Factors Affecting de Novo Methylation of Foreign DNA in Mouse Embryonic Stem Cells*

(Received for publication, April 28, 1999)

Jennifer M. Hertz‡, Gudrun Schell, and Walter Doerfler§

From the Institute of Genetics, University of Cologne, D-50931 Koeln, Germany

Integration of foreign DNA into an established host genome can lead to changes in methylation in both the inserted DNA and in host sequences and potentially alters transgene and cellular transcription patterns. This work addresses the questions of what factors influence de novo methylation, and whether the integration site or inserted DNA can affect de novo methylation. Homologous recombination was used to integrate foreign DNA into a specific gene, B lymphocyte kinase (BLK), in mouse embryonic stem (ES) cells. Two plasmids were chosen for integration; one contained the adenovirus type 2 E2AL promoter upstream of the luciferase reporter gene, and the second carried the early SV40 promoter. The methylation patterns were analyzed using HpaII and MspI restriction endonucleases for both homologously recombined and randomly integrated foreign DNA in the ES cell clones.

Upon homologous reinsertion of the BLK gene into the genome of mouse ES cells, methylation patterns in this gene were reestablished. In DNA segments adjoined to the BLK gene, the de novo patterns of DNA methylation depended on the viral sequences in these clones and on the locations of the inserts, i.e. on whether the insertions resulted from homologously recombined or randomly integrated foreign DNA. In homologously recombined DNA, sequences carrying the adenovirus type 2 promoter were heavily methylated, and those with an SV40 promoter and an SV40 enhancer element remained unmethylated or hypomethylated. Upon removal of the enhancer element, these inserted constructs also became heavily methylated. In addition, all randomly integrated constructs were heavily methylated independently of the promoter and enhancer element present in the construct. These results indicate that modes and sites of integration as well as the inserted nucleotide sequence, possibly promoter strength, are factors affecting de novo methylation.

Specific patterns of DNA methylation, an epigenetic modification in the mammalian genome (1), have been shown to affect gene expression (2–5). DNA methylation is involved in a number of regulatory mechanisms, such as imprinting (6, 7), X chromosome inactivation (8), carcinogenesis (9–11), and embryonic development (12). Adenovirus (Ad) DNA has been used as a model system to study the effects of foreign DNA integration and methylation on host and Ad gene expression (2, 13–16). The Ad genome is 34–36 kb in size and consists of a linear, double-stranded DNA molecule with a 55-kDa protein covalently attached at either 5' terminus. When this virus replicates in permissive cells, the replicated genome is not methylated (17, 18). However, when the virus DNA integrates into the host genome, as in Ad-induced tumor cells or in Ad-transformed cells, the viral genome becomes methylated by a de novo mechanism (19–21). Frequently, multiple copies of Ad DNA are integrated at many sites in nontandem arrays in the host genome, but there seems to be a preference for sites with patchy homologies at the junctions between cellular and viral DNA (15, 16, 22, 23). When adenovirus (foreign) DNA is inserted into the mammalian genome, there are changes in the methylation patterns of the flanking sequences due to the position of the integration and/or the inserted DNA sequences. After Ad integration, the methylation in the flanking host sequences can be altered (24) as well as other more remote genomic sequences such as major histocompatibility complex class I, Ig Cκ and intracisternal A particle DNA sequences (25). In addition, when only bacteriophage λ DNA has been stably integrated into cellular DNA, again intracisternal A particle DNA sequences have shown changes in their methylation patterns (26).

Methylation of Ad or any other foreign DNA integration may be affected by the site of integration in the host genome. In addition, position effects on de novo methylation have been shown in several different systems including cell lines containing a variety of transgenes (27, 28) and in transgenic mouse models (29, 30).

Detailed analyses of DNA methylation and expression always reveal a more complex situation, because transgenic cell lines usually contain multiple copies of the gene of interest at different positions or in concatenated forms in the host genome, which is a common result of random integration. In order to understand better the regulation of de novo methylation, homologous recombination was used to insert different foreign DNAs into one site in the genome of mouse embryonic stem cells. The results show that random integration appears to promote de novo methylation. For homologous recombination with one allele of the same insert in the same cell type, the original methylation pattern for the genomic sequence is reestablished. However, for foreign DNA the extent of de novo methylation depends on the newly integrated sequence.

**EXPERIMENTAL PROCEDURES**

**Plasmids Used for Homologous Recombination and for DNA Probes—**AdBLK and SVBLK (Fig. 1) were constructed to perform homologous

*This research was supported by European Union Grant EN AX 101805 and Deutsche Forschungsgemeinschaft Grant SFB274-A1. The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‡ Recipient of a fellowship from the Alexander von Humboldt-Stiftung, Bonn, for part of the time of her stay in Koeln.

§ To whom correspondence should be addressed: Inst. of Genetics, University of Cologne, Weyertal 121, D-50931 Koeln, Germany. Tel.: 49-221-470-2386; Fax: 49-221-470-5163; E-mail: doerfler@scan.genetik.uni-koeln.de.

Received for publication, April 28, 1999

The Journal of Biological Chemistry Vol. 274, No. 34, Issue of August 20, pp. 24232–24240, 1999

Printed in U.S.A.

© 1999 by The American Society for Biochemistry and Molecular Biology, Inc.
Factors Affecting de Novo Methylation in ES Cells

Factors Affecting de Novo Methylation in ES Cells

Several methods were used to determine which clones contained homologously recombined DNA within the BLK gene (provided by Gemma Texido and Alexander Tarakhovsky), since this part of BLK was not present in any of the constructs used for transfection into ES cells. The BLK-5′-ab probe is the fragment between the ApaI site in exon 2 and the BamHI site in exon 3, while the BLK-3′-ab probe is the DNA fragment from the same BamHI site to the XhoI site in exon 4. The Ad luc probe is the XhoI to Sall DNA fragment (3201 bp) of pAd2E2AL-LUX (35), and the same fragment was isolated from the PGL2 control vector (Promega). This enhancer contains both the 27- and 21-bp repeats. A third plasmid, SVpBLK, has been constructed that contains SVLuc without the SV40 enhancer sequences at the same restriction site.

An additional BLK fragment, called BLK-out, was used to determine which clones contained homologously recombined DNA within the BLK gene (provided by Gemma Texido and Alexander Tarakhovsky), since this part of BLK was not present in any of the constructs used for transfection into ES cells. The BLK-5′-ab probe is the fragment between the ApaI site in exon 2 and the BamHI site in exon 3, while the BLK-3′-ab probe is the DNA fragment from the same BamHI site to the XhoI site in exon 4. The Ad luc probe is the XhoI to Sall DNA fragment (3201 bp) of pAd2E2AL-LUX (35), and the same fragment was isolated from the PGL2 control vector (Promega). This enhancer contains both the 27- and 21-bp repeats. A third plasmid, SVpBLK, has been constructed that contains SVLuc without the SV40 enhancer sequences at the same restriction site.

RESULTS

Homologous Recombinant and Randomly Integrated DNA in ES Clones—Two different DNAs, AdBLK or SVBLK, were electroporated into mouse ES cells to obtain homologous recombinants and to investigate whether specific DNA sequences can affect de novo methylation. ES cells were chosen as the host, since they possess a de novo methylation mechanism (43–45). The BLK gene was chosen as the target, and both the genomic BLK DNA fragments and the ES cells (BL6/III) are derived from the same mouse strain, C57BL/6. The plasmids, AdBLK and SVBLK, contain the neomycin and luciferase genes and are flanked by two BLK genomic DNA fragments. Therefore, after HR, the foreign DNA segment is inserted within intron 3 of the BLK gene (Fig. 1B). Two types of HR clones were obtained after electroporation into ES clones, and they differ only within the luciferase region. The A clones carry AdLuc, the Ad2 E2A late promoter 5′ of the luciferase reporter gene (Fig. 1A, Table I, left side). This promoter has been shown to be inhibited by sequence-specific methylation (46, 47). The S clones have SVLuc with the SV40 early promoter 5′ of the luciferase coding sequence and at the 3′-end the SV40 enhancer element containing the 72- and 21-bp repeats (Fig. 1A, Table I, right side).

Genomic DNA was obtained from about 40 different ES clones from each transfection to determine which clones had incorporated the neomycin and luciferase sequences within the expected BLK site by homologous recombination and which clones contained part or all of these sequences in other locations due to random integration events. The genomic DNA was cleaved with HindIII, and the fragments and the expected片段 were separated by electrophoresis on a 0.8% agarose gel for Southern blot analysis. Two different probes were used to analyze the location of the inserts and the presence of the luciferase sequence (Fig. 2A, Table I, right side). This probe has been shown to be inhibited by sequence-specific methylation (46, 47). The S clones have SVLuc with the SV40 early promoter 5′ of the luciferase coding sequence and at the 3′-end the SV40 enhancer element containing the 72- and 21-bp repeats (Fig. 1A, Table I, right side).

Genomic DNA was obtained from about 40 different ES clones from each transfection to determine which clones had incorporated the neomycin and luciferase sequences within the expected BLK site by homologous recombination and which clones contained part or all of these sequences in other locations due to random integration events. The genomic DNA was cleaved with HindIII, and the fragments and the expected fragments were separated by electrophoresis on a 0.8% agarose gel for Southern blot analysis. Two different probes were used to analyze the location of the inserts and the presence of the luciferase sequence (Fig. 2A, Table I, right side). This probe has been shown to be inhibited by sequence-specific methylation (46, 47). The S clones have SVLuc with the SV40 early promoter 5′ of the luciferase coding sequence and at the 3′-end the SV40 enhancer element containing the 72- and 21-bp repeats (Fig. 1A, Table I, right side).

Genomic DNA was obtained from about 40 different ES clones from each transfection to determine which clones had incorporated the neomycin and luciferase sequences within the expected BLK site by homologous recombination and which clones contained part or all of these sequences in other locations due to random integration events. The genomic DNA was cleaved with HindIII, and the fragments and the expected fragments were separated by electrophoresis on a 0.8% agarose gel for Southern blot analysis. Two different probes were used to analyze the location of the inserts and the presence of the luciferase sequence (Fig. 2A, Table I, right side). This probe has been shown to be inhibited by sequence-specific methylation (46, 47). The S clones have SVLuc with the SV40 early promoter 5′ of the luciferase coding sequence and at the 3′-end the SV40 enhancer element containing the 72- and 21-bp repeats (Fig. 1A, Table I, right side). The genomic DNA was cleaved with HindIII, and the fragments and the expected fragments were separated by electrophoresis on a 0.8% agarose gel for Southern blot analysis. Two different probes were used to analyze the location of the inserts and the presence of the luciferase sequence (Fig. 2A, Table I, right side). This probe has been shown to be inhibited by sequence-specific methylation (46, 47). The S clones have SVLuc with the SV40 early promoter 5′ of the luciferase coding sequence and at the 3′-end the SV40 enhancer element containing the 72- and 21-bp repeats (Fig. 1A, Table I, right side).
since these fragments contain both the BLK and the luciferase sequences. When the DNA on the same membrane is probed with left and right BLK fragments, all BLK fragments are present as expected, the endogenous allele (8.5 kb) and the two fragments from the HR allele (5.9 and 7.9 kb).

In addition to the clones that contain foreign DNA homologously recombined within intron 3, a number of ES clones with randomly integrated foreign DNA have been isolated. All RI clones are designated with an asterisk.* No 7.9-kb band is detected in the DNA of these clones when BLK-out is used as the probe. In addition, an off size DNA fragment is visible when the DNA on the membrane is hybridized with the luciferase probe.

Table I lists the Adluc-containing clones, referred to as A clones, and the SVluc-containing clones, referred to as S clones, which have been chosen for further studies. At the time points indicated, genomic DNA was extracted for Southern blot analyses, and in addition, protein extracts were prepared at different times to determine the luciferase activity in the different cell populations at various passages. In this study, we have analyzed a number of different ES clones to ascertain representative sampling.

**Luciferase Expression of ES Clones**—For the protein extracts of all HR A and S clones, the RLU/µg of protein decrease with initial passaging as seen in Fig. 3. Throughout passaging, the S clones exhibit 10–50-fold higher RLU/µg of protein than the A clones (Fig. 3, compare A and B). In addition, the constitutive RLU/µg of protein in later passages is higher for the S clones than for the A clones. Both sets of A* and S* clones with RI foreign DNA exhibit a 100-fold reduction or no luciferase activity at all from early on (passage 3) to the latest passage analyzed (passage 31; data not shown). After an initial drop in luciferase expression, both the A and S clones maintain a constitutive level even to late passages. We have not related these expression levels to the patterns of DNA methylation, because correlations between promoter activity and DNA

---

### Table I

**Survey of all transgenic ES clones used in this study**

| Integrated A (Adluc) clones | Integrated S (SVluc) clones |
|-----------------------------|-----------------------------|
| **Clone** | **Early** | **Middle** | **Late** | **Clone** | **Early** | **Middle** | **Late** |
| A1 | 20 | | | S5 | 6 | 20 | |
| A4 | 5 | 21 | | S11 | 3 | 20 | 30 |
| A5 | 5 | 20 | | S14 | 5 | 20 | 31 |
| A8 | 15 | 9 | 20 | | |
| A15 | 5 | 31 | | So3 | 3 | 20 | |
| A16 | 5 | 21 | | So4 | 6 | 19 | |
| A17 | 7 | 20 | 30 | So5 | 6 | 20 | |
| A18 | 19 | | | So6* | 6 | 23 | |
| Ao1 | 6 | 20 | 30 | S7* | 6 | 20 | 29 |
| Ao8 | 6 | 20 | 30 | So1* | 3 | 20 | 31 |
| Ao12* | 5 | 6 | 20 | So2* | 6 | 20 | 31 |
| Ao2* | 8 | 20 | 31 | So6* | 6 | 20 | |

* ES clones with homologously recombined DNA chosen for further analyses. These clones were derived from cells transfected with a DNA plasmid containing Adluc (see Fig. 1).

* ES clones with homologously recombined DNA chosen for further analyses. These clones were derived from cells transfected with the SVluc-containing plasmid (see Fig. 1).

* Experimental details and clone designations are described under "Experimental Procedures.*

* Early, middle, late, and numbers refer to passages in culture of individual transgenic ES clones.

* Clones carrying randomly integrated DNA are designated with an asterisk.
FIG. 2. Southern blot analyses (40) of the DNAs from ES clones transgenic for AdBLK and SVBLK constructs. 10 µg of genomic DNA from each ES clone was cleaved with HindIII. The fragments were electrophoresed through a 0.8% agarose gel and transferred to a Qiagen Nylon Plus membrane. The DNA probes were radiolabeled with α-32P to detect DNA fragments after hybridization. In all figures, all ES clones that contain Adluc DNA have designations beginning with an A, while those clones that contain the SVluc DNA begin with an S (see Fig. 1). A, BLK-out (map in Fig. 1B) was used as a probe to detect two expected DNA fragments at 8.5 kb, representing the endogenous BLK allele, and 7.9 kb for the homologously recombined allele within intron 3 (see Fig. 1B). C represents the control lane, ES cells that were not transfected. B, the same membranes as in A were reprobed with either Adluc (left panel) or SVluc (right panel) to detect the 7.9-kb fragment representing the luciferase region within BLK intron 3. C, both membranes were reprobed again with BLK-in left and right together to detect all BLK fragments integrated in the A and S clones. Randomly integrated ES clones are marked with an asterisk. A minus symbol represents clones that lack detectable integrated foreign DNA.

Factors Affecting de Novo Methylation in ES Cells

methylation have not been the primary objective of this study. Moreover, in the SV40 promoter there is no HpaII site that would allow investigations by restriction with HpaII and MspI.

Upon Homologous Reintegration, the 5′- and 3′-Flanking BLK Sequences Reestablish 5′-CCGG-3′ Methylation Patterns Similar to Those in the Endogenous BLK Alleles—For the analyses of methylation patterns, all genomic DNA samples were cleaved with BamHI to separate the BLK sequence from that of the foreign DNA (Fig. 1B). In this way, each DNA sequence was analyzed individually. In addition, one of the two restriction endonuclease isoschizomers MspI or HpaII was used to analyze the DNA methylation levels. Unlike MspI, HpaII is unable to cleave the sequence 5′-CCGG-3′ when the internal (3′) C is methylated. Therefore, the difference in DNA fragment patterns generated with these two endonucleases represents differences in DNA methylation at 5′-CCGG-3′ sites. When control genomic DNA (C in Fig. 4) from untransfected cells is cleaved with BamHI (lanes B), a single fragment is detected as expected when BLK-3′xb was used as a probe (Fig. 4, C, lanes B). MspI cuts the DNA into much smaller fragments (lanes M) which are not resolved by electrophoresis on a 0.8% agarose gel. There are larger molecular mass fragments present after BamHI/HpaII cleavage (Fig. 4, lanes H) as compared with BamHI/MspI cleavage (lanes M). This finding demonstrates that many HpaII sites within the BLK 3′-flanking sequence (see Fig. 1B) are methylated. A comparison of the control (C) DNAs to the DNAs from A and S clones reveals that the HpaII patterns are very similar or identical (Fig. 4, H lanes). Moreover, for a given DNA fragment within an H lane of any of the ES clones, the relative intensities are similar as for the corresponding bands in the control DNA H lane (Fig. 4, e.g. C and A17 or S11; H lanes). This similarity of HpaII patterns within the BLK 3′-flanking sequence has been seen for all A and S clones analyzed. With a few exceptions, similar results have also been obtained for the 5′-flanking BLK sequence (data not shown). First, as in the BLK 3′-flanking sequence for the control (C) DNA from untransfected ES cells, there are many larger molecular mass fragments present in the H lanes as compared with the M lanes. This finding again implies that the BLK 5′-flanking sequences are methylated. The HpaII patterns of DNA from one clone, S4, show some differences in banding patterns of the 5′-flanking sequence, both between passages (p6 to p20) and compared with DNA from the parental untransfected ES cells (data not shown). Slight differences in the HpaII patterns may not be detectable, since the 5′- and 3′-BLK probes hybridize to both the endogenous and HR allelic fragments. We conclude that the homologously reintegrated BLK sequences are remethylated in 5′-CCGG-3′ patterns very similar to those of the endogenous BLK gene.

A and S Clones Develop Different Methylation Patterns in the Luciferase Region—The luciferase regions of the A and S clones were analyzed to investigate whether foreign DNA sequences influence de novo methylation after their integration. Fig. 5 shows the autoradiograms of Southern blots of genomic DNA isolated at different passages for a number of the A clones. The entire luciferase region including the Ad promoter was used as the probe (Adluc; Fig. 1A). As expected, there is no detectable hybridization to the DNA in the control lane (C, lanes B, M, and H). However, for the DNAs from the A clones, the expected 5-kb fragment is detected after BamHI cleavage (B lanes). There are 15 5′-CCGG-3′ sites (vertical lines in Fig. 1A) that can be differentially methylated. Different cleavage patterns are generated by MspI versus HpaII (M and H lanes, respectively). Upon HpaII cleavage of the DNA from all passages analyzed, very large HpaII fragments, ranging in size up to the 5-kb BamHI fragment in the DNA of many clones (e.g. A4, A5, A8, A16, and A18) are generated (Fig. 5).

For the S clones, a very different trend is seen. SVluc (Fig. 1A) was used as a probe to determine the methylation patterns of these clones. The DNA samples of only a few clones, such as S5 and S11, exhibit larger HpaII fragments at molecular masses higher than those for the MspI fragments. For the majority of the S clones, the MspI and HpaII cleavage patterns are quite similar (Fig. 6, M versus H lanes). These data demonstrate that the HpaII patterns of DNA from A and S clones are strikingly different. The foreign luciferase gene thus becomes extensively de novo methylated when it is located downstream of the Ad2 promoter but remains unmethylated or
hypomethylated when the same gene is within the context of the SV40 promoter and enhancer.

Both the A and S clones were passaged for an extended period of time to determine the regulation of de novo methylation for a given site in the genome, i.e., within the BLK gene. For the A and S clones analyzed, once the HpaII pattern was established, it remained unchanged with passaging, e.g., A15 and S14 (Figs. 5 and 6). In addition, longer exposures were made, and phosphor images were analyzed that confirmed these results (data not shown).

All Randomly Integrated Sequences Have a High Degree of Methylation in the Luciferase Region—As a control, genomic DNAs from a number of A* and S* clones with randomly integrated DNA were also analyzed (Fig. 7). In contrast to the A and S clones analyzed, the HpaII pattern for the DNAs from the homologously recombinant clones with Ad2 or SV40 promoters in front of the luciferase gene are strikingly dissimilar, the neomycin regions just 5' of the luciferase sequence have also been analyzed in DNA from the same ES clones. Since the nucleotide sequence of the inserted foreign DNA is identical in both sets of these HR clones, the same neomycin probe was used to determine the HpaII cleavage patterns. In the control lanes (Fig. 8, C, lanes B, H, and M), there are no detectable bands at the sizes expected for the neomycin fragments. For the DNAs from the A clones, the HpaII patterns (H lanes) show various intensities of molecular fragments larger than the MspI patterns (M lanes) (e.g., clones A5, A8, and A18) (Fig. 8). In addition, there are clones such as A17 and A15 that have similar MspI and HpaII patterns (M and H lanes). There is a range of HpaII patterns that differ in relative intensities of the larger molecular mass fragments within a given H lane. The A clones with larger HpaII fragments in the luciferase region also have higher amounts of larger HpaII fragments in the neomycin region. Those clones, whose DNAs exhibit intermediate DNA methylation in the luciferase region, have very low to undetectable levels in the

**Factors Affecting de Novo Methylation in ES Cells**

**A Luciferase Expression of S clones**

**B Luciferase Expression of A clones**

**The DNAs in the A and S Clones Differ in Their Methylation Patterns in the Neomycin Gene**—Since the HpaII patterns for the DNAs from the homologously recombinant clones with Ad2 or SV40 promoters in front of the luciferase gene are strikingly dissimilar, the neomycin regions just 5' of the luciferase sequence have also been analyzed in DNA from the same ES clones. Since the nucleotide sequence of the inserted foreign DNA is identical in both sets of these HR clones, the same neomycin probe was used to determine the HpaII cleavage patterns. In the control lanes (Fig. 8, C, lanes B, H, and M), there are no detectable bands at the sizes expected for the neomycin fragments. For the DNAs from the A clones, the HpaII patterns (H lanes) show various intensities of molecular fragments larger than the MspI patterns (M lanes) (e.g., clones A5, A8, and A18) (Fig. 8). In addition, there are clones such as A17 and A15 that have similar MspI and HpaII patterns (M and H lanes). There is a range of HpaII patterns that differ in relative intensities of the larger molecular mass fragments within a given H lane. The A clones with larger HpaII fragments in the luciferase region also have higher amounts of larger HpaII fragments in the neomycin region. Those clones, whose DNAs exhibit intermediate DNA methylation in the luciferase region, have very low to undetectable levels in the
neomycin region. Hence, there appears to be a congruence in the levels of de novo methylation between the luciferase and neomycin regions; the DNAs from clones that are strongly methylated in the luciferase region are also more highly methylated in the neomycin region.

If this correlation holds for the S clones as well, we would expect very low levels of methylation in the neomycin regions, since there is little to no methylation present in the luciferase sequence. As expected for the S clones, the $MspI$ and $HpaII$ patterns are quite similar (Fig. 9, compare the $M$ and $H$ lanes) (i.e. no or very low methylation is detected).

The DNA in All RI Clones Has a High Degree of Methylation in the Neomycin Region—Next, the neomycin region was investigated for differences between the $HpaII$ patterns of the RI $A^*$ and $S^*$ clones. As observed in the luciferase region (Fig. 7), only very large molecular mass fragments are detected in the $HpaII$ patterns of the DNA on Southern blots hybridized with the neomycin probe (data not shown). These data confirm that all
segments of the randomly integrated foreign DNA become extensively de novo methylated both in the luciferase and neomycin genes.

DNA Methylation Patterns and Luciferase Expression of the SVpBLK Clones Lacking the Enhancer Element—Sp1 sites have been implicated in inhibiting de novo methylation of 5'-CG-3'-rich regions (48, 49) and in increasing gene expression (50). The Sp1 sites could negatively affect the de novo methylation of the entire luciferase and neomycin genes within the S clones. To test this hypothesis, new ES clones were established using another plasmid, SVpBLK, which contains SVluc as in the S clones, but without the SV40 enhancer at the 3'-end (see Fig. 1A). The 21- and 72-bp repeats that are part of the enhancer were also removed.

In contrast to the HR clones from the S group, these new SVp clones (HR and RI clones) have genomic DNA that is not completely cleaved with HpaII as seen with the A clones (data not shown). These results indicate that the DNA has become highly de novo methylated in all ES clones lacking the enhancer and the Sp1 sequences. While promoter strength is

FIG. 7. Randomly integrated DNAs in both the A and S clones show a high degree of de novo methylation within the luciferase region. For Southern blot analyses, genomic DNA from each clone was cleaved with BamHI (B lanes) and then with either HpaII (H lanes) or MspI (M lanes). Probes SVluc or Adluc (see Fig. 1, A and B) were used to determine the methylation patterns of the S* or A* clones, respectively, within this sequence of integrated DNA by comparing the M and H lanes within each passage. Randomly integrated ES clones are marked with asterisks. The results for the DNAs from clones S18 and So5, which show low levels of DNA methylation and are derived from HR experiments, are shown for comparison.

FIG. 8. The DNAs from Ad sequence-containing ES clones show a varying degree of de novo methylation within the neomycin region. Membranes carrying DNA from A clones (see Fig. 6) were reprobed with neomycin (neo, Fig. 1B) to determine the methylation patterns in the neomycin sequence of integrated DNA of the ES clones by comparing the M and H lanes.

FIG. 9. SV40 sequence-containing clones show low to no de novo methylation within the neomycin region. Membranes carrying DNA from S clones (see Fig. 7) were reprobed with neomycin (neo, Fig. 1B) to determine the methylation patterns in this sequence of integrated DNA of ES clones by comparing the M and H lanes within each passage.

Factors Affecting de Novo Methylation in ES Cells
probably a factor influencing de novo methylation, these data imply that cis-acting sequences can inhibit de novo methylation over several kb, as suggested elsewhere (51).

The luciferase expression from the SVp HR clones with the highly de novo methylated sequences is at least 100-fold lower (10–25 RLU/μg of protein) as compared with the original S clones (data not shown). Similarly, the SVp RI clones show very low to no luciferase activity (data not shown).

De Novo Methylation after Release and Reestablishment of Selection—Can this experimental mode lead to the high levels of methylation in the AdLuc-containing clones? An early passage of cell clone A8 was used for the release and reselection experiment and passaged either in the presence or the absence of G418. Genomic DNA from both clonal sublines was then isolated after 8 days. A comparison of the two sets of A8 DNA shows that there is no change in the methylation patterns in either the luciferase or neomycin region. This finding suggests that the expression of neomycin does not play a role in the inhibition or activation of de novo methylation of its sequence or of the luciferase gene.

DISCUSSION

De Novo Methylation of Foreign DNA Is Dependent on Several Factors—For many problems in molecular biology, it is of considerable interest to understand the mechanisms and the regulation of de novo methylation of foreign DNA that has been inserted into an established mammalian genome, haphazardly or by experimental design. One of the major unresolved questions about the mechanism of de novo DNA methylation addresses the factors determining the sites and extent of foreign DNA methylation upon insertion into the recipient genome. The following factors could be of relevance: (i) the site of insertion, (ii) the nucleotide sequence of the foreign DNA, (iii) specific motifs in the foreign DNA, (iv) the timing of the insertion event relative to the cell cycle, and (v) promoter strengths and/or elements (e.g. enhancers) in the inserted foreign DNA.

We have devised a series of experiments in which part of the murine BLK gene on chromosome 14 was reinserted into one of its authentic allelic genomic sites by homologous recombination. Mouse ES cell clones with the reinserted DNA in the correct position were selected. ES clones with the BLK DNA inserted into randomly targeted positions were also analyzed as controls. Foreign DNA sequences have been attached to the BLK gene in the constructs used in the electroporation-driven transfection experiments and have been placed inside the endogenous BLK gene (Fig. 1B). Three different constructs were transfected into ES cells to create three types of ES clones. In all of them, the neomycin gene has the same promoter, phosphoglycerate kinase, but the luciferase gene has a different promoter and/or enhancer sequence in each construct. The A clones carry the luciferase gene under the control of the Ad2 E2A late promoter. For the S clones, the SV40 early promoter is located 5’ of the luciferase gene, and the SV40 enhancer sequences are located 3’.

The DNAs in all of the A clones become de novo methylated in both the luciferase and neomycin genes, while the S clones stay hypomethylated even after 30 passages. The third type of ES clone, SVp, is identical to the S clone, except that it contains the promoter without the SV40 enhancer sequences at the 3’-end. Unlike the original S clones, these HR clones become de novo methylated within the luciferase and neomycin sequences at levels as seen for the A clones. These results indicate that the enhancer element, containing the 21- and 72-bp repeats, is required to inhibit de novo methylation within the entire luciferase and neomycin regions. The protection of sequences from de novo methylation may be due to factors binding these sequences, e.g. multiple Sp1 sites that are present in the enhancer sequence. Inhibition may occur at sites where there is an interaction between the promoter and the enhancer element within the transgene. In addition, the strength of the SV40 promoter, as compared with the Ad2 E2A late promoter, could also be a factor.

The DNAs from all of the RI clones in our study, regardless of the foreign sequence integrated, become highly methylated. These sequences can be recognized as foreign and may activate a host defense de novo methylation reaction (52, 53). Thus, the enhancer element that is present in some of the RI clones (data not shown) or promoter strength does not suffice to inhibit de novo methylation after random integration. Topologically correct insertion, as after reintegration by HR, apparently does not activate this defense mechanism. Therefore, our data indicate that genome location plays a major role in the regulation of de novo methylation.

The results presented in this report lead to the following conclusions. (i) The homologously recombined BLK sequence becomes remethylated in patterns very similar to the authentic preexisting cellular patterns of the endogenous alleles. (ii) Randomly integrated foreign DNA becomes very heavily methylated. Thus, position and vicinity effects including sequence and structure of the insert must be of considerable importance in determining de novo methylation patterns. (iii) The extents of de novo methylation of the foreign luciferase and neomycin genes reintegrated homologously with the BLK sequence depend on the sequences present in these constructs and/or on promoter strength. The early SV40 promoter with the enhancer sequences seems to prohibit de novo methylation, and the luciferase gene is expressed in the transgenic ES cells. Constructs under the control of the E2A late promoter of Ad2 DNA or under the SV40 promoter without the enhancer elements become extensively de novo methylated even in the authentic BLK position, and luciferase expression is reduced. However, in randomly integrated constructs the nature of the promoter or presence of the SV40 enhancer element is not decisive for de novo methylation. Thus, their role in determining de novo methylation may not be fundamental, at least not in randomly integrated foreign DNA. (iv) Removal and addition of the selective drug G418 does not alter the methylation levels of the examined DNA sequences in the ES clones that contain the Ad2 E2A late promoter. (v) After homologous recombination into BLK genes, nucleotide sequences in the transgenic foreign DNA can affect de novo methylation. One can ponder the possibility that promoters with their transcriptional potential in transgenic sequences within a specific chromatin arrangement interact with the cellular machinery that is responsible for de novo methylation or its inhibition.

A Working Model—The mechanism of de novo methylation of foreign DNA inserted into an established mammalian genome cannot yet be explained in detail. De novo methylation may be related to an ancient cellular defense system targeted against foreign DNA (52, 53). In the present study, we have investigated some of the factors influencing de novo methylation when an authentic cellular DNA sequence is reinserted into the mouse genome at different locations. The experimental approach chosen exposes the authentic unmethylated BLK sequence to the de novo methylation system in two different ways. In one set of experiments, the BLK DNA has been reinserted into its original chromosomal location by homologous recombination. Alternatively, upon random integration by heterologous recombination, the BLK and adjacent sequences are located at several randomly selected sites where they are apparently recognized as foreign by the de novo methylation machinery. In the latter context, promoter strength and presence or absence of the SV40 enhancer sequence seem to be of
As a working model, we are pursuing the possibility that the enzymatic process of de novo methylation recognizes and is dependent upon the specific chromatin structure at the site of an individual gene or DNA segment. Thus, the endogenous BLK DNA at its authentic site or the same DNA inserted at randomly chosen locations would look very different to the DNA methyltransferase system. When the BLK gene and adjoining sequences in the construct are homologously reinserted into one of the authentic allelic BLK positions on mouse chromosome 14, the preexisting chromatin configuration can be reconstituted. Under these conditions, the de novo methylation system then imprints the methylation pattern specific for the BLK locus that had preexisted prior to the transfection of ES cells. In contrast, after random integration, when the BLK construct arrives in an alien position with a locus-specific, but not BLK-typical, chromatin arrangement, a different pattern of de novo methylation will be imposed upon the integrated DNA that is recognized as foreign in this location with a heterologous sequence context. Of course, this model still leaves the question unanswered of how de novo methylation and chromatin structure are interdependent. Differences in the accessibility of individual 5′-CG-3′ dinucleotides to the DNA methyltransferase system may be one of the important parameters.

Acknowledgments—We thank our colleagues at the Institute of Genetics, Werner Muller, and Alexander Tarakhovsky, for advice and Raul Torres and Ralph Kuhn for help with the ES techniques.

REFERENCES

1. Halliday, R. (1987) Science 238, 163–170
2. Doerfler, W. (1988) Annu. Rev. Biochem. 57, 93–124
3. Razin, A., and Cedar, H. (1991) Microbiol. Rev. 55, 41–45
4. Eden, S., and Cedar, H. (1994) Curr Opin. Genet. Dev. 4, 253–259
5. Kass, S. U., Pruss, D., and Wolffe, A. P. (1997) Trends Genet. 13, 444–449
6. Li, E., Beard, C., and Jaenisch, R. (1993) Genes Dev. 8, 2325–2334
7. Chang, L. S., Jan, H. I., Koh, T. W., Ng, H. H., Xu, G., and Li, E. (1997) Science 277, 1996–2000
8. Laird, P. W. (1997) Mol. Med. Today 3, 223–229
9. Ramchandani, S., Macleod, A. R., Pinard, M., van Hove, E., and Syf, M. (1997) Proc. Natl. Acad. Sci. U. S. A. 94, 684–689
10. Li, E., Bestor, T. H., and Jaenisch, R. (1992) Cell 70, 915–926
11. Sutter, D., and Doerfler, W. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 255–256
12. Krueckel, I., and Doerfler, W. (1982) EMBO J. 1, 409–414
13. Doerfler, W. (1993) Prog. Nucleic Acids Res. Mol. Biol. 46, 1–36
14. Doerfler, W. (1996) BBA Rev. Cancer 1288, 579–599
15. Gunthert, U., Schweiger, M., Stupp, M., and Doerfler, W. (1976) Proc. Natl. Acad. Sci. U. S. A. 73, 3923–3927
16. Sutter, D., Westphal, M., and Doerfler, W. (1978) Cell 14, 569–585
17. Vardimon, I., Neumann, B., Kuhlmann, I., Sutter, D., and Doerfler, W. (1980) Nucleic Acids Res. 8, 2461–2473
18. Orend, G., Knoblauch, M., Kamber, C., Tjia, S. T., Schmitz, B., Linkswitz, A., Meyer zu Altensldhese, G., Maas, J., and Doerfler, W. (1995) J. Virol. 69, 1220–1242
19. Deuring, R., Winterhoff, U., Tamanoi, F., Stabel, S., and Doerfler, W. (1981) Nature 293, 81–84
20. Gahlmann, R., Leisten, R., Vardimon, L., and Doerfler, W. (1982) EMBO J. 1, 1101–1104
21. Lichtenberg, U., Zock, C., and Doerfler, W. (1988) Virus Res. 11, 335–342
22. Heller, K., Kamber, C., Wilgenbus, P., and Doerfler, W. (1995) Proc. Natl. Acad. Sci. U. S. A. 92, 5535–5539
23. Remus, R., Kamber, C., Heller, H., Schmitz, B., Schell, G., and Doerfler, W. (1999) J. Virol. 73, 1010–1022
24. Jahnke, D., and Jaenisch, R. (1985) Mol. Cell. Biol. 5, 2212–2220
25. Guy, L.-G., Kothary, R., and Wall, L. (1997) Nucleic Acids Res. 25, 4400–4407
26. Sasaki, H., Allen, N. D., and Surani, M. A. (1993) in DNA Methylation: Molecular Biology and Biological Significance (Just, J.-P., and Saluz, H. P., eds) pp. 469–486, Birkhaeuser Verlag, Basel
27. Koetsier, P. A., Mangel, L., Schmitz, B., and Doerfler, W. (1996) Transgen. Res. 4, 235–244
28. Kozak, C. A., Dymecki, S. M., Niederhuber, J. E., and Desiderio, S. V. (1991) Genomics 9, 762–764
29. Boer, P. H., Potten, H., Adra, C. N., Jardine, K., Mullhoffer, G., and McBurney, M. W. (1990) Biochem. Genet. 28, 299–308
30. Dymecki, S. M., Zwollo, P., Zeller, K., Kuhajda, F. P., and Desiderio, S. V. (1992) Bio. J. Biol. Chem. 267, 4815–4823
31. Li, E., Beard, C., Li, E., and Jaenisch, R. (1995) Nature 376, 362–365
32. Neumann, B., and Barlow, D. P. (1996) Curr. Opin. Genet. Dev. 6, 159–163
33. Beard, C., Li, E., and Jaenisch, R. (1995) Genes Dev. 9, 2325–2334
34. Koetsier, P. A., Mangel, L., Schmitz, B., and Doerfler, W. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 348–352
35. Southern, E. (1975) J. Mol. Biol. 98, 503–517
36. Lister, R., and Doerfler, W. (1993) BioTechniques 15, 260–261
37. Bradford, M. M. (1976) Anal. Biochem. 72, 248–254
38. Jaenisch, R., Harbers, K., Jahner, D., Stewart, C., and Stuhlmann, H. (1982) J. Cell. Biochem. 20, 331–336
39. McBurney, M. W., Sutherland, L. C., Adra, C. N., Leclaire, B., Rudnicki, M. A., and Jardine, K. (1991) Nucleic Acids Res. 19, 5755–5761
40. Kozak, C. A., Dymecki, S. M., Niederhuber, J. E., and Desiderio, S. V. (1991) Science 251, 254–258
41. Leder, E., and Imai, K. (1997) Cell 131, 247–254
42. Torres, R. M., and Kuhn, R. (1997) in Laboratory Protocols for Conditional Gene Targeting, pp. 73–78, Oxford University Press, Oxford
43. Liljestrom, P., Tus, N., Hulebrosk, D., and Garoff, H. (1991) J. Virol. 65, 4107–4113
44. Mansour, S., Thomas, K., and Caspetic, M. (1988) Nature 336, 348–352
45. Mansour, S., Thomas, K., and Caspetic, M. (1988) J. Virol. 62, 509–513
46. Kozak, C. A., Dymecki, S. M., Niederhuber, J. E., and Desiderio, S. V. (1991) Proc. Natl. Acad. Sci. U. S. A. 88, 915–926
47. Sutter, D., and Doerfler, W. (1980) Proc. Natl. Acad. Sci. U. S. A. 77, 255–256
48. Krueckel, I., and Doerfler, W. (1982) EMBO J. 1, 409–414
49. Doerfler, W. (1983) J. Mol. Biol. 159–163