Nucleic Acids Research, Vol. 20, No. 7 1579–1585

Purification and characterization of the *Msp*I DNA methyltransferase cloned and overexpressed in *E.coli*

Ashok K.Dubey, Beat Mollet+ and Richard J.Roberts*
Cold Spring Harbor Laboratory, PO Box 100, Cold Spring Harbor, NY 11724, USA

Received January 13, 1992; Revised and Accepted February 25, 1992

ABSTRACT

The *Msp*I restriction-modification system, which recognizes the sequence 5'-CCGG-3', has been previously cloned and sequenced (1). We subcloned the methyltransferase gene (M.*Msp*I) downstream of the ptac promoter in the multicopy vector pUC119 and overexpressed it in *E. coli*. Upon induction with IPTG, M.*Msp*I constitutes more than 10% of cellular protein. A scheme has been devised to purify large amounts of biologically active M.*Msp*I to apparent homogeneity from these overexpressing *E. coli* cells. Approximately 0.8 mg of pure M.*Msp*I per gram of cells (wet weight) can be obtained. The apparent molecular weight of M.*Msp*I is 49 kD, by SDS gel electrophoresis and 48 – 54 kD by gel filtration. At low concentrations (less than 0.4 mg/ml), the methyltransferase is a monomer in solution but at higher concentrations (greater than 3.0 mg/ml) it exists predominantly as a dimer. Polyclonal antibodies raised against M.*Msp*I cross-react with the DNA-methyltransferases of several other restriction-modification systems.

INTRODUCTION

DNA methyltransferases are ubiquitous in living cells (2,3), but in only a few cases have specific biological roles been defined. In prokaryotes they often form part of restriction-modification systems (4) or play an essential role in DNA mismatch repair (5). In many plants, large amounts of 5-methylcytosine are present (6), but its function is unknown. The dinucleotide CG is frequently methylated in mammals and other higher eukaryotes (7), and methylation of transcriptional control regions has been correlated with the absence of transcription (8). Whether methylation causes this inhibition of transcription or is the result of transcriptional inactivity is undetermined. Many epigenetic phenomena, such as imprinting (9), are postulated to involve the effects of methylation but rigorous proof is still lacking.

Three kinds of DNA methylation are known, leading to N^6-methyladenine, N^4-methylcytosine or 5-methylcytosine. Much work has been done on the *Eco*RI methyltransferase, which forms N^6-methyladenine (10–12), and some studies have also been carried out on other enzymes forming N^6-methyladenine (13). The *Hha*I (14) and *Eco*RII methyltransferases (15,16), both of which form 5-methylcytosine, have been studied extensively. The reaction mechanism involves the formation of a covalent intermediate between the 6-position of cytosine and a cysteine residue that lies in a well-conserved region of all enzymes forming 5-methylcytosine (m5C-methylases) (17,18). Little is known about the mechanism employed by enzymes forming N^4-methylcytosine.

We have been studying another m5C-methyltransferase, M.*Msp*I, from *Moraxella* species (1). This enzyme recognizes the sequence 5'-CCGG-3' and methylates the outer cytosine residue (19). We are using this enzyme as a model system to understand in detail how m5C-methyltransferases work both from a mechanistic and a structural standpoint. We are especially interested in defining the regions of the methyltransferase responsible for sequence-specific DNA recognition and also to define the mechanistic role of each of the motifs, previously shown to be present in all sequenced m5C-methyltransferases (20,21). In this paper we describe the preparation of a clone that greatly overexpresses M.*Msp*I in *E. coli* and the subsequent purification and preliminary characterization of this methyltransferase.

MATERIALS AND METHODS

Chemicals and Reagents

Restriction and modification enzymes, except *Cla*I and *Afl*III (from Bethesda Research Laboratories), DNA-ligase and bacteriophage λ-DNA were obtained from New England Biolabs and used as recommended by the supplier. Reagents for quantitative assay and SDS-polyacrylamide gel electrophoresis were from BioRad and used as recommended. Nitrocellulose membranes and the Western blotting kit (ProtoBlot, Rabbit) were from Promega. Phosphocellulose P11 was from Whatman. The FPLC columns Mono S and Mono Q were from Pharmacia. Lysozyme, tetracycline, ampicillin, carbenicillin, protamine sulfate and S-adenosyl-L-methionine (AdoMet) were obtained from Sigma. [^{3}H]-AdoMet (specific activity 11–13 C_i/mmole) was from New England Nuclear. Scintillation fluid,

* To whom correspondence should be addressed

+ Present address: Nestlé Research Centre, Nestec Ltd., Vers-chez-les-Blanc, PO Box 44, CH 1000 Lausanne 26, Switzerland

Cytoscint™, was purchased from ICN Biomedicals, Inc. All chemicals used were of the highest purity reagent grade.

Strains and Plasmids

E. coli K12, strain RR1 (22), was used as a host to overexpress M.*Msp*I. Plasmid pNK627, which is a pACYC184 derivative carrying the lacI^Q gene inserted within its chloramphenicol resistance gene, and plasmid ptac12 (23,24) were obtained from Dr. N. Kleckner (Harvard University). Plasmid pM921 which carries the *Msp*I restriction-modification genes was described previously (1). *E. coli* strain RR1, harboring the overexpressing plasmid ps3 (see below) together with pNK627, was grown in super broth (25), containing tetracycline (20 μg/ml) and/or carbenicillin (200 μg/ml). Cultures were induced with 0.1 mM IPTG at OD_{600} approximately 1.2 and growth was continued for 3 hours.

Construction of the *Msp*I overexpression plasmid

Recombinant DNA experiments were performed using established procedures (26). The steps involved in the construction of the overexpression vector, ps3, are outlined in Figure 1.

pBR322 was linearized with *Nde*I and resected to give blunt ends using the Klenow fragment of *E. coli* DNA polymerase I. The linearized plasmid was further treated with *Hin*dIII and a 2.1 kb ×fragment containing the ampicillin resistance gene and the origin of replication was purified on a 1% agarose gel. A 3 kb *Eco*RI to *Hin*dIII fragment, containing the *Msp*I restriction-modification genes, was isolated from pM921 such that the *Eco*RI end had been rendered blunt by resection following the initial cleavage. This 3 kb fragment was ligated to the 2.1 kb fragment described above and was transformed into *E. coli* strain RR1. Several individual transformants isolated on ampicillin plates were tested for the presence of the gene for M.*Msp*I by assaying their sensitivity to digestion with the *Msp*I restriction enzyme (R.*Msp*I) *in vitro*. The resulting 5.1 kb plasmid was designated pn1. DNA from pn1 was linearized with *Hin*dIII and partially digested with *Rsa*I and blunt-ended by treatment with the Klenow fragment of DNA polymerase I. Fragments of approximately 3.5 kb were gel-purified, ligated and transformed into *E. coli*. Individual clones, isolated on ampicillin plates, were screened by restriction enzyme digestion (*Rsa*I, *Dde*I and *Taq*I) for the correct size and for resistance to *Msp*I digestion *in vitro*. One isolate, designated pq8, was used for further studies.

pq8 was cleaved with *Cla*I and blunt-ended by treatment with the Klenow fragment of DNA polymerase I. Following cleavage with *Pst*I, the large *Cla*I to *Pst*I fragment, containing the *Msp*I methyltransferase gene was purified and ligated to the small *Pst*I to *Pvu*II fragment from ptac12, containing the ptac promoter. Following transformation into *E. coli* strain RR1 harboring F'lacI^Q individual colonies growing in ampicillin, were tested for the presence of M.*Msp*I and analyzed by restriction enzyme digestion. One clone, designated pr13, was cleaved with *Eco*RI and *Afl*III and the methyltransferase containing fragment was recloned into pUC119, replacing the *Sma*I to *Eco*RI fragment from the polylinker. The resulting construct, showing ampicillin resistance and M.*Msp*I activity, was designated ps3 and the key elements of its structure confirmed by sequencing.

Purification of M.*Msp*I

M.*Msp*I was purified from *E. coli* strain RR1 carrying the plasmid ps3 by the following procedure. All operations were performed at 4°C unless stated otherwise.

Step 1: Preparation of crude extract and removal of nucleic acids

Frozen cells (5 g, wet weight), from cultures grown and induced at 37°C, were thawed at 4°C and suspended in 20 ml of buffer A (10 mM potassium phosphate (pH 7.4), 1 mM Na_2EDTA, 14 mM β-mercaptoethanol, 10% glycerol) containing 0.3 M NaCl. The cell suspension was incubated with 15 mg lysozyme at 4°C for 30 min and then sonicated. Following centrifugation at 31000×g for 30 minutes, 2.5 ml protamine sulfate (40 mg/ml) was slowly added to precipitate nucleic acids and the supernatant recovered by centrifugation.

Step 2: Chromatography on Phosphocellulose

Freshly prepared crude extract was diluted to a conductivity equivalent to 0.15 M NaCl and applied to a phosphocellulose column (22×53 mm) pre-equilibrated with buffer A containing 0.15 M NaCl. The column was operated at a flow rate of 1 ml/min. After sample loading, the column was washed with 80 ml of the starting buffer (buffer A + 0.15 M NaCl). Elution was with 160 ml of a linear gradient from 0.15 M−1.0 M NaCl. Fractions (4ml) were assayed for M.*Msp*I activity and active fractions were checked by SDS-polyacrylamide gel electrophoresis before pooling. M.*Msp*I eluted between 0.3−0.4 M NaCl (Figure 3).

Step 3: FPLC chromatography on Mono S

The pooled fractions from the phosphocellulose column were adjusted to 0.1 M NaCl, by dilution and loaded on a Mono S column (HR 5/5) previously equilibrated with buffer A containing 0.1 M NaCl and running at a flow rate of 1 ml/min. The column was washed for 10 minutes following loading and eluted with 40ml of a linear gradient from 0.1−0.5 M NaCl. Fractions (1 ml) were assayed for M.*Msp*I activity and peak fractions were finally pooled based on their SDS-polyacrylamide gel profile. The enzyme eluted from the column at a salt concentration slightly above 0.2 M NaCl.

Step 4: FPLC Chromatography on Mono Q

The enzyme pool from the Mono S column was diluted to adjust the conductivity equivalent to 0.1 M NaCl and applied to a Mono Q column, previously equilibrated with buffer A containing 0.1 M NaCl. The column was washed for 10 minutes following loading and eluted with 40 ml of a linear gradient from 0.1−0.4 M NaCl. Fractions were assayed as in Step 3 and the enzyme eluted at a salt concentration slightly below 0.2 M. Pooled fractions were dialyzed against storage buffer (0.1 M NaCl, 50 mM Tris-HCl pH 8.0, 10 mM Na_2EDTA, 1 mM DTT, 50% glycerol) and stored at −20°C.

Polyclonal antibodies against M.*Msp*I

Inclusion bodies were prepared from a 37°C induced culture (50 ml) of *E. coli* RR1, containing ps3 and F'lacI^Q. Cells were suspended in 4 ml of lysis buffer (0.25 M NaCl, 20 mM Tris.HCl pH8.0, 10 mM β-mercaptoethanol, 1 mM Na_2EDTA and 1 mg/ml lysozyme) and stored at 4°C for 30 min and then sonicated. Following centrifugation, the pellet, which contained the inclusion bodies, was washed with lysis buffer and resuspended in 2 ml SDS sample buffer (5% glycerol, 30 mM Tris.HCl pH 6.8, 0.1% SDS, 0.35 M β-mercaptoethanol, 0.02% Bromophenol blue) and boiled for 5 minutes. The denatured

inclusion bodies were separated on a 1.5 mm thick polyacrylamide gel run under denaturing conditions (27). The M.*Msp*I band was localized by Coomassie Blue staining in water, excised and antigen samples were prepared and injected into rabbits according to published procedures (28).

Western blotting, protein sequence analysis and dot-blot analysis

SDS gel electrophoresis was carried out using standard procedures (27). Proteins were electroblotted onto a polyvinylidene difluoride (PVDF) membrane which was stained with Coomassie Blue to locate the protein bands. For N-terminal amino acid sequencing, the desired protein bands were excised from the PVDF membrane and sequencing was performed on an Applied Biosystems 475 protein sequencer. For immunoblot experiments, proteins were separated on a 10% SDS polyacrylamide gel and blotted onto a nitrocellulose membrane. The blot was developed using the Alkaline Phosphatase System from Promega following the procedure provided by the supplier.

Assay for methyltransferase activity

M.*Msp*I activity was measured qualitatively using a protection assay in which bacteriophage λ-DNA was incubated with the methyltransferase and, following methylation, the DNA was challenged with R.*Msp*I. A typical assay reaction was carried out in 20 μl reaction cocktail containing: 100 mM NaCl, 50 mM Tris.HCl (pH8.0), 10 mM Na_2EDTA, 14 mM β-mercapto-ethanol, 80 μM S-adenosyl-L-methionine, 1 μg λ-DNA and 2 μl of column fractions. The reaction mixture was incubated at 37°C for 30 min and then heated at 70°C for 5 min to inactivate the methyltransferase. Following cooling, $MgCl_2$ was added to 10 mM and incubated for 30 min at 37°C with R.*Msp*I (5 units). The samples were analyzed by 0.8% agarose gel electrophoresis.

Quantitative measurements of M.*Msp*I activity were based on the incorporation of tritiated methyl groups into substrate DNA. The methylation reactions were carried out as above, except that 2.2 μM [^{3}H] S-adenosyl-L-methionine was included in the reaction mixtures. Following incubation, samples were withdrawn, adsorbed onto 2.5 mm Whatman DE81 paper circles and immersed in 0.2 M $(NH_4)HCO_3$. Following washes with 0.2 M $(NH_4)HCO_3$, H_2O and 100% ethanol the circles were dried and counted.

The specific activity (units/mg protein) was based on the initial rate of methyl transfer into substrate DNA. One unit of M.*Msp*I incorporates 1 pmole of methyl group per min under standard reaction conditions.

Molecular Weight Determination

The molecular weight of the reduced and denatured protein was determined by 10% SDS-polyacrylamide gel electrophoresis using appropriate molecular weight standards (BioRad). The native molecular weight and state of aggregation of M.*Msp*I in solution were determined by gel filtration chromatography on Superose 12 (Pharmacia, HR 10/30, FPLC). The column was equilibrated with 0.2M NaCl, 50mM Tris-HCl pH 8.0, 1 mM Na_2EDTA, 5% glycerol and then calibrated with standards of known molecular weight; β-amylase, alcohol dehydrogenase, bovine serum albumin, carbonic anhydrase and cytochrome c. The standard plot of molecular weight against partition coefficient (29) was used to determine apparent molecular weights at different protein concentrations.

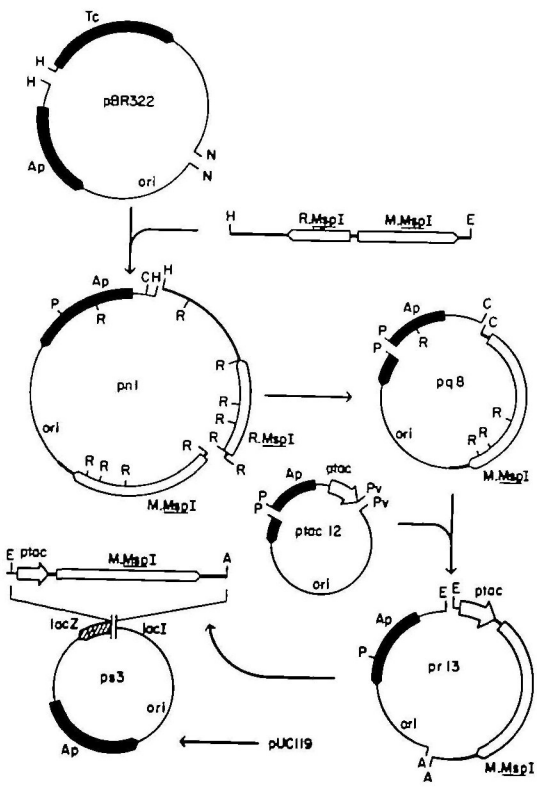

Figure 1. Construction of M.*Msp*I overexpression plasmid. The 3 kb *Eco*RI to *Hin*dIII fragment, containing the *Msp*I restriction-modification genes, was obtained from the original clone (1), subcloned into pBR322 and subsequently manipulated as described in Methods. Restriction sites are abbreviated as follows: H: *Hin*dIII, E: *Eco*RI, N: *Nde*I, R: *Rsa*I, P: *Pst*I, Pv: *Pvu*II, C: *Cla*I, A: *Afl*III.

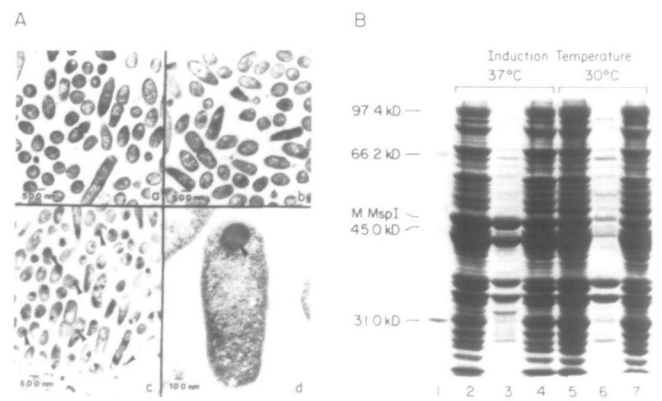

Figure 2. Inclusion body formation by M.*Msp*I. **Panel A**: *E. coli* cells were harvested during logarithmic growth, fixed and examined by transmission electron microscopy. (a) cells with no plasmid; (b) uninduced cells containing ps3; (c) fully induced cells containing ps3; (d) a single cell containing an inclusion body shown at high magnification. **Panel B**: Temperature sensitivity of inclusion body formation. Cells were grown at 37°C until OD_{600} = 1.2, induced with 0.1 mM IPTG and growth continued at either 30°C or 37°C. After 2 hours the cells were harvested and analyzed by SDS polyacrylamide gel electrophoresis. Lane 1: molecular weight markers, lane 2: total cellular protein from cells induced and grown at 37°C cells, lane 3: insoluble protein from 37°C cells, lane 4: soluble protein from 37°C cells. Lanes 5–7 correspond to lanes 2–4 except that cells were induced and grown at 30°.

RESULTS

Overproduction of M.*Msp*I in *E. coli*

From the original clone, pM921, containing the intact *Msp*I restriction-modification system (1) the methyltransferase gene was excised and placed downstream of the ptac promoter as indicated in Figure 1. The final construct, ps3, contains approximately 130 nucleotides between the promoter sequence and the methionine codon at which the translation of the methyltransferase gene begins. When this plasmid is placed in *E. coli* RR1 containing F'lacI^{Q}, expression from the ptac promoter is repressed and little or no M.*Msp*I is detected by SDS gel electrophoresis of total cellular protein. However, M.*Msp*I activity is present, because in the absence of induction, this plasmid DNA is completely resistant to the action of R.*Msp*I. Upon induction, the strain produces high levels of the M.*Msp*I. However, the methyltransferase protein becomes aggregated and is sequestered in a biologically inactive form. Inclusion bodies can be seen inside the cells when checked by electron microscopy (Figure 2, panel A). These inclusion bodies are easily purified and contain large quantities of the methyltransferase as judged by SDS-polyacrylamide gel electrophoresis. Inclusion bodies have been reported with several other overexpression systems (30–32).

It has been reported that the formation of inclusion bodies is often dependent upon the growth temperature of the culture (33). We have found that this is also the case for the M.*Msp*I inclusion bodies. When the cells are grown at 37°C and induced at 37°C, inclusion bodies are readily apparent, whereas if induction is carried out at 30°C they are not produced (Figure 2, panel B). From Figure 2B it can be seen that two prominent protein bands are present in the insoluble fraction of the lysate, which contains the inclusion bodies, and both bands disappear at the lower temperature. The apparent molecular weight of the upper band is 49 kD, in good agreement with the predicted molecular weight of 47,659 calculated from the sequence (1). The lower band, corresponding to an apparent molecular weight of 45 kD, could represent a truncated version of M.*Msp*I. To test this possibility we excised both bands from an SDS gel and subjected them to N-terminal protein sequencing. The 49 kD band gave an

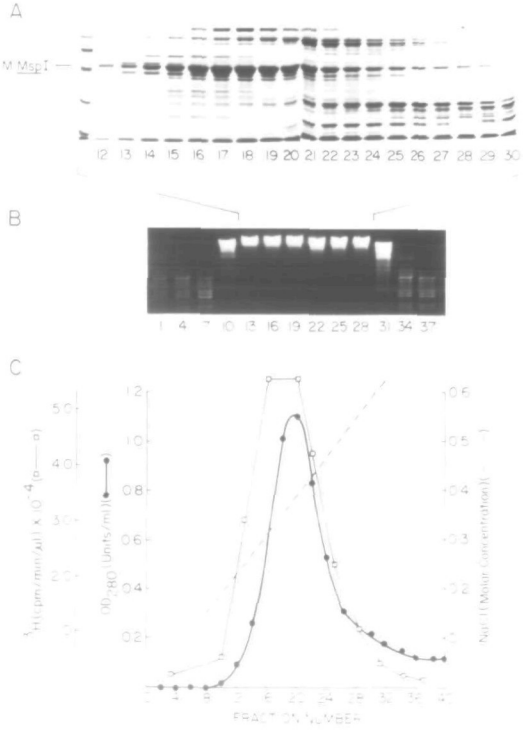

Figure 3. Phosphocellulose chromatography of M.*Msp*I. The crude extract recovered after protamine sulfate precipitation was chromatographed on phosphocellulose and assayed as described in Methods. **Panel A:** SDS polyacrylamide gel analysis. **Panel B:** Protection assay. **Panel C:** Incorporation of ^{3}H-methyl groups and OD profile.

Figure 4. M.*Msp*I purification profile. **Lane 1:** crude extract after protamine sulfate precipitation (30μg). **Lane 2:** phosphocellulose pool (10μg total protein). **Lane 3:** Mono S pool (7μg total protein). **Lane 4:** Mono Q pool (4μg total protein)

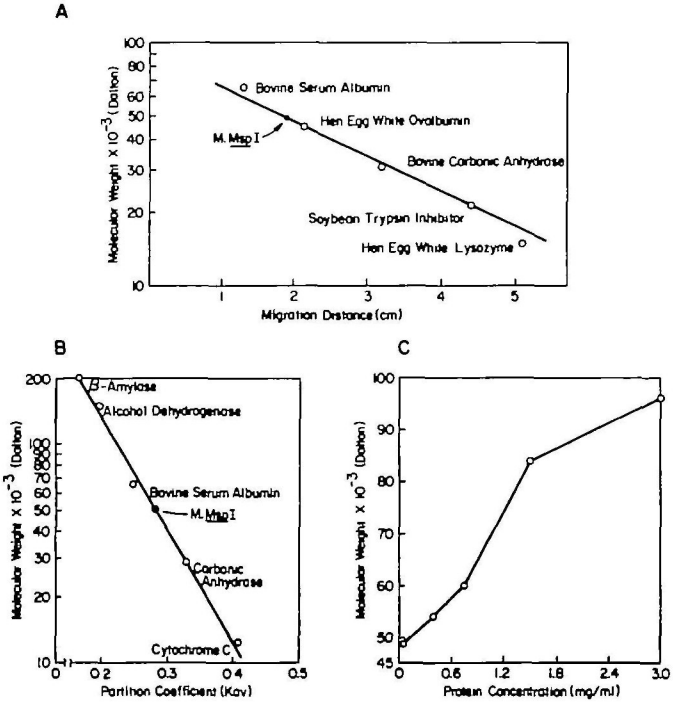

Figure 5. Molecular size determination for M.*Msp*I. **Panel A:** Homogeneous M.*Msp*I and the indicated marker proteins were subjected to 10% polyacrylamide gel electrophoresis in the presence of 0.1% SDS. The mobilities of the marker proteins were used to calibrate the curve and the molecular weight of reduced, denatured M.*Msp*I was calculated by interpolation. **Panel B:** Native M.*Msp*I (loading concentration 0.05 to 0.4 mg/ml) and various marker proteins were subjected to gel filtration chromatography on Superose 12 as described in Methods. The partition coefficients of the protein standards were used to calibrate the curve and the molecular weight of native M.*Msp*I was calculated by interpolation. **Panel C:** M.*Msp*I at various loading concentrations was subjected to Superose 12 gel filtration chromatography and apparent molecular weights were calculated as for Panel B.

unambiguous sequence for 31 amino-acids. This sequence exactly matched the N-terminal sequence predicted from the DNA sequence, including the N-terminal methionine. The lower band also contained an N-terminal methionine, but in this case the sequence of 20 amino-acids, determined experimentally, corresponded to a stretch of sequence beginning at the second methionine in the M.*Msp*I open reading frame. This is located 40 amino acids downstream from the first methionine.

Purification

Protamine sulfate was used to clear the initial cell extracts of nucleic acids and constituted the first step in the purification. Subsequent fractionation of the M.*Msp*I preparation was achieved by phosphocellulose (PC) chromatography (Figure 3) and resulted in an approximately 14-fold increase in specific activity with a recovery of more than 50% of the total activity. Further chromatography by FPLC on Mono S resulted in a further 1.6-fold increase in specific activity with only a slight loss in total activity recovered. The most prominent contaminant at this stage appeared as a low molecular protein band (Figure 4, Lane 3). The last step in the purification was achieved by FPLC on Mono Q which gave a 15% increase in specific activity with only a modest loss of total activity recovered. The final pool from this column appeared homogenous (Figure 4, Lane 4). As summarized in Table 1, a routine purification procedure starting with 5.0 g of fully induced cells yielded 4.0 mg of M.*Msp*I with a specific activity of 7.4×10^4 units/mg. Figure 4 shows the degree of purity attained at each purification step as visualized by Coomassie blue staining of an SDS-polyacrylamide gel.

Molecular Weight and Homogeneity

The molecular weight of M.*Msp*I, under denaturing conditions, was determined to be 49 kD by 10% SDS-polyacrylamide gel electrophoresis (Figure 5, panel A). Attempts to determine the molecular weight of native M.*Msp*I under non-denaturing conditions by polyacrylamide gel electrophoresis were unsuccessful, since we were unable to find conditions that would enable the native protein to migrate into the gel. Other experiments involving native gel systems such as isoelectric focusing encountered similar difficulties because of the anomalous behaviour of M.*Msp*I. These observations prompted us to examine the possibility that M.*Msp*I might form aggregates in solution. This was done by gel filtration chromatography.

M.*Msp*I, in the concentration range 0.05 to 3 mg/ml, was chromatographed on an FPLC-Gel Filtration Column, Superose 12, calibrated with several protein standards (b-Amylase, Alcohol Dehydrogenase, BSA, Carbonic Anhydrase and Cytochrome c (Figure 5, panel B). At low concentrations, from 0.05 to 0.4 mg/ml, M.*Msp*I elutes predominantly as a monomer with an apparent molecular weight of 48 to 54 kd. At higher concentrations up to 3.0 mg/ml M.*Msp*I eluted with an apparent molecular weight lying between the monomeric and dimeric forms suggesting that dimerization was occurring, but that dissociation to monomers took place during the elution. At concentrations above 3.0 mg/ml, M.*Msp*I elutes predominantly as a dimer with an apparent molecular weight of 96kd (Figure 5, panel C). Usually a shoulder was present in the void volume at 3 mg/ml that might represent an oligomeric form of M.*Msp*I (data not shown). Quantitatively, the shoulder accounts for about 5%

Table 1. Purification Summary from 5g Cells

Purification Step	Total Protein (mg)	% Recovery (Protein)	Specific activity (Units/mg Protein)	Fold Purification	Total Activity	% Recovery Activity
Crude	365	100	2.96×10^3	–	10.8×10^5	100
Protamine sulfate	300	82	2.92×10^3	–	8.8×10^5	81
Phosphocellulose	15	4	4.0×10^4	14	6.0×10^5	56
Mono S	6.1	1.7	6.4×10^4	22	4.0×10^5	37
Mono Q	4.2	1.2	7.4×10^4	25	3.1×10^5	29

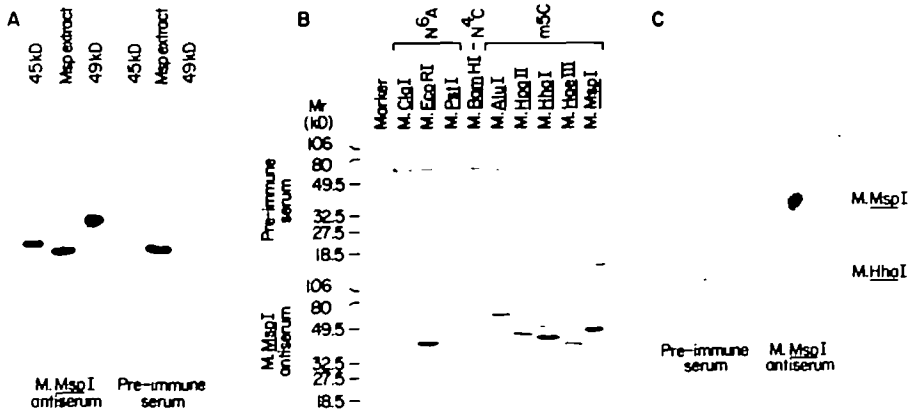

Figure 6. Immunological analysis of M.*Msp*I. **Panel A:** Immunoblot analysis of M.*Msp*I in *Moraxella* sp. A crude extract of *Moraxella* species was analyzed by Western analysis with polyclonal M.*Msp*I antibodies or preimmune serum. **Lane 1:** SDS gel-purified truncated form of M.*Msp*I (45kD). **Lane 2:** *Moraxella* extract. **Lane 3:** SDS gel-purified intact M.*Msp*I (49kD). **Panel B:** Immunological cross reactivity of anti-M.*Msp*I antibodies. Commercial preparations of methyltransferases, as indicated, were fractionated by SDS gel electrophoresis and blotted onto a nitrocellulose membrane. The proteins were reacted with either pre-immune serum (upper panel) or polyclonal M.*Msp*I antibodies (lower panel). **Panel C:** Dot blot analysis of native forms of M.*Msp*I and M.*Hha*I. Active methyltransferases were adsorbed onto a nitrocellulose membrane and developed as for Panel A.

of the total protein. Similar high molecular weight aggregates were also observed by glycerol gradient centrifugation (data not shown).

Antibodies to M.*Msp*I

To prepare polyclonal antibodies against M.*Msp*I we partially purified inclusion bodies and subjected the resulting protein to SDS gel electrophoresis. The 45 and 49 kD bands were excised and used separately to inoculate rabbits. The resulting antisera were tested for their ability to recognize authentic M.*Msp*I and the results for the antiserum raised against the 49 kD protein are shown in Figure 6, panel A. It can be seen that the antiserum reacts with both the full length and the truncated methyltransferase produced in *E. coli*. Similar results were obtained with the antiserum raised against the 45 kD protein and so all further experiments were conducted using the antiserum raised against the 49 kD protein. To determine whether both forms are also produced by the original strain, we examined a cell extract from *Moraxella* sp. by Western blotting with the anti-M.*Msp*I antibodies. This experiment demonstrated that in *Moraxella* species, only the higher molecular weight form is present (Figure 6, panel A). This suggests that the formation of the truncated protein is an artifact of the overexpression of the gene in *E. coli*.

We also checked whether the M.*Msp*I antibodies showed cross-reactivity with other m5C-methylases. For this experiment we used commercial preparations of several methyltransferases, which were extensively purified but were not homogeneous. As shown in Figure 6, panel B, cross-reactivity was seen against proteins of the correct molecular weight in commercial preparations of the M.*Hpa*II, M.*Hha*I, M.*Hae*III and M.*Alu*I, all of which are known to be m5C-methylases (34). No reactivity is detected against M.*Bam*HI, an N4C-methyltransferase (34) or the adenine methyltransferases M.*Cla*I and M.*Pst*I (34). Surprisingly, the adenine methyltransferase M.*Eco*RI (34) does show strong cross-reactivity. It should be noted that the pre-immune serum also showed some cross-reactivity with M.*Eco*RI, although it was much enhanced in the immune serum (Figure 6, panel B).

The distribution of epitopes responsible for antigenicity may differ significantly for a reduced and denatured protein as compared to a properly folded, active protein molecule. Antibodies raised against denatured protein vary in their ability to cross-react with native protein (28). Therefore we tested our antibodies, raised against denatured protein, for their ability to react with native M.*Msp*I and native M.*Hha*I. A strong signal was observed against native M.*Msp*I. Interestingly, a weak, but detectable signal was observed for native M.*Hha*I (Figure 6, panel C).

DISCUSSION

We have constructed a clone that greatly overexpresses *M.Msp*I, so that it accounts for at least 10% of total cellular protein. This has been accomplished by using the strong hybrid promoter, ptac (23,24), to drive transcription in a controllable fashion on the high copy number plasmid ps3, a derivative of pUC119. Because cells expressing M.*Msp*I induce the *mcr*B restriction system (35), it was essential to use an *mcr*B$^-$ strain and *E. coli* RR1 was chosen. This strain also contained an F-factor, carrying lacI^Q, to control the basal level of transcription from the ptac promoter. However, it should be noted that even without induction M.*Msp*I was produced at levels sufficient to render the ps3 DNA completely resistant to the action of R.*Msp*I.

Following induction, M.*Msp*I is greatly overproduced and forms insoluble inclusion bodies. This is in contrast to other methyltransferases, which have been overproduced in soluble form, such as M.*Hha*I (36), M.*Hha*II (37), M.*Eco*RI (38) and M.*Eco*P15I (39). The M.*Msp*I inclusion bodies, as seen under the electron microscope, are circular dense intracellular deposits capable of distorting the cell shape. These inclusion bodies appear morphologically distinct from those of eukaryotic proteins overexpressed in *E. coli* (31). The M.*Msp*I aggregates are highly insoluble and were fairly resistant to detergent treatment. Another interesting anomaly was also observed during M.*Msp*I overexpression. We routinely detected a lower molecular weight band that had an N-terminal sequence beginning at the second methionine in the M.*Msp*I coding sequence. We believe this probably represents aberrant initiation at this codon, but cannot rigorously exclude an N-terminal specific processing event.

One difficulty encountered during the purification of M.*Msp*I was its propensity to aggregate. This property undoubtedly facilitated the formation of inclusion bodies and hampered many of our early attempts at purification. For instance, dialysis of crude or partially purified enzyme preparations caused aggregation and precipitation of M.*Msp*I. For this reason, we devised a procedure that avoided dialysis. This had the bonus of reducing purification time so that the final procedure could be accomplished in 12 to 15 hours from cells to homogeneous protein. Using the protocol described we were able to recover as much as 4% of total cellular protein as soluble and active methyltransferase. This estimate is based on the 25-fold purification necessary to get pure enzyme and represents a lower estimate of the total M.*Msp*I present in fully-induced cells. The scheme described above has also been used successfully to purify M.*Hpa*II to homogeneity from a similar overexpression system (Dubey, unpublished results) and may have general utility.

An interesting finding from this study is the immunological cross-reactivity of the M.*Msp*I antibodies with m5C-methylases from other restriction-modification systems. This may reflect the fundamental structural similarities that have been found based on sequence comparison (20,21). However, it should be noted that antibodies raised against the human DNA methyltransferase do not cross react with bacterial methyltransferases such as *Eco*RI and *Hpa*II (40). This is despite the fact that the C-terminal third of the human enzyme is very similar to a complete bacterial m5C-methylase. One possible reason for this difference could be that the antibodies to the human enzyme were raised against native protein (41), while ours were raised against denatured protein. The extensive N-terminal region, unique to the human enzyme, might be the principal antigen either by its nature or as a result of folding to mask the C-terminal region. In contrast, our antibody shows strong reactivity with native M.*Msp*I and even weak cross-reactivity with native M.*Hha*I. These antibodies might prove valuable in rapidly identifying bacterial strains that carry m5C-methylases.

Most characterized DNA methyltransferases are functional as monomers (42), although recently it has been reported that M.*Rsr*I shows partial dimerization in solution (43) and two methyltransferases from *Diplococcus pneumoniae* appear to exist as dimers (44). In both previous cases the effects could reflect a tendency of the proteins to aggregate rather than having functional significance. From our work it is clear that M.*Msp*I

is functional as a monomer, but forms dimers and higher aggregates as the enzyme concentration increases. This might have been helpful for crystallization if the aggregation had proceeded in an orderly manner, but so far it has proven impossible to obtain crystals of M.*Msp*I either alone or complexed with various DNA substrates (X. Cheng, unpublished observations).

ACKNOWLEDGEMENTS

The authors thank Adrian Krainer, Jim Pflugrath, Jeff Kuret and Dan Marshak for helpful discussions and much useful advice. It is also a pleasure to acknowledge the extraordinary efforts of Xiaodong Cheng, who undertook the crystallization trials. This work was supported by grants from the NIH (GM 46127) and NSF (DMB-8917650). B.M. was the recipient of a grant from the Swiss National Science Foundation (No. 83.391.086). Centralized facility support for protein chemistry and electron microscopy was from a Cancer Center Support grant from the NCI (CA 45508). We thank David Spector for the electron micrographs and help in interpreting them, and Dan Marshak for the protein sequence determination.

REFERENCES

1. Lin, P.M., Lee, C.H and Roberts, R.J. (1989) *Nucl. Acids Res.* **17**, 3001–3011.
2. Razin, A. (1989) *Genetic Eng.* **11**, 1–11.
3. Razin, A. (1984) In *DNA Methylation: Biochemistry and Biological Significance* (A. Razin, R. Cedar, A.D. Riggs, Eds.) pp 127–146, Springer-Verlag, Inc., New York, N.Y.
4. Smith, H.O. and Kelly, S.V. (1984) In *DNA Methylation, Biochemistry and Biological Significance* (A. Razin, R. Cedar, A.D. Riggs, Eds.) pp 39–71, Springer-Verlag, Inc., New York, N.Y.
5. Modrich, P. (1987) *Ann Rev. Biochem.* **56**, 435–466.
6. Vanyushin, B.F. (1984) In *Current Topics in Microbiology and Immunology* (T.A. Trautner, Ed.) Vol. 108, pp 9–114, Springer Verlag, Berlin.
7. Lewis, J. and Bird, A. (1991) *FEBS Letters* **285**, 155–159.
8. Thompson, J.P., Simkevich, C.P., Holness, M.A., Kang, A.H. and Raghow, R. (1991) *J. Biol. Chem.* **266**, 2549–2556.
9. Holliday, R. (1987) *Science* **238**, 163–170.
10. Rubin, R A. and Modrich, P. (1977) *J. Biol. Chem.* **252**, 7265–7272.
11. Everett, E., Falick, A.M. and Reich, N.O. (1990) *J. Biol. Chem.* **265**, 17713–17719.
12. Reich, N.O., Maegley, K.A., Shoemaker, D.D. and Everett, E (1991) *Biochemistry* **30**, 2940–2946.
13. Bergerat, A. and Guschlbauer, W. (1990) *Nucl. Acids Res.* **18**, 4369–4375.
14. Wu, J.C. and Santi, D.V. (1987) *J. Biol. Chem.* **262**, 4778–4786.
15. Friedman, S., Som, S. and Yang, L.F. (1991) *Nucl. Acids Res.* **19**, 5403–5408.
16. Som, S. and Friedman, S. (1991) *J. Biol. Chem.* **266**, 2937–2945.
17. Wu, J.C and Santi, D.V. (1985) In *Biochemistry and Biology of DNA Methylation*, (G.I. Cantoni and A. Razin, Eds) Alan R. Liss Inc., New York, pp 115–129.
18. Chen, L., MacMillan, A.M., Chang, W., Ezaz-Nikpay, K., Lane, W.S. and Verdine, G. (1991) *Biochemistry* **30**, 11018–11025.
19. Walder, R.Y., Langtimm, C.J., Chatterjee, R. and Walder, J.A. (1983) *J. Biol. Chem.* **258**, 1235–1241.
20. Lauster, R., Trautner, T.A. and Noyer-Weidner, M. (1989) *J. Mol. Biol.* **206**, 305–312.
21. Posfai, J., Bhagwat, A.S., Posfai, G. and Roberts, R.J. (1989) *Nucl. Acids Res.* **17**, 2421–2435.
22. Bolivar, F., Rodriguez, R.L., Greene, P.J., Betlach, M.C., Heynecker, H.L., Boyer, H.W., Crosa, J.H. and Falkow, S. (1977) *Gene* **2**, 95–113.
23. Amann, E., Brosius, J. and Ptashne, M. (1983) *Gene* **25**, 167–178.
24. deBoer, H.A., Comstock, L.J. and Vasser, M. (1983) *Proc. Natl. Acad. Sci. USA* **80**, 21–25.
25. Howe, M.M. (1973) *Virology* **54**, 93–101.
26. Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
27. Laemmli, U.K. (1970) *Nature* **227**, 680–685.
28. Harlow, E. and Lane, D. (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
29. Siegel, L.M. and Monty, K.J. (1966) *Biochim. Biophys. Acta* **112**, 346–362.
30. Goldschmidt, R. (1970) *Nature* **288**, 1151–1156.
31. Williams, D.C., Van Frank, R.M., Muth, W.L. and Burnett, J.P. (1982) *Science* **215**, 687–688.
32. Schumacher, G., Sizmann, D., Haug, H. and Bock, A. (1986) *Nucl. Acids Res.* **14**, 5713–5727.
33. Schein, C.H. (1989) *Biotechnology* **7**, 1141–1148.
34. Roberts, R.J. and Macelis, D. (1991) *Nucl. Acids Res.* **19**, 2077–2019.
35. Raleigh, E.A. and Wilson, G.G. (1986) *Proc. Natl. Acad. Sci. USA* **83**, 9070–9074.
36. Wu, J.C. and Santi, D.V. (1988) *Nucl. Acids Res.* **16**, 703–717.
37. Chandrasegaran, S., Wu, L.P., Valda, E. and Smith, H.O. (1988) *Gene* **74**, 15–21.
38. Cheng, S.C., Kim, R., King, K., Kim, S.H. and Modrich, P. (1984) *J. Biol. Chem.* **259**, 11571–11575.
39. Rao, D.N., Page, M.G.P. and Bickle, T.A. (1989) *J. Mol. Biol.* **209**, 599–606.
40. Pfeifer, G.P., Kohlmaier, L., Tomassetti, A., Schleicher, R., Follmann, H., Pfohl-Leszkowicz, A., Dirheimer, G. and Drahovsky, D. (1989) *Arch. Biochem. Biophy.* **268**, 388–392.
41. Kaul, S., Pfeifer, G.P. and Drahovsky, D. (1984) *Eur. J. Cell Biol.* **34**, 330–335.
42 Modrich, P. and Roberts, R.J. (1982) In *Nucleases* (ed. S.A. Linn and R.J. Roberts), p. 109–154. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
43. Kaszubska, W., Aiken, C., O'Connor, C.D. and Gumport, R.I. (1989) *Nucl. Acids Res.* **17**, 10403–10425.
44. de la Campa, A.G., Kale, P., Springhorn, S.S. and Lacks, S.A. (1987) *J. Mol. Biol.* **196**, 457–469.