DNA Demethylase Is a Processive Enzyme*

(Received for publication, January 20, 1999, and in revised form, January 29, 1999)

Nadia Cervoni, Sanjoy Bhattacharya, and Moshe Szyf‡

From the Department of Pharmacology, McGill University, Montreal, Quebec H3G 1Y6, Canada

DNA methylation patterns are generated during development by a sequence of methylation and demethylation events. We have recently demonstrated that mammals bear a bona fide demethylase enzyme that removes methyl groups from methylated cytosines. A general genome wide demethylation occurs early in development and in differentiating cell lines. This manuscript tests the hypothesis that the demethylase enzyme is a processive enzyme. Using bisulfite mapping, this report demonstrates that demethylase is a processive enzyme and that the rate-limiting step in demethylation is the initiation of demethylation. Initiation of demethylation is determined by the properties of the sequence. Once initiated, demethylation progresses processively. We suggest that these data provide a molecular explanation for global hypomethylation.

Several lines of evidence have established that DNA methylation patterns are tissue specific and are generated during development by rounds of de novo methylation and demethylation events (1–4). Two modes of demethylation have been documented: site-specific demethylation that coincides in many instances with the onset of gene expression of specific genes (5, 6) and a general genome wide global demethylation that occurs during early development in vivo (7, 8), during cellular differentiation (9–11), and in cancer cells (12). The global demethylation is consistent with the hypothesis that a general demethylase activity that is activated at specific points in development or oncogenesis exists (13). However, until recently the identity of the enzymes responsible for demethylation has been a mystery. It has been generally assumed that an activity that can transform a methylated cytosine to a cytosine by direct removal of a methyl group does not exist. Therefore, it has been proposed previously that demethylation involves either a glycosylase and repair activity (10, 13, 14) or a nucleotide excision repair activity (15). Whereas a demethylation process that involves a sequence of glycosylase, nuclelease, repair, and ligase enzymatic activities could possibly be involved in site specific demethylation, it is hard to understand how it can catalyze genome wide global demethylation. Global nicking and repair would have a serious impact on the integrity of the genome. What enzymatic activity is responsible for global demethylation?

We have recently identified and cloned from human cancer cells a bona fide DNA demethylase that catalyzes the hydrolytic removal of methyl residues from methyl cytosine in DNA (16). DNA demethylase demethylates both fully methylated and hemimethylated DNA, shows dinucleotide specificity, and can demethylate mCpG sites in different sequence contexts (16). Since this enzyme can remove methyl groups from DNA without damaging the DNA, it is a candidate to be involved in global hypomethylation. One essential property of an enzyme that removes methylation from wide regions of the genome could be processivity.

In this paper we have used sodium bisulfite DNA mapping to determine whether purified demethylase demethylates DNA in a processive or distributive manner in an isolated system. Bisulfite treatment that is followed by PCR amplification, cloning, and sequencing of individual molecules of DNA allows one to determine the state of methylation of a single DNA molecule at a time, at single base resolution (17). This allows for the analysis of an interaction of an enzyme with a single substrate molecule. We test here whether methylated murine and bacterial gene sequences are processively or distributively demethylated by purified demethylase and whether this process is affected by the properties of the methylated sequence.

The processive nature of DNA demethylase in vitro demonstrated in this manuscript provides a model to explain the mechanism involved in genome wide hypomethylation.

MATERIALS AND METHODS

Purification of Demethylase—Nuclear extracts were prepared from A549 (ATCC: CCL 185) cultures at near confluence as described previously (18). A freshly prepared nuclear extract (1 ml, 6 mg) was diluted to a conductivity equivalent to 0.2 m NaCl in buffer L (10 mM Tris-Cl, pH 7.5, 10 mM MgCl₂) and applied onto a DEAE-Sephadex A-50 column (Amersham Pharmacia Biotech) (2.0 × 1 cm) that was pre-equilibrated with buffer L at a flow rate of 1 ml/min. Following a 15-ml wash with buffer L, proteins were eluted with 5 ml of a linear gradient of NaCl (0.2–5.0 M). 0.5-ml fractions were collected and assayed for demethylase activity. Demethylase eluted between 4.9 and 5.0 M NaCl. Active DEAE-Sephadex column fractions were pooled, adjusted to 0.2 M NaCl by dilution, and loaded onto an SP-Sepharose column (Amersham Pharmacia Biotech) (2.0 × 1 cm). Following washing of the column as described above, the proteins were eluted with 5 ml of a linear NaCl gradient (0.2–5.0 M). 0.5-ml fractions were collected and assayed for demethylase activity. Demethylase activity eluted around 3.0 M NaCl. Active fractions were pooled, adjusted to 0.2 M NaCl by dilution, and applied onto a Q-Sepharose (Amersham Pharmacia Biotech) column (2.0 × 1 cm), and proteins were eluted as described above. The demethylase activity eluted around 4.8–5.0 M NaCl. The pooled fractions of Q-Sepharose column were loaded onto a 2.0 × 2.0 cm DEAE-Sephalac column (Amersham Pharmacia Biotech) and eluted with 10 ml of buffer L. The activity was detected at fraction 4, which is very near the void volume. Demethylase is purified 500,000-fold using this protocol.

Demethylase activity was assayed by measuring the conversion of methyl-dCMP (mdCMP)³ in a poly(dC)-(PtdG)₉ double-stranded DNA strand to dCMP as described previously (18). One unit of demethylase is defined as the activity required to transform 1 pmol of mdCMP to dCMP in 1 h at 37 °C (16).

In Vitro Methylation of Substrates—P梅Cat² plasmid or pBlue-script SK(+)- plasmid were methylated in vitro by incubating 5 µg of plasmid DNA with 8 units of SssI CpG DNA methyltransferase (19) (New England Biolabs Inc.) in a buffer recommended by the manufacturer containing 320 µM S-adenosylmethionine, at 37 °C for 2 h. After

* This work was supported by the National Cancer Institute Canada.
† To whom correspondence should be addressed. Tel.: 514-398-7107; Fax: 514-398-6690; E-mail: mszyf@pharma.mcgill.ca.
‡ From the Department of Pharmacology, McGill University, Montreal, Quebec H3G 1Y6, Canada.

³ The abbreviations used are: mdCMP, methyl-dCMP; PCR, polymerase chain reaction; CAT, chloramphenicol acetyltransferase; bp, base pair(s).
DNA Demethylase Is a Processive Enzyme

RESULTS

Two Possible Modes of Demethylation—Enzymes interacting with DNA can act either processively or distributively. A processive mechanism implies that once the demethylase lands on, or centers itself on, a DNA molecule, it slides along the DNA molecule and demethylates a stretch of CpGs in cis as it progresses along the DNA strand (see Fig. 1 for model). The alternative, distributive mechanism implies that each demethylation event requires an independent interaction of the enzyme with the DNA. The enzyme is dislodged from the DNA after demethylation of each CpG site and will then land on any accessible methylated CpG site on any molecule of DNA in the reaction mixture in no specific order. These two models can be differentiated if one can monitor the intermediate stages of a demethylation reaction of a sequence of methylated CG sites on individual molecules of DNA. A distributive mechanism predicts that the demethylated sites will be interspersed within methylated sites and will be randomly distributed in the different DNA molecules (Fig. 1A). On the other hand, if demethylation is progressive, two classes of molecules will be identified at an intermediate stage, molecules that bear stretches of nonmethylated sites and others that are fully methylated (Fig. 1B).

Restriction Enzyme Analysis of the Pattern of Demethylation in Vitro of a CpG-methylated Plasmid—pBluescript (+) plasmid was fully methylated in vitro and incubated at 37 °C with 1.4 units of demethylase for either 25 min or 3 h. Following demethylation, the plasmid was subjected to digestion with HpaII, which cleaves the sequence CCGG when the internal C is not methylated. As shown in Fig. 2, demethylation of the plasmid is completed after 3 h, as indicated by the complete digestion of the plasmid with HpaII and the appearance of all the predicted HpaII fragments. At 25 min the plasmid is incompletely demethylated as indicated by the presence of plasmid DNA that is resistant to HpaII cleavage. However, the pattern of cleavage by HpaII is consistent with a processive mechanism of demethylation. If demethylation is distributive, the demethylated CpGs should be randomly spread, resulting in a gradient of sizes of HpaII partial fragments. However, the results in Fig. 2 show two classes of bands, fully digested HpaII fragments and plasmids that are completely resistant to HpaII.

Bisulfite Mapping of in Vitro Demethylated Mouse dnmt1 Sequences—Whereas the restriction enzyme pattern of digestion shown above suggests a processive mechanism of demethylation, it only allows us to look at a population of plasmids at a given time. To fully demonstrate a processive mechanism of demethylation, one has to be able to follow the fate of methylated CG sites on a single molecule of DNA. Bisulfite mapping, which is followed by PCR amplification, cloning, and sequencing of individual clones allows one to determine the state of methylation of a single DNA molecule at a time, at single base resolution (17). Fully methylated pMetCAT+ plasmid DNA, which has been described previously (18) (Fig. 3A and 4A, zero time) bears a genomic fragment of the dnmt1 gene and a bacterial chloramphenicol acetyltransferase gene on the same DNA molecule. Thus, it allows us to follow the demethylation of a stretch of CpG sites located in a low density CG region, which is characteristic of many vertebrate genes (dnmt1) as well as a dense CG
region in the bacterial CAT (Fig. 4). In vitro methylated pMetCAT was incubated with purified demethylase for different time intervals as indicated in Fig. 3 and Fig. 4. The demethylated DNA was treated with sodium bisulfite (which results in the conversion of nonmethylated cytosines to thymidine, whereas methylated cytosines are protected), amplified by PCR using appropriate primers for the indicated region of dnmt1, and subcloned. Subclones, each of which represents a single molecule of DNA treated with demethylase for a specific amount of time, were sequenced.

We first calculated the average methylation at each site in a representative population of in vitro demethylated pMetCAT plasmids. This allowed us to monitor the general progression of the demethylation reaction and to identify sites that exhibited either specific resistance or sensitivity to demethylation. Fig. 3A shows the map of the CG sites in the dnmt1 region and the average percentage methylation of each site in a representative population. The 265-bp region amplified contains 5 CG sites or 1 CG site in 50. At 10 s 11 out of 14 clones observed remain fully methylated (78%); however, methylation decreases with increased time. At 30 s, 8 clones remain fully methylated out of
In Vitro Demethylation of a Dense CG Region Residing within the CAT—Bisulfite mapping of the state of methylation of a second region in the same pMetCAT+ plasmid allowed us to confirm our initial observations regarding the processivity of the reaction and to determine whether demethylation is dependent on the properties of the sequence. We amplified a 250-base pair sequence within the bacterial CAT gene, with a CpG density of 5.2% (13 CpGs within a region of 250 bp), which is more than two times greater than the density of the dnmt1 region. Fig. 4A shows a physical map of the CAT gene and the location of the CpG sites that were mapped. The average methylation of the CAT sites at different time intervals through the demethylation reaction was calculated, and it demonstrates that the demethylation of the CAT region which is co-linear with the dnmt1 region is slower (Fig. 4A). This is consistent with the slower demethylation of the Sk plasmid, which is also CG-rich relative to the dnmt1 region (Fig. 3). At zero time, all 13 CpGs are fully methylated and remain so at 30 s, unlike the pattern seen in the dnmt1 region (Fig. 3A). Even at 2 h, the demethylation reaction is not complete, 2 clones out of 8 are still methylated. When the state of methylation of independent molecules is examined (Fig. 4B), the distribution of demethylated sites is not random among the different plasmid molecules. Some molecules are fully methylated whereas others bear a stretch of demethylated CGs. This again is consistent with a processive mechanism. Once demethylation is initiated in this region it progresses in cis.

DISCUSSION

Global changes in demethylation require enzymatic activities that can transform methylated cytosines, which are spread over large segments of DNA to nonmethylated cytosines. One attractive explanation might be that the demethylase enzyme is a processive enzyme. The mechanisms responsible for demethylation have been unknown, since the identity of the enzymatic activities responsible for demethylation was a mystery. Our identification of a bona fide demethylase (16) allows us to address the mechanisms of demethylation for the first time. In this manuscript we use the bisulfite mapping method (17), which provides a unique opportunity to dissect the interaction of an enzyme with a single substrate molecule at a time. Using this method, we demonstrate that the demethylase is a processive enzyme. We suggest that the critical step in demethylation is the interaction of the enzyme with the substrate. Once demethylation is initiated, it will proceed uninterrupted for at least 250 bp at a high rate, since no partial demethylation of the dnmt1 region or the CAT region is detected even after 30 s at single molecule resolution.

The presence of regions exhibiting different density of methylated CGs on the same molecule of DNA allowed us to determine whether these differences will affect demethylation and its processivity. It is obvious that the processivity of the demethylase does not extend to the CAT region from the dnmt1 region, since the CAT region (Fig. 4B), which is co-linear with the dnmt1 region (Fig. 3B), is demethylated more slowly. This is consistent with the demethylase proceeding to a certain distance before being dislodged. However, even within the CAT region, once demethylation is initiated, it proceeds in cis, since we have not identified molecules that are partially methylated in this region. The difference in the demethylation of the CAT and dnmt1 regions must reflect a lag in initiation of demethylation in these two different regions. These data are consistent with the interaction of demethylase with the substrate as being the rate-limiting step. We suggest that the interaction of demethylase with DNA can vary based on the properties of the region; some regions such as the dnmt1 sequence on pMetCAT show a high propensity to interact with the enzyme, whereas others such as the CAT regions display a low affinity. Once the demethylase lands on a specific region, it migrates processively to a certain distance until it recognizes a stop signal or dislodges from the DNA.

It has been proposed previously that demethylation in vivo proceeds in cis from centers of demethylation (20), resulting in demethylation of certain stretches of the genome. The data presented here are consistent with this model.

Whereas additional experiments are required to establish that demethylation in vivo progresses processively from centers of methylation as suggested here, the data presented here demonstrate that the demethylase is a processive enzyme and is consistent with this model.

Acknowledgment—We thank Dr. J. David Knox for critical reading of the manuscript and his thoughtful comments.

REFERENCES

1. Razin, A., and Riggs, A. D. (1980) Science 210, 604–610
2. Brandesia, M., Ariel, M., and Cedar, H. (1995) Bioessays 15, 709–713
3. Kafri, T., Ariel, M., Brandesia, M., Shemer, R., Urven, L., McCrarry, J., Cedar, H., and Razin, A. (1992) Genes Dev. 6, 705–714
4. Ariel, M., Cedar, H., and McCrarry, J. (1994) Nat. Genet. 7, 59–63
5. Salsar, H. P., Jirinny, J., and Jost, J. P. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 7167–7171
6. Benveniste, N., Mencher, D., Meyuhas, O., Razin, A., and Reshef, L. (1985) Proc. Natl. Acad. Sci. U. S. A. 90, 10558–10562
7. Mann, M., Boulbelik, M., and Lehnert, S. (1987) Development (Camb.) 99, 371–382
8. Razin, A., Webb, C., Szyf, M., Yisraeli, J., Rosenthal, A., Naveh-Man, Y., Szyf-Gallili, N., and Cedar, H. (1984) Proc. Natl. Acad. Sci. U. S. A. 81, 2275–2279
9. Razin, A., Szyf, M., Kafri, T., Roll, M., Giloh, H., Sarkan, S., Carotti, D., and Cantoni, G. L. (1986) Proc. Natl. Acad. Sci. U. S. A. 83, 2827–2831
10. Szyf, M., Eliaison, L., Manno, V., Klein, G., and Razin, A. (1985) Proc. Natl. Acad. Sci. U. S. A. 82, 8096–8094
11. Feinberg, A. P., Gehrke, C. W., Kuo, K. C., and Ehrlich, M. (1988) Cancer Res. 48, 1159–1161
12. Szyf, M. (1994) Trends Pharmacol. Sci. 7, 233–238
13. Vairapandi, M., and Duke, N. J. (1993) Nucleic Acids Res. 21, 5323–5327
14. Weiss, A., Keshet, J., Razin, A., and Cedar, H. (1996) Cell 87, 709–718
15. Bhattacharya, S., Ramchandani, S., Carotti, D., and Szyf, M. (1999) Nature 397, 579–583
16. Clark, S. J., Harrison, J., Paul, C. L., and Frommer, M. (1994) Nucleic Acids Res. 22, 2990–2997
17. Szyf, M., Theberge, J., and Zonovic, V. (1995) J. Biol. Chem. 270, 12690–12696
18. Nur, I., Szyf, M., Razin, A., Glaser, G., Rottem, S., and Razin, S. (1985) J. Bacteriol. 164, 19–24
19. Szyf, M. (1991) Biochem. Cell Biol. 69, 764–767