Volume 12 Number 13 1984 Nucleic Acids Research

Analysis of the variations in proviral cytosine methylation that accompany transformation and morphological reversion in a line of Rous sarcoma virus-infected Rat-1 cells

Siân Searle, David A.F.Gillespie[1], David J.Chiswell and John A.Wyke

Imperial Cancer Research Fund Laboratories, St. Bartholomew's Hospital, Dominion House, Bartholomew Close, London EC1A 7BE, UK

Received 17 May 1984; Accepted 12 June 1984

ABSTRACT

Cells of the A11 lineage of Rat-1 contain a single complete Rous sarcoma provirus. Variation in the activity of this provirus accompanies fluctuations in the lineage between normal and transformed phenotypes. Increased proviral cytosine methylation of the doublet CpG in the tetranucleotide CCGG correlates with transcriptional inactivity and this pattern of cytosine hypermethylation is stable, even when the cells are transformed by another virus. However, transformation can also be induced by 5-azacytidine (but not by other mutagens) and in these transformants reduced proviral cytosine methylation is accompanied by increased proviral transcription. Differences in CCGG methylation between normal and transformed cells are found mainly in the 3' half of the provirus; sites near and within the src gene are heavily methylated only when the provirus is transcriptionally inactive. On the other hand, both transformed and normal A11 derivatives show little, if any, cytosine methylation of CCGG sequences in and flanking the 5' portion of the provirus.

INTRODUCTION

Fluctuations in provirus transcription can determine morphological transformation and reversion in Rous sarcoma virus (RSV) infected rat cells. Such fluctuations are either spontaneous (1) or induced by fusion of infected to uninfected cells (2) and they provide a versatile model for studying transcriptional regulation of eukaryotic gene expression. We previously showed that reduced expression is associated with changes in the integrated provirus, as judged by decreased sensitivity to nuclease digestion, increased methylation and an increased propensity to be detached from nuclear "cage" preparations treated with restriction enzymes (3,5 and Dyson et al., submitted for publication). These three types of change are thought to reflect alterations in chromatin structure, but it is not known how they relate to one another, nor to the modulations in gene activity that they accompany.

Cytosine methylation is the most extensively studied of these phenomena and it is usually (but not always) decreased when a quiescent gene becomes

active (for recent reviews see 6,7) However, contention persists over whether altered cytosine methylation is a prerequisite for gene activation and, if so, whether methylation at specific sites or regions in and around the gene is crucial to this control. This lack of consensus reflects 1) the plethora of genes studied and the possibility that not all are affected in the same way by base modifications: 2) the gross alterations in cytosine methylation seen in certain genes and 3) the relatively crude means of analysing and and manipulating this phenomenon (6).

Our previous work hinted that RSV-infected rat cells might offer certain advantages in examining the role of cytosine methylation (3). We had isolated a large family of related clones of transformed or normal morphology, thus permitting extensive comparisons to be made between methylation and gene activity in a single lineage. Moreover, although the RSV provirus in morphologically normal lines was more methylated than in transformed clones, these differences seemed relatively subtle, suggesting that potential regulatory modulations in methylation might be identified without recourse to *in vitro* manipulation. We now report that provirus hypermethylation in a lineage of such cells correlates precisely with transcriptional inactivity. The region whose methylation is associated with transcriptional silence lies close to the *src* gene and not at the 5' end of the provirus where transcription is initiated.

MATERIALS AND METHODS

Cells

The genealogy of the cells used in this study is shown in Figure 1. They are all derived from a single clone, A11, a B77-infected Rat-1 cell which, although originally morphologically normal, quickly segregated transformed derivatives in culture (1,8). Stable single cell clones of both normal and transformed morphologies were derived from this line. These daughter clones were identified by Roman numerals with a suffix indicating their morphologies (N-normal or T-transformed). A further series of subclones was isolated from the transformed lines IVT and VIT. These granddaughter clones were identified by Arabic numerals, again with a suffix denoting morphology. Being the daughters of a transformed cell, the normal granddaughter clones are true morphological revertants. Transformed subclones were obtained from some normal lines after treatment with 5-azacytidine (Sigma, St.Louis, Mo.) (see below). These are identified by the prefix 'Aza'. Other transformed subclones were derived after infection with Kirsten murine sarcoma virus and

given the prefix 'MuSV'.

Cloning was performed by transferring single cells in suspension with a fine-drawn capillary pipette to individual microplate wells, where the morphology of developing colonies was monitored.

Analytical procedures

DNA purification, restriction enzyme digestion, gel electrophoresis, transfer and filter hybridization followed standard techniques as used in this laboratory (1,8). The probes used were derived from pSRA2, a cloned DNA provirus of the Schmidt-Ruppin strain of Rous sarcoma virus (9), as shown in Figure 4. Molecular cloning of the integrated B77 provirus in clone VIT and of the corresponding cellular integration site in uninfected Rat-1 are described in Gillespie et al., (submitted for publication).

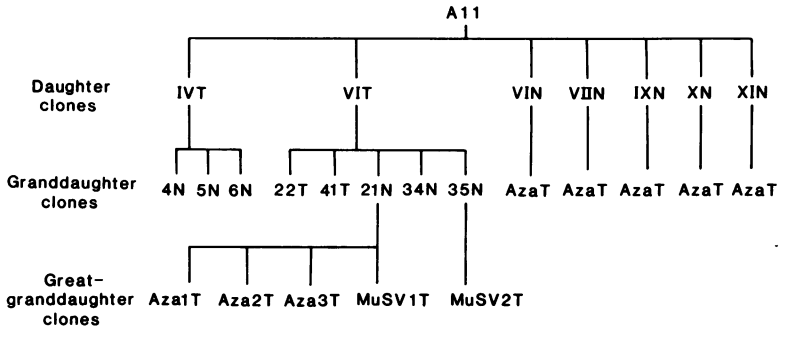
RESULTS

Provirus methylation in RSV-infected Rat-1 cells

The major modified base in eukaryotic DNA is 5-methylcytosine (5me-C), usually found in the doublet 5me-CpG, although a minority are found as 5me-CpC. One method of detecting the presence of this base is to use restriction endonucleases whose recognition sequence includes cytosine and that are sensitive to the presence of the modified base. Hpa II is one such enzyme which will digest at its recognition sequence (CCGG) only if the internal C is unmethylated. This is a particularly useful enzyme because it can be compared with its isoschizomer, Msp I, which will digest both CCGG and C5me-CGG, but not 5me-CCGG.

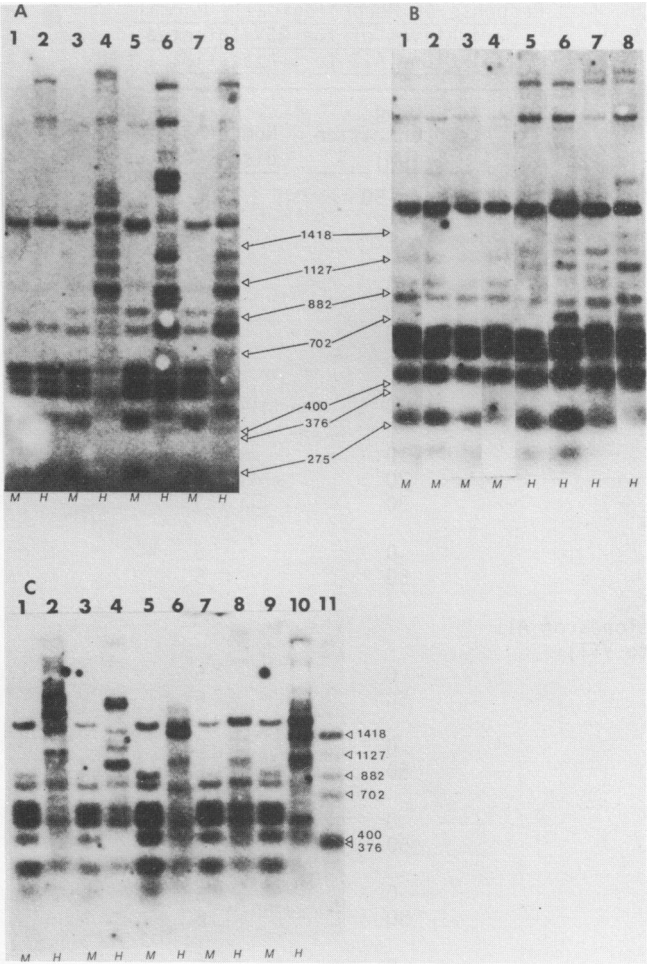
A limitation of these enzymes is that only one in sixteen of all CpG doublets are represented in the recognition sequence. Conversely, in a large DNA such as the 10 kilobase (kb) RSV provirus these enzymes cut frequently and many of the resultant fragments either comigrate or are too small to detect by transfer hybridization. Nonetheless, this technique can be used to estimate levels of methylation of the provirus in general and also methylation of more specific areas for which probes are available.

DNA was extracted from the cell lines derived from A11 (see Figure 1), digested with either Msp I or Hpa II, electrophoresed and analysed by Southern transfer, hybridizing with pSRA2, which recognizes the whole provirus. Figure 2, Panel A shows the digestion patterns observed for the transformed clone VIT and three of its revertant subclones, Panel B shows the transformed clone IVT and three revertant subclones, whilst Panel C



Legend to Figure 1. The Genealogy of the Derivatives of the B77-infected Rat-1 clone, A11. Daughter clones are identified by Roman numerals, granddaughters by Arabic numbers. Suffix N indicates normal morphology, T, transformed. The prefix Aza identifies transformed clones obtained after treating normal cells with 5-azacytidine, the prefix MuSV denotes those obtained after Kirsten murine sarcoma virus infection (see Results).

illustrates five morphologically normal siblings of both VIT and IVT that had probably never been transformed. The Msp I digests of all thirteen cell lines are similar whether morphology is transformed or normal (best seen in Panel B where the Msp I digests of transformed parent and revertant daughters are side by side). The only discernible variation in these Msp I digests is in the intensity of a fragment of just over 900 base pairs (bp). In contrast, the Hpa II digestions are very variable. Those of the transformed clones VIT and IVT (Panel A,lane 2 and Panel B,lane 5) are the same as their counterpart Msp I digests, with the exception of two high molecular weight fragments that are present in normal Rat-1 cells (data not shown) and presumably represent c-src sequences. However, Hpa II digests of all eleven morphologically normal clones show a disappearance or reduced intensity of some smaller fragments (800 bp or less) and the appearance,at variable intensities, of novel higher molecular weight bands. Each additional band represents at least one, and possibly more, recognition sequences of the enzyme that contains an internal 5me-C, their variation indicating that these methylated sites differ from clone to clone. Thus all the morphologically normal clones contain more methylated CpG doublets than their transformed brethren,but the level detected varies from the clone XN, where the only obvious novel fragment is one of about 1050 bp (Figure 2, Panel c, lane 8), to the revertants shown in Figure 1, Panel A, where there are a number of additional bands.



Legend to Figure 2. Patterns of proviral DNA methylation in normal and transformed All derivatives as revealed by paired digestion with either Msp I or Hpa II. All digested samples were electrophoresed and probed with nick-translated pSRA2, that recognises all proviral sequences and (depending on stringency) rat cellular src gene elements. Panel A: lanes 1 and 2, VIT; lanes 3 and 4, 35N; lanes 5 and 6, 34N; lanes 7 and 8, 21N; odd numbered lanes are DNA digested with Msp I, even numbered lanes show Hpa II digestions. Panel B: lanes 1 and 5, IVT; lanes 2 and 6, 4N; lanes 3 and 7, 5N; lanes 4 and 8, 6N; lanes 1 to 4 show Msp I digestions, lanes 5-8 Hpa II digestions. Panel C: lanes 1 and 2, VIN; lanes 3 and 4, VIIN; lanes 5 and 6, IXN; lanes 9 and 10, XIN; odd numbered lanes show Msp I digestions, even numbered lanes Hpa II digestions. Panel C, lane 11 shows marker DNA. The markers used throughout are Hpa II digested polyoma virus DNA. Their positions are shown by arrowheads with their sizes in base pairs (10).
For ease of reference italicized letters at the bottom of each panel identify Msp I digests (M) and Hpa II digests (H).

Table 1. Response of Morphologically Normal Derivatives of the RSV-Infected Rat-1 Clone A11 to 5-Azacytidine

	Drug concentration (μM)	Number of 90mm dishes tested	Average number of foci per dish
Uninfected Rat-1	5-50	56	< 0.02
Revertant subclones of VIT:			
21N	0	9	0.2
	5	2	14.0
	10	2	76.0
	20	2	93.5
	30	2	310.0
35N	0	14	0.07
	5	2	8.0
	10	3	7.6
	20	2	14.5
	30	3	19.0
34N	0	6	1.3
	50	5	5.7
Normal subclones of A11 (siblings to VIT):			
VIN	0	6	< 0.2
	50	6	18.0
VIIN	0	6	0.2
	50	5	5.6
IXN	0	6	< 0.2
	50	6	1.9
XN	0	6	< 0.2
	50	6	2.5
XIN	0	5	< 0.2
	50	6	6.5

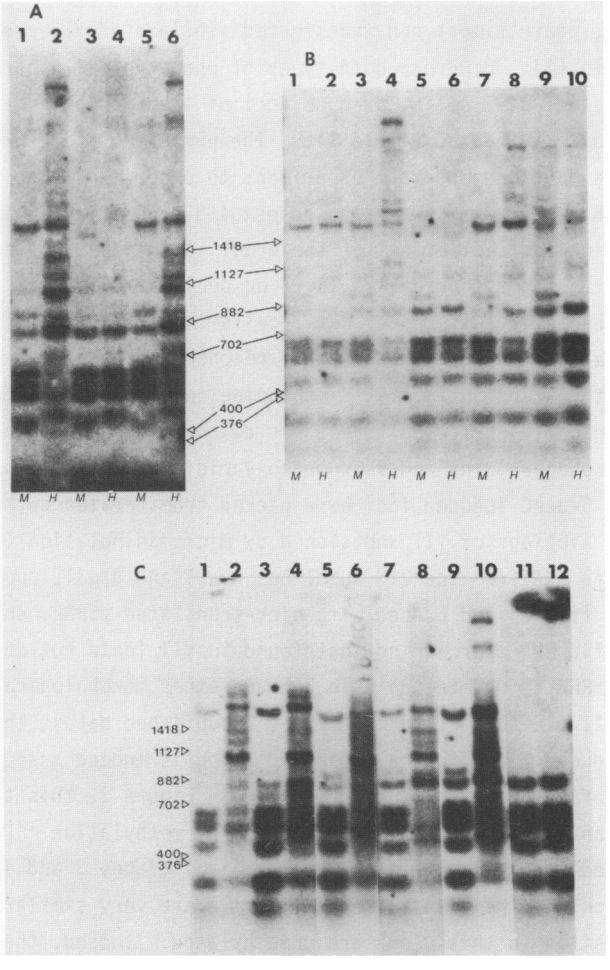
The role of methylation in provirus expression

The universal occurrence of provirus hypermethylation in cell lines that are not expressing the provirus genes does not prove that methylation causes this reduced expression. Accordingly, we tried to alter the level of methylation of the provirus using the analogue 5-azacytidine (5-azaC), that is incorporated into DNA in place of cytosine but is not recognized by the cell's methylation machinery. The effect, if any, of 5-azaC on provirus transcription was screened by testing its ability to induce transformation in morphologically normal cultures.

Morphological revertants of the clone VIT show a low level of spontaneous transformation, whereas most non-transformed siblings of VIT are morphologically more stable (1). However, both types of phenotypically normal cells, but not uninfected Rat-1 cells, produce foci of transformation a week after a 24 hour exposure to 5-azaC (Table 1). The numbers of foci produced vary from clone to clone and from experiment to experiment but, where tested, they increase with increasing doses of 5-azaC up to a maximum at 30 μM, suggesting a direct effect of the drug. At higher drug doses the incidence of transformation remains level or declines slightly, although the viability of monolayer cultures of Rat-1 cells is not seriously impaired until concentrations of 500 μM 5-azaC are reached (data not shown). Unlike 5-azaC, the methylation antagonist L-ethionine induces no transformation in these morphologically normal cultures, nor do the mutagens ethane methane sulphonate,nitrosoguanidine and 5-bromodeoxyuridine (data not shown)

Individual 5-azaC induced foci were picked from treated cultures of 21N and the normal siblings of VIT and cloned by micromanipulation (Figure 1). Total and polyA+ RNA from the 21N subclones Aza1T and Aza3T, spotted onto nitrocellulose filters and probed with nick-translated pSRA2, showed levels of virus-specific RNA approaching that found in VIT (data not shown). Since virus-specific RNA is undetectable in 21N and other morphologically normal subclones of A11 (1; D.J.C. and D.A.F.G., unpublished data), this suggests that 5-azaC induced morphological transformation is indeed associated with re-expression of the B77 provirus. As shown in Figure 3, this transformation is also accompanied by reduced proviral cytosine methylation. Panel A compares 21N (lanes 1 and 2) with its subclone Aza3T (lanes 3 and 4). The Msp I and Hpa II tracks of the transformed derivative are very similar, showing that the CCGG sites in this clone are unmethylated. Indeed, the disappearance of a fragment of about 1600 bp that is present with both Msp I and Hpa II digestion of VIT (Figure 2, Panel A, lanes 1 and 2), suggests that Aza3T DNA is even less methylated that that of its transformed grandparent. The independently derived Aza1T (Figure 3, Panel C, lanes 11 and 12) and Aza2T (data not shown) resemble Aza3T and it is interesting that in all three 21N derivatives the high molecular weight putative c-src fragments are also undetectable and presumably hypomethylated.

The 5-azaC induced transformants of the VIT siblings are also undermethylated in comparison to their normal progenitors (compare Figure 3,Panel B with Figure 2, Panel C). The 1600 bp fragment that disappears in the 21N transformants may be absent, as in IXN AzaT (Figure 3, Panel B, lanes 5 and

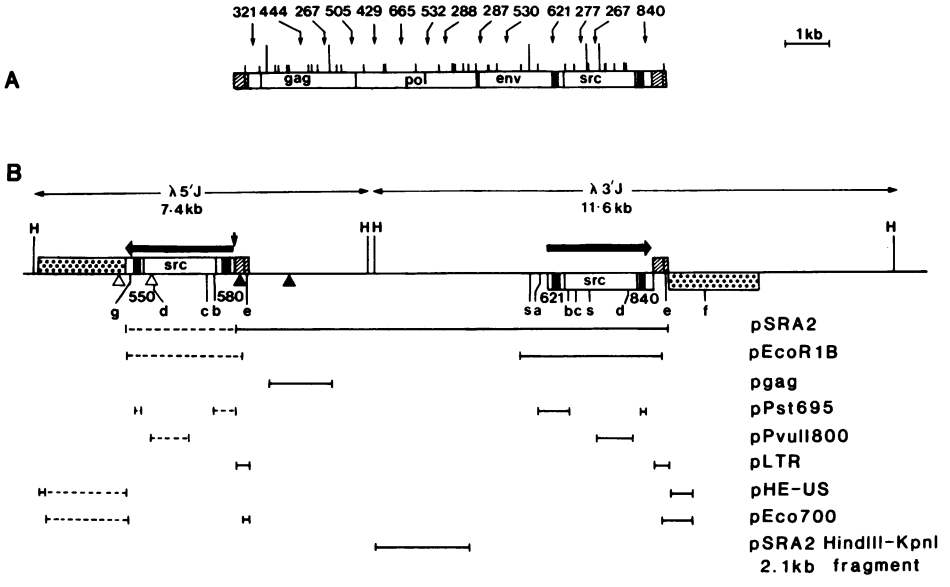


Legend to Figure 3. Proviral DNA methylation patterns in transformed clones derived from morphologically normal A11 derivatives after treatment with 5-azaC or infection with Ki-MuSV. The panels show paired digestions with Msp I (odd numbered lanes) or Hpa II (even numbered lanes), with probe and markers as in Figure 2. Panel A: lanes 1 and 2, 21N; lanes 3 and 4, 21N Aza3T; lanes 5 and 6, 21N 3 weeks after treatment with 30μM 5-azaC. Panel B: lanes 1 and 2, VIN AzaT; lanes 3 and 4, VIIN AzaT; lanes 5 and 6, IXN AzaT; lanes 7 and 8, XN AzaT; lanes 9 and 10, XIN AzaT. Compare these tracks in Panel B with the corresponding parental clones in Figure 2, Panel C. Panel C: lanes 1 and 2, 35N (pattern in October, 1981); lanes 3 and 4, 35N (pattern in March, 1982); lanes 5 and 6, KiMuSV transformed 35N (35N MuSV2T); lanes 7 and 8, 21N; lanes 9 and 10, KiMuSV transformed 21N (21N MuSVIT); lanes 11 and 12, 21N AzaIT.

6), or may persist, as in other transformants. However, VIIN AzaT (Figure 3, Panel B, lanes 3 and 4) seems to display a level of methylation that differs from VIIN (Figure 2, Panel C, lanes 3 and 4) only in the intensity, rather than in the complexity, of the higher molecular weight fragments in the Hpa II digest. Moreover, the low level of cytosine methylation at Hpa II sites already noted for the clone XN (Figure 2, Panel C, lanes 7 and 8) is not appreciably altered in the transformed subclone XN AzaT (Figure 3, Panel B, lanes 7 and 8). Thus, the correlation between 5-azaC-induced provirus expression and undermethylation, although strong is not absolute and it is possible that the dose-related transformation by 5-azaC reflects mechanisms other than reduction in cytosine methylation. Firstly, the analogue may both activate provirus expression and decrease provirus cytosine methylation, but by independent mechanisms. However, Figure 3, Panel A, lanes 5 and 6 show that a culture of 21N maintained for three weeks after treatment with a concentration of 5-azaC that induces maximal transformation retains the Hpa II digestion pattern of untreated 21N (Figure 3, Panel A, lanes 1 and 2). Thus, proviral hypomethylation is seen only in the minority of cells that become transformed (Panel A, lanes 3 and 4). The smaller fragments of 500-700 bp appear more prominent in the Hpa II digest in lane 6 than in lane 2, but over a hundred foci of transformed cells can appear in 21N after 3 weeks culture in 5-azaC (Table 1), and this could well explain these fragments against a background of otherwise hypermethylated DNA.

Secondly, 5-azaC induced transformation may be independent of provirus activation, and reduced provirus methylation may be a consequence rather than a cause of transformation. The failure of 5-azaC to transform uninfected Rat-1 argues against this hypothesis (Table 1), but the possibility that unrelated transforming events lead to changes in provirus methylation was investigated further by infecting both 21N and 35N with Kirsten murine sarcoma virus (KiMuSV). Both revertants are transformed by KiMuSV as readily as is normal Rat-1, presumably through the agency of the viral v-Ki-ras gene. Several such transformants were cloned and the methylation of their B77 proviruses examined (Figure 3, Panel C). Transformants of both 35N (lanes 5 and 6) and 21N (lanes 9 and 10) retained hypermethylation of their proviruses as in the parental revertants (lanes 1 to 4 and 7 and 8 respectively). Thus cell transformation on its own is not responsible for provirus hypomethylation.

These arguments and the data presented in Figures 2 and 3 strongly favour a causal connection between hypermethylation and reduced provirus activity.



Legend to Figure 4. Diagrams of proviral genome structures. Diagram A: The Prague strain RSV subgroup C provirus, showing coding assignment of genome and locations of restriction enzyme sites (from virus nucleotide sequence of reference 11). Hatched boxes denote the proviral long terminal repeat (LTR) sequences, solid boxes show the direct repeats that flank the src gene. Markers above the boxes show the Msp I/Hpa II recognition sequences (CCGG). Long markers show the subset of sequences that are also sites for Sma I restriction (CCCGGG). Numbered arrows locate fragments of more than 300 bp that would be generated by Msp I restriction. Many fragments of the same size and position also occur in the B77 provirus as determined by restriction mapping of cloned DNA using the panel of probes depicted in diagram B, but the only restriction site we have confirmed by direct DNA sequencing is the Msp I site in the U5 region of the LTR (Gillespie et al., submitted for publication).
Diagram B: The structure of the integrated B77 provirus in the A11 lineage (Gillespie et al., submitted for publication). A single complete integrated provirus is aligned with the diagram A RSV provirus. Immediately 5' to the integrated provirus are duplicated sequences that comprise 1) 3' proviral elements including part of the env gene, the direct repeats and the whole of src (open and solid boxes); 2) cellular sequences that originate from the region immediately 3' to the integrated provirus (stippled boxes). Some rearrangement has occurred in these duplicated sequences but the orientation of the viral sequences, as shown by the heavy arrows above the boxes, is reversed in the 5' duplication. The orientation of the duplicated cell sequences is not known. Note that the 3' proviral LTR is not duplicated. The point of recombination between the 5' LTR and the inverted viral sequences (vertical arrow) occurs between nucleotide 6840 in env and nucleotide + 26 in the U3 portion of the 5' LTR (Gillespie et al., submitted for publication). The point of recombination in the cellular sequences has not been determined. Above the diagram are shown the positions of the Hind III sites in and flanking the integrated provirus (H), cleavage at which generated fragments that were both separated for analysis of methylation (see

text) and cloned as 5' "junction" (5'J) and 3' "junction" (3'J) clones. Arrowheads below the diagram show previously determined sites of sensitivity to DNase I (3). Open arrowheads are sites found in both transformed and revertant cells, solid arrowheads are those found in transformed cells only. Also below the diagram are shown the position of Msp I sites determined by restriction mapping of cloned DNA. This mapping agrees with the sequence data of Schwartz et al., (11), so the precise location of the sites and the exact size of the fragments they generate are calculated from these data. Numbers indicate the major fragments discussed with Figure 5 and letters identify individual Msp I sites discussed in the text. Sites b to e are present both in "template" provirus and in the duplicated region; site f is the nearest site in 3' flanking DNA; site g does not have an identifiable counterpart at the 3' virus/host junction, so its existence must reflect rearrangement of the duplicated DNA; s indicates Sma I sites, investigated in the 3' half only. At the bottom of the diagram are shown the extent and the designation of subclones of pSRA2 and clones from VIT and normal Rat-1 used to produce probes. Solid lines show the sequences homologous to these probes in the integral provirus and cell "template" regions, pecked lines show regions in the 5' duplication that should show homology with the probes.

However, although cultural studies show that cell morphology, either normal or transformed, is a relatively stable phenotypic trait, it might be argued that proviral methylation patterns are very labile and only fortuitously correlate with cell morphology. This appears very unlikely for two reasons. 1) Upon subcloning cells (Figure 1) there is a very consistent correlation between proviral methylation patterns and the presumed or demonstrated levels of provirus activity (Figures 2 and 3, all panels). 2) Repeated assessments of the methylation pattern of the same cell clone remain similar as long as the phenotype is stable. This has been a consistent finding over the four years of this investigation and an example is shown in Figure 3, Panel C, lanes 1 to 4, in which it can be seen that the methylation of 35N is unchanged during five months of continuous culture.

Locating the methylated sites in morphologically normal derivatives of A11

Although the B77 provirus in normal derivatives of A11 is hypermethylated, even Hpa II digestion seldom yields virus-specific fragments of greater than 2 kb and most are considerably smaller than this (Figures 2 and 3). Since the integrated B77 provirus is nearly 10 kb in length this means that a number of proviral CCGG sequences must be unmethylated. Moreover, although both the size and intensity of Hpa II fragments differ in each independently-derived untransformed line, there are some common features, notably a band of about 1050 bp that is a major fragment in most clones. Similarity in size does not, of course, imply similarity in content, but taken together these considerations suggest that, unlike many other systems studied, provirus transcription in the A11 lineage, may be regulated by limited changes in

cytosine methylation. It may thus be possible to identify crucial methylated regions that operate in a naturally-occurring modulation of gene expression.

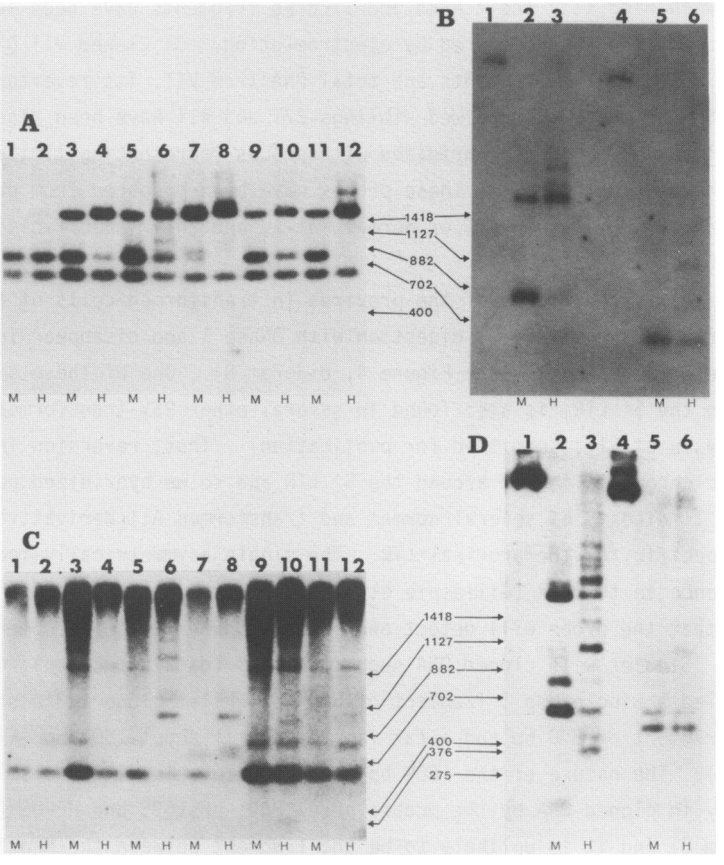
However, even in such a favourable system as these A11 derivatives, the identification of postulated regulatory regions meets with a number of problems. Firstly, although methylation-sensitive restriction enzymes should detect regions of, say, a couple of hundred base pairs in which most CpG residues are methylated, if only one or two methylated doublets are crucial their detection is unlikely. Conversely, the recognition sequence CCGG is far more frequent in RSV than in most eukaryote DNA; Prague C strain RSV (11) contains 42 of these sites (Figure 4, diagram A). B77 virus, although it has yet to be fully sequenced, is likely to share many of these sites with Prague C RSV, and it is thus clear that the gels shown in Figures 2 and 3 will not resolve many of these virus-specific fragments when probed with pSRA2. Thirdly, although the A11 series contains only one complete B77 provirus, the integration of this virus in the host DNA is not a simple insertion. In common with the majority of other B77 transformed Rat-1 clones that we have studied, the DNA flanking the integral provirus comprises duplicated and rearranged sequences of both virus and host cell origin. The structure of the provirus in the A11 lineage is summarized in Diagram B of Figure 4 and described fully in Gillespie et al.,(submitted for publication). Since proviral sequences are duplicated 5' to the integral provirus they,too, could contain methylation sensitive restriction enzyme sites that would be detected by virus-specific probes. This will complicate the interpretation of the investigations to be described, but it has the advantage of extending our analysis into DNA 5' to the integral provirus and, since 5' rearrangements are a frequent concomitant of B77 transformation of Rat-1, they probably play a role in this transformation (Gillespie et al.,submitted for publication).

Three approaches have been used to dissect this situation. 1) The integrated provirus in VIT has been digested with Hind III and cloned in bacteriophage λ (Gillespie et al.,submitted for publication). One clone contains a 7.4 kb insert that extends from the viral Hind III site at nucleotide 2740 into 5' flanking DNA, and includes all the 5' duplicated viral and host sequences. The sequences that are represented in these 5' duplications are contained in their un-rearranged "template" configuration in a second clone of 11.6 kb, extending 3' from the viral Hind III site at nucleotide 2870 (Figure 4, Diagram B). 2) DNA from the VIT revertant 21N has been similarly

digested with Hind III, the 7.4 kb and 11.6 kb fragments have been separated by electrophoresis and recovered by electroelution. 3) Cloned VIT DNA, separated 21N Hind III fragments and total DNA from VIT, its revertants 21N, 34N and 35N and their transformed siblings 22T and 41T have been digested with Msp I and Hpa II and hybridized with probes representing limited regions of the viral genome. These probes have been prepared from pSRA2, the integrated provirus in VIT or normal Rat-1, and their specificities are shown in Figure 4, Diagram B.

Two sites at the 5' end of the provirus in transformed cells of the A11 lineage are hypersensitive to digestion with DNase 1 and disappear in morphological revertants (3 and see Figure 4, diagram B). One of these sites, located in the 5' LTR, is also found in several other RSV-transformed Rat-1 clones (Dyson et al., submitted for publication). Thus, reversion involves changes in chromatin in and around the 5' LTR and so we hybridized paired Msp I/Hpa II digests of several normal and transformed A11 derivatives with a probe specific for the proviral LTR. The single asymmetrically-located CCGG sequence in the LTR (Gillespie et al., submitted for publication) predicts that the probe will detect one intense and one faint fragment from each LTR. Studies with cloned DNA supported this (data not shown). Clone λ 5'J revealed a single Msp I fragment of 580 bp, whilst clone λ 3'J showed an intense fragment of 840 bp and a far fainter one of about 1600 bp (Figure 4, Diagram B). The nature of the 1600 bp band is uncertain. It is also detected faintly in cloned DNA by the probes pPvuII800, pPst695 and HE-US (Figure 4, Diagram B) and it is unlikely to be the fragment between the Msp I site in the LTR (e in Figure 4, Diagram B) and the adjacent cell site, for the latter, as detected with the probe HE-US (that recognizes only cell sequences) is only 900 bp away (f in figure 4, Diagram B). We believe that the 1600 bp fragment is probably due to failure of digestion, for an unknown reason, at site e, generating a fragment spanning sites d to f (Figure 4, Diagram B)

Figure 5, Panel A shows the presence of these fragments in A11 lineage genomic DNA. The 5' 580 bp fragment varies little in the different clones, whichever enzyme is used. In contrast, the 840 bp 3' band varies in intensity in different clones and shows a consistently reduced intensity in digests with Hpa II, this being accompanied by a broadening of the 1600 bp fragment or the appearance of larger bands. However, the extent of this methylation is irrespective of cell phenotype, occurring in both normal and transformed siblings. Moreover, the 1600 bp fragment is far more obvious than in cloned DNA and is seen also after Msp I digestion (the



Legend to Figure 5. Proviral methylation patterns investigated with sub-genomic probes and on separated regions of the genome. Panels A and B are probed with nick-translated pLTR, Panels C and D with nick-translated pPst695 (see Figure 4, diagram B). Panels A and C: lanes 1 and 2, A11 before subcloning; lanes 3 and 4, VIT; lanes 5 and 6, 21N; lanes 7 and 8, 35N; lanes 9 and 10, 22T; lanes 11 and 12, 41T. Odd numbered lanes are digested with Msp I, even numbered lanes with Hpa II. Note the very faint 1600 bp fragments in A11 before subcloning. Panels B and D show 21N DNA: lanes 1-3, the separated 3' 11.6 kb Hind III fragment, undigested (lane 1) or digested with Msp I (lane 2) or Hpa II (lane 3); lanes 4-6 the separated 5' 7.4 kb Hind III fragment, undigested (lane 4) or digested with Msp I (lane 5) or Hpa II (lane 6). Note that in Panel D, lane 4 the undigested fragment contains some contaminating DNA that might account for some minor bands in lanes 5 and 6.

original A11 line is an exception; like some 5-azaC induced transformants in Figure 3, it lacks an obvious 1600 bp fragment). Digestion of separated 21N 5' and 3' Hind III fragments confirms these findings (Figure 5, panel B).

The 840 bp 3' band is heavily methylated on at least one of its flanking CCGG residues, generating bands of 1600 bp or more, whilst the 5' 580 bp fragment is only slightly methylated.

What CCGG methylation there is in and around the 5' LTR generates a faint band of about 1050 bp. However, the weak intensity of this band suggests that is is not the major novel 1050 bp fragment that results from proviral methylation (Figures 2 and 3). Probes for other portions of the provirus revealed partial methylation at some sites but likewise failed to locate the principal revertant-specific fragments (data not shown) until we used clone pPst695, derived from nucleotides 6500 to 7200 approximately. Mapping of cloned DNA (not shown) reveals four regions of homology with this probe (Figure 4, diagram B). In the "template" integral provirus it recognises, in its region of origin, a 620 bp Msp I fragment, with faint detection of flanking fragments of about 200 bp. Since the probe spans the 5' member of the direct repeat sequences that bracket src (solid boxes in Figure 4, diagram B), it also detects the 3' member of the repeat, located in the 840 bp fragment that is also detected by the LTR probe. Both these regions are found in the 5' duplication of proviral sequences and here the probe recognises the 580 bp fragment, also detected by the LTR probe, and a 550 bp band.

These fragments are seen in the genomic DNA of various cell clones as a broad, heavy band of about 600 bp and fainter bands at 840 bp and about 1600 bp (Figure 5, panel C). However, in the revertants 21N and 35N, novel fragments are seen, the most obvious at 1050 bp, with a concomitant reduction of the 600 bp complex. Analysis of the separated Hind III fragments of 21N (Figure 5, panel D) shows that most methylation is in the 3' fragment, with little in the 5' portion. Thus, the 3' bands of approximately 200 bp and the major 620 bp fragment are much fainter in the Hpa II digests, giving way to novel fragments of 400 bp or more, the most intense being at 1050 bp. Since digestion with Sma I, followed by hybridization with pPst695 (data not shown) indicates little, if any, methylation of the flanking Sma I sites (s in Figure 4, diagram B), the major methylated sites are probably those at a, b and c in Figure 4, diagram B (the fragments of more than 1600 bp probably originate from methylation of the 840 bp band: compare Figure 5, panels D and B). In contrast, the 5' Hind III fragment shows no evidence for methylation of the 550 bp band, and only slight methylation of the 580 bp band to yield a 1050 bp fragment in accordance with the data of Figure 5, panel B.

In conclusion, we find little if any methylation of the CCGG residues

b to e and g in and around the 5' end of the provirus, whereas the same and other sites in the 3' half of the provirus are markedly methylated (Figure 4, diagram B). The methylation specific to the revertants seems concentrated in a region near the 5' end of the src gene, as judged by methylation of the sites, a, b and c.

DISCUSSION

An association between hypermethylation of cytosine residues and gene inactivity is common in higher eukaryotes. A similar association is found in their viruses (12), including retroviruses (13-16) and has also been reported in studies on reversion of retrovirus transformed cells (3, 17, 18). Our findings conform to these general principles, but show some distinctive features. 1) Many CCGG sites in and around the provirus remain unmethylated even when the provirus is inactive. 2) Methylation levels show reversible fluctuations in successive clonal generations of the A11 lineage. 3) Sites in the "template" region of the provirus can be methylated, whilst apparently identical sites in the nearby duplicated proviral region remain unmodified. 4) Increased cytosine methylation specific to reduced provirus expression is not detected in the 5' region of the integrated provirus but occurs in the 3' half near the beginning of the src gene.

Trivial explanations for these findings are unlikely. Many methylated cytosine residues have probably gone undetected but our data suggest that regional rather than site specific changes in methylation are important. Not only do we detect CCGG methylation in blocks (as in the sites a, b and c shown in the 3' half of the provirus in Figure 4) but the same pattern is seen in two separate revertants (2IN and 35N; Figure 5) and the frequency of restriction fragments of common size suggests that the pattern is repeated in other clones containing inactive proviruses (Figs. 2 and 3). Such consistency in different clones also makes it unlikely that the alterations in methylation are unrelated to changes in gene activity. The data shown in Table 1 and Figure 3 provide further evidence for the importance of methylation in specific gene expression; the methylation antagonist 5-azaC, but not other mutagens, can activate RSV transformation in association with proviral hypomethylation, whereas transformation by an independent agent (KiMuSV) does not affect RSV methylation.

Comparisons between various inactive and active genes have shown different patterns of change in the levels and sites of methylation, for reasons that are not clear. The role of cytosine methylation may differ between

organisms and between genes in an organism (19) but, even if causally related to gene activity, the extent of methylation may vary with different patterns of gene expression. Thus, some regions of the genome may become methylated and perpetually silent (as with much X chromosome inactivation; 20). Other genes may be permanently active (21), at least in some cell lineages, whilst yet others are active at only certain times. A frequent finding with these latter two classes of genes is that hypomethylation in and around the 5' region of the gene is necessary, but not sufficient, for transcription and methylation of these regions in vivo or in vitro is associated with transcriptional silence (see, for example, 21, 22). The example we report here represents another category in which fluctuations in transcription do not occur from a stable hypomethylated region but are concomitant with reversible changes in methylation. This phenomenon has also been reported in other examples of reversion of virally transformed cells (17, 18) and it can be induced by fusing transformed to normal cells (Dyson et al., submitted for publication). It is thus possible that such modulations in methylation are common in eukaryote genomes, the variations in activity of randomly-integrated proviruses identifying regions where it occurs.

This lability of methylation levels is associated in our study with an unusual distribution of methylated residues, the 5' region of the transcriptional unit remaining hypomethylated. We do not know whether these two features are related but studies on in vitro methylation of cloned Moloney murine sarcoma viruses have shown that methylation of the v-mos gene inhibited transformation more effectively than methylation of the LTR. Such in vitro studies are subject to caveats, for although in vitro methylation may diminish gene activity this may not reflect in vivo patterns (22). Nonetheless, these findings are an interesting in vitro precedent for the distribution of methylation we observe in vivo in the A11 lineage.

Whether or not the distribution and low level of methylation of the A11 provirus reflects the reversibility of its inactivation, the mechanisms by which methylation occurs must distinguish sites in the "template" region of the provirus, that become methylated, from identical sites less than 10 kb upstream that remain unmodified. This discrimination, moreover, does not correlate with regions of altered chromatin structure as judged by changes in nuclease hypersensitivity, for hypersensitive sites that disappear upon morphological reversion are located in the 5' end of the provirus, whereas no such sites exist in the 3' half. The spatial separation of changes in methylation and nuclease sensitivity accords with the failure to show clear

mechanistic relationships between methylation and altered chromatin structure (24). It indicates that they are manifestations of distinct phenomena that are both important for gene expression, but may affect it independently.

ACKNOWLEDGEMENTS

We are grateful to Drs. M. Bissell and G. Peters for comments on this manuscript and to Mrs. I. Esmond and Mrs. J. Newton for typing.

[1]Present address: Amersham International Limited, White Lion Road, Amersham, Bucks HP7 9LL, UK

REFERENCES

1. Chiswell, D.J., Enrietto, P.J., Evans, S., Quade, K. and Wyke, J.A. (1982a) Virology 116, 428-440.
2. Dyson, P.J., Quade, K. and Wyke, J.A. (1982) Cell 30, 491-498.
3. Chiswell, D.J., Gillespie, D.A. and Wyke, J.A. (1982b) Nucl. Acids Res. 10, 3967-3980.
4. Cook, P.R., Lang, J., Hayday, A., Lania, L., Fried, M., Chiswell, D.J. and Wyke, J.A. (1982) Embo J. 1, 447-452.
5. Nicolas, R.H., Wright, C.A., Cockerill, P.N., Wyke, J.A. and Goodwin,G.H (1983) Nucl. Acids Res. 11, 753-772.
6. Bird, A.P. (1984) Nature 307, 503-504.
7. Trautner, T.A. (Ed.) (1984) Current Topics Microbiol. Immunol. 108.
8. Wyke, J.A. and Quade, K. (1980) Virology 106, 217-233.
9. De Lorbe, W.J., Luciw, P.A., Goodman, H.M., Varmus, H.E. and Bishop, J.M. (1980) J. Virol. 36, 50-61.
10. Soeda, E., Arrand, J.R., Smolar, N., Walsh, J.E. and Griffin, B.E. (1980) Nature 283, 445-453.
11. Schwartz, D.E., Tizard, R. and Gilbert, W. (1983) Cell 32, 853-869.
12. Doerfler, W., Kruczek, D., Eick, L., Vardimon, L. and Kron, B. (1982). Cold Spring Harbor Symp. Quant. Biol. 47, 593-603.
13. Groudine, M., Eisenman, R.and Weintraub, H. (1981) Nature 292, 311-317.
14. Stuhlmann, H., Jähner, D. and Jaenisch, R. (1981) Cell 26, 221-232.
15. Hoffman, J.W., Steffen, D., Gusella, J., Tabin, C., Bird, S., Cowing, D. and Weinberg, R.A. (1982) J. Virol. 44, 144-157.
16. Hu, W.-S., Fanning, T.G. and Cardiff, R.G. (1984) J. Virol. 49, 66-71.
17. Groffen, J., Heisterkamp, N., Blennerhassett, G. and Stephenson, J.R. (1983) Virology 126, 213-227.
18. Mathey-Prevot, B., Shibuya, M., Samarut, J. and Hanafusa, H. (1984) J. Virol. 50, 325-334.
19. Tanaka, K., Appella, E. and Jay, G. (1983) Cell 35, 457-465.
20. Shapiro, L.J. and Mohandas, T. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 631-637.
21. Stein, R., Sciaky-Gallili, N., Razin, A. and Cedar, H. (1983) Proc. Natl. Acad. Sci. USA 80, 2422-2426.
22. Busslinger, M., Hurst, J. and Flavell, R.A. (1983) Cell 34, 197-206.
23. McCready, M.L., Jhappan, C., Ascione, R. and Vande Woude, G.F. (1983) Mol. Cell Biol. 3, 305-314.
24. Felsenfeld, G., Nichol, J., Behe, M., McGhee, J. and Jackson, D. (1982) Cold Spring Harbor Symp. Quant. Biol. 47, 577-584.