DNA binding properties in vivo and target recognition domain sequence alignment analyses of wild-type and mutant Rssl [N6-adenine] DNA methyltransferases

Sandra S. Szegedi and Richard I. Gumport*

Department of Biochemistry and College of Medicine, 600 South Mathews Avenue, University of Illinois at Urbana-Champaign, Urbana, IL 61801, USA

Received June 8, 2000; Revised July 10, 2000; Accepted August 4, 2000

ABSTRACT

A genetic selection method, the P22 challenge-phage assay, was used to characterize DNA binding in vivo by the prokaryotic β class [N6-adenine] DNA methyltransferase M·RsrI. M·RsrI mutants with altered binding affinities in vivo were isolated. Unlike the wild-type enzyme, a catalytically compromised mutant, M·RsrI (L72P), demonstrated site-specific DNA binding in vivo. The L72P mutation is located near the highly conserved catalytic motif IV, DPPY (residues 65–68). A double mutant, M·RsrI(L72P/D173A), showed less binding in vivo than did M·RsrI(L72P). Thus, introduction of the D173A mutation deleteriously affected DNA binding. D173 is located in the putative target recognition domain (TRD) of the enzyme. Sequence alignment analyses of several β class MTases revealed a TRD sequence element that contains the D173 residue. Phylogenetic analysis suggested that divergence in the amino acid sequences of these methyltransferases correlated with differences in their DNA target recognition sequences. Furthermore, MTases of other classes (α and γ) having the same DNA recognition sequence as the β class MTases share related regions of amino acid sequences in their TRDs.

INTRODUCTION

Prokaryotic type II DNA methyltransferase (MTase) enzymes provide an attractive opportunity to study DNA–protein interactions. Rssl [N6-adenine] DNA methyltransferase (M·Rssl) from Rhodobacter sphaeroides recognizes the duplex DNA sequence GAATTC and catalyzes the transfer of a methyl group from S-adenosylhomocysteine (AdoHcy) (1). We used the highly sensitive Salmonella bacteriophage P22 challenge-phage assay (2,3), which monitors cell viability as a function of site-specific DNA binding, to study M·Rssl–DNA interactions in vivo. The challenge-phage system assays the ability of wild-type and mutant M·Rssl enzymes to serve as pseudo-repressors of transcription of the reporter gene ant encoded on an engineered, infective P22 bacteriophage genome. An EcoRI pseudo-operator site is placed downstream (3’) of the promoter region of the ant gene. Successful repression of ant expression by an M·Rssl protein bound to the EcoRI pseudo-operator site prevents production of the Ant protein, which if expressed leads ultimately to a phage lytic cycle and death of the Salmonella typhimurium host. Thus, successful, site-specific DNA binding by M·Rssl represses ant transcription and promotes the establishment of lysogeny, which is quantified as a kanamycin-resistant colony.

We isolated a weakly catalytically active mutant, M·Rssl (L72P), that, unlike wild-type M·Rssl, bound in vivo to the EcoRI pseudo-operator site with enough affinity to give a robust DNA binding response in the challenge-phage assay. L72 is located in a loop (4) following catalytic motif IV, DPPY (residues 65–68). A double mutant, M·Rssl(L72P/D173A), had less site-specific DNA binding affinity compared to the L72P single mutant in vivo, suggesting that D173 is involved in target DNA recognition. The single mutant M·Rssl (D173A) retained catalytic activity comparable to the wild-type enzyme but failed to bind in vivo. In the crystal structure of M·Rssl, D173 forms an ion pair with R176 and is located in a 310-helix (4) that is part of the putative target recognition domain (TRD) of the enzyme (5).

TRD sequence alignment analyses of M·Rssl and several β class MTases revealed that D173 and R176 are moderately and highly conserved (≥70 and ≥90%, respectively). Conservation of D173 and R176 in the TRD of the β class MTases recognizing different DNA sequences suggests a role in maintaining the structural integrity of this region of the enzymes. Analogously, the ‘methylase fold’ forms the catalytic domain of all DNA MTases (6). Phylogenetic analysis of the aligned β class enzymes showed that MTases having the same DNA recognition sequence were more related than members of the same class recognizing other DNA sequences. In agreement with the phylogenetic analysis, β class MTases having the same or similar target recognition sequences displayed increased conservation of particular arrays of amino acid residues in certain regions of the TRD, apart from those conserved over the whole class.
These results imply a more direct role for these sequence arrays in DNA sequence recognition.

Furthermore, comparison of the amino acid sequences of the β class TRDs with MTases of other classes (α and γ) recognizing similar DNA sequences revealed common stretches of amino acid sequence similarity. Thus, amino acids conserved throughout the TRDs of the β class, which comprises enzymes recognizing a variety of target sequences, may be involved in forming the nucleus of a general recognition structure. Conversely, the amino acids in the TRDs that differ over the entire β class but are conserved in subgroups that recognize similar targets may be more directly involved in DNA sequence recognition.

MATERIALS AND METHODS

Materials

Unpurified or OPC-purified oligodeoxyribonucleotide primers for PCR mutagenesis were purchased from Gibco BRL. Cloned Pfas DNA polymerase was from Stratagene. All other restriction or DNA modification enzymes were from Promega or New England Biolabs. λ DNA and all restriction, modification and PCR enzymes were purchased from Gibco BRL, Promega, New England Biolabs or Stratagene.

Plasmids, bacteriophage and bacterial strains

The Escherichia coli strain DHI (F sup E44 hsdK17 recA1 endA1 gyrA96 thi-I relA1) was routinely used for transformation and expression of plasmids. Salmonella typhimurium strains MS1868 [lexA414(Am) Fels ’ hsdSB (r-m')] and MS1883 [lexA414(Am) Fels’ hsdSB (r-m’) supE] were used for challenge-phage assays and propagation of P22 bacteriophage 10-C containing a GAAATTC pseudo-operator and P22-1000 lacking this site, respectively. P22 bacteriophage 10-C and P22-1000 were a gift of J. F. Gardner (University of Illinois, Urbana-Champaign, IL) and were prepared using the protocols of Maloy et al. (3).

A derivative of pTrc99A (Amersham-Pharmacia) containing the R·RsrI (E117Q) gene, pTrc6#, was prepared by Yonghong Huan (University of Illinois, Urbana-Champaign, IL). P22 10-C bacteriophage containing methylated DNA was generated by isolating it from an MS1868 host culture containing 10 µM isopropyl-β-D-thiogalactoside (IPTG)-induced M·RsrI. Bacterio DNA from the induced cultures was prepared as described (3) and found to be completely resistant to R·RsrI and P22 bacteriophage 10-C and P22-1000 were a gift of J. F. Gardner (University of Illinois, Urbana-Champaign, IL) and were prepared using the protocols of Maloy et al. (3).

Challenge-phage assays to assess DNA binding in vivo

Colonies of S. typhimurium MS1868 containing a plasmid (a pTrc99A derivative) for the IPTG-inducible expression of wild-type (7) and mutant M·RsrI proteins and R·RsrI (E117Q) were grown overnight in 5 ml of LB and ampicillin (Ap; 100 µg/ml) cultures. An aliquot (50 µl) of these overnight cultures was inoculated into fresh broth (5 ml) of the same composition. The cultures were grown in a shaking water bath at 37 °C to an OD 650 = 0.5–0.6 as measured in 1 cm quartz cuvettes with a Hewlett Packard 8451 diode array spectrometer. Portions (500 µl) of this culture were added to several tubes containing 1.5 ml of LB and Ap (100 µg/ml) and 1, 10, 100 or 1000 µM IPTG. These subcultures were incubated with shaking at 37 °C for ~1 h to induce enzyme production. Challenge phage (10-C, methylated 10-C or P22-1000) were added to 100 µl of the induced subcultures at a multiplicity of infection (m.o.i.) of 25 and incubated at ambient temperature for 1 h. The uninfected and infected subcultures were then serially diluted 10-fold to 1 × 10–10 in LB broth. To determine the viable cell concentration, dilutions of the uninfected subcultures were spotted (5–10 µl) onto LB, agar and Ap (100 µg/ml) containing the same concentration of IPTG used for induction and grown at 37 °C for 48–72 h to allow colony formation. Dilutions of infected subcultures were plated onto LB, agar, Ap (100 µg/ml) and Kn (50 µg/ml) with the proper concentration of IPTG to determine the lysogens/ml. The absence of a lysogen was defined as one colony and the lowest frequency (~1 × 10–8–1 × 10–9) is given as the reciprocal of the viable cell count at each IPTG concentration. The frequency of lysogeny is expressed as the number of lysogens divided by the number of viable cells at each IPTG concentration and is plotted as the log of the lysogeny frequency as a function of the log of the IPTG concentration.

Random PCR mutagenesis

Forced nucleotide random PCR mutagenesis of the rsrIM gene using Taq DNA polymerase, one dNTP in excess of the other three and high MgCl 2 concentrations in the presence of MnCl 2 was used to generate a population of mutated PCR products (8). The forward and reverse PCR primers (d CCTGCAATGAAGATGATCTCA) and (d AAGGTTCGGAATTTCTATGAAAGACATCTCC) were used to amplify the rsrIM gene from template DNA. The NcoI, EcoRI and BamHI sites in each primer are underlined. Four 100 µl PCR reactions were performed, each containing 300 nM forward and reverse primer, 200 µM dNTPs with one dNTP added to a final concentration of 3.2 mM in each reaction, 1× PCR buffer supplied by the manufacturer (Gibco BRL), 9.5 mM MgCl 2, 0.5 mM MnCl 2, 5 U Taq DNA polymerase (Gibco BRL) and ~0.5 µg DNA template. A three-step PCR thermocycling protocol was used to amplify the rsrIM gene: (i) 94 °C for 5 min; (ii) 30 cycles of 94 °C for 1 min, 25 °C for 1 min and 72 °C for 1 min; (iii) 72 °C for 10 min. For each reaction, the PCR product of the correct size was confirmed by agarose gel electrophoresis. Each product was isolated using a Qiagen QIAquick PCR purification kit following the protocol of the manufacturer. The combined reactions were digested with R·NcoI and R·BamHI and concentrated by organic extraction and ethanol precipitation. After suspension in sterile water, the mutagenized PCR fragments were ligated into R·NcoI and R·BamHI digested dephosphorylated pTrc99A. To maximize the number of clones the ligation reactions were first transformed into E·coli DH1, because the transformation efficiency of ligated DNA in S. typhimurium MS1868 was at least 3-fold less efficient. Clones from the DH1 transformation plates were combined by washing the plates with 10 ml of LB and plasmid DNA was isolated using a Qiagen Qiaprep Spin Plasmid Kit following the protocol of the manufacturer. Plasmid DNA was then transformed into MS1868.

Challenge-phage assays and isolation of pTrc99A:: rsrIM (L72P/D173A)

Approximately 1000 colonies were washed from three MS1868 transformation plates by suspension in 10 ml of LB
supplemented with 50 µg/ml Ap. IPTG was added to the 10 ml cell suspension to 0.5 mM and the bacteria grown at 37°C for 2 h with shaking. A portion of this culture was diluted to \( A_{600} = 0.6 \) (~\( 5 \times 10^8 \) cells/ml) in LB supplemented with 50 µg/ml Ap and 0.5 mM IPTG. Two aliquots (500 µl) of these diluted cultures were infected with either 10-C or P22-1000 bacteriophage at an m.o.i. of 25 at ambient temperature for 1 h. Cells were collected by brief centrifugation and 400 µl of the supernatant removed. The pelleted cells were resuspended and the remaining 100 µl was plated onto LB, IPTG (0.5 mM), Ap (100 µg/ml) and Kn (40 µg/ml) plates and grown overnight at 37°C. Only two of several lysogens picked and inoculated into 5 ml of LB broth supplemented with IPTG (0.5 mM), Ap (100 µg/ml) and Kn (40 µg/ml) grew into cultures overnight at 37°C. Plasmids isolated from these two cultures were re-transformed into MS1868 for further challenge-phae-ages. Plasmids were also transformed into DH1 for restriction digestion analysis with R-\textit{NcoI} and R-\textit{BamHI} to confirm the presence of the gene (7) and for sequence analysis. Sequence analysis revealed that both plasmids contained genes encoding the double mutation (L72P/D173A).

Site-directed PCR mutagenesis to generate M-\textit{Rsr}I single mutants L72P and D173A

We separated the L72P and D173A mutants by whole plasmid site-directed PCR mutagenesis (based on the Quick Change Site-Directed Mutagenesis Kit from Stratagene). A 50 µl reaction contained 850 nM each of the OPC purified forward mutagenic primers LeutoProForward d(CCTCCTACAAA-CATCATGCGCCGGACTGGATGCAC) or AsptoAlaForward d(GTATTTCTTCGATCTTGCAGCTGTGCGGACC-TATG) and their reverse complements, 50 ng template plasmid pTe99A::\textit{rsmLM}, \textit{Pfu} DNA polymerase (Stratagene) and 200 µM dNTPs. A three-step PCR thermocycling protocol was used: (i) 95°C for 3 min; (ii) 18 cycles of 95°C for 30 s, 55°C for 1 min and 68°C for 12 min; (iii) 68°C for 20 min. Subsequently, 1 µl (10 U) of \textit{R-DpnI} was added directly to the PCR reaction and allowed to incubate at 37°C for 1 h to digest any remaining parental DNA. Portions of each PCR reaction were then electroporated into DH1 and subsequently the plasmid DNA from several clones was isolated for restriction analysis. A single R-\textit{NgoIV} site was introduced into the plasmid as a result of incorporation of the proline mutation and a third R-P\textit{vuII} site was introduced by the Ala173 mutation. The sequences of clones showing the correct restriction digestion patterns were confirmed.

Preparation of cellular extracts of IPTG-induced MS1868 cells containing wild-type or mutant M-\textit{Rsr}I for SDS–PAGE analysis

A portion (200 µl) of an overnight culture of MS1868 harboring a plasmid for the IPTG-inducible expression of wild-type or mutant M-\textit{Rsr}I proteins was added to 100 ml aliquots of LB broth supplemented with 200 µg/ml Ap and subsequently grown at 37°C to an optical density of 0.15–0.2 at 650 nm and induced by the addition of IPTG to final concentrations of 1, 10, 100 or 1000 µM. Cells were grown to an OD of 0.4–0.6 at 650 nm, which reflects the OD after approximately 1 h of IPTG induction of M-\textit{Rsr}I proteins, before addition of challenge-phaeages. Subsequently, cells were harvested by centrifugation for 10 min at 4°C. The wet cell paste was resuspended in 500 µl of sonic buffer (20 mM Tris–HCl, pH 8.0, 10% glycerol, 20 mM β-mercaptoethanol, 1 mM EDTA, pH 8.0, 50 mM KCl, 0.4 mM phenylmethylsulfonyl fluoride) and sonicated using a Heat Systems Ultrasonic Processor W-385 microtip at 25% amplitude for 4 times for 30 s interspersed with 1 min incubations on ice to prevent overheating. The resultant sonicates were centrifuged immediately at 10 000 g at 4°C for 50 min. The clarified extracts were collected, the protein quantified using the Bradford assay (9) and the solutions frozen on dry ice and stored at ~8°C. Equivalent amounts (1–10 µg) of IPTG-induced MS1868 extracts were subsequently analyzed for M-\textit{Rsr}I induction using either 20% homogeneous Phastsystem minigels using the Phastsystem (Amersham Pharmacia Biotech) or 12.5% pre-cast gels from Bio-Rad. SDS–PAGE, Coomassie Blue staining and gel densitometric analysis was performed as described (7).

Sequence alignments and phylogenetic analysis

Searches within the non-redundant database containing 457 798 sequences were performed using PSI-BLAST (10), available at the National Center for Biological Information (NCBI). Amino acid sequences of 17 β MTases having significant similarity (PSI-BLAST scores > 40) with M-\textit{Rsr}I were aligned with Clustal W (11) without manual adjustment using BioEdit Sequence Alignment Editor v.8.4.5 (12). A phylogenetic tree created from the Clustal W alignment data was generated using TreeView 1.5.2 (13). Amino acid sequences contained in the TRDs of β MTases were also aligned with other MTase sequences using Clustal W. Chemical similarity analysis was performed using the GeneDoc program v.2.5 (developed by K. B. Nicholas at the Pittsburgh Supercomputing Center and available at http://www.cris.com/~Ketchup/genedoc.shtml ). The following classification of chemically similar amino acids was used: 1, basic (K, R and H); 2, acidic (E and D); 3, polar (N, Q, S, T and C); 4, aromatic (F, Y and W); 5, non-polar aliphatic (L, I, M, V and A); 6, proline (P); 7, glycine (G).

RESULTS

Challenge-phae assays

The challenge-phae assay measures DNA binding by an IPTG-induced, plasmid-encoded protein expressed in \textit{S. typhimurium}. Protein binding to an altered promoter region of the infecting P22 bacteriophage genome prevents expression of an anti-repressor (\textit{ant}) gene (2). The \textit{ant} protein binds and inactivates the primary \textit{c2} repressor protein, which allows expression of phage lytic genes. Upon successful binding of the test protein, \textit{ant} gene expression is repressed and the P22 phage genome instead integrates into the host chromosome by site-specific recombination to generate a lysogen. In the P22 challenge phage used in this study an M-\textit{Rsr}I (\textit{EcoRI}) binding site (GAATTC) was engineered downstream (3') of the \textit{Pant} promoter. The bacteriophage carries a kanamycin resistance (\textit{katR}) gene, so that binding of M-\textit{Rsr}I to the \textit{EcoRI} site gives rise to a colony on an agar plate containing kanamycin.

Prior to infection, growth of cells harboring the \textit{rsmLM} gene with increasing IPTG concentrations produced increasing amounts of M-\textit{Rsr}I. A challenge phage lacking the \textit{EcoRI} pseudo-operator sequence served as a negative control. The
catalytically inactive RsrI endonuclease R-RsrI (E117Q), known to show a site-specific binding response in the P22 challenge-phage assay (14), was used as a positive control. DNA binding was quantified by counting the number of KnR lysogens formed as a function of the number of viable cells at a given IPTG concentration.

Methylated EcoRI pseudo-operator site-containing P22 bacteriophage was prepared to determine whether or not M-RsrI binding in vivo was sensitive to methylation of the GAATTC binding site. Methylated bacteriophage preparations were generated by cultivating bacteriophage in a strain containing IPTG-induced M-RsrI to provide site-specific methylation. DNA from unmethylated or methylated bacteriophage preparations was then purified and subjected to digestion with R-EcoRI. DNA isolated from a bacteriophage preparation derived from an M-RsrI-containing Salmonella host was completely resistant to cleavage by R-EcoRI (data not shown). On the other hand, bacteriophage DNA isolated from preparations from cells lacking M-RsrI was completely susceptible to R-EcoRI digestion.

DNA binding analysis in vivo

Wild-type M-RsrI failed to elicit a binding response with either site-specific or control challenge phage at any of the IPTG concentrations tested. Thus, low concentrations of wild-type M-RsrI in the S. typhimurium cells did not bind with enough affinity to act as a repressor of ant expression. At higher concentrations of IPTG, production of the wild-type enzyme was toxic (Fig. 1, line D). Using PCR random mutagenesis (8), we selected a double mutant, M-RsrI (L72P/D173A), that showed IPTG-inducible, sequence-specific binding (Fig. 1, line J). Subsequently, we used site-directed mutagenesis to separate the mutants. Only the mutants that contained the L72P substitution bound DNA better than line H. At higher concentrations of IPTG the difference in binding became negligible, suggesting that the crippled endonuclease binds a methylated site when sufficiently concentrated. For M-RsrI (L72P), ~1200-fold less binding

Table 1. Frequency of lysogeny with unmethylated or methylated EcoRI site-containing P22 phage using M-RsrI variants

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>[IPTG] (µM)</th>
<th>Viable cell count (cells/ml)</th>
<th>Lysogen frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Unmethylated phage</td>
</tr>
<tr>
<td>M-RsrI (L72P)</td>
<td>1</td>
<td>6.6 (3.5) × 10⁸</td>
<td>1.8 (1.3) × 10⁻⁶</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>2.6 (3.4) × 10⁸</td>
<td>7.8 (0.4) × 10⁻⁶</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>8.4 (10) × 10⁻⁵</td>
<td>5.6 (0.4) × 10⁻¹</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>2.7 (2.3) × 10⁻¹</td>
<td>5.6 (2.8) × 10⁻¹</td>
</tr>
<tr>
<td>M-RsrI (L72P/D173A)</td>
<td>1</td>
<td>7.2 (4.5) × 10⁴</td>
<td>1.4 × 10⁻⁹</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>6.8 (3.4) × 10⁴</td>
<td>1.5 × 10⁻⁹</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>5.7 (4.3) × 10⁴</td>
<td>3.3 (1.7) × 10⁻⁹</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.4 (1.8) × 10⁴</td>
<td>1.1 (0.7) × 10⁻¹</td>
</tr>
<tr>
<td>R-RsrI (E117Q)</td>
<td>1</td>
<td>9.5 (17) × 10⁻⁷</td>
<td>3.5 (1.6) × 10⁻⁶</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>5.9 (19) × 10⁻⁷</td>
<td>1.5 (0.2) × 10⁻⁵</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>5.1 (10) × 10⁻⁷</td>
<td>4.8 (0.8) × 10⁻²</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>4.5 (2.5) × 10⁻¹</td>
<td>1.5 (0.2) × 10⁻¹</td>
</tr>
<tr>
<td>M-RsrI</td>
<td>1</td>
<td>3.3 (8.4) × 10⁴</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>3.4 (6.7) × 10⁴</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>2.5 (2.0) × 10⁴</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>1.8 (1.1) × 10⁴</td>
<td>ND</td>
</tr>
<tr>
<td>M-RsrI (D173A)</td>
<td>1</td>
<td>1.9 (6.6) × 10⁸</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>7.6 (3.2) × 10⁸</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>100</td>
<td>9.1 (7.8) × 10⁸</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>1000</td>
<td>3.2 (4.2) × 10⁴</td>
<td>ND</td>
</tr>
</tbody>
</table>

aValues in parentheses represent 1 SD from the mean of three individual experiments.

bND, no lysogenic response was detected.
The differing lysogenic frequencies of the L72P-containing mutants could result from differential protein production. To test this possibility, cultures of Salmonella MS1868 containing wild-type, L72P/D173A, L72P or D173A were prepared, induced with varying IPTG concentrations and cellular extracts prepared. The induction conditions were identical to those used for challenge-phage assays immediately before addition of phage. Equivalent amounts of total protein from each extract were subjected to SDS-PAGE, Coomassie Blue staining and densitometric analysis. Only cultures induced with 0.1 and 1.0 mM IPTG showed readily detectable overproduction of the M·RsrI proteins (Fig. 2). All the M·RsrI proteins were similarly expressed at these IPTG concentrations, exhibiting at most a 2-fold variation in amount as analyzed by densitometry. Thus, gross differences in protein concentration cannot explain the differences in the challenge phage binding responses. The 170 000-fold decrease in binding to unmethylated phage at 100 µM IPTG for the double mutant (L72P/D173A) compared to the single mutant (L72P) is therefore likely due to the D173A mutation.

**Catalytic activity of mutant RsrI MTases**

Plasmid DNA isolated from both uninduced and 1 h 1 mM IPTG-induced cultures containing the wild-type or D173A mutant proteins were completely resistant to digestion in vitro by R·EcoRI, indicating that the EcoRI sites contained at least one methyl group (data not shown). In contrast, plasmid DNA isolated from 1 mM IPTG-induced cultures of either the L72P or L72P/D173A mutants were ~75% susceptible to digestion with R·EcoRI, indicating that unmethylated EcoRI sites remained. Whereas purified wild-type M·RsrI completely protected λ DNA from EcoRI digestion in λ DNA protection assays (7) at an enzyme concentration of 17 nM, purified M·RsrI (L72P) displayed only partial protection of λ DNA substrate at the highest concentration tested, 1.1 µM (data not shown). Therefore, the M·RsrI (L72P) protein is at least 65-fold less catalytically active than the wild-type protein. λ DNA protection assays with a cellular extract containing the D173A mutant, which was refractory to purification, showed ~3- to 5-fold lower activity than extracts containing wild-type enzyme (data
not shown). Whereas the L72P/D173A double mutant has less DNA binding affinity than the single mutant L72P as measured by the challenge-phage assay, the presence of the D173A mutation alone has little effect on the catalytic activity of the enzyme in vitro. These combined results suggest that the effects of the D173A mutation are limited to direct, canonical DNA binding rather than to an indirect effect due to hindered cofactor binding.

DNA binding analysis in vitro

M·RsrⅠ(L72P) and M·RsrⅠ(L72P/D173A) have been purified to >90% homogeneity, allowing DNA binding analysis in vitro with the electrophoretic gel retardation assay used to characterize wild-type M·RsrⅠ(7). Table 2 lists the apparent DNA dissociation constants for wild-type and mutant M·RsrⅠproteins on unmethylated, hemimethylated and control DNA substrates in the presence of sinefungin, which promotes site-specific binding (7). The cofactor binding affinity of the L72P mutants has not been determined, excluding a direct comparison of the apparent DNA dissociation constants with the wild-type enzyme. However, both L72P-containing mutants displayed dissociation constants in the nanomolar range and, like the wild-type enzyme, preferred hemimethylated to unmethylated substrates. Generally, both wild-type and M·RsrⅠ(L72P) bound DNA substrates with decreasing affinities: abasic > hemimethylated > unmethylated > dimethylated > control. Thus, M·RsrⅠ(L72P) shares with the wild-type enzyme similar methylation-sensitive discrimination in canonical DNA binding. However, M·RsrⅠ(L72P) showed markedly less discrimination for dimethylated DNA relative to the wild-type enzyme. Thus, Leu72 may help mediate discrimination between substrate and product DNA. The enhanced affinity of M·RsrⅠ(L72P) for an abasic DNA substrate suggests that, like the wild-type enzyme, the mutant can flip the target adenine base.

In agreement with the challenge-phage assays, the double mutant L72P/D173A bound less well than the single mutant L72P, having 11- and 4.3-fold less affinity for unmethylated and hemimethylated target DNA substrates, respectively. The double mutant M·RsrⅠ(L72P/D173A) had less affinity for canonical DNA relative to M·RsrⅠ(L72P) in vitro, providing evidence that the presence of the D173A mutation hinders DNA binding. However, the possibility exists that DNA binding by the D173A mutant may be affected by hampered cofactor binding. This explanation seems unlikely because D173 residues in the TRD and not the cofactor binding region of the protein. Additionally, the fungiform structure of the enzyme shows D173 lying in the stalk far removed from the cofactor binding site (4). Collectively, these results support those showing decreased DNA binding in vivo by the L72P/D173A double mutant relative to the L72P single mutant as determined by challenge-phage assays.

Sequence alignments

We performed an amino acid sequence alignment of M·RsrⅠwith several β class N4mC and N6mA MTases to examine the possible conservation of amino acid residues in the regions containing the L72 and D173 mutations. Seventeen β class N6 and N4 MTases having significant similarity with M·RsrⅠas determined by a PSI-BLAST (10) search were then aligned using Clustal W sequence analysis (11). Leu72 is four amino acids C-terminal to the highly conserved catalytic motif IV, DPPY (residues 65–68). Both this Clustal W alignment and those of Malone et al. (5) and Bujnicki and Radlinska (15) reveal no preservation of an alphatic amino acid at this position. In contrast, we found that the acidic D173 residue is moderately conserved in a part of the putative TRD of β MTases identified by sequence alignments by Malone et al. (5) as a region between amino acid sequence motifs involved in catalysis (IV–VIII) or cofactor binding (I–III and X) (Fig. 3, region 2). However, other sequence alignments by Gong et al. (6) and Bujnicki and Radlinska (15) have identified region 1 (containing β-strand 7) in our alignment (Fig. 3) as part of catalytic motif VIII. Thus, the function of residues contained in region I in catalysis or DNA recognition is unclear. The alignment in Figure 3 details the conservation of chemically similar residues. The β MTase TRD is centrally located between the N-terminal catalytic segment containing motifs IV–VIII and the C-terminal cofactor binding portion containing motifs I–III and X. Two other residues are highly conserved in TRD region 2: an aromatic and a basic residue (positions 170 and 176, respectively, in M·RsrⅠ). Both D173 and R176 are contained within the 3␤-helix F1 in the structure of M·RsrⅠand form an ion pair (4). Other highly conserved residues throughout the β MTase TRD alignment correlate with particular structural features of M·RsrⅠ(Fig. 3, regions 1 and 3–5). For instance, residues in TRD regions 1 and 4 are contained within β-strands 7 and 8, respectively, in M·RsrⅠ. Sequence alignments by Bujnicki and Radlinska (15) identified a weakly conserved region in the TRD region of N4mC MTases (motif IX-N4, N/Q/D-V/I-W/N-E/D-I/V), forming a short β-strand, which corresponds to β-strand 8 (Fig. 3, region 4) of M·RsrⅠin the

Table 2. Apparent DNA dissociation constants for wild-type and mutant M·RsrⅠproteins in the presence of sinefungin

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>5 nM unmethylated (TB)</th>
<th>5 nM hemimethylated (*TB)</th>
<th>0.5 nM abasic hemimethylated (±TB)</th>
<th>5 nM dimethylated (±TB)</th>
<th>5 nM control (NS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>M·RsrⅠ(L72P)</td>
<td>33.8 (8.7)b</td>
<td>9.7 (3.0)</td>
<td>0.7 (0.1)</td>
<td>44.5 (11.7)</td>
<td>&gt;&gt;250</td>
</tr>
<tr>
<td>M·RsrⅠ(L72P/D173A)</td>
<td>374.9 (70.6)</td>
<td>41.5 (12.1)</td>
<td>ND*</td>
<td>ND</td>
<td>ND</td>
</tr>
<tr>
<td>M·RsrⅠ</td>
<td>17 (1.5)</td>
<td>6 (1.6)</td>
<td>2 (0.8)</td>
<td>156 (22.6)</td>
<td>&gt;&gt;800</td>
</tr>
</tbody>
</table>

*T, top strand; B, bottom strand; ab, abasic; m, methylated.
*Apparent dissociation constants are in nM and values in parentheses represent ± 1 SD from the mean values from three experiments. Purification, DNA sequences and gel retardation assay conditions were performed as described (8).
*ND, not done.
alignment reported here. The Clustal W alignment shows that the second and third amino acid residues of motif IX-N4 (V and W) are ≥90% conserved (Fig. 3, region 4), showing that this motif (comprising a β-strand) is present in N6mA MTases as well as N4mC MTases. Furthermore, TRD regions 3 and 5 correspond to β-turns. The retention of chemically similar amino acid residues in the β- MTase TRD alignment among N4mC and N6mA MTases recognizing different DNA sequences suggests their roles in maintaining the structure of the TRDs. The N4mC β MTase M·PvuII has been crystallized in a complex with AdoMet (6). A large portion of the TRD of this enzyme is disordered (residues 179–216), precluding a direct comparison of structural features of the TRD of M·PvuII with M·RsrI. However, amino acids comprising structural elements β-strand 7 (in region 1) and β-strand 8 (in region 4) in M·RsrI and M·PvuII align moderately well in the Clustal W alignment (Fig. 3, nos 1 and 17). Furthermore, residues 172–176 of M·PvuII comprise a 310-helix that may be structurally equivalent to the 310-helix in region 2 of M·RsrI (Fig. 3). Thus, most of the TRDs of β MTases in the Clustal W alignment may have common structural features.

The lack of conservation of some parts of the β MTase TRDs suggests their possible participation in site-specific DNA recognition. We find that β MTases sharing similar target DNA sequences have linear conservations of chemically similar amino acid residues throughout their TRDs. Analysis of a 15 amino acid residue sequence of the TRD from region 2 (outlined by a dashed black box in Fig. 3) details this observation (Fig. 4). For instance, in TRD region 2 the adenine MTases DpnII B and LlaDCHI B recognize GATC (Fig. 4, nos 6 and 7) and contain the sequence RHYYNY(D/E)LMKE(L/F)NDG. BabI, SmeI and HinFII, which recognize GANTC (Fig. 4, nos 9–11), contain (F/Y)TFNY(E/K) or α-helix F in M·RsrI (Fig. 3). Thus, the less conserved portions of TRD region 2 (the β-turn and α-helix F in M·RsrI) may share similar structural elements composed of chemically dissimilar amino acid functionalities that mediate sequence-specific DNA recognition. Other parts of the β MTase TRD are conserved only within certain subgroups in the alignment shown in Figure 3, including segments between regions 2 and 3 and 4 and 5. The guide tree generated as a result of Clustal W analysis of the 17 β class MTases showed that MTases sharing the same DNA recognition sequence were more closely related than enzymes recognizing less similar targets (Fig. 5). Phylogenetic analyses of α class MTases (16) and N4mC MTases (15) have also revealed a similar correlation between DNA

Figure 3. Relationships of DNA TRD amino acid sequences of several β class N4mC and N6mA MTases. The position of the conserved acidic residue D173 in TRD region 2 of M·RsrI is indicated by the red triangle. The target A or C base in the target recognition sequence is indicated in bold. The methylatable base of M·Pac25I (residue 15) has not been identified. Chemically similar amino acid residues (see Materials and Methods) having a high degree of conservation (≥90%) are red; moderately conserved amino acid residues (≥70%); less conserved amino acid residues are green (≥50%). The secondary structure in which each amino acid of the TRD of M·RsrI participates is indicated: B, β-strand; H, helix; t, β-turn; ( ), loop; ( ), gaps generated as a result of alignment. The names of some structural elements are indicated along the top in blue (5). A segment of region 2, outlined within a stippled box, is expanded in Figure 4.

Figure 4. Amino acid sequence segment from β MTase TRD region 2 (Fig. 3). The classification scheme is the same as in Figure 3.
recognition sequence specificities and amino acid sequence similarities in their TRDs.

Portions of amino acid sequences from various β MTase TRDs are also present in the TRDs of MTases of other classes having similar or identical DNA recognition specificities. For instance, similarity exists between the GATC- and GANTC-recognizing β MTase and the α and γ class, GATC- and GANTC-recognizing MTases (Fig. 6, A–C.). The TRDs are located in the central and C-terminal regions of α and γ enzymes, respectively. Alignment analyses of α MTases recognizing GATC and GANTC by several groups (16–18) have identified the segments of residues shown in Figure 6A and C as part of the TRD of these enzymes. The C5 MTase M·LlaR2I, which methylates the cytosine in the DNA sequence GATC, shares a linear array of chemically similar amino acids in its TRD with the β MTase M·LlaCHIB, which methylates an adenine (Fig. 6B). The GAATTC-recognizing TRD of the β MTase M·RsrI is also evident in the TRD of the γ MTases M·EcoRI and M·Van91II, which, like M·RsrI, recognize the DNA sequence GAATTC (Fig. 6D). Similarity also occurs in the M·RsrI TRD motif and the unclassified MTase XmrI, which recognizes the interrupted EcoRI DNA sequence GAA(N4)TTC (Fig. 6E). The conserved residues present in these alignments suggest that a specific linear arrangement of chemically similar functionalities contribute to forming a structure suitable for site-specific DNA recognition of similar sequences.

**DISCUSSION**

**Failure of wild-type M·RsrI to bind DNA in vivo**

Wild-type M·RsrI and M·RsrI (D173A) failed to bind DNA site specifically in the challenge-phage assay at lower concentrations of IPTG; higher concentrations of the wild-type enzyme killed the cells. The absence of lysogens at low concentrations is presumably because the enzymes methylate the target recognition sequence and dissociate. In vitro DNA binding studies using gel retardation assays show that M·RsrI bound weakly to fully methylated GAATTC ($K_D = 117$ nM; 7).

**Figure 6.** Clustal W alignments of partial amino acid sequences of the β class TRD with other MTases. (A) Comparison of the β class GATC motif with GATC-recognizing α class MTases. Chemically similar amino acid residues (see Materials and Methods) having a high degree of conservation (≥80%) are red; moderately conserved amino acid residues are blue (≥60%). (B) Comparison of the β class LlaDCHIB GATC motif with GATC-recognizing C5 MTase LlaR2I. (C) Comparison of the β class HincFI GANTC motif with α class CviBI. (D) Comparison of M·RsrI with M·EcoRI and M·Van91II. (E) Comparison of M·RsrI TRD with M·XmrI.
Higher levels of wild-type M·RsrI proteins could be toxic (Fig. 1, line D) because of ectopic methylation at functionally critical sites, as observed with M·EcoRl (19), or because of increased occupancy of important sites on the host genome. The L72P mutant also displays decreased cell viability, but requires higher concentrations of IPTG induction than the wild-type enzyme (Fig. 1, lines D and E). M·RsrI (L72P) is catalytically compromised, suggesting that excessive methylation may contribute to the decreased viability of wild-type enzyme (Fig. 1, lines D and E). M·RsrI (D173A) production was not toxic (Fig. 1, line A), although it displays activity in vitro comparable to the wild-type enzyme. Thus, the D173A mutation appears to disrupt DNA binding enough to decrease toxicity. M·RsrI (D173A) may also act like M·EcoRl (H235N), a bending-deficient mutant, which displays enhanced discrimination against non-canonical DNA relative to the wild-type enzyme (20).

DNA binding by M·RsrI mutants

M·RsrI (L72P) bound site-specifically in vivo. The binding is not likely due to grossly differential levels of protein expression. The L72P mutation lies four amino acids C-terminal to the highly conserved catalytic motif, DPPY (residues 65–68), that is found in all amino M·Tsases. The crystal structure of DpnM in complex with AdoMet (16) as well as mutational studies with EcoDam (21) and T4 Dam (22) support the role of motif IV in both cofactor binding and catalysis. Studies with two DPPY mutants of M·EcoRV, D193N and D193A, showed that both lost catalytic activity but retained some degree of AdoMet and DNA binding (23). The residual catalytic activity of purified M·RsrI (L72P) in λ DNA protection assays demonstrates that the enzyme can bind both AdoMet and DNA functionally.

Both the DPPY motif and the L72P mutation lie in a loop in the structure of M·RsrI (4). The introduction of a Pro at position 72 in the loop may affect cofactor binding or catalytic activity without impairing site-specific DNA recognition. The mutant remains sensitive to methylation of the recognition site, suggesting that the mutation does not grossly perturb the protein–DNA recognition interface. Since the mutation is located near the active site of M·RsrI, its effects may include (i) interfering with base extrusion, (ii) aligning AdoMet and the flipped target base improperly so as to impair methyl transfer or (iii) trapping the target base in the enzyme to slow product release. The L72P mutation may result in the accumulation of a DNA-bound enzyme intermediate that is normally only transiently present, allowing the detection of site-specific DNA binding in vivo.

For DNA M·Tsases DNA binding is tightly coupled to catalysis in a process likely involving conformational changes such as base extrusion and DNA bending. Solution studies on M·HhaI using electrophoretic gel retardation and 1H-NMR spectral analysis show that the enzyme can exist in a ‘closed’ state, with the target cytosine locked in the active site of the enzyme (24). Structural studies on the ternary complex of M·HhaI–AdoHey–canonical DNA show that an active site pocket is formed only after the target base is extruded from the DNA. The formation of the pocket around the base results from large structural changes in the catalytic loop (25). The introduction of a Pro at position 72 in M·RsrI may reduce the conformational flexibility of the catalytic loop, resulting in altered ability to flip the target base into or release it from the active site of the enzyme.

Studies of DNA binding both in vivo and in vitro and sequence alignment analysis suggest that Asp173 is likely a part of the DNA TRD of the enzyme. D173 is located in the TRD of β class enzymes identified by Malone et al. (5). It is evident from Figure 1 that the double mutant M·RsrI (L72P/D173A) binds DNA less well than does the single mutant M·RsrI (L72P). Although weakened in binding, it retains the ability to discriminate between unmethylated and methylated GAATTC sites. Gel retardation assays in vitro also confirm that M·RsrI (L72P/D173A) binds target DNA substrates less well than M·RsrI (L72P). Furthermore, the single mutant D173A retains catalytic activity comparable to the wild-type enzyme in vivo and has ~3-fold less activity in vitro. The relatively slight impairment of the catalytic activity suggests that the mutation must subtly affect the gross structure of the enzyme. Challenge-phage assays using the L72P mutant and selecting for weaker binding may identify other amino acids functionally involved in site-specific DNA recognition.

Structure-based alignment analysis

Clustal W alignment analysis of several β class M·Tsases revealed that D173 resides within a cluster of conserved amino acids within the putative TRD (Fig. 3). The TRD, based on previous alignments (5,6,15), is defined as the region between motifs involved in cofactor binding and catalysis. These residues may participate directly in DNA recognition through interactions with base pairs or indirectly by supporting a structural framework of amino acid residues that themselves contact the bases. The crystal structure of M·RsrI reveals that D173 and R176 interact to form an ion pair (4). Disruption of this interaction may destabilize the helix in which it participates, thereby affecting DNA binding.

Sequence comparisons with other M·Tsases

Clustal W sequence analysis revealed sequence similarities between the β M·Tsase TRD and those of other M·Tsases sharing the same or similar DNA recognition sequences (Fig. 6). Whereas conserved residues within all the β M·Tsase TRD motifs imply a function in maintenance of the structure of the TRD, shared residues between different classes of M·Tsases recognizing the same DNA sequence suggest a possible role in site-specific DNA recognition.

Four amino acids within the α M·Tsase TRD sequence motif of enzymes recognizing GATC are missing in the β M·Tsases (Fig. 6A). This missing stretch corresponds to residues 131–134 of M·DpnM, which has been crystallized in complex with AdoMet (16). The TRD of M·DpnM forms a highly structured loop. In the loop, residues 131–134 form part of a β-turn. Thus, this part of M·DpnM may serve to stabilize the structural framework of the TRD of α M·Tsases. Residues in the TRD of other classes of M·Tsases showing a lack of conservation when aligned with β M·Tsase TRD motif sequences may serve to preserve a structural element not found in the β M·Tsase TRD or contribute additional amino acid residues to mediate site-specific DNA recognition.

The Clustal W alignment in Figure 3 suggests that the DNA recognition specificity of β M·Tsases might be changed using site-directed mutagenesis. For example, the less conserved amino acid sequences in GANTC-recognizing β M·Tsases (Fig. 3, nos 5–8) in regions 2 or 4 could be mutated to residues contained in GANTC-recognizing sequences (Fig. 3, nos 9–11)
and the mutant enzymes tested to see if the DNA specificity has been changed from GATC to GANTC. M-HpyI, a CATG-recognition enzyme (Fig. 3, no. 8), might similarly be converted to recognize GATC by conservative changes in the TRD using M-LaDCHIB (Fig. 3, no. 7) as a guide for mutagenesis.

Conclusions

We isolated a mutant, M-RsrI (L72P), that binds DNA in vivo site-specifically and in a methylation-sensitive manner as assessed by the challenge-phage assay. The introduction of a Pro at position 72 may deleteriously affect cofactor binding or alter the geometry of the active site during catalysis, leading to the observed impairment in catalytic activity. The mutation may also alter the flexibility of the loop in which it resides, thereby altering conformational changes associated with catalysis, such as DNA bending or base extrusion. The L72P mutation is near the highly conserved catalytic motif, DPPY, proposed to directly interact with the expelled target base. This mutant provides a starting point for an analysis in vivo of other amino acids involved in DNA recognition, as evidenced by the isolation of a double mutant M-RsrI (L72P/D173A). Addition of the D173A mutation weakened site-specific DNA binding in the challenge-phage assay, suggesting that D173, which is located in the target recognition domain, is involved in DNA binding. Residues that are conserved within the TRD of several β MTases are probably involved in maintaining their structures, as well as possibly contributing to site-specific DNA recognition. Residues conserved in the TRDs of different classes of MTases sharing the same DNA recognition sequence specificity may have a more direct role in site-specific DNA recognition.

ACKNOWLEDGEMENTS

The authors wish to thank Dr Arvydas Janulaitis for supplying a clone of the Van911I R-M system and Dr Mei Ge for sequencing of the M-Van911I gene. We are indebted to Dr Yonghong Huan for preparation of the R-RsrI (E117Q)-expressing-plasmid and Ms Nicole Collins for purification of M-RsrI (L72P). We also thank Drs Mair Churchill, Bernard Connolly, Sergei Degtyarev, Jeff Gardner, John Gerlt, Saulius Klimasauskas, Norbert Reich and Shuang-yong Xu for their helpful comments and discussions during preparation of the manuscript. This work was supported, in part, by NIH grant GM25621 to R.I.G., the Department of Biochemistry and a grant from the University of Illinois Research Board.

REFERENCES