form a covalent interaction with the reactive cysteine of such enzymes, led to an enhancement of *S*-adenosyl-L-methionine photolabelling which suggests that, in contrast with results obtained with *EcoRII* DNA methyltransferase [Som and Friedman

(1991) *J. Biol. Chem.* **266**, 2937–2945], methylcysteine is not the photolabelled product. The implications of the results obtained with this mechanism-based inhibitor are discussed with respect to other C-5-specific DNA methyltransferases. Gel-retardation assays in the presence of cognate oligodeoxynucleotides that contain the reactive pyrimidinone base in place of the deoxycytidine target base are described. These demonstrate that most probably a stable covalent bond is formed between the methyltransferase and this oligodeoxynucleotide. However, the alternative of extremely tight non-covalent binding cannot be rigorously excluded. Furthermore, the results from these experiments indicate that the reaction mechanism proceeds in a manner similar to that of *HhaI* DNA methyltransferase with sequence-specific DNA binding being followed by addition of *S*-adenosyl-L-methionine and concomitant isomerization of the ternary complex leading to methyl transfer. *S*-Adenosyl-L-homocysteine appears to inhibit the reaction pathway as a result of either competition with the methyl donor and potentiation of a high-affinity interaction between the enzyme and DNA in an abortive ternary complex or through an allosteric interaction.

INTRODUCTION

Type II restriction and modification enzymes are encoded by two separate genes. These genes normally lie adjacent to each other and can be either chromosomal or plasmid-borne [1]. The product of the restriction gene is a sequence-specific DNA endonuclease whereas the modification gene encodes a DNA methyltransferase (Mtase). Each component of a given type-II restriction and modification system recognizes the same DNA target sequence. The action of the DNA Mtase renders the host genome refractory to restriction by the corresponding endonuclease. The genes encoding the type-II *MspI* restriction and modification enzymes have been cloned and their nucleotide sequences determined [2–4]. Whereas there is no detectable sequence similarity between most restriction endonucleases, extensive similarities occur between DNA C-5 Mtases [5,6]. Thus the *MspI* DNA Mtase (*M.MspI*) belongs to a group of enzymes that catalyse the transfer of methyl groups from *S*-adenosyl-L-methionine (AdoMet) to the 5 position of deoxycytidine, and in this particular case the outer deoxycytidine of the DNA sequence 5'-CCGG-3' is modified.

Recently, Dubey and Roberts [7] have reported high-level expression and purification of recombinant *M.MspI*. The enzyme is active as a monomer, although it is predominantly isolated as

an aggregated species. Using a series of ion-exchange chromatography steps, soluble active homogeneous enzyme was obtained as 4% of the total cellular protein. This preparation was subsequently used to explore the sequence-specific interactions between *M.MspI* and DNA using gel retardation and DNAase footprinting techniques. In conclusion, the enzyme has a K_d for specific unmodified DNA of around 10^{-6} M and shows a higher affinity for hemimethylated DNA but a weaker affinity for fully methylated DNA [8]. In the presence of the cofactor or the derivatives *S*-adenosyl-L-homocysteine (AdoHcy) and sinefungin, the affinity for cognate DNA over non-cognate DNA is stimulated by 100-fold. In addition, DNAase footprinting revealed an intrinsic asymmetry in the interaction between *M.MspI* and the sequences flanking its recognition site.

Using steady-state kinetic methods, in conjunction with isotope labelling experiments, Santi and his colleagues have determined the order of substrate addition in the reaction pathway of *HhaI* C-5 Mtase, *M.HhaI* [9]. This enzyme catalyses methyl transfer to DNA via the transient covalent association of the enzyme with the target deoxycytidine, after which the binding of cofactor leads to completion of the methyl-transfer reaction. By analogy with related methyl-transfer enzymes such as thymidilate synthase [10,11] it has been proposed that the active site of C-5 Mtases contains a Pro-Cys dipeptide motif through which

Abbreviations used: Mtase, methyltransferase; *M.MspI*, *MspI* DNA methyltransferase; AdoMet, *S*-adenosyl-L-methionine; AdoHcy, *S*-adenosyl-L-homocysteine; d⁴H₂C, 2-pyrimidinone-1- β -D-deoxyribofuranoside; GST, glutathione *S*-transferase; f.a.b. m.s., fast atom bombardment mass spectrometry; IPTG, isopropyl β -thiogalactoside.

† To whom correspondence should be addressed.

covalent adduct formation takes place. Confirmation of the role of the reactive cysteine in methyl transfer has come from chemical-modification studies, site-directed mutagenesis of multi-specific C-5 Mtases [12] and more recently from the use of a mechanism-based inhibitor, an oligodeoxynucleotide containing 5-fluorodeoxycytidine [13].

We describe here the synthesis of a novel mechanism-based inhibitor of C-5 DNA methylation. This analogue, 2-pyrimidinone-1- β -D-2'-deoxyribofuranoside ($d^{4H}C$) in which the 4-amino group of deoxycytidine is replaced by hydrogen, when incorporated into synthetic oligodeoxynucleotides can be used to trap the covalent enzyme-DNA complex in a sequence-dependent manner and therefore facilitates the analysis of the reaction mechanism of C-5 Mtases. Unlike oligodeoxynucleotides containing 5-fluorodeoxycytidine [13], covalent complex-formation between C-5 Mtases and the $d^{4H}C$ -substituted DNA takes place in the absence of AdoMet. Therefore the mechanism-based inhibitor we describe exerts its effect at a stage distinct from the β -elimination event. In this paper, we also report the cloning, overexpression, single-step purification and characterization of an active fusion of *M.MspI* with glutathione *S*-transferase (GST-*M.MspI*) and the subsequent application of oligodeoxynucleotides containing $d^{4H}C$ to explore the mechanism of action of *M.MspI*. In all of the experiments carried out, the fusion enzyme behaves in a manner identical with the wild-type enzyme [8] and follows a reaction mechanism similar to that described for *M.HhaI*. In view of the high-level expression and ease of purification of the GST-*M.MspI* fusion, it should prove to be an excellent test bed system for studying the structure and function of C-5 DNA methyl transfer using mutagenesis strategies.

MATERIALS AND METHODS

Reagents and analytical methods

[γ - ^{32}P]ATP (3000 Ci/mmol) was purchased from Amersham, AdoHcy, AdoMet, sinefungin, GSH and GSH-agarose were from Sigma. All enzymes were purchased from NBL except for *M.MspI* which was obtained from New England Biolabs. Unless otherwise stated, all methods followed either the manufacturer's instructions or were as described by Sambrook et al. [14]. All chemicals for oligodeoxynucleotide synthesis (including very NH_2 -labile deoxynucleoside phosphoramidites; sold under the trade name FOD [15-17] phosphoramidites) were obtained from Applied Biosystems, Warrington, Lancs., U.K. 2'-Deoxyuridine was purchased from Sigma, Poole, Dorset, U.K. The sources of all the other reagents (and any necessary purification or drying procedures) used in the chemical and oligodeoxynucleotide syntheses have been given. 1H n.m.r. spectroscopy and fast atom bombardment mass spectrometry (f.a.b. m.s.) were performed as described [18]. The h.p.l.c. systems used have also been previously described. Analytical size (20 cm $\times$ 0.45 cm) C_{18} reversed-phase columns (Apex-1 octadecylsilyl, 5 μm particle size) supplied by Jones Chromatography (Llanbradach; Glamorgan, Wales, U.K.) were used. Columns were run at 1 ml/min and at 50 °C for oligodeoxynucleotide purification and characterization or at room temperature for deoxynucleoside-composition analysis. H.p.l.c. buffer A was 0.1 M triethylamine acetate, pH 6.5, containing 5% acetonitrile and B was 0.1 M triethylammonium acetate, pH 6.5, containing 65% acetonitrile. Deoxynucleoside-composition analysis using snake venom phosphodiesterase and alkaline phosphatase was performed as previously reported [18,19]. The four normal bases were detected at 254 nm (for absorption coefficients see references [18,19]). The modified base

$d^{4H}C$ was detected at 303 nm using an absorption coefficient of $6.47 \times 10^3 M^{-1} \cdot cm^{-1}$ [20].

Cloning and expression of the *M.MspI* gene

The isolation of clones expressing the *MspI* restriction and modification genes was accomplished using the 'Hungarian trick' (see [1] for references) in an identical manner with that reported by Walder et al. [3] and Nwankwo and Wilson [4]. Constructs harbouring the restriction and modification genes were used as templates for the PCR to generate the modification gene alone. Plasmid pNAF3 contains the *MspI* restriction and modification genes encoded on a 2.9 kb *EcoRI*-*HindIII* restriction fragment inserted into pBS plasmid (Stratagene) cleaved with the same restriction enzymes.

The oligodeoxynucleotides used for PCR amplification of the modification gene were 5'-TGACGGCGCATGCAACCTGA-3' and 5'-GTAAACGACGGCCAGT-3'. After amplification, the *M.MspI* gene was ligated first into the *SmaI* site of the multipurpose vector pBS (Stratagene) to produce the recombinant plasmid pNAF8. This step facilitated excision of the *M.MspI* gene as an *EcoRI* restriction fragment which was subsequently ligated into the GST expression vector pGEX-3X (Pharmacia) to give pNRT1. This construct was subsequently digested with *SmaI* and *SphI*, the resultant 3' recess was subsequently filled in using DNA polymerase and the recombinant religated to produce pNRT2. After transformation of *Escherichia coli* RR1 with pNRT2 and induction of liquid cultures in the presence of isopropyl β -thiogalactoside (IPTG), an in-frame fusion of the GST and the *M.MspI* polypeptides was expressed at high level.

High-level expression and purification of *M.MspI*

Since we have found many GST fusion proteins to be largely insoluble when expressed at high level in *E. coli* (C. Taylor, K. Ford, B. Connolly and D. Hornby, unpublished work), and expression of *M.MspI* in *E. coli* led to the accumulation of largely insoluble material [7], we examined the solubility of the GST-*M.MspI* fusion in cell-free extracts prepared from 5 ml cultures grown under a variety of conditions. The most suitable regime for extraction of soluble material involved growth at 30 °C in Luria broth. Extracts were prepared by sonication of cells which had been resuspended in 0.1 vol. PBS (pH 7.3). Large-scale (500 ml) cultures were grown from an overnight 5 ml inoculum to an A_{595} of 0.4, whereupon IPTG was added to give a final concentration of 0.05 mM and the cells were allowed to grow for a further 2 h at 30 °C. Harvested cells were resuspended in PBS (pH 7.3) and sonicated on ice. After brief centrifugation [40000 *g* (18000 rev./min) in a Sorvall SS34 rotor] to pellet insoluble material, the supernatant was applied directly to a GSH-agarose column pre-equilibrated in a buffer of 50 mM Tris/HCl, pH 8, and 1 mM EDTA. After 4 column vol. washes with the loading buffer, the fusion protein was eluted with 3 ml of 10 mM GSH in 50 mM Tris/HCl, pH 8.0, 1 mM EDTA. The purity of the GST-*M.MspI* fractions (0.5 ml) was analysed by SDS/PAGE. Despite several attempts, it proved impossible to cleave *M.MspI* from GST using Factor Xa.

Restriction protection assays

In order to determine the activity of the purified fusion enzyme, a restriction protection assay was performed. Samples (12 μl ; 0.6 μg of protein) from the affinity-column fractions were incubated with 3 μg of λ DNA and 80 μM AdoMet in Mtase buffer (50 mM NaCl, 50 mM Tris/HCl, pH 7.5, 10 mM EDTA

Table 1 Sequences of the oligodeoxynucleotides used in this study

Oligonucleotide	Sequence
ODN-SP1	5'-AACCCGGGCCGCCGGGTT-3'
ODN-NSP(16)	5'-AGCATTCTAGAATGCT-3'
ODN-SP2	5'-AGCATTCCGGAATGCT-3'
ODN-NSP(33)	5'-ATGGGGCGGAGAATGGGCGGAAGTGGGCGGAGT-3'
ODN-MSP2A	5'-AGCATTd ^{4H} CCGGAATGCT-3'
ODN-MSP2B	5'-AGCATTd ^{4H} CGGAATGCT-3'

and 5 mM 2-mercaptoethanol) at 37 °C for 1 h. The reaction was stopped by heating the samples to 65 °C for 10 min. Each reaction mixture was then subjected to digestion with 10 units of *MspI* restriction enzyme (NBL) in a final volume of 30 µl for 1 h at 37 °C under the conditions described by the manufacturer. Digestion products were analysed on a 1 % agarose gel.

Photocross-linking assay

Purified GST-*MspI* (0.6 µg) was incubated with 10 µM [*methyl*-³H]AdoMet (76 Ci/mmol; Amersham) in Mtase buffer in a total volume of 20 µl, in the wells of a microtitre dish. Incubations in the presence of DNA involved addition of 0.1 µg of either specific (ODN-SP1 or ODN-SP2), modified specific (ODN-MSP2A, ODN-MSP2B) or non-sequence-specific duplex oligodeoxynucleotides [ODN-NSP(16)], the sequences of which are shown in Table 1. After exposure to long-wavelength (312 nm) u.v. irradiation for 1 h at 4 °C at a distance of 5 cm from the reaction mixture, the samples were boiled in the presence of SDS and electrophoresed on a 10 % SDS/polyacrylamide gel. After electrophoresis, the section of the gel containing molecular-mass markers and purified GST-*MspI* was stained with Coomassie Blue R. The remainder of the gel containing the photocross-linked samples was fixed for 30 min in propan-2-ol/water/acetic acid (5:13:2, by vol.) and then agitated for 45 min in the fluoror for Amplify (Amersham). This was then dried under vacuum and exposed to X-ray film at -70 °C for 8-10 days.

Gel retardation

ODN-SP1, ODN-SP2, ODN-NSP(16) and the 33 mer oligodeoxynucleotides were synthesized on an Applied Biosystems 381 A DNA synthesizer. The sequences of these oligodeoxynucleotides are given in Table 1. All conventional oligodeoxynucleotides were used without further purification apart from those containing d^{4H}C which were prepared and purified as described below. The oligodeoxynucleotides ODN-MSP2A and ODN-MSP2B contain a d^{4H}C base at the position shown in Table 1. All oligodeoxynucleotides used in this study were self-complementary. Annealing was promoted by incubation of the oligodeoxynucleotides at 90 °C for 5 min with subsequent cooling to room temperature at a rate of 1 °C/min. The oligodeoxynucleotides were labelled at the 5' end with [γ -³²P]ATP using T4 polynucleotide kinase (NBL). Unincorporated label was removed by gel filtration using Sephadex G-25 resin (Pharmacia) and the spun column procedure [14]. Typically, 0.1 µg of fusion protein (final concentration 0.038 µM, assuming the fusion protein is dimeric) was incubated in Mtase buffer with

1 ng of labelled oligodeoxynucleotide (final concentration 0.004 µM) and 160 µM AdoMet in a final reaction volume of 20 µl. Incubation was for 35 min on ice, unless stated otherwise. The reaction mixture was subsequently loaded on to a 4 % non-denaturing polyacrylamide (19:1) gel in 0.5×TBE buffer (50 mM Tris/0.4 M boric acid/1.25 mM EDTA) and electrophoresed for 2-3 h at 100 V and 15 °C. Protein-DNA complexes were visualized after autoradiography with intensifying screens at -70 °C, overnight.

Preparation of 5'-dimethoxytrityl-4-thiadeoxyuridine

Dry deoxyuridine (2.28 g; 10 mmol) was suspended in 50 ml of dry pyridine, and then benzoyl chloride (2.55 ml, 22 mmol) was added. After stirring overnight, t.l.c. [chloroform/methanol (19:1, v/v)] showed complete conversion of deoxyuridine (R_f 0.05) into 3',5'-dibenzoyldeoxyuridine (R_f 0.6). The solution was poured into 2.5 litres of ice-cold water with vigorous stirring. The precipitated product was collected by filtration, washed with ice-cold water and dried. A near-quantitative yield was obtained. ¹H n.m.r. ([²H]chloroform) 8.75 (s, 1H, NH), 8.1-7.4 (m, 11H, benzoyl protons and H6), 5.58 (d, 1H, H5) plus standard sugar spectrum [6.45 (d, 1H, H1'), 5.63 (m, 1H, H3'), 4.85-4.54 (m, 3H, H5', H5'', H4'), 2.81 (m, 1H, H2'), 2.35 (m, 1H, H2'')]. This pattern was seen (with minor variations) in all the derivatives prepared and is from now on simply referred to as standard sugar spectrum. F.a.b. m.s. 437 ($M+H^+$, 80%), 325 (dibenzoyldeoxyribose+, 100%), 113 (heterocycle+2H⁺, 95%). The dried 3',5'-dibenzoyldeoxyuridine (4.14 g, 9.5 mmol) produced above was dissolved in 50 ml of dry dioxan, and 4.85 g of Lawesson's reagent [2,4-bis-(4-methoxyphenyl)-1,3-dithia-2,4-diphosphetane-2,4-disulphide] was added. After refluxing for 2 h, t.l.c. [chloroform/methanol (97:3, v/v)] showed complete conversion of dibenzoyldeoxyuridine (R_f 0.4) into dibenzoyl-4-thiadeoxyuridine (R_f 0.7). The mixture was cooled and, after addition of 10 ml of 5 % NaHCO₃, evaporated to dryness. After work up (product dissolved in 50 ml of chloroform and extracted with 2 × 50 ml of 5 % NaHCO₃ plus 50 ml of saturated NaCl; the chloroform solution dried with Na₂SO₄ and evaporated) the product could be used without further purification. ¹H n.m.r. ([²H]chloroform) 10.35 (s, 1H, NH), 8.13-7.45 (m, 10H, benzoyl protons), 7.35 (d, 1H, H6), 6.27 (d, 1H, H5) plus standard sugar spectrum. F.a.b. m.s. 453 ($M+H^+$ 20%), 325 (dibenzoyldeoxyribose+, 25%), 129 (heterocycle+2H⁺, 100%). The dibenzoyl-4-thiadeoxyuridine was debenzoylated by dissolving in 50 ml of dry methanol and adding sodium methoxide (25 mmol). After 30 min, t.l.c. [chloroform/methanol (97:3, v/v)] showed complete debenzoylation (R_f of product 0.05). Acetic acid (1.43 ml, 25 mmol) was added and the mixture evaporated to dryness. The product was dissolved in water (50 ml) and extracted with 3 × 50 ml of chloroform. The aqueous layer was retained and evaporated to dryness. The 4-thiadeoxyuridine produced was extracted from inorganic salts by the addition of 100 ml of acetone and stirring. The product was dissolved in acetone, and, after filtration, to remove salts was obtained by evaporation. The 4-thiadeoxyuridine produced was not characterized but immediately converted into its 5'-dimethoxytrityl derivative in the usual manner [21]. This compound was purified by flash chromatography using chloroform/methanol (97:3, v/v) (containing 0.1 % pyridine) as eluent. The 5'-dimethoxytrityl-4-thiadeoxyuridine was pure by t.l.c. [chloroform/methanol (19:1, v/v); R_f 0.5] and usually obtained in yields of 50 % commencing from deoxyuridine (four steps). ¹H n.m.r. ([²H]chloroform) 10.12 (s, 1H, NH); 7.66 (d, 1H, H6), 7.44-6.83 (m, 13H, aromatic dimethoxytrityl), 6.12 (d, 1H, H5),

3.84 (s, 6H, CH₃O-) plus standard sugar spectrum. F.a.b. m.s. 547 (*M*+H⁺, 25%), 303 (dimethoxytrityl+, 100%).

Preparation of 5'-dimethoxytrityl-2-pyrimidinone-1-β-D-2'-deoxyribofuranoside

5'-Dimethoxytrityl-4-thiideoxyuridine (2.6 g; 5 mmol) was dissolved in 100 ml of ethanol containing 5 ml of water. Hydrazine hydrate (**caution: extremely toxic**) (1.4 ml, 25 mmol) was added. After 1 h at reflux, t.l.c. [chloroform/methanol (19:1, v/v)] showed complete conversion of the 4-thiideoxyuridine derivative into the 4-hydrazino compound (*R_F* 0.05–0.15 streaks). After evaporation and work up the product was pure enough for the next stage. ¹H n.m.r. ([²H]chloroform) 7.5–6.3 (m, 15H, aromatic dimethoxytrityl, H6, H5) 3.73 (s, 6H, CH₃O-) plus standard sugar spectrum. F.a.b. m.s. 545 (*M*+H⁺, 30%), 303 (dimethoxytrityl+, 100%). This hydrazino derivative was dissolved in 50 ml of ethanol containing 0.5 ml of triethylamine, and Ag₂O (2.3 g, 10 mmol) was added. After 2 h at reflux, t.l.c. [chloroform/methanol (9:1, v/v)] showed disappearance of the starting material (*R_F* 0.2–0.3) and the appearance of the pyrimidinone product (*R_F* 0.6) (several side products were also visible). The mixture was evaporated to dryness, suspended in 200 ml of chloroform/10% aq. NaI (1:1, v/v) and filtered through celite. The chloroform layer was retained and extracted sequentially with 10% NaI, 2 × 5% Na₂S₂O₃, water and then evaporated to dryness. The product was purified by flash chromatography using 500 ml each of ethyl acetate (containing 0.5% triethylamine) and sequentially 1, 3 and 5% methanol. The product was eluted with the last solvent and was obtained in yields of 50% from 5'-dimethoxytrityl-4-thiideoxyuridine. It was pure as assessed by t.l.c. [chloroform/methanol (9:1, v/v); *R_F* 0.6]. ¹H n.m.r. ([²H]chloroform) 8.57 and 8.43 (both m and 1H, H4 and H6 unassigned), 7.44–6.83 (m, 13H, aromatic dimethoxytrityl), 6.05 (m, 1H, H5), 3.69 (s, 6H, CH₃O-) plus standard sugar spectrum. F.a.b. m.s. 515 (*M*+H⁺, 15%), 303 (dimethoxytrityl+, 100%). This product could be phosphorylated in the usual way to give the 3'-cyanoethyl-phosphoramidite [22].

Oligodeoxynucleotide synthesis

Oligodeoxynucleotides were prepared on a 1 μmol scale using NH₃-labile (FOD) deoxynucleoside phosphoramidites [15–17]. The phosphoramidite of 5'-dimethoxytrityl-2-pyrimidinone-1-β-D-2'-deoxyribofuranoside was dissolved in anhydrous (DNA synthesis grade) acetonitrile at the usual concentration of 0.1 M. This solution was filtered through Millex poly(tetrafluoroethylene) 0.5 μm filters (Millipore) to remove particulate matter before use. The standard 1 μmol synthesis programme was used and all preparations were performed trityl-on. After synthesis, deblocking was carried out using 35% aq. NH₃ at room temperature overnight.

Oligodeoxynucleotide purification

Oligodeoxynucleotides were initially purified trityl-on by reversed-phase h.p.l.c. The gradient used was: *t* = 0 min, 75% h.p.l.c. buffer A, 25% h.p.l.c. buffer B; *t* = 20 min, 35% h.p.l.c. buffer A, 65% h.p.l.c. buffer B. After detritylation of the product oligodeoxynucleotides with aqueous acetic acid, a further round of h.p.l.c. purification was undertaken. For d(GACGATAT-

[d⁴H]GTC) a gradient consisting of *t* = 0 min 100% A, 0% B, *t* = 25 min, 85% A, 15% B was used. With d(AGCATT-[d⁴H]CGGAATGCT) and d(AGCATT[d⁴H]GGAATGCT) the gradient was *t* = 0 min, 100% A, 0% B, *t* = 25 min 75% A, 25% B. The purified products were concentrated to a small volume and desalted by passage over Pharmacia pre-packed NAP-25 gel-filtration columns (DNA grade) eluted with 25 mM triethylammonium bicarbonate. This salt was removed by several coevaporations from methanol (Analar) and the purified products were dissolved in 200 μl volumes of distilled water and stored frozen at -20 °C.

Investigation of the NH₃ lability of oligodeoxynucleotides containing d⁴H

About 10 A₂₆₀ units of d(GACGATAT[d⁴H]GTC) was heated at 50 °C in 35% aq. NH₃. The reaction was examined by h.p.l.c. (using the buffer system described above for the final purification of this oligodeoxynucleotide). When the reaction was almost complete (after 22 h), the major products were purified and analysed by base composition analysis.

RESULTS

Synthesis and properties of pyrimidinone-containing DNA

The preparation of a 2-pyrimidinone-1-β-D-2'-deoxyribofuranoside derivative suitable for incorporation into deoxynucleotides using the phosphoramidite approach is given in Figure 1. This synthesis is based on our previously reported preparation of 5-methyl-2-pyrimidinone derivatives [18] but commences from deoxyuridine rather than thymidine. The 4-keto oxygen atom is converted into hydrogen via the 4-thio and the 4-hydrazino derivatives. These procedures were originally developed by Holy and co-workers for the preparation of free pyrimidinone nucleosides [23,24]. A similar synthesis to that given here has been reported with the 4-hydrazino derivative being prepared by a bisulphate-catalysed transamination of a deoxycytidine derivative. Previous incorporations of 2-pyrimidinones into oligodeoxynucleotides have used the phosphotriester approach [25,26] and more recently the phosphoramidite method [18,20].

The 5'-dimethoxytrityl-3'-phosphoramidite of 2-pyrimidinone (VIII, Figure 1) coupled in 100% yields as assessed by trityl cation release during oligodeoxynucleotide synthesis. However, after deblocking by mild ammonia treatment (room temperature, overnight), trityl-specific reversed-phase h.p.l.c. purification and detritylation, analysis by h.p.l.c. showed that the oligodeoxynucleotides were still not pure. This is illustrated in Figure 2 for three d⁴H-C-containing oligomers. In each case the desired product is the largest late eluted peak but considerable amounts of earlier eluted side products are visible. In contrast, oligodeoxynucleotides that only contain the four standard bases appear pure after trityl-specific purification alone. All d⁴H-C-containing products were further purified by reversed-phase h.p.l.c. using shallow acetonitrile gradients. As shown in Figure 2, this resulted in a large increase in the degree of purity, although in one case (Figure 2e and 2f) contaminants were visible after the second purification step. The reason for the lack of purity of d⁴H-C-containing oligodeoxynucleotides is undoubtedly the enhanced reactivity of 2-pyrimidinones as compared with the four usual bases [18,20,25–27]. Pyrimidinones have an extremely acid-sensitive glycosidic bond and H⁺ treatment leads to loss of the heterocycle and in oligodeoxynucleotides to abasic sites [27]. This reaction would occur to some degree during the acid-catalysed detritylation step in oligodeoxynucleotide synthesis

Figure 1 Preparation of a 2-pyrimidinone-1-β-D-2'-deoxyribofuranoside derivative suitable for incorporation into oligodeoxynucleotides by the phosphoramidite approach

Reagents: (1) benzoyl chloride; (2) Lawesson's reagent; (3) sodium methoxide; (4) dimethoxytritylchloride; (5) hydrazine hydrate; (6) Ag₂O; (7) 2-cyanoethyl *N,N*-diisopropylchlorophosphoramidite. Dmt, dimethoxytrityl.

resulting in a drop in purity of the product. Perhaps more importantly 2-pyrimidinones are easily attacked by NH₃ leading initially to a saturated derivative. In oligodeoxynucleotides containing this base, further reactions lead to chain cleavage (see below for details). Indeed the main side products visible in Figure 2 (in all cases the second largest peak that is eluted just before the desired product) appear to be NH₃ adducts at the pyrimidinone base (see below). Using FOD-protected bases and overnight treatment with NH₃ at room temperature [15–17], significant decomposition at the pyrimidinone base is seen (Figure 2). Nevertheless it still proved possible to purify the desired products to reasonable homogeneity. It should be noted that extremely low levels of product were observed when the standard NH₃ deblocking protocol (overnight at 50 °C) was used and so the use of FOD with bases is mandatory. We have previously observed similar effects with 5-methyl-2-pyrimidinones [18] but general lability and NH₃ reactivity is much more pronounced with 2-pyrimidinones, presumably due to less steric hindrance. It may be possible to improve the synthesis of oligodeoxynucleotides containing 2-pyrimidinones using deoxynucleosides protected with functions removable via a β-elimination with non-nucleophilic bases [28–30] or by alteration of the synthesis protocol to remove the need for base protection and consequent NH₃ treatment [31]. Unfortunately the appropriate deoxynucleoside phosphoramidites required are not yet commercially available.

The identity of all the oligodeoxynucleotides prepared was verified by base-composition analysis using snake venom phosphodiesterase and alkaline phosphatase. This is especially important in view of the reactivity of d⁴H³C and the lack of homogeneity of trityl-purified oligomers. As shown in Figure 3 for d(GACGATAT[d⁴H³C]GTC), a correct base-composition analysis was obtained, with one equivalent of the modified base

d⁴H³C being visible. Similar results (not shown) were obtained with the other two oligodeoxynucleotides.

We have investigated the NH₃ lability of 2-pyrimidinone-containing oligodeoxynucleotides as this is relevant to their reaction with C-5 Mtases. Figure 4 shows that treatment of d(GACGATAT[d⁴H³C]GTC) with NH₃ rapidly gives rise to a new product (50 % reaction in about 2 h). This unstable initial product probably arises from NH₃ addition across the 5,6 double bond of the pyrimidinone giving an NH₂ adduct at C-6 with concomitant double-bond saturation. NH₃ does not attack the 5,6 double bond of the usual pyrimidines thymidine, uridine and cytidine. However, stronger nucleophiles such as hydrazine and hydroxylamine can attack this bond under vigorous conditions and the initial product is usually a C-6 adduct with a saturated double bond [32]. The 5,6 double bond in 2-pyrimidinones is clearly more reactive than that of the normal pyrimidines. This is to be expected as substituents that donate electrons into the pyrimidine ring system (e.g. the 4-amino group of cytidine and the 4-oxygen atoms of thymidine and uridine) tend to increase the aromatic character of the heterocycle [33]. Removal of these substituents as occurs with 2-pyrimidinones would therefore render the 5,6 double bond less aromatic and more like an isolated double bond. This would be expected to increase its reactivity particularly in addition reactions. The initial adduct formed between NH₃ and oligodeoxynucleotides containing d⁴H³C undergoes further reactions in the presence of NH₃ giving chain cleavage (Figure 4). This reaction has parallels with Maxam–Gilbert DNA sequencing where initial addition of hydrazine across the 5,6 double bond of the pyrimidines is followed by subsequent reactions leading to strand breakage [34]. We have previously observed a similar set of reactions with oligodeoxynucleotides containing 5-methyl-2-pyrimidinones [18]. However, as mentioned above, 2-pyrimidinone-substituted

Figure 2 Reversed-phase h.p.l.c. traces of d(GACGATAT[d⁴H]GTC) after trityl-on h.p.l.c. purification (a) and after a further round of h.p.l.c. purification using a shallow acetonitrile gradient (b)

(c) and (d) show analogous traces for d(AGCATT[d⁴H]CGGAATGCT) and (e) and (f) give these traces for d(AGCATT[d⁴H]CGGAATGCT).

Figure 3 Deoxynucleoside-composition analysis obtained for d(GACGATAT-[d⁴H]GTC)

Note that [d⁴H] was detected at 303 nm (not 254 nm used with the other bases) and the detector sensitivity was increased by a factor of 2 for this base. The mole fraction of the bases observed were: dC (2.15); d⁴H (1.2); dG (3); T (3.07) and dA (2.95).

oligomers are far more susceptible to attack by NH₃. It should also be noted that the initial product formed by the action of NH₃ on d⁴H-containing oligodeoxynucleotides (before chain cleavage) is in every case identical with the major impurity seen by h.p.l.c. after trityl-specific purification alone. This confirms that NH₃ susceptibility is at least partially responsible for the difficulties seen in the preparation of these oligodeoxynucleotides.

Solubility of the GST-*MspI* fusion protein

It can be clearly seen from Figure 5 that induction of *E. coli* RR1 harbouring the recombinant plasmid pNRT2 with IPTG leads to high-level expression of a protein of molecular mass 70 kDa. The molecular mass of the GST-*M. MspI* fusion protein estimated from the amino acid sequence is 71.6 kDa which is in good agreement with the estimated molecular mass of the induced protein species. Initial inductions of 5 ml cultures expressing pNRT2 at 37 °C revealed that most of the fusion protein partitioned to the insoluble fraction after extraction and centrifugation (see lanes 5 and 6 in Figure 5). However, when the growth temperature was reduced to 30 °C, the fusion enzyme was almost quantitatively obtained in the soluble phase (lane 7, Figure 5). Addition of Triton X-100 to the cells grown at 37 °C after sonication did not significantly improve solubility of the fusion enzyme (see lanes 9 and 10 in Figure 5).

Figure 6 shows an SDS/polyacrylamide gel of fractions before and after affinity chromatography on GSH-agarose. This single chromatography step, which takes around 1 h to perform, is therefore sufficient to generate essentially homogeneous enzyme in almost 100 % yield. This compares favourably with the method recently described by Dubey et al. [7] which involves three ion-exchange steps and takes 12–15 h to produce active homogeneous protein as 4 % of total cellular protein. The availability of both fusion and native enzyme in a pure form should now lead to a rapid accumulation of data on the physical properties of this enzyme and selected mutants.

Activity of the GST-*MspI* fusion enzyme

In order to establish whether the overexpressed fusion protein retained biological activity it was used in a restriction protection assay (see Figure 7). This assay is based on the biological function of the methylation reaction and involves preincubation of DNA with the fusion enzyme in the presence of AdoMet before the addition of *MspI* restriction endonuclease: the GST-*M. MspI* fusion protein is clearly fully functional *in vitro*. The enzyme is also active *in vivo* since when pNRT2 plasmid DNA is isolated from the host strain, it is refractory to *MspI* restriction. The lack of non-specific nuclease activity often associated with preparations of DNA-binding proteins in general is a fortuitous consequence of the purification strategy which circumvents the use of DNA-like affinity resins such as heparin-agarose and phosphocellulose.

The interaction of AdoMet with the GST-*MspI* fusion enzyme

The results of photocross-linking tritiated AdoMet to the fusion protein are shown in Figure 8(a) which reveals that the fusion enzyme can bind AdoMet in the absence of DNA. Similar results have been obtained with other AdoMet-binding enzymes, most notably *M. EcoRII* [35]. This finding is perhaps somewhat unexpected in the light of the mechanism of *M. HhaI* where it has been suggested that AdoMet only interacts significantly with the enzyme-DNA complex, does not bind free enzyme at all and binds poorly to the enzyme-methylated DNA complex [9]. Figure

Figure 4 Effect of treatment of d(GACGATAT[d⁴H]GTC) with NH₃

(a) $t = 0$ h; (b) $t = 2$ h; (c) $t = 22$ h. Peak 1 is d(GACGATAT[d⁴H]GTC). As mentioned in the text, peak 2 probably contains a 5,6 saturated pyrimidinone with NH₂ adducted at C-6. However, this product was not rigorously characterized. Peak 3 is GACGATAX (where X is a T derivative) and peak 4 is pGTC. Both peaks 3 and 4 were characterized by base-composition analysis and the expected results were obtained.

Figure 5 SDS/PAGE analysis of induced pNRT2 and pGEX transformants

Lanes 1 and 2, soluble and insoluble phases of crude extracts of *E. coli* RR1 (pGEX-3X) respectively after induction at 37 °C. Lanes 3 and 4, uninduced extracts of *E. coli* RR1 (pNRT2), soluble and insoluble respectively grown at 37 °C. Lanes 5 and 6, as 3 and 4 but after induction with IPTG. Lanes 7 and 8, as 5 and 6 but cells were grown at 30 °C. Lanes 9 and 10, as 7 and 8 but cells were lysed in the presence of 0.1% Triton X-100. Lane M, molecular-mass markers, and S and I beneath the lanes indicate soluble or insoluble fraction.

8(b) shows the results of a cross-linking experiment performed in the presence of various oligodeoxynucleotide duplexes. A brief exposure of the gel reveals no comparable signal for the enzyme in the presence of the cognate (ODN-SP1 and ODN-SP2) oligodeoxynucleotides, although a weak signal may be masked (see lanes 1 and 4, Figure 8b). However, cross-linking of AdoMet to the enzyme was observed in the presence of the oligodeoxynucleotides ODN-MSP2A, ODN-MSP2B and ODN-NSP(16) (lanes 2, 3 and 5, Figure 8b). Whereas ODN-MSP2A and ODN-MSP2B contain a modified *MspI* site, ODN-NSP(16) is a non-cognate oligodeoxynucleotide. Interestingly, despite the absence of an AdoMet-protein complex in the presence of the cognate oligodeoxynucleotides (ODN-SP1 and ODN-SP2), these oligodeoxynucleotides are methylated by the enzyme and therefore appear as tritiated bands at the foot of the gel (lanes 1 and 4, Figure 8b). The differences in intensity of these two cognate oligodeoxynucleotides reflect the number of *MspI* sites (three for ODN-SP1 and one for ODN-SP2) and their lengths (ODN-SP1

Figure 6 SDS/PAGE analysis of the *M.MspI* fusion protein before (lane 1) and after (lane 2) purification on GSH-agarose

Molecular-mass standards were run in lane M.

is a 20mer whilst ODN-SP2 is a 16 mer). No methylation of the oligodeoxynucleotides ODN-MSP2A, ODN-MSP2B or ODN-NSP(16) was observed. This demonstrates that removal of the 4-NH₂ group from the target deoxycytidine renders it unable to accept a methyl group from AdoMet. Interestingly, removal of the 4-NH₂ group from the internal deoxycytidine also prevents methyl addition to the target base. As expected, the non-cognate oligodeoxynucleotide ODN-NSP(16) could not be methylated.

Gel-retardation assays

GST-*M.MspI* fusion protein-DNA interactions were further studied using a gel mobility-shift assay. Initial experiments were carried out to compare the relative affinities of wild-type *M.MspI* and the fusion enzyme under the conditions shown in Figure 9. Binding of the fusion and native Mtases to the cognate (ODN-SP1) oligodeoxynucleotide was not evident in the absence of AdoMet (lanes 1 and 5, Figure 9). Dubey and Roberts [8] were also unable to demonstrate complex-formation at these

Figure 7 Protection of bacteriophage λ DNA against *MspI* restriction by the *M.MspI* fusion protein

Each lane contains 3 μ g of λ DNA. Lane 1, unmethylated λ DNA incubated with 10 units of *MspI* restriction enzyme for 1 h at 37 °C. Lanes 2–8, λ DNA incubated with 12 μ l samples of GST–*MspI* Mtase (successively eluted from a GSH–agarose affinity column) for 1 h at 37 °C and subsequently incubated with 10 units of *MspI* restriction enzyme for 1 h at 37 °C. Lane 9, uncut λ DNA.

Figure 8 SDS/PAGE analysis of the products of u.v. cross-linking between GST–*M.MspI* and AdoMet and the effect that DNA has on the cross-linking

(a) GST–*M.MspI* and tritiated AdoMet. (b) All incubations were as in (a) with the addition of ODN-SP1 (lane 1), ODN-MSP2A (lane 2), ODN-MSP2B (lane 3), ODN-SP2 (lane 4) or ODN-NSP(33) (lane 5). Lane 6, 32 P-labelled ODN-SP2 alone with no added protein.

concentrations of protein and DNA. With the addition of AdoMet, a substantial retardation effect was seen with the *M.MspI* fusion (lane 2, Figure 9). The interaction between the ODN-MSP2A oligodeoxynucleotide (–d⁴HCCGG–) was signif-

Figure 9 Comparison of the DNA-binding properties of GST–*M.MspI* and native *M.MspI*

Each reaction contained, where appropriate, 1 ng of oligodeoxynucleotide, 100 ng of GST–*M.MspI* or 10 units of *M.MspI* in reaction buffer (see the Materials and methods section). All cofactors were added to a final concentration of 160 μ M. Lane 1, GST–*M.MspI* + ODN-SP1; lane 2, GST–*M.MspI* + ODN-SP1 + AdoMet; lane 3, GST–*M.MspI* + ODN-MSP2A; lane 4, GST–*M.MspI* + ODN-MSP2A + AdoMet; lane 5, *M.MspI* + ODN-SP1; lane 6, *M.MspI* + ODN-SP1 + sinefungin; lane 7, *M.MspI* + ODN-SP1 + AdoMet; lane 8, *M.MspI* + ODN-SP1 + AdoHcy; lane 9, *M.MspI* + ODN-MSP2A; lane 10, *M.MspI* + ODN-MSP2A + sinefungin; lane 11, *M.MspI* + ODN-MSP2A + AdoMet; lane 12, *M.MspI* + ODN-MSP2A + AdoHcy.

antly greater in the presence of AdoMet than that seen with unmodified ODN-SP1 DNA (lane 4, Figure 9). Most striking was the appearance of retarded bands for both the fusion enzyme and *M.MspI* with ODN-MSP2A in the absence of AdoMet, under conditions in which we and others do not detect complex-formation with normal B-DNA (lanes 3 and 9, Figure 9). Differences in mobility seen throughout this experiment reflect differences both in protein size (GST–*MspI* fusion versus native) and the relative lengths of the oligodeoxynucleotides used.

A role for the thiol side chain of the Pro-Cys dipeptide in forming a sulphur bridge to the C-6 position of the target deoxycytidine has been postulated [36]. The implication of the above result is that a cognate DNA duplex sequence in the absence of cofactor is sufficient for orientation of this cysteine in the correct position for interaction to occur with the target deoxycytidine. Indeed, although unmodified DNA containing an *MspI* site has a much weaker affinity for the enzyme than ODN-MSP2A, increasing the protein and DNA concentrations by 10–100-fold does lead to a mixture of specific and non-specific protein–DNA interactions in the absence of AdoMet (see ref. [8]). Furthermore, we have observed that removal of the 4-NH₂ group of the internal deoxycytidine (ODN-MSP2B) abolishes DNA binding completely by the GST fusion enzyme which suggests that the 4-amino position of this base is critical for sequence-specific binding to the enzyme. The ordered binding of DNA in advance of AdoMet is consistent with the mechanism described for *M.HhaI* [9].

Addition of AdoMet and the C-5 Mtase inhibitor sinefungin potentiate specific protein–DNA interactions by GST–*M.MspI* and the native enzyme [8]. Additionally, AdoHcy also promotes an increase in binding to the ODN-MSP2A oligodeoxynucleotide for both fusion and native Mtases (see Figure 9). Relative band intensities suggest that sinefungin and AdoHcy promote higher-affinity DNA binding relative to that of AdoMet under the conditions used in these experiments. This may be a reflection of the action of both sinefungin and AdoHcy as inhibitors, whereas AdoMet is only able to act as a substrate under these conditions.

Figure 10 Gel-retardation analysis of AdoMet (160 μ M) effects on the kinetics of GST-M.*MspI*-DNA interactions

Lane 1, ODN-MSP2A alone; lane 2, ODN-MSP2A + GST-M.*MspI*; lane 3, ODN-MSP2A + GST-M.*MspI* + AdoMet, after 10 min incubation in the presence of AdoMet; lane 4, ODN-SP1; lane 5, ODN-SP1 + GST-M.*MspI*; lanes 6–10, ODN-SP1 + GST-M.*MspI* after incubation in the presence of AdoMet for 10, 20, 30, 40 and 50 min respectively. Samples were loaded at time zero.

Figure 11 Time-dependent binding of GST-M.*MspI* to ODN-SP1 in the presence of sinefungin (160 μ M)

Lanes 1–5, ODN-SP1 + GST-M.*MspI* after incubation in the presence of sinefungin for 15, 30, 45, 60 and 75 min respectively. Samples were loaded at time zero.

In the presence of AdoMet, a certain proportion of the normal DNA will have been methylated, and complex dissociation may well have occurred, resulting in a reduction in the quantity of retarded oligodeoxynucleotide.

Figure 12 Oligodeoxynucleotide competition experiments in the presence of AdoMet

Lanes 1 and 3, ODN-MSP2A and ODN-SP1 in the presence of GST-M.*MspI* respectively; lane 2, GST-M.*MspI* + ODN-MSP2A with a 1000-fold excess of unlabelled competitor ODN-SP1; lane 4, GST-M.*MspI* + ODN-SP1 with a 10-fold excess of unlabelled ODN-SP1. All incubations were for 30 min at 4 °C.

In order to determine the nature of the interaction of the Mtase with AdoMet, a series of timed incubations was set up in which the interaction of the fusion enzyme with AdoMet and DNA was monitored. The results shown in Figures 10 and 11 add a kinetic dimension to the interaction of GST-M.*MspI* with ODN-SP1 and ODN-MSP2A oligodeoxynucleotides in the presence of AdoMet and the interaction with ODN-SP1 in the presence of sinefungin. Clearly, the addition of AdoMet to the enzyme-ODN-SP1 complex leads to a low rate of complex-formation (30–35 min; see lanes 8 and 9 in Figure 10), whereas addition of AdoMet to the enzyme-ODN-MSP2A duplex leads to a rapid formation of the high-affinity complex (less than 10 min; see lane 3 in Figure 10), reflecting the reactivity of this modified oligodeoxynucleotide. Complex-formation is also attained more rapidly with ODN-SP1 after the addition of sinefungin (less than 15 min; lane 1 in Figure 11), compared with the results obtained using AdoMet for the same duplex (Figure 10). The persistence of the retarded complex with sinefungin (up to 75 min; Figure 11) moreover suggests that an abortive ternary complex is formed with sinefungin.

In order to investigate the nature of the interaction between the oligodeoxynucleotides ODN-MSP2A and ODN-SP1 and GST-M.*MspI*, a series of competition experiments was performed in the presence of AdoMet. Figure 12 clearly shows that the complex between ODN-MSP2A and GST-M.*MspI* cannot be displaced significantly by a 1000-fold excess of ODN-SP1, whereas the ODN-SP1 complex can be significantly displaced by a 10-fold excess of unlabelled ODN-SP1. We have shown the overexposure of lanes 1 and 2 in order to emphasize the difference in the affinity shown by GST-M.*MspI* for the substituted and unsubstituted forms of the cognate duplex.

In a further experiment, titration of the ODN-SP1-GST-M.*MspI* complex with the non-specific 33mer oligodeoxynucleotide [ODN-NSP(33)] in the presence of AdoMet shows that complex-formation persists in the presence of a 1000-fold excess of competitor DNA (Figure 13). Clearly, the complex

Figure 13 Non-specific DNA binding of GST-*M.MspI* revealed by competition experiments in the presence of AdoMet

Lane 1, ODN-SP1 + GST-*M.MspI* + AdoMet; lanes 2–7, as lane 1 but with increasing amounts of the competitor ODN-NSP(33) in the following excesses 100-, 200-, 300-, 500-, 700- and 1000-fold. All incubations were for 30 min at 4 °C.

Figure 14 Stability of the GST-*M.MspI*-ODN-MSP2A complex formed in the presence of AdoMet

After 40 min incubation of the fusion protein, DNA and AdoMet at 4 °C, SDS was added to the incubation for 15 min before electrophoresis at room temperature. Final SDS concentrations were: 0%, 0.08%, 0.16%, 0.24%, 0.32%, 0.4%, 0.6% and 0.8% in lanes 2–9 respectively. Lane 1 contains ODN-MSP2A alone.

formed between the enzyme and both ODN-SP1 and ODN-MSP2A are sequence specific and in addition that formed between the modified oligodeoxynucleotide is of a significantly higher

affinity. AdoMet therefore appears to potentiate sequence-specific binding by the Mtase to its cognate site.

In order to explore the nature of the interaction between the ODN-MSP2A oligodeoxynucleotide and the enzyme, preformed ODN-MSP2A-GST-*M.MspI* complexes (with AdoMet present) were exposed to a range of SDS concentrations. The gradual emergence of two retarded bands (indicated with arrows in Figure 14) suggests a population of folded and denatured protein-DNA complexes. Both complexes are stable up to a level of 0.6% (w/v) SDS, suggesting a strong possibly covalent interaction between the enzyme and DNA. The slowly migrating species is still present at a concentration of 0.8% (w/v) SDS. However, these complexes are not as stable as those observed when 5-fluorodeoxycytidine is incorporated into the DNA substrate since up to 95% covalent complex can be obtained in the presence of 1% SDS after boiling with this inhibitor [13].

DISCUSSION

We have used PCR to tailor the gene encoding *M.MspI* for high-level expression as a fusion protein with GST. The biological activity of the purified fusion enzyme is indistinguishable from that of the wild-type enzyme, and our results are fully consistent with those of Dubey and Roberts [8] who used the native enzyme in gel-retardation experiments. Moreover, the one-step affinity purification procedure should simplify the expression and purification of mutants, in particular, chimeric DNA Mtases produced by exchange of the non-conserved TRD region of *M.MspI* with related C-5 Mtases [12].

The purified fusion enzyme has been used to explore the substrate binding by a typical C-5 Mtase. The results we have obtained are therefore likely to apply to all related Mtases, including the vertebrate CpG Mtases which share sequence similarities with the bacterial enzymes [37]. In addition to considering the interaction of *M.MspI* with its natural substrates DNA, AdoMet and additionally AdoHcy, we have also used two inhibitors. Sinefungin is a competitive inhibitor with respect to AdoMet and has been widely used to explore the mechanisms of a variety of Mtase enzymes. We have also synthesized a novel mechanism-based inhibitor, 2-pyrimidinone-1- β -D-2'-deoxyribofuranoside, which appears to act in a manner analogous to 5-azacytidine and 5-fluorodeoxycytidine in trapping the enzyme on the target deoxycytidine through a stable covalent interaction [13]. The detailed mechanism of azacytidine-mediated inhibition has yet to be determined, since the synthesis of oligodeoxynucleotides containing this analogue has not been reported. However, it appears that 5-fluorodeoxycytidine, when incorporated into oligodeoxynucleotides, inhibits the regeneration of active enzyme before Mtase-catalysed β -elimination [13]. Using gel-retardation techniques, we have demonstrated that targeted substitution of oligodeoxynucleotides with 2-pyrimidinone (e.g. ODN-MSP2A) considerably enhances the formation of sequence-specific protein-DNA complexes with *M.MspI*. [Similar experiments with *M1.HgaI*, *M2.HgaI* and *M.HpaII* support the generality of this conclusion (D. P. Hornby, unpublished work).] Moreover, although AdoMet, AdoHcy and sinefungin all potentiate sequence-specific DNA binding by *M.MspI*, the presence of the cofactor is not essential for covalent complex-formation. This observation demonstrates that the 2-pyrimidinone-1- β -D-2'-deoxyribofuranoside analogue inhibits C-5 DNA Mtases before methyl transfer and β -elimination and therefore in a manner distinct from that of the 5-fluorodeoxycytidine analogue.

As discussed in the Introduction, the mechanism of action of the C-5 DNA Mtases involves the initial formation of a covalent

bond between an enzyme cysteine residue and the C-6 atom of the target deoxycytidine to give a 5,6 saturated pyrimidine. This generates a carbanion at C-5 which can attack the methyl group of AdoMet resulting in methylation at C-5. The cysteine thiol pyrimidine bond is broken in a β -elimination involving abstraction of a proton from the C-5 of the pyrimidine ring by a basic amino acid on the enzyme. Initial attack of the enzyme cysteine thiol at the C-6 position of the 2-pyrimidinone analogue is identical with the mechanism proposed for the decomposition of 2-pyrimidinone-containing oligodeoxynucleotides with ammonia. As the 5,6 double bond of 2-pyrimidinones is activated towards addition reactions, in comparison with the usual pyrimidines, one would expect facile formation of the initial protein oligodeoxynucleotide adduct. Indeed we observe sequence-specific adduct formation in the presence and absence of AdoMet, and moreover no methyl transfer takes place, which therefore suggests that pyrimidinone inhibition precedes that demonstrated by 5-fluorocytidine, which requires AdoMet addition for stable formation of the methylated covalent adduct [13]. Presumably in the 2-pyrimidinone-1- β -D-2'-deoxy-ribofuranoside adduct, a proton becomes attached to the C-5 carbanion. The stability of the complex formed between enzyme and d^4H C-containing oligomers implies that the final β -elimination does not take place, or is at least very slow. Two reasons for this observation are possible. First the C-5 proton in the d^4H C intermediate could be less acidic than that of the deoxycytidine species. This would make proton abstraction by the base more difficult. Unfortunately no information is available concerning the relative acidities of the two protons. Second, the change from deoxycytidine to 2-pyrimidinone-1- β -D-2'-deoxyribofuranoside with consequent loss of the 4-amino group may lead to subtle changes in position of the active-site amino acids. This could alter the position of the critical basic residue such that it was no longer optimally oriented for proton abstraction.

The photocross-linking experiments with tritiated AdoMet and GST-M.*MspI* follow the same pattern as those described by Som and Friedman [35] with the C-5 Mtase M.*EcoRII*. In addition we have shown that the presence of certain DNA sequences can potentiate the formation of cross-linked product. The oligodeoxynucleotides that stimulate adduct formation [ODN-MSP2A, ODN-MSP2B and ODN-NSP(16)] are all unable to accept a methyl group from AdoMet. However, these oligodeoxynucleotides differ in several respects. The stimulation caused by ODN-NSP(16) is explained by a non-specific association of the duplex with the enzyme to form a ternary complex lacking sequence-specific protein-DNA interactions. Similarly it appears that removal of the 4-NH₂ from the internal deoxycytidine in the recognition site (CCGG) abolishes sequence-specific interactions and therefore highlights the importance of the amino group in the interaction between M.*MspI* and its cognate site.

Since ODN-MSP2A contains a reactive d^4H C base at the site of methylation, the DNA must form a sequence-specific complex with ODN-MSP2A. Indeed the gel-retardation experiments with ODN-MSP2A confirm this. In contrast with the results of Som and Friedman [35], it is unlikely that the formation of methylcysteine occurs in the M.*MspI* photocross-linking reaction since the nucleophilic cysteine in the Pro-Cys dipeptide is thought to form the point of attachment to the reactive deoxycytidine. Experiments with M.*EcoRII* [35] showed that the major cross-linked product formed in the absence of DNA was methylcysteine. However, other labelled species were obtained but not characterized in detail. It seems possible therefore that, in the presence of the inhibitory oligodeoxynucleotide ODN-MSP2A, the enzyme primarily forms a covalent AdoMet adduct which is

only a minor product in the absence of DNA. Indeed, competition for the cysteine between the reactive d^4H C and AdoMet would be expected to diminish cross-linking if methylcysteine was the product. Although AdoMet photocross-linking provides a useful means of exploring the active site of Mtase enzymes, it is unlikely that in the absence of DNA this reaction is a significant event in catalysis [35]. It may of course imply that C-5 Mtases, like the Dam enzyme [38], bind AdoMet in an allosteric manner. The results presented here cannot distinguish between a catalytic or an allosteric (or possibly both) interaction between M.*MspI* and AdoMet.

Although sequence-specific DNA recognition appears to be the first step in the reaction pathway of M.*MspI*, the results shown in Figure 9 clearly demonstrate the enhanced binding of DNA in the ternary complex. The addition of either AdoMet, AdoHcy or sinefungin leads to potentiation of sequence-specific DNA-protein interactions. The mechanism proposed by Wu and Santi for M.*HhaI* [9] involves the addition of AdoMet after DNA binding followed by the release of AdoHcy and finally methylated DNA. AdoHcy is also a potent competitive inhibitor with respect to AdoMet. Our binding data are entirely consistent with this mechanism. However, elucidation of the reaction pathway of any palindromic DNA Mtase is complicated by the fact that the product of the first round of methyl transfer is hemimethylated DNA. It is possible that for a given DNA Mtase, the affinity for hemimethylated product is different from that for the unmodified substrate. The ability of DNA C-5 Mtases to discriminate between these two substrates is of course central to the biological function of the vertebrate enzyme. We are currently conducting a detailed analysis of the reaction pathway of M.*HgaI* hemimethyltransferase in order to circumvent the problems associated with hemimethylation.

The reaction pathway leading to formation of the ternary complex that is consistent with our binding studies suggests that an isomerization of the ternary complex occurs. Interestingly, the results of a time course of incubation of M.*MspI* with AdoMet in the presence of the cognate oligodeoxynucleotides ODN-SP1 and ODN-MSP2A show that the potentiation of DNA-protein interactions take significantly longer with AdoMet than with either AdoHcy or sinefungin. This suggests that the latter compounds are able to lock the enzyme into a high-affinity DNA-binding mode and perhaps explains their inhibitory effects. AdoMet on the other hand in acting as a methyl donor must balance high-affinity DNA binding with catalytic release of the methylated product. Wu and Santi [9] demonstrated that the non-linearity of M.*HhaI* progress curves was a result of the accumulation of AdoHcy during the reaction. Clearly, the strong interaction between DNA that is promoted by AdoHcy in our gel-retardation experiments supports their conclusion. These results are also consistent with an allosteric role for the cofactor [38]. Nevertheless, M.*MspI* appears to bind DNA weakly in the absence of AdoMet but, after the addition of AdoMet, a slow isomerization of the ternary complex precedes release of products. The accumulation of AdoHcy, which has a higher affinity than AdoMet for the enzyme-DNA binary complex, leads to the formation of a kinetically significant abortive complex.

The application of gel-retardation assays using oligodeoxynucleotides modified by removal of the 4-NH₂ group of the target deoxycytidine to explore m^5C DNA Mtase enzymes is an extremely powerful approach. The substitution of deoxycytidine by a pyrimidinone base results in the loss of one hydrogen bond with the corresponding guanosine. Therefore in a duplex with two such substitutions a local distortion of DNA structure may occur. It is therefore possible that the ODN-MSP2A duplex binds to the enzyme with high affinity in part because it mimics

the transition state of the reaction. Moreover, the high level expression of a functional *MspI* Mtase as a fusion with GST offers the opportunity to investigate the enzymology of DNA methylation using site-directed mutagenesis and domain-swap strategies with this test bed system.

We are grateful for the support of the Yorkshire Cancer Research Campaign and the Science and Engineering Research Council. We also thank Rich Roberts for communicating results before publication and for several helpful discussions. C.T. was the recipient of an SERC CASE studentship in collaboration with Northumbrian Biologicals, U.K. We also thank Phil Eastlake for many helpful discussions. B.A.C. is a Research Fellow of the Lister Institute of Preventive Medicine.

REFERENCES

- Wilson, G. G. (1991) *Nucleic Acids Res.* **19**, 2539–2566
- Walder, R. Y., Langtim, C. J., Chatterjee, R. and Walder, J. A. (1983) *J. Biol. Chem.* **258**, 1235–1241
- Walder, R. Y., Walder, J. A. and Donelson, J. E. (1984) *J. Biol. Chem.* **259**, 8015–8026
- Nwankwo, D. O. and Wilson, G. G. (1988) *Gene* **64**, 1–8
- Lauster, R. (1989) *J. Mol. Biol.* **206**, 313–321
- Posfai, J., Bhagwat, A. S., Posfai, G. and Roberts, R. J. (1989) *Nucleic Acids Res.* **17**, 2421–2435
- Dubey, A. K., Mollet, B. and Roberts, R. J. (1992) *Nucleic Acids Res.* **20**, 1579–1585
- Dubey, A. K. and Roberts, R. J. (1992) *Nucleic Acids Res.* **20**, 3167–3173
- Wu, J. C. and Santi, D. V. (1987) *J. Biol. Chem.* **262**, 4778–4786
- Matthews, D. A., Appelt, K., Oatley, S. J. and Xuong, Ng. H. (1990) *J. Mol. Biol.* **214**, 923–936
- Matthews, D. A., Villafranca, J. E., Janson, C. A., Smith, W. W., Welsh, K. and Freer, S. (1990) *J. Mol. Biol.* **214**, 937–948
- Balganesh, T. S., Reiners, L., Lauster, R., Noyer-Weidner, M., Wilke, K. and Trautner, T. A. (1987) *EMBO J.* **6**, 3543–3549
- Chen, L., MacMillan, A. M., Chang, W., Ezaz-Nikpay, K., Lane, W. S. and Verdine, G. L. (1991) *Biochemistry* **30**, 11018–11025
- Sambrook, J., Fritsch, E. F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd edn., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
- McBride, L. J., Kierzek, R., Beaceage, S. L. and Caruthers, M. H. (1986) *J. Am. Chem. Soc.* **108**, 2040–2048
- Schulhof, J. C., Molko, D. and Teole, R. (1987) *Nucleic Acids Res.* **15**, 397–416
- Vu, H., McCollum, C., Jacobson, K., Thiessen, P., Vinayak, R., Spiess, E. and Andrus, A. (1990) *Tetrahedron Lett.* **31**, 7269–7272
- Connolly, B. A. and Newman, P. C. (1989) *Nucleic Acids Res.* **17**, 4957–4974
- Connolly, B. A. (1991) in *Oligonucleotides and their Analogues, A Practical Approach* (Eckstein, F., ed.), pp. 155–183, IRL Press, Oxford
- Gildea, B. and McLaughlin, L. W. (1989) *Nucleic Acids Res.* **17**, 2261–2281
- Jones, R. A. (1984) in *Oligonucleotide Synthesis, a Practical Approach* (Gait, M. J., ed.), pp. 23–34, IRL Press, Oxford
- Sonha, N. D., Biernat, J., McManus, J. and Koster, H. (1984) *Nucleic Acids Res.* **12**, 4539–4557
- Wightman, R. and Holy, A. (1973) *Collect. Czech. Chem. Commun.* **38**, 1381–1397
- Cech, D. and Holy, A. (1977) *Collect. Czech. Chem. Commun.* **42**, 2246–2260
- Altermatt, R. and Tamm, C. (1985) *Helv. Chim. Acta* **68**, 475–483
- Rappaport, H. P. (1988) *Nucleic Acids Res.* **16**, 7253–7267
- locono, J. A., Gildea, B. and McLaughlin, L. W. (1990) *Tetrahedron Lett.* **31**, 175–178
- Himelsbach, F., Schultz, B. S., Trichtinger, T., Charubala, R. and Pfeiderer, W. (1984) *Tetrahedron* **40**, 59–72
- Stengele, K. P. and Pfeiderer, W. (1990) *Tetrahedron Lett.* **31**, 2549–2552
- Webb, T. R. and Matteucci, M. D. (1986) *Nucleic Acids Res.* **14**, 7661–7674
- Gryaznov, S. M. and Letsinger, R. L. (1991) *J. Am. Chem. Soc.* **113**, 5876–5877
- Kochetkov, N. K. and Budovskii, E. I. (1972) in *Organic Chemistry of Nucleic Acids Part B*, pp. 381–448, Plenum Press, London
- Hurst, D. T. (1980) in *An Introduction to the Chemistry and Biochemistry of Pyrimidines, Purines and Pteridines*, pp. 15–63, John Wiley and Sons, Chichester
- Maxam, A. M. and Gilbert, W. (1980) *Methods Enzymol.* **65**, 499–560
- Som, S. and Friedman, S. (1991) *J. Biol. Chem.* **266**, 2937–2945
- Santi, D. V., Garrett, C. E. and Barr, P. J. (1983) *Cell* **33**, 9–10
- Bestor, T., Laudano, A., Mattaliano, R. and Ingram, V. (1988) *J. Mol. Biol.* **203**, 971–983
- Bergerat, A., Guschlbauer, W. and Fuzakerley, G. V. (1991) *Proc. Natl. Acad. Sci. U.S.A.* **88**, 6394–6397