JOURNAL OF VIROLOGY, May 1987, p. 1368–1374
0022-538X/87/051368-07$02.00/0

Vol. 61, No. 5

# Significance of DNase I-Hypersensitive Sites in the Long Terminal Repeats of a Moloney Murine Leukemia Virus Vector

JENNIFER A. RASMUSSEN[1,2]† AND ELI GILBOA[1]‡*

*Department of Molecular Biology, Princeton University, Princeton, New Jersey 08544,[1] and Department of Human Genetics, School of Medicine, Yale University, New Haven, Connecticut 06510[2]*

Recieved 27 August 1986/Accepted 14 January 1987

**A Moloney murine leukemia virus-derived retroviral vector (N4) carrying the bacterial neomycin resistance gene (*neo*) was used to study the chromatin configuration of integrated proviral DNA in NIH 3T3-derived cell lines containing one copy of the vector DNA per cell. Three independently obtained cell lines were examined. In two of these cell lines, the vector was introduced by viral infection, while in the third the construct was introduced by DNA transfection. Such transfected cell lines (including the one examined) usually express 10- to 50-fold less virus-specific RNA than do cell lines obtained by viral infection. All three cell lines exhibited similar patterns of DNase I-hypersensitive (HS) sites. Two strong DNase I HS sites were detected in the 5′ long terminal repeat, which contains signals required for proper and efficient initiation of viral transcription. One of these sites was found to overlap the viral enhancer sequences, while the other site mapped very close to the start site for viral transcription. A third HS site was detected in nearby internal viral sequences. Only one HS site was found in the 3′ long terminal repeat, which contains the signal(s) required for proper addition of a poly(A) tail to viral transcripts. This HS site was located in the region of the viral enhancer. Several weak DNase I HS sites were also found in the cellular sequences adjacent to the integration sites, at different locations in each cell line analyzed. No common pattern of cellular DNase I HS sites was found. These observations suggest that (i) the 5′ and 3′ long terminal repeats of integrated retroviral proviruses exhibit different chromatin conformations, possibly reflecting the different functions encoded by the otherwise identical sequences, and (ii) the DNase I HS sites detected in these studies reflect only a potential for transcription and are not a reflection of the high transcriptional activity characteristic of retroviruses.**

Numerous studies have demonstrated that active genes are contained in a chromatin conformation that is different from that of inactive genes (for reviews, see references 12, 28, and 48). One reflection of the structural differences between active and inactive chromatin is a differential sensitivity to digestion by DNase I. Genes that are actively being transcribed reside in a region of chromatin that is more sensitive to DNase I digestion than is bulk chromatin (14, 46). In addition, chromatin containing active or potentially active genes has been shown to possess short stretches (50 to 400 base pairs [bp] of DNA) that exhibit an even greater sensitivity to DNase I (11, 49). While these DNase I-hypersensitive (HS) sites have been found 3′ to or even within the body of certain genes (5, 35, 40), they are most often detected at the 5′ end of active genes (5, 23, 29, 35, 40, 47, 50). In many cases, the presence of DNase I HS sites in the 5′ region of active genes has been correlated with the expression of those genes (5, 29, 35, 47, 50), and in several instances DNase I HS sites have been found in or near sequences shown to be important in regulating the expression of ajacent genes (4, 32, 51). Although the precise nature of DNase I HS sites is unknown, the fact that regions of chromatin containing DNase I HS sites have occasionally also been shown to be sensitive to digestion by other nucleases (25, 29, 41) suggests that these regions of chromatin are in a particularly "relaxed" or "open" conformation.

* Corresponding author.
† Present address: Department of Human Genetics, School of Medicine, Yale University, New Haven, CT 06510.
‡ Present address: Department of Molecular Biology and Virology, Memorial Sloan-Kettering Cancer Center, New York, NY 10021.

Retroviruses provide a useful system for studying eucaryotic gene expression. As part of its replication cycle (reviewed in reference 45), the RNA genome of a retrovirus is reverse transcribed to produce a DNA copy which becomes stably integrated into the host chromosomal DNA of the host cell. Although DNA sequencing and restriction enzyme analysis have failed to provide evidence for site-specific integration (2, 18, 36), it remains possible that retroviruses integrate preferentially into active chromatin regions. Because of the unique mechanism of retroviral reverse transcription, the provirus contains a long sequence that is repeated at each end (long terminal repeats [LTRs]). Viral RNA transcription initiates in the 5′ LTR and terminates in the 3′ LTR (reviewed in reference 8). Thus the retroviral provirus represents a complete transcription unit containing two identical sequences which encode different functions. This fact, coupled with the fact that retroviruses are capable of integrating at a large number of sites in the chromosomal DNA of the host cell, makes the retrovirus-infected cell an interesting system for studying the chromatin structure of an actively transcribed set of genes. Several investigators have detected DNase I HS sites in the 5′ or 3′ LTRs of integrated retroviral proviruses (7, 9, 16, 33, 43, 44). In some of these studies, the presence of these HS sites has been correlated with the activity of the corresponding proviruses (7, 9, 16).

In these studies, a retrovirus-derived vector (N4) carrying the bacterial *neo* gene was introduced into NIH 3T3 cells by either DNA transfection or retroviral infection, and the presence of DNase I HS sites within the LTRs was investigated by using the technique of indirect end labeling (49). Cell lines containing one copy of the vector DNA per cell were analyzed and high-resolution DNase I mapping was achieved. Several questions were addressed: (i) do cellular

sequences at the proviral integration site have a characteristic chromatin configuration; (ii) do the functionally distinct 5′ and 3′ LTRs have similar or different chromatin configurations; and (iii) are differences in the transcriptional activity of individual viral templates reflected in the chromatin configuration of the 5′ LTRs?

## MATERIALS AND METHODS

**Cells.** TX-1 is a line of NIH 3T3 cells transfected with the recombinant retroviral vector N4. IF-1 and IF-2 are two independently derived cell lines obtained by infection of NIH 3T3 cells with N4 virus. All cells were grown in Dulbecco modified Eagle medium containing 10% fetal bovine serum.

**Isolation of nuclei.** Cells were trypsinized, washed three times with 1× SSCP (0.15 M NaCl, 0.015 M cacodylic acid, pH 7.2), and lysed in 1× SSCP containing 0.25% Nonidet P-40. Nuclei were pelleted and then washed once with 1× SSCP.

**DNase I digestions.** Nuclei were suspended in DNase I digestion buffer {10 mM PIPES [piperazine-*N*,*N*′-bis(2-ethanesulfonic acid)], pH 7.0, 0.17 M KCl, 11% sucrose, 3 mM $MgCl_2$} at 10 $A_{260}$ units per ml. DNase I (Sigma Chemical Co.) was added to various concentrations, and the reactions were incubated at 37°C for 25 min. The reactions were terminated by the addition of EDTA to 20 mM, proteinase K to 100 μg/ml, and sodium dodecyl sulfate to 0.5%. The samples were incubated at 37°C overnight and then subjected to three phenol-chloroform extractions followed by two chloroform extractions. RNase A was added to 10 μg/ml and the samples were incubated at 37°C for 45 min. High-molecular-weight DNA was purified by two phenol-chloroform extractions, one chloroform extraction, and ethanol precipitation.

**Gel electrophoresis and Southern blot hybridization.** High-molecular-weight DNA was digested by appropriate restriction endonucleases, fractionated on agarose gels, and then transferred to nitrocellulose filters as described by Southern (38). Uniformly labeled $^{32}P$ probes were prepared by nick translation (34). Blots were prehybridized for several hours at 42°C in 50% formamide–5× SSPE, 5× Denhardt solution (10)–100 μg of denatured salmon sperm DNA per ml. Hybridizations were carried out overnight at 42°C in a fresh aliquot of the same solution containing $^{32}P$-labeled probe at a concentration of $10^6$ cpm/ml. Blots were washed several times with 2× SSC (SSC = 0.15 M NaCl plus 0.015 M sodium citrate)–0.1% sodium dodecyl sulfate at room temperature, once with the same solution at 68°C, and once or twice with 0.2× SSC–0.1% sodium dodecyl sulfate at 68°C. Autoradiography was carried out at −80°C with one intensifying screen.

**Isolation of cytoplasmic RNA and Northern blot hybridization.** Cells were lysed with 10 mM Tris (pH 7.5)–10 mM NaCl–3 mM $MgCl_2$–0.5% Nonidet P-40. The concentration of NaCl was adjusted to 0.2 M, and nuclei were pelleted. The supernatant was extracted twice with equal volumes of phenol-chloroform-isoamyl alcohol (1:1:0.02) and RNA was ethanol precipitated. Poly(A)$^+$ mRNA was purified by oligo(dT)-cellulose chromatography (1), fractionated on formaldehyde-agarose gels (26), and subsequently transferred to nitrocellulose membranes as described by Thomas (42). Prehybridizations, hybridizations, and washings were performed as described above for Southern blot hybridization.

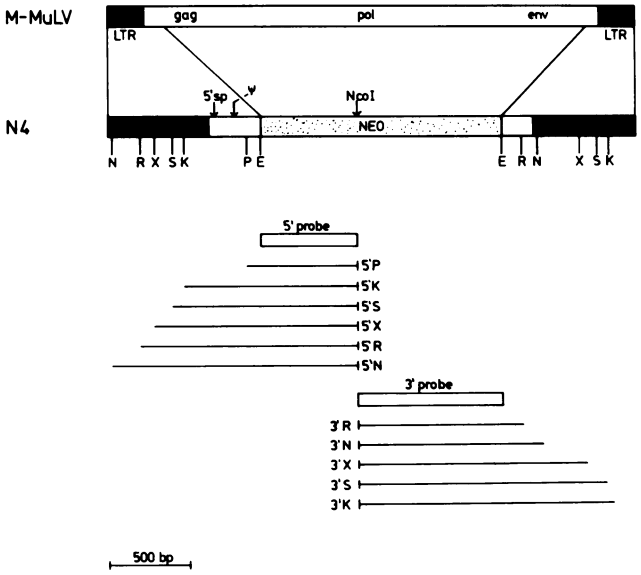


FIG. 1. Structure of M-MuLV-derived vector N4 and DNase I mapping strategy. The genomic organization of M-MuLV is shown at the top (for review, see reference 8). In N4, most of the M-MuLV-coding sequences have been replaced with the bacterial *neo* gene. The 5′ M-MuLV-derived portion of N4 includes an LTR and approximately 350 bp of adjacent downstream sequences. Thus the vector retains the 5′ donor splice site (5′sp) and the viral packaging signal (ψ). The 3′ M-MuLV-derived portion of N4 includes an LTR and 190 bp of adjacent upstream sequences. The *Nco*I site of the *neo* gene was used as the reference restriction site for all mapping experiments. The 0.6-kb *Eco*RI-*Nco*I fragment of the *neo* insert was used as a hybridization probe to detect HS sites in and around the 5′ LTR. The 0.9-kb *Nco*I-*Eco*RI fragment of the *neo* insert was used to detect HS sites in the region of the 3′ LTR. Hybridizing size markers were generated by digesting N4 DNA with *Nco*I plus one of several other restriction enzymes. The positions in the N4 provirus of these restriction sites are indicated. In cases where more than one recognition site for a given restriction enzyme occurs in the provirus, only that fragment which is recognized by either the 5′- or the 3′-specific probe is included in the set of fragments drawn under that probe. Abbreviations: N, *Nhe*I; R, *Eco*RV; X, *Xba*I; S, *Sac*I; K, *Kpn*I; P, *Pvu*I; E, *Eco*RI.

## RESULTS

Figure 1 shows the structure of the Moloney murine leukemia virus (M-MuLV)-derived vector (N4) used in these studies. In this vector, most of the internal M-MuLV sequences have been removed and have been replaced by the bacterial *neo* gene derived from Tn*5* (22). The structures of similar vectors carrying the *neo* gene (37) and other selectable genes (3, 27, 30, 31) were previously reported.

N4 DNA was used to transfect NIH 3T3 cells or was first converted to virus (by transfection followed by infection with wild-type helper M-MuLV [19]) and then used to infect the same cells. Cells harboring the N4 DNA and expressing the *neo* gene were selected with G418, and independent colonies were expanded to cell lines. Several cell lines obtained by transfection (TX) or infection (IF) were analyzed for the intactness and number of copies per cell of vector DNA. Figure 2 shows a DNA blot analysis of three cell lines: TX-1, which was obtained by DNA transfection, and IF-1 and IF-2, which were obtained by infection. All three cell lines were found to contain one intact copy of the vector DNA per cell.

DNase I HS sites were mapped by the indirect end-labeling method of Wu et al. (49). In this procedure, nuclei are digested lightly with DNase I, and then high-molecular-weight DNA is isolated and redigested with a "reference" restriction enzyme chosen so that the region containing the putative HS site(s) lies close to, but does not itself contain, a recognition site of that restriction enzyme. When a Southern blot containing such doubly digested DNA is hybridized with a short probe prepared from DNA sequences lying very close to the reference restriction site, subbands may appear. Each subband is defined on one side by a DNase I-generated double-stranded cut in the chromosomal DNA and on the other side by the reference restriction site. Thus, the size of a DNase I-generated subband can be used to estimate the distance between the reference restriction site and the DNase I HS site.

The mapping strategy used in these studies is illustrated in Fig. 1. *Nco*I, which cuts near the middle of the 3.3-kilobase (kb) N4 recombinant provirus (within the *neo* gene) was chosen as the reference restriction enzyme for the mapping of DNase I HS sites in both the 5′ and 3′ LTRs. A 0.6-kb DNA fragment prepared from sequences lying immediately upstream from the *Nco*I site was used as a hybridization probe to map DNase I HS sites in the 5′ portion of the provirus. HS sites in the 3′ portion of the provirus were detected with a 0.9-kb probe prepared from sequences immediately downstream from the *Nco*I site. Since both probes consist entirely of *neo* sequences, they do not cross-hybridize with helper M-MuLV or with endogenous retroviral sequences; i.e., they are specific for the newly introduced N4 provirus. To facilitate determining the location of DNase I HS sites within the N4 provirus, a set of size markers was generated (Fig. 1). Each marker was prepared by digesting N4 plasmid DNA with *Nco*I (the reference restriction enzyme) and then with a second restriction enzyme that cuts in the region of the 5′ or 3′ LTR or both. The marker DNAs were electrophoresed and transferred to nitrocellulose in parallel with the *Nco*I DNase I-digested genomic DNAs, and the resulting blots were hybridized with either the 5′- or the 3′-specific probe. The position of a particular *Nco*I DNase I subband relative to the marker bands allowed for the mapping of that DNase I HS site relative to the marker restriction sites.

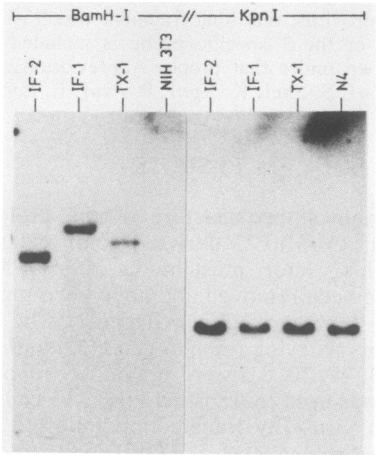


FIG. 2. Southern blot analysis of DNA in cells transfected or infected with vector N4. Cell line TX-1 was derived from transfection of NIH 3T3 cells with plasmid N4. Cell lines IF-1 and IF-2 were derived from infection of NIH 3T3 cells with N4 virus. Cellular DNA was digested with *Bam*HI, which has no recognition site within N4, and hybridized with a $^{32}$P-labeled *neo*-specific probe to determine the number of copies per cell of the newly introduced vector DNA. Cellular DNA and N4 plasmid DNA were digested with *Kpn*I and hybridized with a $^{32}$P-labeled *neo*-specific probe to confirm the intactness of the integrated vector DNA. *Kpn*I has two recognition sites within the LTRs, as shown in Fig. 1.

**Mapping DNase I HS sites in the 5′ and 3′ LTRs of cells infected with N4 virus.** Nuclei prepared from two independently derived N4-infected cell lines, IF-1 and IF-2, were incubated with various concentrations of DNase I. High-molecular-weight DNA was purified, cleaved with *Nco*I, and electrophoresed on agarose gels. Several size markers (described in the legend to Fig. 1) were run in parallel. The DNA was transferred to nitrocellulose and the resulting blot was hybridized with the 0.6-kb 5′-specific probe described in the legend to Fig. 1. The results of this analysis are shown in Fig. 3 (top). The major band in the first few lanes of each cell line corresponds to a DNA fragment defined on its 3′ side by the *Nco*I site of the N4 *neo* gene and on its 5′ side by an *Nco*I site in flanking cellular DNA. That this fragment is approximately the same size in IF-1 and IF-2 suggests that in each cell line the provirus has by chance integrated approximately the same distance downstream from a cellular *Nco*I site (evidence that they are two independent integration events will be presented shortly). As expected, the intensity of this band decreases as the concentration of DNase I increases (except for the 1.6-μg/ml lane of IF-1, where less DNA was loaded onto the gel). Several lanes in Fig. 3 contain bands of higher molecular weight than the major *Nco*I fragment. The origin of these bands is unclear and may represent cellular cross-hybridizing sequences or partially single-stranded DNA molecules, since they are predominant after digestion with DNase I. Several DNase I-generated subbands were also detected in both cell lines. Cell line IF-1 yielded two prominent subbands which began to appear at DNase I concentrations of about 0.4 or 0.8 μg/ml and became most visible at a DNase I concentration of 3.2 μg/ml (Fig. 3, top, HSI and HSII). At a DNase I concentration of 6.4 μg/ml, the slower migrating of these two subbands had almost disappeared, while the faster-migrating subband remained. A third subband began to appear at approximately 3.2 μg of DNase I per ml and became most visible at about 6.4 μg/ml (HSIII). According to the size markers described in the legend to Fig. 1, HS site I lies in the 5′ LTR between the *Nhe*I site and the *Xba*I site (overlapping the *Eco*RV site), HS site II maps near the *Sac*I site of the 5′ LTR, and HS site III lies between the *Kpn*I site of the 5′ LTR and a downstream *Pvu*I site. The results for cell line IF-2 (Fig. 3, top) were very similar to those obtained for cell line IF-1, except that the DNase I concentrations at which each HS site was detected were slightly different. These differences may be real or, more probably, they may merely reflect experimental variations encountered when DNase I digestions are performed on different preparations of nuclei.

Figure 3 (bottom) shows the mapping of DNase I HS sites in the 3′ LTRs of the two N4-infected cell lines, using the 0.9-kb probe shown in Fig. 1. The major band in the first several lanes of each cell line corresponds to a DNA fragment defined on its 5′ end by the *Nco*I site in the *neo* gene and on its 3′ end by an *Nco*I site in the downstream cellular sequences. That the mobility of this band was different for IF-1 and IF-2 (unlike in Fig. 3, top) confirms that the two cell lines were derived from independent integration events. Upon limited digestion with DNase I, a major HS site (HSIV) was detected in both IF-1 and IF-2. According to the size markers described in the legend to Fig. 1, this site

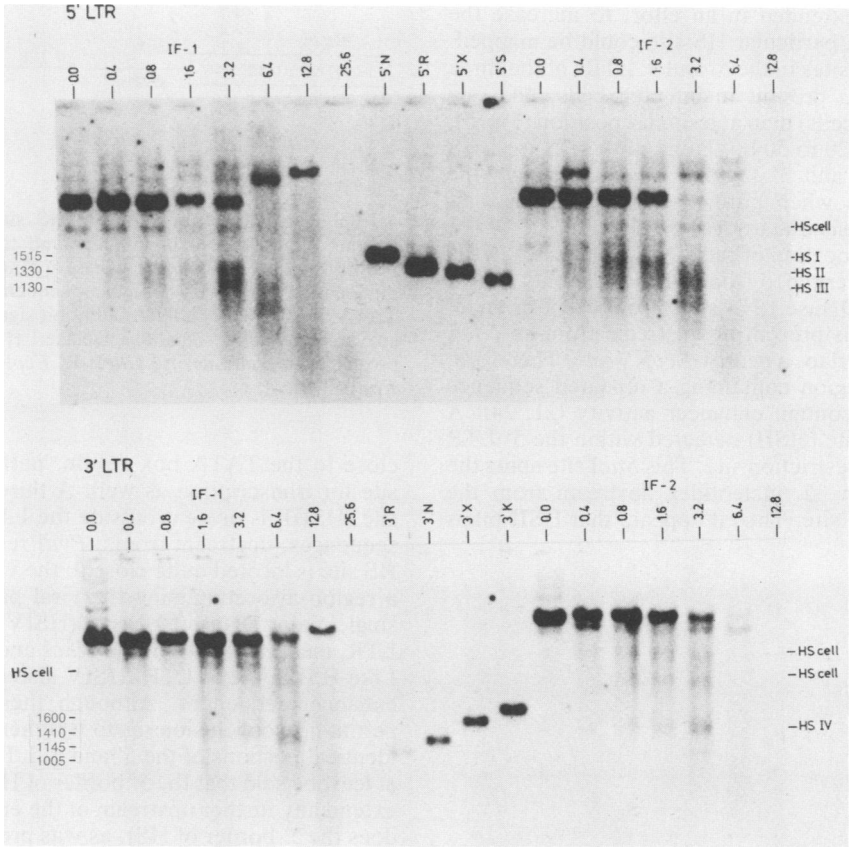


FIG. 3. Detection of DNase I HS sites in the 5′ and 3′ portions of N4 in infected cells. IF-1 and IF-2 nuclei were digested with increasing concentrations of DNase I (0 to 2.56 μg/ml, as indicated). High-moleuclar-weight DNA was purified, digested with *Nco*I, and subjected to Southern blot hybridization analysis, using either the 5′- or the 3′-specific probe described in the legend to Fig. 1. Hybridizing size markers (also detected in the legend to Fig. 1) were run in parallel and length (base pairs) is indicated. (Top) DNase I HS sites in and around the 5′ LTR of N4. Several DNase I-generated subbands (HSI, HSII, HSIII) were detected in the 5′ portion of the vector in each cell line. Cell line IF-2 contains an additional DNase I-generated subband (HS cell), located in upstream cellular sequences. (Bottom) DNase I HS sites in and around the 3′ LTR region of N4. A single DNase I-generated subband (HSIV) was detected in each cell line. In addition, several downstream cellular HS sites (HS cell) were observed (at different positions for each cell line).

maps within the 3′ LTR, between the *Nhe*I and *Xba*I restriction sites.

It should be noted that, in addition to the provirus-containing HS sites described above, each cell line yielded additional DNase I-generated subbands (denoted HS-cell in Fig. 3). These subbands represent cellular DNase I HS sites occurring just outside the proviruses, at different positions for each cell line.

**HS sites in the 5′ and 3′ LTR regions of N4 in transfected cells.** It has been demonstrated by our laboratory (19) and by others (31) that retroviral constructs are more efficiently expressed when they are introduced into cells by viral infection as compared to calcium phosphate-mediated DNA transfection. Figure 4 shows Northern blot analysis of cytoplasmic poly(A)$^+$ RNA prepared from cell lines that have been either infected (IF-1 and IF-2) or transfected (TX-1) with the retroviral construct N4. The infected cell lines clearly contain 5- to 10-fold more vector-specific RNA than does the transfected cell line. To determine whether the reduced transcriptional activity of the transfected DNA is manifested at the level of chromatin configuration, the DNase I sensitivity of the 5′ and 3′ LTRs was determined and compared with that of cell lines IF-1 and IF-2. For this analysis, a higher percentage agarose gel was used and

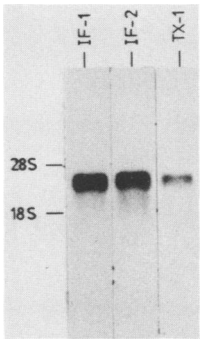


FIG. 4. Cytoplasmic RNA in cell lines obtained either by infection or transfection with retroviral vector N4. Cytoplasmic poly(A)$^+$ RNA isolated from infected cell lines IF-1 and IF-2 and from transfected cell line TX-1 was fractionated on a 1% agarose formaldehyde gel, transferred to nitrocellulose, and hybridized with a $^{32}$P-labeled *neo*-specific probe. rRNA was run in parallel for size comparison. The mobilities of 28S (5 kb) and 18S (2 kb) rRNA are indicated.

separation time was extended in an effort to increase the accuracy with which a particular HS site could be mapped. All major DNase I HS sites in the 5′ and 3′ LTRs of the three proviruses tested (two present in infected cells, and one present in transfected cells) map at a similar position (Fig. 5), within an accuracy of 20 to 50 bp. None of the DNase I HS sites found in the 5′ and 3′ regions was detected in the corresponding regions when purified IF-I DNA was subjected to similar analysis (data not shown).

Figure 6 shows the location of the DNase I HS sites in the 5′ and 3′ LTRs, as determined from the experiments shown in Fig. 3 and 5. Two DNase I HS sites are present in the 5′ LTR. One site, HSI, is present downstream from an *Nhe*I restriction site and overlaps a nearby *Eco*RV site. Therefore, HSI maps within a region containing a repeated sequence shown previously to contain enhancer activity (21, 24). A second DNase I HS site (HSII) detected within the 5′ LTR coincides with a *Sac*I restriction site. This *Sac*I site abuts the "TATA" box present 32 nucleotides upstream from the transcription initiation site. Thus it appears that HSII maps close to the TATA box region, perhaps covering the start site for transcription as well. A third, weaker DNase I HS site (HSIII) is present outside the LTR in downstream viral sequences, upstream from a *Pvu*I restriction site. Thus this HS site is located quite close to the viral donor splice site, in a region also containing the viral packaging signal (27). A single major DNase I HS site (HSIV) was detected in the 3′ LTR, mapping between the *Nhe*I and *Xba*I restriction sites. Like HSI in the 5′ LTR, HSIV also coincides with the viral enhancer sequences. Although these experiments do not permit firm conclusions as to whether HS sites I and IV span identical portions of the 5′ and 3′ LTRs, respectively, it can at least be said that the 5′ border of HSIV does not appear to extend any further upstream of the enhancer sequences than does the 5′ border of HSI, as was previously suggested (43).

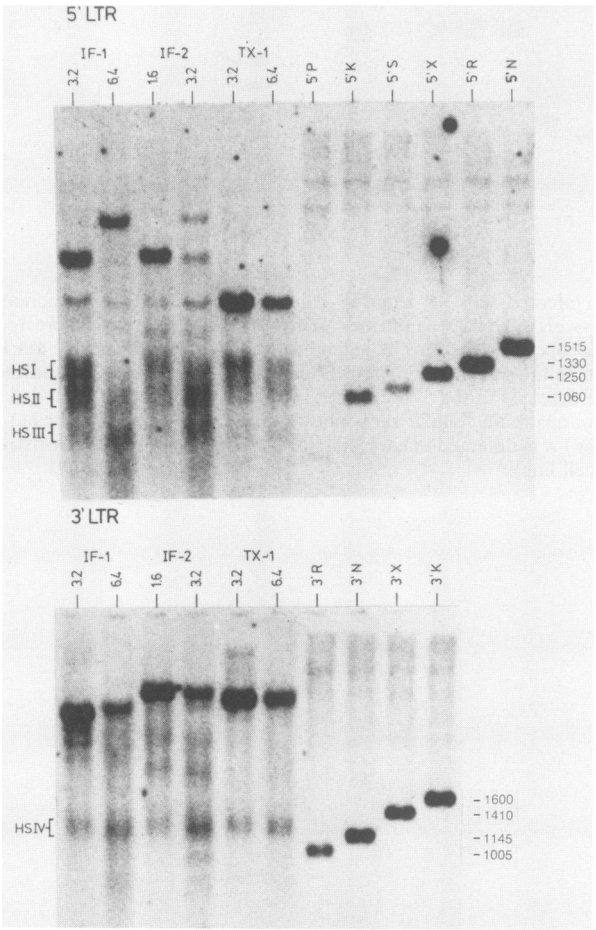


FIG. 5. Direct comparison of the DNase I HS sites in the 5′ and 3′ LTR regions of N4 in cell lines IF-1, IF-2, and TX-1. For each cell line, nuclei were digested with the two concentrations of DNase I that were shown previously to give the best visualization of DNase I HS sites. High-molecular-weight DNA was isolated, cut with *Nco*I, and subjected to Southern blot hybridization analysis, using either the 5′- or the 3′-specific probe described in the legend to Fig. 1. Hybridizing size markers (also described in the legend to Fig. 1) were run in paralled and length (in base pairs) is indicated.

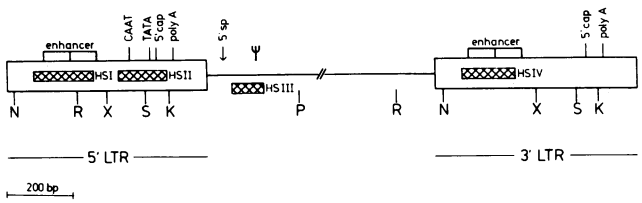


FIG. 6. Summary of DNase I HS sites in the 5′ and 3′ LTR regions of vector N4 in infected and transfected cell lines. The positions of marker restriction sites are indicated. Also indicated are the locations of the viral enhancer sequences, "CAAT" sequences, TATA box, cap site, poly(A) addition signal, 5′ donor splice site (5′ sp), and packaging signal (ψ). DNase I HS sites are represented by boxes. Abbreviations: N, *Nhe*I; R, *Eco*RV; X, *Xba*I; S, *Sac*I; K, *Kpn*I; P, *Pvu*I.

## DISCUSSION

In these studies, the technique of indirect end labeling to map DNase I HS sites was used to probe the chromatin structure of an M-MuLV-derived recombinant retrovirus vector. One question which these studies have addressed is whether the different functions encoded by the 5′ and 3′ LTRs of a single provirus are accompanied by distinct chromatin configurations. If so, this would suggest that a particular chromatin structure is of functional significance. Analysis of two cell lines harboring one integrated provirus per cell revealed that two DNase I HS sites are present in the 5′ LTR and a third site is present in adjacent downstream viral sequences. The HS sites in the 5′ LTR overlap the viral enhancer sequences (HSI) and the start site for transcription (HSII). A single HS site (HSIV) was detected in the 3′ LTR. Like HSI in the 5′ LTR, HSIV coincides with the enhancer sequences. The presence in the 5′ LTR of an HS site (HSII) overlapping the 5′ cap site and its absence in the 3′ LTR suggest that the chromatin configuration associated with this HS site is relevant to the process of initiation of transcription.

A similar pattern of DNase I HS sites in the 5′ and 3′ regions of wild-type M-MuLV proviruses has been described recently (43). In the present studies, a recombinant retroviral vector (N4) was used to examine M-MuLV chromatin structure. This approach has several advantages over the use of wild-type M-MuLV. First, because N4 carries a selectable marker (i.e., the bacterial *neo* gene), cell lines containing a single integrated provirus per cell are easily obtained. Therefore, all of the DNase I HS sites detected in a given cell line can be attributed to an individual provirus. In contrast, because most cell lines infected with wild-type retroviruses

harbor several copies of the viral genome, the pattern of DNase I HS sites found in a wild-type M-MuLV-infected cell line must be considered an average for all of the proviruses within that cell line (43). Another advantage of using recombinant retroviral vector N4 is that it contains DNA sequences (i.e., the *neo* gene) not found in retroviral or host cell DNA which can therefore be used as a vector-specific hybridization probe. Finally, since the length of the entire N4 provirus is only approximately 3.3 kb, the *Nco*I site used as a reference for DNase I HS mapping (Fig. 1) is quite close to the border of each LTR (i.e., ~950 bp from the 3′ border of the upstream LTR and ~1,100 bp from the 5′ border of the downstream LTR). Therefore, DNase I-induced cleavage in the LTRs generates subfragments whose sizes can be determined with considerable accuracy.

The use of a recombinant retroviral vector to study the chromatin structure of M-MuLV proviruses also has potential disadvantages. In particular, it is possible that internal viral sequences may also play an essential role in retroviral gene expression, for example, by encoding an enhancerlike function (as is the case in the expression of immunoglobulin genes) or by conferring increased stability to viral transcripts. Obviously, a DNase I HS site associated with such a function could be missed by using retroviral vectors such as N4.

In an M-MuLV-infected cell containing one to several proviruses, about 1% of the total poly(A)$^+$ cytoplasmic RNA is of viral origin (8). An important objective of these studies was to determine whether this high transcriptional activity of the M-MuLV genome is associated with a specific chromatin configuration. We took advantage of a previous observation that M-MuLV LTR-directed transcription is more efficient when the retroviral genome is introduced into cells by viral infection compared with DNA transfection (19). Indeed, the transcriptional activity of the N4 provirus is 5- to 10-fold higher in infected cells than in cells transfected with the N4 plasmid DNA (Fig. 4). The pattern of DNase I HS sites in the transfected cell line is indistinguishable from that of the two cell lines obtained by infection with the corresponding virus (Fig. 3 and 5). Thus, the difference in the transcriptional activity between the two identical DNA templates is not associated with an observable difference in chromatin configuration. A similar pattern of DNase I HS sites was found in the LTRs of M-MuLV, the replication-competent parent virus from which the N4 vector was derived (43). These findings are consistent with the observation that genes expressed at different levels exhibit the same degree of overall DNase I sensitivity (6, 15). When considered along with the observations that DNase I sensitivity sometimes precedes the activation of certain genes (5, 39), and remains after the turning off of certain other genes (17), our findings support the argument that overall DNase I sensitivity as well as the appearance of distinct HS sites reflects only a predisposition for gene activity. It is possible, however, that changes in chromatin configuration are involved in mediating the level of gene expression, but that these changes are not reflected by differences in DNase I sensitivity. Consistent with this possibility is the recent observation that differences in the levels of mRNA that accumulated after activation of four *Drosophila* heat shock genes were correlated with the relative sensitivities of those genes to cleavage by MPE.Fe(II), a reagent apparently able to detect more subtle changes in chromatin structure than is the much larger and more sterically hindered DNase I molecule (6).

Although integration of retroviruses into the host cell chromosome does not show any specificity at the nucleotide level or as determined by restriction enzyme analysis (2, 18, 36), it is thought that retroviruses might integrate preferentially into regions of open chromatin (perhaps nearby strong cellular enhancers), accounting for the strong transcriptional activity of the viral genome (13, 20). One question addressed in these studies was whether the cellular sequences adjacent to the proviral integration site are contained in a specific chromatin configuration that can be detected by the presence of discrete DNase I HS sites. Several weak DNase I HS sites are present in the cellular sequences adjacent to the proviral integration sites of the infected cell lines IF-1 and IF-2, but not in the transfected cell line TX-1 (Fig. 3). However, no common cellular DNase HS sites were found upstream or downstream from the integrated proviruses. Although this analysis did not reveal the existence of a chromatin configuration unique to the proviral integration site, a general conclusion cannot be drawn from these observations because only 1 to 3 kbp of cellular sequences were examined and, as described above, some features of chromatin structure may not be reflected in DNase I hypersensitivity.

## ACKNOWLEDGMENTS

We thank Nancy Plum for typing the manuscript.

This work was funded by American Cancer Society grant MV-116C to E.G.

## LITERATURE CITED

1. **Aviv, H., and P. Leder.** 1972. Purification of biologically active globin messenger RNA by chromatography on oligothymidilic acid cellulose. Proc. Natl. Acad. Sci. USA **69**:1408–1412.
2. **Bacheler, L. T., and H. Fan.** 1979. Multiple integration sites for Moloney murine leukemia virus in productively infected mouse fibroblasts. J. Virol. **30**:657–667.
3. **Bandyopadhyay, P. K., and H. M. Temin.** 1984. Expression of complete chicken thymidine kinase gene inserted in a retrovirus vector. Mol. Cell. Biol. **4**:749–754.
4. **Bier, E., Y. Hashimoto, M. I. Greene, and A. M. Maxam.** 1985. Active T-cell receptor genes have intron deoxyribonuclease hypersensitive sites. Science **229**:528–534.
5. **Burch, J. B. E., and H. Weintraub.** 1983. Temporal order of chromatin structural changes associated with activation of the major chicken vitellogenin gene. Cell **33**:65–76.
6. **Cartwright, I. L., and S. C. R. Elgin.** 1986. Nucleosomal instability and induction of new upstream protein-DNA associations accompany activation of four small heat shock protein genes in *Drosophila melanogaster*. Mol. Cell. Biol. **6**:779–791.
7. **Chiswell, D. J., D. A. Gillespie, and J. A. Wyke.** 1982. The changes in proviral chromatin that accompany morphological variation in avian sarcoma virus-infected rat cells. Nucleic Acids Res. **10**:3967–3980.
8. **Coffin, J.** 1982. Structure of the retroviral genome, p. 261–368. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
9. **Conklin, K. F., J. M. Coffin, H. L. Robinson, M. Groudine, and R. Eisenman.** 1982. Role of methylation in the induced and spontaneous expression of the avian endogenous virus ev-1: DNA structure and gene products. Mol. Cell. Biol. **2**:638–652.
10. **Denhardt, D. J.** 1966. A membrane-filter technique for the detection of complementary DNA. Biochem. Biophys. Res. Commun. **23**:641–646.
11. **Elgin, S. C. R.** 1981. DNase I-hypersensitive sites of chromatin. Cell **27**:413–415.
12. **Elgin, S. C. R.** 1982. Chromatin structure, DNA structure. Nature (London) **300**:402–403.
13. **Feinstein, S. C., S. R. Ross, and K. R. Yamamoto.** 1982. Chromosomal position effects determine transcriptional potential of integrated mammary tumor virus DNA. J. Mol. Biol. **156**:549–565.
14. **Garel, A., and R. Axel.** 1976. Selective digestion of transcrip-

tionally active ovalbumin genes from oviduct nuclei. Proc. Natl. Acad. Sci. USA 73:3966–3970.
15. **Garel, A., M. Zolan, and R. Axel.** 1977. Genes transcribed at diverse rates have a similar conformation in chromatin. Proc. Natl. Acad. Sci. USA 74:4867–4871.
16. **Groudine, M., R. Eisenman, and H. Weintraub.** 1981. Chromatin structure of endogenous retroviral genes and activation by an inhibitor of DNA methylation. Nature (London) 292:311–317.
17. **Groudine, M., and H. Weintraub.** 1982. Propagation of globin DNase I-hypersensitive sites in absence of factors required for induction: a possible mechanism for determination. Cell 30:130–139.
18. **Hughes, S. H., P. R. Shank, D. H. Spector, H.-J. Kung, J. M. Bishop, H. E. Varmus, P. K. Vogt, and M. L. Breitman.** 1978. Proviruses of avian sarcoma virus are terminally redundant, co-extensive with unintegrated linear DNA and integrated at many sites. Cell 15:1397–1410.
19. **Hwang, L. H. S., and E. Gilboa.** 1984. Expression of genes introduced into cells by retroviral infection is more efficient than that of genes introduced into cells by DNA transfection. J. Virol. 50:417–424.
20. **Jaenisch, R., D. Jahner, P. Nobis, I. Simon, J. Lohler, K. Harbers, and D. Grotkopp.** 1981. Chromsomal position and activation of retroviral genomes inserted into the germ lines of mice. Cell 24:519–529.
21. **Jolly, D. J., A. C. Esty, S. Subramani, N. T. Friedman, and I. M. Verma.** 1983. Elements in the long terminal repeat of murine retroviruses enhance stable transformation by thymidine kinase gene. Nucleic Acids Res. 11:1855–1872.
22. **Jorgensen, R. A., S. J. Rothstein, and W. S. Reznikoff.** 1979. A restriction enzyme cleavage map of Tn5 and location of a region encoding neomycin resistance. Mol. Gen. Genet. 177:65–72.
23. **Keene, M. A., V. Corces, K. Lowenhaupt, and S. C. R. Elgin.** 1981. DNase I hypersensitive sites in Drosophila chromatin occur at the 5′ ends of regions of transcription. Proc. Natl. Acad. Sci. USA 78:143–146.
24. **Laimins, L. A., G. Khoury, C. Gorman, B. Howard, and P. Gruss.** 1982. Host-specific activation of transcription by tandem repeats from simian virus 40 and Moloney murine sarcoma virus. Proc. Natl. Acad. Sci. USA 79:6453–6457.
25. **Larsen, A., and H. Weintraub.** 1982. An altered DNA conformation detected by S1 nuclease occurs at specific regions in active chick globin chromatin. Cell 29:609–622.
26. **Lehrach, H., D. Diamond, J. M. Wozney, and H. Boedtker.** 1977. RNA molecular weight determinations by gel electrophoresis under denaturing conditions, a critical reexamination. Biochemistry 16:4743–4751.
27. **Mann, R., R. C. Mulligan, and D. Baltimore.** 1983. Construction of a retrovirus packaging mutant and its use to produce helper-free defective retrovirus. Cell 33:153–159.
28. **Mathis, D., P. Oudet, and P. Chambon.** 1980. Structure of transcribing chromatin. Prog. Nucleic Acid Mol. Biol. 24:1–55.
29. **McGhee, J. D., W. I. Wood, M. Dolan, J. D. Engel, and G. Felsenfeld.** 1981. A 200 base pair region at the 5′ end of the chicken adult β-globin gene is accessible to nuclease digestion. Cell 27:45–55.
30. **Miller, A. D., D. J. Jolly, T. Friedmann, and I. M. Verma.** 1983. A transmissible retrovirus expressing human hypoxanthine phosphoribosyltransferase (HPRT): gene transfer into cells obtained from humans deficient in HPRT. Proc. Natl. Acad. Sci. USA 80:4709–4713.
31. **Miller, A. D., M. F. Law, and I. M. Verma.** 1985. Generation of helper-free amphotropic retroviruses that transduce a dominant-acting, methotrexate-resistant dihydrofolate reductase gene. Mol. Cell. Biol. 5:431–437.
32. **Mills, F. C., L. M. Fisher, R. Kuroda, A. M. Ford, and H. J. Gould.** 1983. DNase I hypersensitive sites in the chromatin of human μ immunoglobin heavy-chain genes. Nature (London) 306:809–812.
33. **Montandon, P.-E., F. Montandon, and H. Fan.** 1982. Methylation state and DNase I sensitivity of chromatin-containing Moloney murine leukemia virus DNA in exogenously infected mouse cells. J. Virol. 44:475–486.
34. **Rigby, P. W. J., M. Dieckmann, C. Rhodes, and P. Berg.** 1977. Labeling of DNA to high specific activity by nick translation. J. Mol. Biol. 113:237–258.
35. **Shermoen, A. W., and S. K. Beckendorf.** 1982. A complex of interacting DNase I-hypersensitive sites near the drosophila glue protein gene, Sgs4. Cell 29:601–607.
36. **Shimotohno, D., and H. M. Temin.** 1980. No apparent nucleotide sequence specificity in cellular DNA juxtaposed to retrovirus proviruses. Proc. Natl. Acad. Sci. USA 77:7357–7361.
37. **Sorge, J., A. E. Cutting, V. D. Erdman, and J. W. Gautsch.** 1984. Integration-specific retrovirus expression in embryonal carcinoma cells. Proc. Natl. Acad. Sci. USA 81:6627–6631.
38. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. 98:503–517.
39. **Stalder, J., M. Groudine, J. B. Dogson, J. D. Engel, and H. Weintraub.** 1980. Hb switching in chickens. Cell 19:973–980.
40. **Stalder, J., A. Larsen, J. D. Engel, M. Dolan, M. Groudine, and H. Weintraub.** 1980. Tissue-specific DNA cleavages in the globin chromatin domain introduced by DNase I. Cell 20: 451–460.
41. **Sweet, R. W., M. V. Chao, and R. Axel.** 1982. The structure of the thymidine kinase promoter: nuclease hypersensitivity correlates with expression. Cell 31:347–353.
42. **Thomas, P. S.** 1980. Hybridization of denatured RNA and small DNA fragments transferred to nitrocellulose. Proc. Natl. Acad. Sci. USA 77:5201–5205.
43. **Thompson, T., and H. Fan.** 1985. Mapping of DNase I-hypersensitive sites in the 5′ and 3′ long repeats of integrated Moloney murine leukemia virus proviral DNA. Mol. Cell. Biol. 5:601–609.
44. **van der Putten, H., W. Quint, I. M. Verma, and A. Berns.** 1982. Moloney murine leukemia virus-induced tumors: recombinant proviruses in active chromatin regions. Nucleic Acids Res. 10:577–592.
45. **Varmus, H., and R. Swanstrom.** 1982. Replication of retroviruses, p. 369–512. *In* R. Weiss, N. Teich, H. Varmus, and J. Coffin (ed.), RNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
46. **Weintraub, H., and M. Groudine.** 1976. Chromosomal subunits in active genes have an altered conformation. Science 193: 848–856.
47. **Weintraub, H., A. Larsen, and M. Groudine.** 1981. β-Globin-gene switching during the development of chicken embryos: expression and chromosome structure. Cell 24:333–344.
48. **Weisbrod, S.** 1982. Active chromatin. Nature (London) 297: 289–295.
49. **Wu, C., P. M. Bingham, K. J. Livak, R. Holmgren, and S. C. R. Elgin.** 1979. The chromatin structure of specific genes. I. Evidence for higher order domains of defined DNA sequence. Cell 16:797–806.
50. **Wu, C., and W. Gilbert.** 1981. Tissue-specific exposure of chromatin structure at the 5′ terminus of the rat preproinsulin II gene. Proc. Natl. Acad. Sci. USA 78:1577–1580.
51. **Zaret, K. S., and K. R. Yamamoto.** 1984. Reversible and persistent changes in chromatin structure accompany activation of a glucocorticoid-dependent enhancer element. Cell 38:29–38.