Mutation of CpGs in the Murine Stem Cell Virus Retroviral Vector Long Terminal Repeat Represses Silencing in Embryonic Stem Cells

Received for publication, August 18, 2003, and in revised form, September 18, 2003
Published, JBC Papers in Press, October 14, 2003, DOI 10.1074/jbc.M309128200

C. Scott Swindle‡, Hyung G. Kim§, and Christopher A. Klug¶
From the Department of Microbiology, Division of Development and Clinical Immunology, University of Alabama at Birmingham, Birmingham, Alabama 35294

Although DNA methylation and transcriptional repression are generally associated, a causal role for DNA methylation in silencing of retroviral vectors has not been established. The newer generation murine stem cell virus retroviral vector (MSCV) lacks many of the repressive cis-acting DNA sequences identified in Moloney murine leukemia virus but remains sensitive to transcriptional silencing in various cell types. To determine the contribution of cytosine methylation to MSCV silencing, we mutated CpG dinucleotides located in the MSCV long terminal repeat (LTR) that are clustered in the U3 region and directly spanning the transcription start site in the R region. Effects of the CpG mutations on MSCV silencing were assessed in murine embryonic stem cells. An analysis of numerous clonal proviral integrants showed that mutation of CpGs in both clusters eliminated proviral integrants that were completely silenced. Variegated expression was shown to represent a substantial component of intracellular silencing and was independent of the presence of CpGs in the LTR. Treatment of transduced cells with 5-azadeoxycytidine delayed establishment of the silenced provirus but had only a modest effect on expression of some proviral integrants at late times post-transduction. These results are direct evidence for a causal contribution of DNA methylation in the LTR to MSCV silencing and define the promoter region CpGs as a repressive element in embryonic stem cells. Furthermore, distinct mechanisms are suggested for establishment and maintenance of the silenced proviral state.

Because of their ability to stably integrate into their host cell genomes, murine oncoretroviruses are among the most commonly used vectors for gene delivery into mammalian cells. Although retroviral vectors are used extensively for many applications including gene therapy, their utility remains limited by their sensitivity to transcriptional silencing by host cell-repressive mechanisms. Cytosine DNA methylation and deacetylation of nucleosomal histone proteins are two biochemical processes that are widely associated with transcriptional repression of both cellular genes and retroviruses (1–5). Cytosine methylation is directed by DNA methyltransferases primarily to CpG dinucleotides, although cytosine methylation in the context of CpNpG has been demonstrated in some cases (6, 7). When acetylated, specific lysine residues in the amino-terminal tails of histones H3 and H4 are substrates for deacetylation by specific histone deacetylases (HDACs)1 (8). The common association of hypoacetylated histones with methylated DNA and the presence of HDACs as components of transcriptional co-repressor complexes that specifically associate with methyl-CpG binding domain (MBD) proteins suggest a mechanism for targeting histone deacetylation to methylated DNA (9, 10).

Although reactivation of silenced provirus by demethylation with the cytosine methylation inhibitors 5-azacytosine or 5-azadeoxycytidine (collectively 5-azaC) is perhaps the strongest evidence for a causal role of DNA methylation in retroviral silencing, the effects of these agents on silenced provirus are often minimal (3, 11, 12). Furthermore, silencing of proviral expression in embryonic stem (ES) cells has been shown to occur prior to detection of methylated proviral DNA (13, 14) and was unaffected by null mutations in the de novo methyltransferase genes, dnm15a and dnm3b (2), which were shown to abolish methylation of newly integrated provirus (15). Thus, despite the general association between DNA methylation and retroviral silencing, their causal relationship has not been established.

Distinct mechanisms have been suggested for establishment and maintenance of silenced provirus. In mouse erythroblasts (MEL) cells, silenced rather than expressed provirus was hypermethylated and specifically associated with hypoacetylated histone H3 and MeCP2 (1, 7), the latter being a MBD protein known to associate with HDAC-containing co-repressor complexes (10). This finding suggests a methylation-dependent mechanism for retroviral silencing in MEL cells that involves specific recruitment of HDACs to methylated DNA through MBD proteins. Interestingly, at early times post-transduction, silenced provirus was sensitive to both 5-azaC and the HDAC inhibitor, trichostatin-A (TSA), but sensitivity to both agents diminished over time concomitant with the accumulation of cytosine methylation in proviral DNA (7). This suggests a dis-

1 The abbreviations used are: HDAC, histone deacetylase; 5-azadC, 5-azadeoxycytidine; ES, embryonic stem; FACS, fluorescence-activated cell sorting; GFP, green fluorescence protein; HSC, hematopoietic stem cell; LTR, long terminal repeat; MBD, methyl-binding domain; MEL, mouse erythroblast; MMLV, Moloney murine leukemia virus; MSCV, murine stem cell virus; TSA, trichostatin A; SNARF, seminaphthorhodafluor.
Methylation Dependence of Retroviral Silencing

Distinct mechanism for maintenance of the silenced proviral state in MEL cells that is associated with increased methylation density.

Some trans-acting repressive factors are known to recruit HDACs to DNA independently of MBD proteins or DNA methylation (16–18). Early studies with Moloney murine leukemia virus (MMLV) identified a number of repressive cis-acting elements in the LTR and untranslated region of the virus (19). These include binding sites for the embryonal LTR-binding protein (20) and the YY1 nuclear factor, the latter of which is known to associate with HDAC-containing transcriptional corepressor complexes (18). Although the newer generation MSCV retroviral vector investigated in this study lacks many of the known cis-acting repressive elements present in MMLV (21), it remains sensitive to transcriptional silencing in ES cells and hematopoietic stem cells (HSC) (2, 3, 22) and has been shown to harbor repressive cis-acting sequences in the LTR and untranslated region (23).

Repressive effects of DNA methylation have been reported to be most potent when located within or near promoter regions (24, 25). In the MSCV LTR, two major clusters of six and seven CpG dinucleotides are located within proximity of each other, completely within the U3 region and the other directly overlapping the transcription initiation site in the R region. To directly determine the contribution of DNA methylation in the promoter region of the LTR to retroviral silencing, we mutated the CpGs in each of these clusters independently or in toto and assessed effects of the mutations on silencing of MSCV expression in murine ES cells. An analysis of numerous single copy proviral integrants showed that mutation of CpGs in both clusters eliminated proviral integrants that were completely silenced, thus defining the promoter region CpGs as a repressive element in ES cells. Irrespective of the presence of CpGs in the mutated promoter region, many of the integrants exhibited variegated expression, which is defined by expression (or silencing) of provirus in a fraction of cells within a clonal population. Unlike completely silenced integrants, the variegated phenotype was shown to reflect a reversible transient state of silencing. However, similar to completely silenced clones, the frequency of silencing within each variegated clonal population was variable over time and an apparent property of the genomic site of integration. Treatment with 5-azadC delayed the kinetics of establishing the silenced proviral state but resulted in only modest, if any, reactivation of completely silenced or variegated clones at late times post-transduction. These results are direct evidence for a causal contribution of DNA methylation in the LTR to retroviral silencing and suggest distinct mechanisms for establishment and maintenance of the silenced state.

EXPERIMENTAL PROCEDURES

Retroviral Vectors—The green fluorescent protein (GFP) coding sequence optimized for human codon usage (26) was inserted into the MSCV retroviral vector (21) in place of the phosphoglycerate kinase-neomycin cassette. For introducing the CpG mutations into the 3'-LTR, primer spanning the XbaI site and 3'-primer containing the U3 and R region mutations spanning the SacI and KpnI sites, respectively. Amplified DNA was sequenced with the primer 5'-GATTTGAAAATGATTTTGTG-3' and 5'-TTTTTAGGGTGTTTTAAGGATTTG-3'.

RESULTS

Mutation of CpG Clusters within the U3 and R Regions of MSCV—The promoter region of the MSCV LTR contains two clusters of six and seven CpGs in the U3 and R regions, respectively (Fig. 1A). These CpGs account for all but a few of the CpGs present in the MSCV LTR. To determine the contribution of these sequences to MSCV silencing, MSCV vectors were constructed with the CpGs mutated in either cluster (M-U3 or M-R) or both clusters (M-U3/R) (Fig. 1B). Mutations in the R region were introduced such that base pairing was maintained in a stem-loop structure previously shown important for retroviral expression (Fig. 1B) (30). As a reporter for transcriptional activity, the coding sequence for GFP was placed under control of the LTR (Fig. 1C).

Distinct Expression Patterns among Individual Proviral Integrants—The murine embryonic day 14.1 ES cell line was transduced with the unaltered MSCV vector and each of the three MSCV mutants (M-U3, M-R, and M-U3/R). The MMLV-based vector, MFG (31), which contains additional known repressive cis-acting sequences in the U3 region, was included as a benchmark for retroviral silencing in this study. ES cells were transduced with limiting amounts of viral supernatants to minimize transductants containing multiple copies of integrated provirus. At 24 h post-transduction, transduced ES cells were sorted for proviral GFP expression by fluorescence-activated cell sorter (FACS) on a Becton Dickinson FACStar Plus cell sorter. For ES cell cloning, cells were sorted as single cells into separate wells of a 96-well plate.

Integration Site Analysis—Genomic DNA purified from ES cells clones was digested with BamHI, which is present as a unique restriction site in MSCV located 3' of the GFP-coding sequence, and analyzed for the presence of proviral DNA by Southern blotting using a radiolabeled GFP sequence probe.

Statistical Analysis—Individual integrants were categorized based on expression frequency into one of three categories: silenced (0–5% GFP-positive cells); variegated (5–75% GFP-positive cells); and sustained (75–100% GFP-positive cells). At each time point, the distribution of integrants of each vector between the three categories was compared with that of the unmodified MSCV vector by χ2 analysis (28). Confidence levels of difference (P) between distributions were determined based on the χ2 distribution and two degrees of freedom.

5-azadC and TSA Treatments—For treatments with 5-azadC and TSA, ES cells were cultured in complete media containing 5-azadC or TSA at concentrations of 50 and 10 nM, respectively. Media containing either agent were changed daily during treatments.

SNARF-1 Labeling of Cells—For tracking cell divisions, transduced ES cells were sorted for proviral GFP expression as described above, incubated in PBS containing 5 μM seminaphthorhodafluor-1 at 37 °C for 15 min (Molecular Probes, Eugene, OR), and re-introduced into culture following removal of free seminaphthorhodafluor-1.

Bisulfite Sequencing—Genomic DNA was isolated from ES cells, and bisulfite-treated as described by Clark et al. (29). The LTR was amplified from bisulfite-treated DNA by nested PCR using the following primer sets: first round, 5'-GGAAGAATGAGGAGATGAG-3' and 5'-CCTCAAATCTACCAATATGTT-3'; and second round, 5'-TTTTAGGGTGTTTTAAGGATTTG-3' and 5'-CCTCCAAAAATCAAATATGGTGTTC-3'.

SNARF-1 Labeling of Cells—For tracking cell divisions, transduced ES cells were sorted for proviral GFP expression as described above, incubated in PBS containing 5 μM seminaphthorhodafluor-1 at 37 °C for 15 min (Molecular Probes, Eugene, OR), and re-introduced into culture following removal of free seminaphthorhodafluor-1.

RESULTS

Mutation of CpG Clusters within the U3 and R Regions of MSCV—The promoter region of the MSCV LTR contains two clusters of six and seven CpGs in the U3 and R regions, respectively (Fig. 1A). These CpGs account for all but a few of the CpGs present in the MSCV LTR. To determine the contribution of these sequences to MSCV silencing, MSCV vectors were constructed with the CpGs mutated in either cluster (M-U3 or M-R) or both clusters (M-U3/R) (Fig. 1B). Mutations in the R region were introduced such that base pairing was maintained in a stem-loop structure previously shown important for retroviral expression (Fig. 1B) (30). As a reporter for transcriptional activity, the coding sequence for GFP was placed under control of the LTR (Fig. 1C).

Distinct Expression Patterns among Individual Proviral Integrants—The murine embryonic day 14.1 ES cell line was transduced with the unaltered MSCV vector and each of the three MSCV mutants (M-U3, M-R, and M-U3/R). The MMLV-based vector, MFG (31), which contains additional known repressive cis-acting sequences in the U3 region, was included as a benchmark for retroviral silencing in this study. ES cells were transduced with limiting amounts of viral supernatants to minimize transductants containing multiple copies of integrated provirus. At 24 h post-transduction, transduced ES cells were sorted for proviral GFP expression by FACS and deposited as single cells into separate wells of a 96-well plate. Clonogenic cultures derived from each transduced clone were monitored for GFP expression by flow cytometry at 2, 4, and 10 weeks post-transduction with some clones maintained for 16 weeks. The first time point at 2 weeks post-transduction was the minimal time required to obtain a sufficient number of cells...
Methylation Dependence of Retroviral Silencing

**Fig. 1.** CpG mutations introduced into the MSCV LTR. A. CpG dinucleotides (●) mutated in the M-U3, M-R, and M-U3/R vectors are indicated in brackets. The transcription start site is indicated by the arrow. B. Nucleotide sequence of the region in the MSCV LTR affected by the mutations. Top and bottom strands are indicated below the sequence. The two strands of the stem-loop structure are indicated by the boxes with the loop region in between. C. Complete proviral structure of the GFP-coding retroviral vectors.

in a single ES cell colony for flow cytometric analysis. Integration site analysis was performed on all of the clones by Southern blotting, and only clones containing a single proviral copy were included in the analysis (Fig. 2A). This allowed the transcriptional activity of each individual proviral integrant to be measured independently so that the absolute overall silencing frequency among individual integrants could be determined for each vector. The methylation status of the LTRs from nine MSCV integrants (17 total CpGs including those targeted for mutation) was determined by bisulfite sequence analysis (29) of proviral DNA at 10 weeks post-transduction. Significant levels of cytosine methylation were observed at each of the CpGs analyzed for all nine integrants.

At 2 weeks post-transduction, the frequency of GFP expression within each clonal population of cells varied among the integrants. Cells were considered GFP-positive if their GFP fluorescence intensity by FACs analysis was greater than that of non-transduced ES cells. Data for a representative set of MSCV integrants are shown in Fig. 2B where the expression pattern of each MSCV integrant over the course of the experiment is displayed. For some integrants, GFP expression was observed in either none or all of the cells within each clonal population, indicating complete silencing or maintenance of expression by those integrants, respectively (see Fig. 2B, MSCV clones 4 and 14). The remaining integrants (~40% for all of the vectors excluding MFG) exhibited variegated expression in that GFP was observed in only a fraction of the cells within each clonal population (see Fig. 2B, MSCV clone 6). This result is similar to that previously observed for a MMLV-based vector in a murine fibroblast cell line when several transduced clones were analyzed for proviral expression (11). For all of the vectors, the expression pattern and mean fluorescence intensity of GFP observed for each clone at 2 weeks post-transduction were maintained through subsequent time points with limited variability and no trends of diminishing percentages of GFP-positive cells were evident among individual clones (Fig. 2B and data not shown). Thus, the small degree of variability observed between time points for many of the integrants does not represent progressive silencing following the 2-week time point.

**Contribution of CpGs in the LTR to Retroviral Silencing**

The expression patterns of all of the integrants for each vector are displayed in Fig. 3, enabling comparison between vectors in their overall transcriptional activities. A difference was apparent between MSCV and the M-U3/R mutant in how individual integrants were distributed according to the percentage of GFP-positive cells within each clonal population. Individual integrants were categorized as silenced (0–5% GFP-positive cells), variegated (5–75% GFP-positive cells), or sustained (75–100% GFP-positive cells), and the overall silencing frequency of each vector was statistically compared with the unmodified MSCV vector at each time point as described under “Experimental Procedures” (Table I). The majority of MFG integrants were completely silenced with 24/32 MFG integrants (75%) completely silenced on at least two of the three time points. Only one MFG integrant (3%) exhibited a sustained pattern on two of the three points, and three (9%) displayed a variegated pattern on all three time points. For MSCV, 30 of 57 integrants (53%) displayed a sustained pattern on at least two time points. Nine integrants (16%) were completely silenced, and the remaining integrants were variegated. For the M-U3/R vector, 30 of 48 integrants (63%) exhibited a sustained expression pattern on at least two time points. Most importantly, no completely silenced integrants were observed for the M-U3/R vector at the 4- and 10-week time points and only two (4%) were observed at the 2-week time point. Statistical comparison between the M-U3/R and unmodified MSCV vectors by χ² analysis showed that the frequency of MSCV silencing was reduced significantly by mutation of CpG dinucleotides in both U3 and R regions combined (p < 0.05 at 4 and 10 weeks). These data are direct evidence for a causal contribution of cytosine methylation in the LTR to MSCV silencing in murine ES cells. Furthermore, failure of CpG mutations in the U3 or R region alone to reduce the silencing frequency to a significant extent suggests that methylation density across the LTR is important in MSCV silencing.
Variegated Expression: a Reversible and Transient Silenced State—The variegated expression patterns observed for many of the integrants at 2 weeks post-transduction was stable through subsequent time points with limited variability between time points in the fraction of GFP-positive cells within each clonal population (Fig. 2B). The variegated phenotype could represent an early irreversible silencing event, loss of provirus in a fraction of cells within each clonal population, or a transient state of silencing in which transcriptional activity of the same proviral integrant varies stochastically between cells. To distinguish between these possibilities, GFP-positive and negative cells were sorted from six variegated MSCV clones, re-plated separately into culture, and monitored for GFP expression for 20 days. An irreversible silencing event or loss of provirus in a fraction of cells in a clonal population would be indicated by complete maintenance of the silenced state in the sorted GFP-negative population. Conversely, a reversible transient state of silencing would be indicated by re-establishment of the original variegated pattern in both GFP-negative and positive populations. For each MSCV clone tested, variegated expression patterns were re-established from both GFP-positive and GFP-negative sorted populations within 7 days of sorting and were stable through 20 days. The data for MSCV clone six are displayed in Fig. 4. At day 7 following the sort, GFP-positive and negative cells were again sorted from the variegated cultures that were re-established
The silencing frequency among individual MSCV proviral integrants was significantly reduced by mutation of CpGs in the MSCV LTR

| Vector | Time | No. of silenced (%) | No. of variegated (%) | No. of sustained (%) | $P$ values (vs. WT) |
|--------|------|---------------------|----------------------|---------------------|-------------------|
| MFG    | 2 ws | 25 (78)             | 6 (19)               | 1 (3)               | 0.000             |
|        | 4    | 24 (75)             | 8 (25)               | 0 (0)               | 0.000             |
|        | 10    | 19 (59)             | 12 (38)              | 1 (3)               | 0.000             |
| MSCV   | 2    | 7 (13)              | 21 (38)              | 27 (49)             | 55 (1.000)        |
|        | 4    | 7 (12)              | 20 (35)              | 30 (53)             | 57 (1.000)        |
|        | 10   | 9 (16)              | 20 (35)              | 28 (49)             | 57 (1.000)        |
| M-U3   | 2    | 3 (5)               | 30 (46)              | 32 (49)             | 65 (0.247)        |
|        | 4    | 3 (5)               | 24 (37)              | 38 (55)             | 65 (0.393)        |
|        | 10   | 4 (6)               | 19 (29)              | 26 (45)             | 65 (0.120)        |
| M-R    | 2    | 9 (12)              | 35 (45)              | 33 (43)             | 77 (0.704)        |
|        | 4    | 7 (9)               | 36 (47)              | 33 (43)             | 76 (0.560)        |
|        | 10   | 11 (14)             | 34 (44)              | 32 (42)             | 77 (0.567)        |
| M-U3/R | 2    | 2 (4)               | 18 (30)              | 26 (47)             | 46 (0.326)        |
|        | 4    | 0 (0)               | 20 (42)              | 28 (58)             | 48 (0.042)        |
|        | 10   | 0 (0)               | 18 (38)              | 30 (63)             | 48 (0.015)        |

from the initial sort and re-plated separately as before. Again, variegated expression patterns were re-established in both GFP-positive and negative populations by 7 days and stable through 15 days (Fig. 4). Purity of the sorted populations was confirmed by FACs analysis immediately after the secondary sort (Fig. 4). These data show that the variegated expression patterns reflect a reversible transient state of silencing that is a property of the genomic site of integration. Furthermore, the large fraction of integrants exhibiting this pattern (~40% for all of the vectors excluding MFG) identifies variegated expression as a significant component of silencing observed in non-clonal retroviral transductions in undifferentiated ES cells.

Establishment and Maintenance of Retroviral Silencing: Dependence on DNA Methylation, Histone Deacetylation, and Replication—To determine the involvement of cytosine methylation or histone deacetylation in maintaining the completely silenced and variegated states, subsets of clones for each vector were treated at 10 weeks post-transduction with either 5-azadC or TSA for 4 days. Both 5-azadC and TSA were used at maximal concentrations, which in independent titration experiments (data not shown) resulted in greater than 75% cell viability and had no effect on the proliferation rate of ES cells for up to 8 days. For all of the vectors, only modest re-activation was observed for some integrants using either agent with changes in GFP expression that appeared within the small range of variability among time points. Data for the treated MSCV clones are displayed in Fig. 5 and are characteristic of the silenced state, these data suggest that the initiation of proviral silencing is a methylation-dependent process.

DNA methylation and chromatin-remodeling/modification are known to be replication-dependent processes. The gradual extinction of proviral silencing shown in Fig. 6 suggests that multiple rounds of replication are necessary for establishment of the silenced state. Alternatively, only one or two cell divisions (or rounds of replication) may be necessary to remodel the chromatin configuration of proviral DNA and silence expression. Those cells still expressing provirus in Fig. 6 may represent cells that had yet to transit one or two cell divisions, which would result from heterogeneity among ES cells in their proliferation rates. To directly correlate the kinetics of silencing with the number of cell divisions, MFG and MSCV-transduced ES cells were sorted for proviral GFP expression immediately after transduction and labeled with the fluorescent cell-tracking dye SNARF-1, which is spectrally distinct from the proviral GFP reporter. Transduced cells were monitored simultaneously for extinction of GFP expression and the number of completed cell divisions by flow cytometry (Fig. 7). Successive cell divisions are indicated by a 2-fold reduction in SNARF-1 fluorescence intensity. By tracking the SNARF-1 mean fluorescence intensity at consecutive time points, an average cell division rate of ~2.5/day was calculated for both transduced and untransduced cells. This proliferation rate was uniform for the entire ES cell population as indicated by the narrow distribution of SNARF-1 fluorescence intensity at each time point (Fig. 7). These data show that silencing of proviral expression continued to occur through at least 10 cell divisions following transduction and that one or two rounds of replication were not sufficient to silence all of the proviral integrants.

**DISCUSSION**

For both MFG and MSCV, cytosine methylation of CpGs in the LTR has been correlated with transcriptional silencing (7, 23, 32). The vast majority of CpGs present in the MSCV LTR are located within the two clusters targeted for mutation, one within the U3 region and the other directly spanning the transcription initiation site in the R region (Fig. 1A). The repressive potential of sequences in this region of the MSCV LTR was previously suggested by the ability of a restriction fragment containing these CpG clusters to repress expression of a cis-linked globin locus control region in transgenic mice (23). Because the CpGs in the MMLV LTR are similarly organized and

state is also independent of histone acetylation, assuming that histone acetylation was reduced by the TSA treatments. Nonetheless, the lack of response by the silenced and variegated clones to either 5-azadC or TSA was independent of the presence of CpGs (or methylation density) in the promoter region.

To specifically investigate a role for cytosine methylation or histone deacetylation in the establishment of proviral silencing, ES cells transduced with MFG or MSCV were treated with 5-azadC or TSA immediately after transduction and monitored for 8 days in the presence of 5-azadC or TSA for a loss of GFP expression. Analysis beyond 8 days was prohibited because of cumulative cytotoxic effects of the drugs. For both MFG and MSCV, progressive silencing was observed over the course of 8 days following transduction (Fig. 6). TSA had only modest effects on silencing of either vector. However, 5-azadC resulted in a significant reduction in the silencing of MFG between 4 and 8 days post-transduction (Fig. 6A) and had a modest effect on MSCV silencing (Fig. 6B). The reduction in silencing observed between 4 and 8 days post-transduction (Fig. 6A) was not an artifact of reduced numbers of cell divisions imparted by 5-azadC cytotoxicity, because the concentration used for 5-azadC was shown to have no effect on the proliferation rate of ES cells for up to 8 days (data not shown). In contrast to maintenance of the silenced state, these data suggest that the initiation of proviral silencing is a methylation-dependent process.
located outside of the region in the LTR where the repressive cis-acting elements were identified, these CpG clusters may be primarily responsible for the repressive activity that remains in MSCV. Repressive effects of cytosine methylation have been reported to be most potent when the methylation is located within or near the promoter region (24, 25), which was further precedent for investigating the repressive potential of CpGs in this region.

When individual integrants were categorized into one of three categories according to expression pattern (silenced, variegated, or sustained) and the vectors were compared by statistical analysis for distribution of their integrants among the three categories, the silencing frequency of MSCV was shown to be significantly reduced by mutation of CpGs in both clusters in the M-U3/R vector (Table I). Failure to reduce silencing to a similar extent by CpG mutations in either cluster alone is consistent with a silencing mechanism dependent on methylation density. These results also suggest that the reduced silencing frequency observed for the M-U3/R vector (Table I) did not result from disruption of cis-acting repressive elements in either region. Otherwise, the silencing frequency of MSCV would have been significantly lowered by mutation of just one of the CpG clusters. One of the major elements important for expression of C-type retroviruses is a conserved stem-loop structure located at the 5' end of the R region (30, 33). Mutations in this region have been shown to be tolerated as long as the base pairing in the stem-loop is maintained (33). Thus, the R region mutations were introduced accordingly (Fig. 1B). Because no significant loss of transcriptional activity was observed for the M-R or M-U3/R vector, the R region mutations did not likely

![Figure 4](image-url)

**Fig. 4.** Variegated expression patterns reflect a reversible and transient silenced state. GFP-positive and negative cells of MSCV clone 6 were sorted by FACS and re-plated separately into culture. The original variegated expression pattern (time 0 pre-sort) was re-established in both GFP-positive and negative cell populations within 7 days of culture (day 7 post-sort). Re-establishment of the variegated expression pattern was recapitulated in GFP-positive and negative cells after a secondary sort of the re-established variegated cultures at day 7. Purity of the sorted populations was confirmed by FACS analysis immediately after the secondary sort (time 0 post-sort). Percentages of GFP-positive cells are indicated above the gates. Nonviable cells were excluded by propidium iodide staining.
for establishment of the silenced state. Transduced ES cell clones were treated with 5-azadC or TSA for 4 days at 10 weeks post-transduction and analyzed for re-activation of proviral GFP expression by flow cytometry. Data are from a subset of MSCV clones selected for treatment and are representative of observations made for all of the other vectors. Each point represents the frequency of GFP expression from a single proviral integrant within a clonal population of cells.

FIG. 5. Dependence on methylation and histone deacetylation for maintenance of the silenced state. Transduced ES cell clones were treated with 5-azadC or TSA for 4 days at 10 weeks post-transduction and analyzed for re-activation of proviral GFP expression by flow cytometry. Data are from a subset of MSCV clones selected for treatment and are representative of observations made for all of the other vectors. Each point represents the frequency of GFP expression from a single proviral integrant within a clonal population of cells.

Fig. 5. Dependence on methylation and histone deacetylation for maintenance of the silenced state. Transduced ES cell clones were treated with 5-azadC or TSA for 4 days at 10 weeks post-transduction and analyzed for re-activation of proviral GFP expression by flow cytometry. Data are from a subset of MSCV clones selected for treatment and are representative of observations made for all of the other vectors. Each point represents the frequency of GFP expression from a single proviral integrant within a clonal population of cells.

For the MFG vector, initiation of silencing was delayed in the presence of 5-azadC at a dose that did not affect DNA replication (Fig. 6A), suggesting methylation dependence for establishment of the silenced state. This finding contrasts with the lack of response to 5-azadC observed for all of the vectors at 10 weeks post-transduction. These data are consistent with previous observations of MFG silencing in MEL cells in which sensitivity of silenced provirus to 5-azadC was restricted to early times post-transduction before accumulation of cytosine methylation in proviral DNA (7). Interestingly, 5-azadC had only a modest effect on the initiation of silencing of the MSCV vector (Fig. 6B), which could be because of the lower overall silencing frequency for MSCV in ES cells.

Specific and potent restrictions to retroviral expression have been described in undifferentiated stem cells from which silenced expression is inherited in differentiated progeny (3, 5, 36). Despite significant progress in development of retroviral vectors for gene therapy, evidence is scarce for any retroviral vectors that are completely resistant to silencing when present at low copy number in HSC-transplanted mice. Because the risk of insertional mutagenesis by retroviral integration is proportional to copy number, vectors with substantial resistance to silencing at low copy number should be critical for HSC retroviral gene therapy (37, 38). Further evaluation of the M-U3/R vector in HSC-transplanted mice may identify promoter CpG elimination as a general strategy for increasing resistance of retroviral vectors to silencing.

Most evidence for a causal role of DNA methylation in transcriptional repression has been correlative and indirect. By

would probably not have been re-established in the sorted GFP-negative and positive populations in such a rapid manner (7 days). This would require rapid kinetics of methylation and demethylation at these integration sites or the ability of small changes in methylation density to influence transcriptional activity. That the silencing observed for the variegated clones is reversible could suggest a silencing mechanism distinct from that responsible for the completely and irreversibly silenced clones. Alternatively, the same silencing mechanism could be responsible for both classes of integrants, only active at a greater degree in the completely silenced than the variegated integrants. In this context, the probability of expression from any given provirus would depend on the site of integration but would be greater for the U3/R vector at any given integration site. For example, the M-U3/R vector might exhibit variegated expression at sites that would otherwise be completely silenced for the wild-type MSCV vector.

In a previous analysis of a MMLV-based vector, a large fraction of variegated integrants was similarly observed in a murine fibroblast cell line (11). This suggests that variegated expression is responsible for a substantial component of retroviral silencing in multiple cell types. However, in that study, variegated expression was methylation-dependent based on reactivation of one integrant by 5-azacytidine treatment. The difference between this observation and our result could be attributed to the higher concentration of 5-azacytidine (4 μM) used in that study, which was beyond the concentration tolerable by ES cells (50 nM) in our study. The reversible nature of the silencing occurring in the variegated clones (Fig. 4) is similar to observations made for the interleukin-4 gene where expression of interleukin-4 from any given allele was shown to occur with a defined probability irrespective of its previous expression (34). Variegated expression is also characteristic of integration near or into pericentromeric heterochromatin where a “fluid” chromatin environment results from localized spreading of heterochromatin along the chromosome in these regions (35).

For the MFG vector, initiation of silencing was delayed in the presence of 5-azadC at a dose that did not affect DNA replication (Fig. 6A), suggesting methylation dependence for establishment of the silenced state. This finding contrasts with the lack of response to 5-azadC observed for all of the vectors at 10 weeks post-transduction. These data are consistent with previous observations of MFG silencing in MEL cells in which sensitivity of silenced provirus to 5-azadC was restricted to early times post-transduction before accumulation of cytosine methylation in proviral DNA (7). Interestingly, 5-azadC had only a modest effect on the initiation of silencing of the MSCV vector (Fig. 6B), which could be because of the lower overall silencing frequency for MSCV in ES cells.

Specific and potent restrictions to retroviral expression have been described in undifferentiated stem cells from which silenced expression is inherited in differentiated progeny (3, 5, 36). Despite significant progress in development of retroviral vectors for gene therapy, evidence is scarce for any retroviral vectors that are completely resistant to silencing when present at low copy number in HSC-transplanted mice. Because the risk of insertional mutagenesis by retroviral integration is proportional to copy number, vectors with substantial resistance to silencing at low copy number should be critical for HSC retroviral gene therapy (37, 38). Further evaluation of the M-U3/R vector in HSC-transplanted mice may identify promoter CpG elimination as a general strategy for increasing resistance of retroviral vectors to silencing.

Most evidence for a causal role of DNA methylation in transcriptional repression has been correlative and indirect. By
Methylation Dependence of Retroviral Silencing

**Fig. 7.** Correlation between the kinetics of proviral silencing and number of completed cell divisions. ES cells transduced with MFG or MSCV were sorted for proviral GFP expression immediately after transduction, labeled with the cell-tracking dye SNARF-1, and monitored simultaneously for extinction of proviral expression (GFP) and number of completed cell divisions (SNARF-1) by FACS. Nonviable cells were excluded by propidium iodide staining.

Mutating CpG dinucleotides in the LTR of MSCV and by quantitative evaluation of large numbers of single proviral integrants, we have directly assessed the contribution of DNA methylation to retroviral silencing. Our results provide direct evidence for a causal role of DNA methylation in retroviral silencing that is dependent on methylation density in the retroviral promoter region and the site of proviral integration.

**Acknowledgments**—We thank Miriam Spell for performing the FACS sorting and members of the Division of Developmental and Clinical Immunology for their support.

**REFERENCES**

1. Lorincz, M. C., Schubeler, D., and Groudine, M. (2001) Mol. Cell. Biol. 21, 7913-7922
2. Pannell, D., Osborne, C. S., Yao, S., Sukonnik, T., Pasceri, P., Kariakiskis, A., Okane, M., Li, E., Lipshitz, H. D., and Ellis, J. (2000) EMBO J. 19, 5884-5894
3. Cherry, S. R., Biniszewicz, D., van Parijs, L., Baltimore, D., and Jaenisch, R. (1994) Nature 369, 625-628
4. Challita, P. M., and Kohn, D. B. (1994) Proc. Natl. Acad. Sci. U. S. A. 91, 2567-2571
5. Clark, S. J., Harrison, J., and Frommer, M. (1995) Nat. Genet. 10, 20-27
6. Lorincz, M. C., Schubeler, D., Goede, S. C., Walter, M., Groudine, M., and Martin, D. I. (2000) Mol. Cell. Biol. 20, 842-850
7. Richards, E. J., and Elgin, S. C. (2002) Cell 108, 489-500
8. Jones, P. L., Veenstra, G. J., Wade, P. A., Vermaak, D., Kass, S. U., Landsberger, N., Strouboulis, J., and Wolff, A. (1998) Nat. Genet. 19, 187-191
9. Nan, X., Ng, H. H., Johnson, C. A., Loherty, C. D., Turner, B. M., Eisenman, R. N., and Bird, A. (1998) Nature 393, 386-389
10. Hobein, R. C., Mighielsen, A. A., van der Jagt, R. C., van Ormondt, H., and van der Eb, A. J. (1991) J. Biol. Chem. 266, 904-912
11. Stewart, C. L., Stuhlmann, H., Jahnner, D., and Jaenisch, R. (1982) Proc. Natl. Acad. Sci. U. S. A. 79, 4098-4102
12. Gautsch, J. W., and Wilson, M. C. (1993) Nature 361, 32-37
13. Niwa, O., Yokota, Y., Ishida, H., and Sugahara, T. (1983) Cell 32, 1105-1113
14. Okano, M., Bell, D. W., Haber, D. A., and Li, E. (1999) Cell 99, 247-257
15. Nagy, L., Kan, H. Y., Chakravarti, D., Lin, R. J., Haassig, C. A., Ayer, D. E., Schreiber, S. L., and Evans, R. M. (1997) Cell 89, 373-380
16. Laherty, C. D., Yang, W. M., Sun, J. M., Davie, J. R., Seto, E., and Eisenman, R. N. (1997) Cell 89, 349-356
17. Yang, W. M., Yao, Y. L., Sun, J. M., Davie, J. R., and Seto, E. (1997) J. Biol. Chem. 272, 28001-28007
18. Swindle, C. S., and Klug, C. A. (2002) J. Hematother. Stem Cell Res. 11, 449-456
19. Tsukiyama, T., Niwa, O., and Yokoro, K. (1989) Mol. Cell. Biol. 9, 4670-4676
20. Hawley, R. G., Lieu, F. H., Fong, A. Z., and Hawley, T. S. (1994) Gene Ther. 1, 136-138
21. de Guzman, C. G., Warren, A. J., Zhang, Z., Garland, L., Erickson, P., Drabkin, H., Hiebert, S. W., and Klug, C. A. (2002) Mol. Cell. Biol. 22, 5506-5517
22. Osborne, C. S., Pasceri, P., Singal, R., Sukonnik, T., Ginder, G. D., and Ellis, J. (1999) J. Virol. 73, 5490-5496
23. Curradi, M., Izzo, A., Badaracco, G., and Landsberger, N. (2002) Mol. Cell. Biol. 22, 3157-3173
24. Drabkin, H., Hiebert, S. W., and Klug, C. A. (2002) Mol. Cell. Biol. 22, 5517-5527
25. Pear, W. S., Nolan, G. P., Scott, M. L., and Baltimore, D. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 5260-5264
26. Anderson, M. T., Tjose, I. M., Lorincz, M. C., Parkes, D. R., Herzenberg, L. A., and Nolan, G. P. (1996) Proc. Natl. Acad. Sci. U. S. A. 93, 8508-8511
27. Pear, W. S., Nolan, G. P., Scott, M. L., and Baltimore, D. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 8392-8396
28. Daniel, W. W. (1995) Biostatistics, 6th Ed., John Wiley & Sons, Inc., New York
29. Clark, S. J., Harrison, J., Paul, C. L., and Frommer, M. (1994) Nucleic Acids Res. 22, 2990-2997
30. Trubetskoy, A. M., Okenquist, S. A., and Lenz, J. (1999) J. Virol. 73, 3477-3483
31. Dranoff, G., Jaffe, E., Lazenby, A., Golombek, P., Levitsky, H., Brose, K., Jackson, V., Hamada, H., Pardoll, D., and Mulligan, R. C. (1993) Proc. Natl. Acad. Sci. U. S. A. 90, 3539-3543
32. Lorincz, M. C., Schubeler, D., Hutchinson, S. R., Dickerson, D. R., and Groudine, M. (2002) Mol. Cell. Biol. 22, 7572-7580
33. Cupelli, L., Okenquist, S. A., Trubetskoy, A., and Lenz, J. (1998) J. Virol. 72, 7807-7814
34. Hu-Li, J., Pannetier, C., Guo, L., Lohning, M., Gu, H., Watson, C., Assemacher, M., Radbruch, A., and Paul, W. E. (2001) Immunity 14, 1-11
35. Milot, E., Strouboulis, J., Trumborn, T., Wijgerde, M., de Beer, E., Langeveld, A., Tan-Un, K., Vergeer, W., Yannoutsos, N., Grosveld, F., and Fraser, P. (1996) Cell 87, 105-114
36. Klug, C. A., Cheshier, S., and Weissman, I. L. (2000) Blood 96, 894-901
37. Li, Z., Fehse, B., Schiedeirer, B., Dullmann, J., Frank, O., Zander, A. R., Ostertag, W., and Baum, C. (2002) Leukemia 16, 1655-1663
38. Marshall, E. (2002) Science 298, 34-35