

The RNA moiety of chick embryo 5-methylcytosine-DNA glycosylase targets DNA demethylation

Jean-Pierre Jost*, Marc Frémont, Michel Siegmann and Jan Hofsteenge

Friedrich Miescher Institute, PO Box 2543, CH-4002 Basel, Switzerland

Received August 5, 1997; Revised and Accepted September 26, 1997

ABSTRACT

We have previously shown that DNA demethylation by chick embryo 5-methylcytosine (5-MeC)-DNA glycosylase needs both protein and RNA. RNA from enzyme purified by SDS-PAGE was isolated and cloned. The clones have an insert ranging from 240 to 670 bp and contained on average one CpG per 14 bases. All six clones tested had different sequences and did not have any sequence homology with any other known RNA. RNase-inactivated 5-MeC-DNA glycosylase regained enzyme activity when incubated with recombinant RNA. However, when recombinant RNA was incubated with the DNA substrate alone there was no demethylation activity. Short sequences complementary to the labeled DNA substrate are present in the recombinant RNA. Small synthetic oligoribonucleotides (11 bases long) complementary to the region of methylated CpGs of the hemimethylated double-stranded DNA substrate restore the activity of the RNase-inactivated 5-MeC-DNA glycosylase. The corresponding oligodeoxyribonucleotide or the oligoribonucleotide complementary to the non-methylated strand of the same DNA substrate are inactive when incubated in the complementation test. A minimum of 4 bases complementary to the CpG target sequence are necessary for reactivation of RNase-treated 5-MeC-DNA glycosylase. Complementation with double-stranded oligoribonucleotides does not restore 5-MeC-DNA glycosylase activity. An excess of targeting oligoribonucleotides cannot change the preferential substrate specificity of the enzyme for hemimethylated double-stranded DNA.

INTRODUCTION

Recent published work suggests that the active DNA demethylation reaction may involve RNA alone (1) or RNA in combination with protein(s) (2). In the first case it is thought that the catalytic function of the RNA is to remove methylcytosine and to transfer it to the RNA. The reaction is apparently resistant to proteinase K and only sensitive to RNases (1). In the second case the reaction is carried out by a 5-methylcytosine (5-MeC)-DNA glycosylase which is associated with a RNA (2). Both protein and RNA are necessary for the reaction since the activity of the enzyme is

abolished by proteinase K or RNase treatment (2). The involvement of both protein and RNA in the catalysis of enzymic reactions is well documented. Such an association is exemplified by the ribosome particles and by the aminoacyl-tRNAs with the translation initiation factor(s) (3). Group II intron RNAs associate with a DNA endonuclease and cleave one strand of the DNA duplex while the protein associated with the DNA cleaves the other strand in a site-specific manner (4). RNase P is also a complex between an RNA and a protein. In this particular case the RNA alone is capable of processing precursor tRNAs in the presence of high concentrations of Mg^{2+} , whereas at low Mg^{2+} concentrations it requires the presence of the protein (5). In the case of telomerase the RNA does not have any catalytic function, but rather serves as a primer for the telomerase reaction (6,7). More recently it has been shown that methylation of the ribose moiety of rRNA also needs short sequences of antisense snRNAs for targeting of the methylation reaction. In this case 10–20 bases complementary to the rRNA associate with the enzyme protein (8–10). Circumstantial evidence suggests that DNA methyltransferases may also require RNA for *de novo* DNA methylation (11,12). Similarly, we show here that the role of cloned RNA prepared from purified 5-MeC-DNA glycosylase is to guide the enzyme to the demethylation site.

MATERIALS AND METHODS

RNA purification

Chick embryo 5-MeC-DNA glycosylase was purified as previously published (2,13). The purified enzyme fraction was loaded onto a 10% SDS-polyacrylamide gel. The band showing 5-MeC-DNA glycosylase activity, which migrated at 52 kDa, was cut out and the enzyme extracted from the gel by incubation with 0.15 M NaCl, 20 mM HEPES, 1 mM EDTA, 0.1% SDS, 5 mM DTT for 1 h at room temperature. The sample was filtered through a Millipore Ultrafree-MC filter and precipitated with 4 vol acetone and 100 μ g Dextran-T 70 at -80°C . After centrifugation and three washes with acetone (90% v/v in water) the samples were resuspended in 100 μ l 20 mM HEPES, 1 mM EDTA, 100 mM NaCl and digested for 30 min at 37°C with 50 μ g proteinase K. The RNA was extracted with phenol:chloroform and precipitated with ethanol. An aliquot of the acetone-precipitated enzyme was denatured in guanidinium hydrochloride and slowly renatured by dialysis (13). The renatured enzyme was tested as previously described (13).

*To whom correspondence should be addressed. Tel: +41 61 697 6688; Fax: +41 61 721 4091; Email: jost@fmi.ch

RNA cloning

Two different methods were used to obtain cDNA from the purified RNA.

The first method involved ligation of a phosphorylated 24mer oligoribonucleotide (R1, 5'-GGUACCCUCGAGGAAUUCGCG-ACG-3') to the 3'-end of the RNA using T4 RNA ligase (Biolabs). The ligation was performed for 30 min at 37°C in the presence of 40 U RNase inhibitor (Boehringer) and 2.5% PEG 8000. The ligation product was used as template for cDNA synthesis with an oligodeoxyribonucleotide complementary to R1 (P1, 5'-CGCGAA-TTCCTCGAGGGTACC-3'). Reverse transcription was performed for 15 min at 70°C using 5 U rTth DNA polymerase (Perkin Elmer) in the presence of 40 U RNase inhibitor. The RNA was removed by incubation for 30 min at 60°C in 0.2 M NaOH followed by ethanol precipitation. A phosphorylated oligodeoxyribonucleotide (P2, 5'-GGTACCCTCGAGGAATTCGCGACG-3') was then ligated at the 3'-end of the cDNA using T4 RNA ligase. The ligation product was used as template for PCR amplification and the PCR product cloned using the TA Cloning Kit (Invitrogen). Individual clones were sequenced using the Sequenase kit (Amersham).

A second method involved direct cDNA synthesis with rTth DNA polymerase and random hexamers. The reaction was first incubated at 37°C for 20 min with 5 U rTth DNA polymerase, 50 ng random hexamers and 40 U RNase inhibitor and then transferred to 70°C to resolve any RNA secondary structures. After 5 min 2.5 U rTth polymerase were added and the reaction was incubated at 70°C for a further 15 min. RNA was removed as described above, then the deoxyoligonucleotides P2 and P3 (5'-CGTAGGATCCGCGGCCGCGAG-3') were ligated onto the 3'- and the 5'-ends of the cDNA respectively. The ligation product served as template for PCR and the PCR product was analyzed as above. The cloned cDNA was then sequenced using either the Sequenase kit (Amersham) or an automated sequencer (Perkin Elmer 377 DNA sequencer). Oligonucleotides were synthesized by the phosphoramidite approach using an Applied Biosystems 392 DNA/RNA synthesizer.

Total RNA isolation and dot blot hybridization

Total RNA from 12 day old chicken embryos was isolated by the guanidinium thiocyanate procedure of Chomczynski and Sacchi (14). Poly(A)-containing mRNA was isolated from total RNA by chromatography on oligo(dT)-cellulose type 7 (Pharmacia). Nuclei were purified through sucrose gradients according to Sierra (15). Nuclear RNA was isolated by digesting the nuclei in 10 mM Tris, pH 7.5, 10 mM EDTA, 10 mM NaCl, 0.5% SDS (w/v), 2% (v/v) 2-mercaptoethanol containing 1.5 mg/ml proteinase K. After 3 h incubation at 50°C DNA was sheared by several passages through a 22 gauge needle. RNA was sedimented through a cushion of 5.7 M CsCl (16,17). Dot blot hybridization of purified RNA was carried out on nylon membranes (Micron Separations Inc.). Hybridization was at 37°C for 24 h in 0.25 M sodium phosphate, pH 7.2, 1 mM EDTA and 7% (w/v) SDS. Filters were washed in 0.2× SSC, 0.02% SDS (w/v) at 37°C. The labeled probes (concentration 4×10^6 c.p.m./ml) were 5'-GCTT-ATTTTTCATTTTGGCGACTATGTGTAAAGTCGTC-3' for clone 1 (probe 1) and 5'-GGGTTGTCTAGTATCTCGTTCCGT-CACCGTGATTGCC-3' for clone 4 (probe 4).

Standard assay and complementation test of 5-MeC-DNA glycosylase

The standard assay for 5-MeC-DNA glycosylase was carried out as previously described (13). The labeled DNA substrate was a double-stranded hemimethylated oligonucleotide with a methylated lower strand (5'-TCACGGGATCAATGTGTTCTTTTCAGCTCm-CGGTACGCTGACCAGGAATACC-3'). All reaction products were analyzed on 20% polyacrylamide-urea sequencing gels. The gels were exposed for 30–60 min to X-ray films at -80°C.

Two different versions of the complementation test were used.

First version. In a total volume of 50 µl containing 20 mM EDTA, 20 mM EGTA, 20 mM HEPES, pH 7.5, 50 mM NaCl, 4 mM Pefabloc (Boehringer), 50 µg enzyme grade bovine serum albumin (BSA) and 30 µg post-heparin-Sepharose 5-MeC-DNA glycosylase fraction were consecutively assembled. Where indicated the enzyme was inactivated by incubating for 20 min at 37°C in the presence of 0.5 µg heat-treated pancreatic RNase A. At the end of the preincubation all tubes were put on ice and each sample treated with RNase A received a 1/10 vol 50 mM dithiothreitol (DTT), 1000 U RNasin (100 U/µl), 5–10 µg appropriate recombinant RNA or oligoribonucleotides and 10 ng ³²P-labeled, double-stranded hemimethylated oligonucleotide. Before onset of preincubation the positive controls received a final concentration of 5 mM DTT, 1000 U RNasin. After incubation at 37°C for 45 min all samples were diluted with 150 µl H₂O and extracted with phenol:chloroform. The supernatant fractions were then ethanol precipitated, dissolved in 95% formamide-dye, denatured for 5 min at 95°C and separated on a 20% polyacrylamide-urea sequencing gel.

Second version. In a total volume of 50 µl containing 20 mM HEPES, pH 7.5, 50 mM NaCl, 5 mM CaCl₂, 4 mM Pefabloc, 50 µg enzyme grade BSA and 30 µg post-heparin-Sepharose 5-MeC-DNA glycosylase fraction were consecutively added. Where indicated the enzyme was inactivated by preincubation at 37°C for 7 min in the presence of 50 U micrococcal nuclease. At the end of preincubation all of the tubes were chilled on ice and each sample received EDTA and EGTA at final concentrations of 20 and 25 mM respectively. In addition, each tube received a 1/10 vol 50 mM DTT, 200 U RNasin, 5–10 µg appropriate oligoribonucleotide and 10 ng ³²P-labeled double-stranded hemimethylated substrate. The positive controls received EDTA, EGTA, DTT and RNasin prior to preincubation. The experiment was continued as for the first version of the complementation test. Recombinant porcine RNase inhibitor was produced in large scale as described by Neumann *et al.* (18).

Chemicals and enzymes

Benzamidine was purchased from Fluka AG (Buchs/SG, Switzerland). Phenylmethylsulfonyl fluoride, Pefabloc, proteinase K, RNase A (DNase-free) were obtained from Boehringer Mannheim, *Micrococcus* nuclease was from Promega. Polynucleotide kinase and restriction enzymes were purchased from Biofinex (Praroman, Switzerland). [α -³²P]dATP and [γ -³²P]ATP triethylammonium (3000 Ci/mmol) were purchased from Amersham. Some oligonucleotides, DNA and RNA were synthesized by Microsynth (Balgach, Switzerland).

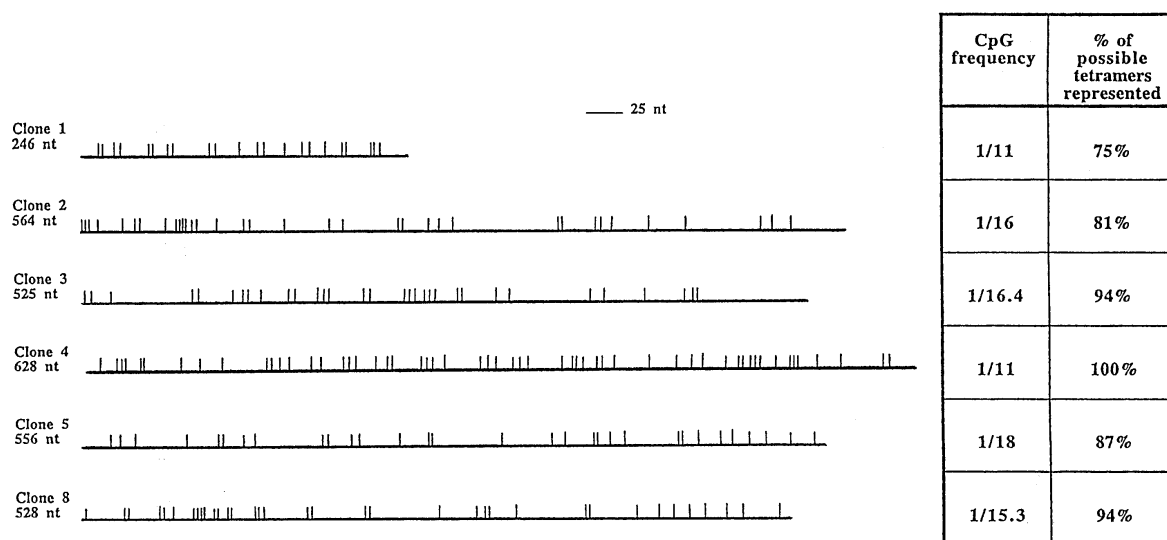


Figure 1. Map of the cloned RNA isolated from gel purified 5-MeC-DNA glycosylase. The small vertical bars indicate the position of the CpGs. The percent of possible tetramers represented in the CpG sites refers to the percentage of all possible XpCpGpY sites, where X and Y are either A, T, G or C.

RESULTS

Cloning and sequencing RNA from highly purified 5-MeC-DNA glycosylase

We have recently shown that RNA isolated from gel purified 5-MeC-DNA glycosylase could restore the activity of the enzyme after it had been inactivated with RNase A (2). It was therefore of interest to clone and characterize the RNA present in the purified enzyme. RNA was cloned as described in Materials and Methods. The clones have inserts ranging from 200 to 600 bp. The sequencing of these clones shows that they are all different, with no significant homology. Accession numbers to the EMBL nucleotide sequences library are: clone 1, Y14827; clone 2, Y14828; clone 3, Y14829; clone 4, Y14830; clone 5, Y14831; clone 8, Y14832. A computer analysis of the sequences showed no significant homology with any other known RNA. However, as shown in Figure 1, all the clones have a high density of CpGs. On average they have one CpG per 14 bases and the average ratio of CpG/GpC is 1.1. The dot blot hybridization shown in Figure 2 indicates that total RNA, nuclear RNA or poly(A)-containing mRNA prepared from 12 day old chicken embryos hybridized with probes 1 and 4 (derived from clones 1 and 4). Northern blot hybridization of probe 4 with total RNA and poly(A)-containing mRNA shows a specific signal just below the 18 S rRNA band (preliminary results, not shown).

Complementation test with a recombinant RNA and the RNase-inactivated 5-MeC-DNA glycosylase

Sense or antisense RNA encoded by clone 1 of Figure 1 was produced by using either T7 or Sp6 RNA polymerase respectively. Purified RNA was then tested with the complementation assay as described in Materials and Methods. Figure 3 (upper) shows that treatment of 5-MeC-DNA glycosylase with pancreatic RNase A (lane 2) completely abolished enzyme activity. However, if an excess of RNase inhibitor is added following preincubation of the enzyme with pancreatic RNase together with the recombinant RNA in the sense (lane 4) or antisense (lane 3) orientation, part

of the original activity can be restored. This shows that the recombinant RNA from one single clone has similar properties to the total RNA purified from the enzyme (2). The recombinant RNA obtained from clone 4 of Figure 1 gives similar results in the complementation test (data not shown). A look at the sequences of the sense and antisense RNAs of clone 1 of Figure 1 reveals that the sense RNA contains the six bases CTCCGG complementary to the target methylated site of the labeled DNA substrate, whereas the antisense RNA has only the four bases CCGG complementary to the same demethylation site and the CCGG sequence is situated in a double-stranded structure of the RNA. Figure 3 (lower) shows that recombinant RNA, as previously shown for total RNA (2) isolated from the purified enzyme, is totally inactive when incubated with labeled DNA substrate in the absence of protein. Identical results were obtained with recombinant RNA obtained from clone 4 of Figure 1 (data not shown).

Short synthetic oligoribonucleotides complementary to the target demethylation site can restore the activity of the RNase-inactivated 5-MeC-DNA glycosylase

It was shown above that short stretches of recombinant RNA were complementary to the target of DNA demethylation. In order to test the possibility that one of the functions of the RNA was targeting of the demethylation reaction, a series of oligoribonucleotides were tested in the complementation assay. The oligoribonucleotide GCUCCGGUCAC is complementary to the non-methylated CpG present in the hemimethylated DNA duplex, whereas GUGACCGGAGC is complementary to the methylated site of the same substrate. The oligoribonucleotide CUCUCUCUCUU is not complementary to the labeled DNA substrate. The results presented in Figure 4 (upper, lane 4) show clearly that only the oligoribonucleotide GUGACCGGAGC, complementary to the methylated site in the DNA duplex, is able to restore enzyme activity in the complementation test. The other ribonucleotides had no effect (lanes 3 and 5). Additional controls show that the oligoribonucleotide GUGACCGGAGC in double-stranded form is incapable of restoring 5-MeC-DNA glycosylase activity (data

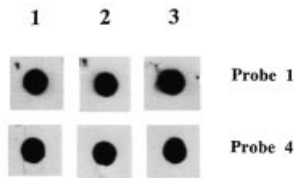


Figure 2. Dot blot hybridization of total embryonic RNA (lane 1) nuclear RNA (lane 2) and total poly(A)-containing mRNA (lane 3) from 12 day old chicken embryos. RNA (3 µg/spot) was denatured and spotted onto nylon membranes and hybridized with probes derived from clones 1 and 4 as outlined in Materials and Methods.

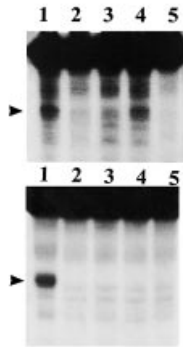


Figure 3. (Upper) Complementation assay using recombinant RNA (clone 1 of Fig. 1) and RNase-inactivated 5-MeC-DNA glycosylase. The reaction product was separated on a 20% polyacrylamide-urea gel. The arrowhead shows the correct position of the cleavage product. Lane 1 is the positive control. Lane 2 is the enzyme treated with RNase A. Lane 3 is as lane 2 except that preincubation with nuclease was followed by consecutive addition of RNase inhibitor and 10 µg antisense recombinant RNA. This was followed by incubation at 37°C for 40 min. Lane 4 is as lane 3 except that incubation was continued in the presence of 10 µg sense recombinant RNA. Lane 5 is incubation of the labeled probe alone. (Lower) Incubation of recombinant RNA with labeled hemimethylated substrate in the absence of the enzyme. Lane 1, positive control with the intact enzyme; lane 2, incubated with 10 µg tRNA; lane 3, 10 µg antisense RNA; lane 4, 10 µg sense recombinant RNA; lane 5, labeled substrate incubated alone.

not shown). Since the present results rely heavily on the reliability of the complementation assay, the possibility that the results obtained were due to competition between the added oligoribonucleotides and endogenous RNA (from 5-MeC-DNA glycosylase) in the presence of an incompletely inactivated RNase had to be ruled out. Figure 4 (upper, lane 5) shows clearly that 10 µg CUCUCUCUCCU, which was not complementary to the DNA substrate, did not restore 5-MeC-DNA glycosylase activity, thus ruling out a non-specific effect of the oligoribonucleotides. In addition, labeled oligoribonucleotides incubated in the presence of 50 U *Micrococcus* nuclease, 5 mM CaCl₂ and 25 mM EGTA showed very little degradation of labeled RNA (data not shown).

Figure 4 (lower) shows that in the presence of single-stranded methylated DNA substrate and the RNase-inactivated enzyme there is no complementation of 5-MeC-DNA glycosylase with GUGACCGGAGC, indicating that full enzyme activity can only be obtained with a hemimethylated double-stranded DNA, RNA complementary to the target methylated site and the protein moiety of the enzyme.

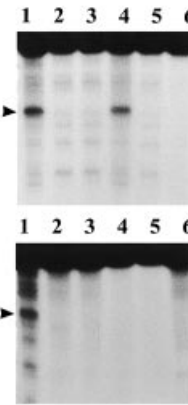


Figure 4. Complementation assay with short oligoribonucleotides (11 bases) and the *Micrococcus* nuclease-inactivated 5-MeC-DNA glycosylase. The reaction product was separated on a 20% polyacrylamide-urea gel. The arrowhead shows correct position of the cleavage product. (Upper) Lane 1 is the positive control with the intact enzyme. Lane 2 is the enzyme treated with *Micrococcus* nuclease. Lane 3 is as lane 2 except that following preincubation with the nuclease, EGTA, EDTA and oligoribonucleotides were consecutively added and incubation continued for 40 min. The oligoribonucleotide (10 µg) GCUCCGGUCAC is complementary to the unmethylated strand of the DNA duplex used as substrate. Lane 4 is as lane 3 except that incubation was continued in the presence of 10 µg GUGACCGGAGC complementary to the methylated strand of the DNA duplex. Lane 5 is as lane 3 except that incubation was continued in the presence of 10 µg CUCUCUCUCCU, which is not complementary to the labeled hemimethylated DNA substrate. Lane 6 is the labeled DNA substrate incubated alone. (Lower) Complementation assay with the same short oligoribonucleotides as in the upper panel and with single-stranded methylated DNA substrate and *Micrococcus* nuclease-inactivated 5-MeC-DNA glycosylase. Lane 1 is a positive control, as in the upper panel. In the other lanes the labeled DNA substrate was a single-stranded methylated oligodeoxyribonucleotide.

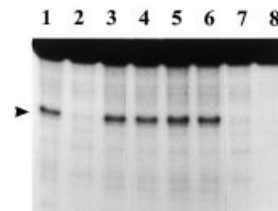


Figure 5. Number of complementary bases necessary for complementation assay of the *Micrococcus* nuclease-inactivated 5-MeC-DNA glycosylase activity. The reaction product was analyzed on a 20% polyacrylamide-urea gel. The arrow shows correct position of the cleavage product. Lane 1 is the positive control and lane 2 is the enzyme treated with *Micrococcus* nuclease. For the complementation test shown in lanes 3-7 the following oligoribonucleotides were used (bases complementary to the target sequence are underlined): lane 3, GUGACCGGAGC; lane 4, ACGACCGGAGU; lane 5, UCUACCGGAUG; lane 6, UCAUCCGGU-AU; lane 7, ACAUCCGUCUA; lane 8, substrate incubated alone.

What is the minimum number of complementary bases necessary to restore activity of the nuclease-inactivated 5-MeC-DNA glycosylase?

A series of oligoribonucleotides (11 bases long) containing different numbers of bases complementary to the target sequence of demethylation were synthesized. They were (complementary bases underlined): GUGACCGGAGC (11 bases), ACGACCGGAGU (8 bases), UCUACCGGAUA (6 bases), UCAUCCGGUAU (4 bases) and ACAUCCGUCUA (3 bases). When

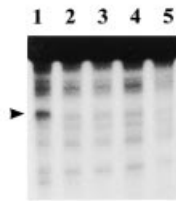


Figure 6. Complementation assay with short oligodeoxyribonucleotides (11 bases) and RNase-inactivated 5-MeC-DNA glycosylase. The reaction products were separated on a 20% polyacrylamide-urea gel. The arrow indicates the correct position of the cleavage product. Lane 1 is the positive control incubated with the intact enzyme. Lane 2 is the enzyme treated with RNase A. Lanes 3 and 4 are as lane 2 except that DTT, RNasin and the oligodeoxyribonucleotides (DNA) were added consecutively after preincubation with RNase and incubation was continued for 40 min. Lane 3 received 10 μ g GCTCCGGTCAC (complementary to the unmethylated strand) and lane 4 had 10 μ g GTGACCGGAGC (complementary to the methylated strand). Lane 5 is the labeled DNA probe incubated alone.

tested in the complementation assay the oligoribonucleotides containing between 4 and 11 bases complementary to the target sequence of the DNA substrate give a similar level of reactivation of the *Micrococcus* nuclease-inactivated 5-MeC-DNA glycosylase (care has to be taken that the other bases of the 11 base long oligoribonucleotides are not complementary to either DNA strand of the labeled substrate) (Fig. 5, lanes 3–6). However, the presence of only 3 bases complementary to the demethylation site (lane 7) is insufficient to restore enzyme activity. This means that at least 4 bases, one mCpG and two adjacent bases, are required for efficient targeting of the demethylation reaction. The total number of possible combinations of CpG flanked with A, T, G or C is 16. Figure 1 (last column) shows that the cloned RNAs have between 75 and 100% of all of the possible 16 combinations.

Specificity of RNA in the complementation and the standard 5-MeC-DNA glycosylase assays

To address the question of whether or not the oligoribonucleotides could be replaced by the same sequence of oligodeoxyribonucleotides the complementation assay was carried out using the two oligodeoxyribonucleotides GCTCCGGTCAC (complementary to the non-methylated strand) and GTGACCGGAGC (complementary to the methylated strand of the hemimethylated DNA duplex). Figure 6 (lanes 3 and 4) shows clearly that neither of the oligodeoxyribonucleotides were able to complement RNase-inactivated 5-MeC-DNA glycosylase, thus demonstrating the requirement for RNA in restoring enzyme activity.

We have shown previously that 5-MeC-DNA glycosylase cleaved hemimethylated DNA preferentially and that non-methylated or symmetrically methylated DNA sequences were very poor substrates (13). By saturating the enzyme with RNA complementary to the methylated target sequence is it possible to modify the substrate specificity? Figure 7 (left, lane 2) shows that an excess (10 μ g) of the targeting oligoribonucleotide GUGACCGGAGC does not modify activity of the intact enzyme towards hemimethylated DNA. The same experiment carried out with symmetrically or unmethylated DNA (lane 2 of the middle and right hand panels) indicates that an excess of the targeting RNA cannot modify specificity of the enzyme towards symmetrically methylated or unmethylated DNA substrates. Attempts to target a site other than CpG were unsuccessful (data not shown), thus confirming our previously published results (19).

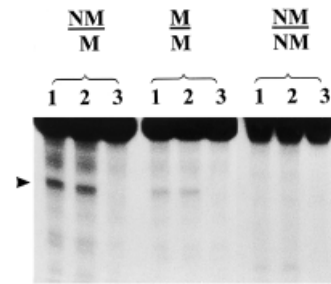


Figure 7. Effect of an excess of targeting oligoribonucleotide GUGACCGGAGC on the specificity of 5-MeC-DNA glycosylase. The reaction products were analyzed on a 20% polyacrylamide-urea gel. The arrowhead indicates the correct position of the cleavage product. NM/M stands for non-methylated DNA substrate. Lanes 1 are positive controls of the enzyme incubated with the indicated DNA substrates. Lanes 2 are reaction mixture supplemented with 10 μ g targeting oligoribonucleotide GUGACCGGAGC. Lanes 3 are the labeled substrates incubated alone.

DISCUSSION

As shown by sequencing, the RNA present in the purified 5-MeC-DNA glycosylase is highly heterogeneous and all clones tested so far have different sequences (altogether 14 different clones have been characterized). However, these clones have a common feature: they are all very rich in CpGs (Fig. 1). On average they have one CpG per 14 bases. In addition, the ratio CpG/GpC is on average 1.1 for the six clones shown in Figure 1 (bulk DNA has a ratio of 0.2). This is a strong indication that the RNA linked to 5-MeC-DNA glycosylase may be transcribed from CpG islands (20). Dot blot hybridization carried out with total RNA, nuclear RNA and poly(A)-containing mRNA from 12 day old chicken embryos shows that RNA associated with purified 5-MeC-DNA glycosylase is indeed present in the mRNA fraction. We are presently testing whether the presence of one of these CpG-rich RNA associated with 5-MeC-DNA glycosylase influences *in vivo* demethylation of its coding DNA. As we have seen in Figure 5, we need a minimum of 4 bases, including the CpG, for recognition of the demethylation site (this in the absence of any additional base complementary to the opposite DNA strand). Therefore, due to their heterogeneity, it appears that these RNAs do not represent a universal targeting sequence for DNA demethylation. Clone 4 (Fig. 1), which is 618 bases long, contained all 16 of the possible combinations of CpG flanked by A, T, C or G. This RNA by itself should be sufficient to serve as a universal targeting sequence. So why are so many different, unrelated RNA sequences required? These different RNAs tightly linked to 5-MeC-DNA glycosylase possibly represent transcripts from CpG islands which should remain unmethylated. Should one of the CpGs in a CpG island become methylated during DNA replication it could form a hemimethylated substrate. Since both DNA methyltransferase and 5-MeC-DNA glycosylase prefer hemimethylated DNA as substrate, a specific mechanism should exist to determine whether or not the hemimethylated site becomes fully methylated or demethylated. Different strategies can be envisaged. For example, additional regulatory protein(s) or RNA(s) could favor one or the other reaction and/or the molar ratio between the two enzymes at a precise time point of

replication could decide whether a given site is methylated or demethylated. It is conceivable that the RNA tightly bound to 5-MeC-DNA glycosylase may have a dual function: targeting of demethylation and inhibition of DNA methyltransferase, thus favoring the demethylation reaction. Preliminary experiments have shown that DNA methyltransferase purified from HeLa cells is strongly inhibited by recombinant RNA (Thiry, Frémont and Jost, unpublished results). The use of 4 bases as a recognition sequence could possibly also serve as a very efficient way to target demethylation sites in the non-coding region of a gene and the high density of CpGs in the RNA could increase the probability of reaction. Some observations made in different cell systems show that there is a positive correlation between presence of unmethylated CpG islands and presence of active 5-MeC-DNA glycosylase. Conversely, cells with heavily methylated CpG islands have no trace of 5-MeC-DNA glycosylase activity (Jost, unpublished results). However, the causality between the presence of unmethylated CpG islands and activity of 5-MeC-DNA glycosylase remains to be demonstrated. Keeping the CpG islands free of methylation may require both *cis*- and *trans*-acting elements. For example, it is known that Sp1 binding sites flanking CpG islands are essential to keep them methylation free (21,22). Moreover, there is evidence that proteins like NF- κ B may be involved in active demethylation of specific genes in B cells (23). One limitation of RNA as an efficient targeting molecule for *in vitro* DNA demethylation is its secondary structure. The sequences of the cloned RNAs as shown by computer analysis all showed complex secondary structures. In addition, we know from our *in vitro* experiments that double-stranded oligoribonucleotides cannot target demethylation under our experimental conditions (data not shown). This could explain why the CCGG present in a double-stranded RNA region (antisense RNA) of clone 1 was inefficient in complementing the nuclease-inactivated 5-MeC-DNA glycosylase (see Fig. 3, upper, lane 3). Therefore, one could speculate that *in vivo* such secondary structures may be selectively destabilized by proteins.

As we have shown previously, the population of RNA tightly associated with the active enzyme has no trace of catalytic activity when incubated under various conditions in the presence of the DNA substrate but in the absence of the protein moiety (2). Similarly, pure recombinant RNA in the sense orientation can only complement 5-MeC-DNA glycosylase that has been inactivated by RNase (Fig. 3) but is without catalytic activity when incubated alone with the labeled DNA substrate in the absence of protein (Fig. 3, lower). However, at this stage one cannot completely rule out that full-size RNA may have some catalytic activity. The nature of the tight association of the RNAs

with 5-MeC-DNA glycosylase and how the enzyme performs its reactions are still unknown. These questions will be addressed in the near future, when we have cloned the protein moiety of this enzyme.

ACKNOWLEDGEMENTS

We would like to thank Mrs Yan-Chim Jost for her technical assistance and for typing the manuscript. We are also grateful to Drs E.J.Oakeley and J.Paszkowski for their critical reading of this manuscript. We would also like to thank Novartis for financial support of this project.

REFERENCES

- Weiss,A. Keshet,I., Razin,A. and Cedar,H. (1996) *Cell*, **86**, 709–718.
- Frémont,M., Siegmann,M., Gaulis,S., Matthies,R., Hess,D. and Jost,J.P. (1997) *Nucleic Acids Res.*, **25**, 2375–2380.
- Nierhus,K.H., Franceschi,F., Subramanian,A.R., Erdmann,V.A. and Wittmann-Liebold,B. (eds) (1992) *The Translational Apparatus, Structure, Function, Regulation, Evolution*. Plenum Press, New York, NY, p. 746.
- Zimmerly,S., Guo,H., Eskes,R., Yang,J., Perlman,P.S. and Lambowitz,A.M. (1995) *Cell*, **83**, 529–538.
- Guerrier-Takada,C., Gardiner,K., Marsh,T., Pace,N.R. and Altman,S. (1983) *Cell*, **35**, 849–857.
- Greider,C.W. and Blackburn,E.H. (1987) *Cell*, **51**, 887–898.
- Landblad,V. and Wright,W.E. (1996) *Cell*, **87**, 369–375.
- Kiss-László,Z., Henry,Y., Bachelierie,J.P., Caizergues-Ferrer,M. and Kiss,T. (1996) *Cell*, **85**, 1077–1088.
- Cavaillé,J., Nicoloso,M. and Bachelierie,J.P. (1996) *Nature*, **383**, 732–735.
- Tollervey,D. (1996) *Science*, **273**, 1056–1057.
- Angell,S.M. and Baulcombe,D.C. (1997) *EMBO J.*, **16**, 3675–3684.
- Wassenegger,M., Heimes,S., Riedel,L. and Sängler,H.L. (1994) *Cell*, **76**, 567–576.
- Jost,J.P., Siegmann,M., Sun,L. and Leung,R. (1995) *J. Biol. Chem.*, **270**, 9734–9739.
- Chomczynski,P. and Sacchi,N. (1987) *Anal. Biochem.*, **162**, 156–159.
- Sierra,F. (1990) In Azzi,A., Polak,J.M. and Saluz,H.P. (eds), *A Laboratory Guide to In Vitro Transcription*. Birkhäuser, Basel, Switzerland, pp. 30–44.
- Glisin,V., Crkvenjakov,R. and Byus,C. (1974) *Biochemistry*, **13**, 2633–2637.
- Chirgwin,J.M., Przybyla,A.E., MacDonald,R.J. and Rutter,W.J. (1997) *Biochemistry*, **18**, 5294–5299.
- Neumann,U., Hofsteenge,J., Arkena,A.H. and Dijkstra,B.W. (1993) *J. Mol. Biol.*, **231**, 505–508.
- Jost,J.P. (1993) *Proc. Natl. Acad. Sci. USA*, **90**, 4685–4688.
- Antequera,F. and Bird,A. (1993) In Jost,J.P. and Saluz,H.P. (eds), *DNA Methylation: Molecular Biology and Biological Significance*. Birkhäuser, Basel, Switzerland, pp. 169–185.
- Macleod,D., Charlton,J., Mullins,J. and Bird,A.P. (1994) *Genes Dev.*, **8**, 2282–2292.
- Brandeis,M., Frank,D., Keshet,I., Siegfried,Z., Mendelsohn,M., Nemes,A., Temper,V., Razin,A. and Cedar,H. (1994) *Nature*, **371**, 435–438.
- Kirilov,A., Kistler,B., Mostoslavsky,R., Cedar,H., Wirth,T. and Bergman,Y. (1996) *Nature Genet.*, **13**, 435–441.