Proc. Natl. Acad. Sci. USA
Vol. 78, No. 7, pp. 4485–4489, July 1981
Genetics

Specific transcriptional initiation *in vitro* on murine type C retrovirus promoters

(AKR murine leukemia virus/Harvey murine sarcoma virus/long terminal redundancy/RNA polymerase II)

MICHAEL C. OSTROWSKI, DIANA BERARD, AND GORDON L. HAGER

Laboratory of Tumor Virus Genetics, National Cancer Institute, National Institutes of Health, Bethesda, Maryland 20205

Communicated by Herschel L. Roman, April 6, 1981

ABSTRACT **We have investigated the ability of molecularly cloned murine type C retroviral DNA to direct accurate initiation of RNA synthesis when added to cell-free extracts. Two different cloned proviruses were used. The first was derived from an integrated molecule of AKR murine leukemia virus and contains adjacent host information. The origin of the second was an unintegrated permuted copy of Harvey murine sarcoma virus. We found that the leukemia virus cloned provirus, as predicted by structural considerations, contained two functional RNA polymerase II promoters located in the U3 region present at either end of the molecule. These promoters initiate transcription at equal rates *in vitro*. We also found that the permuted sarcoma virus clone contained an RNA polymerase II promoter in the U3 region. Removal of viral sequences 49 bases upstream of the *in vitro* sarcoma virus initiation site by restriction cleavage results in loss of specific transcription, indicating a role for this information in *in vitro* promotion. The 5′ ends of *in vitro* and *in vivo* viral RNA were compared by nuclease mapping techniques and found to be identical. Based on this evidence, we conclude that murine retroviral genomes contain sufficient information to initiate transcription independent of any host information *in vitro* and that these viral promoters are probably also active *in vivo*. In addition to the promoter in U3, Harvey murine sarcoma virus contains a second promoter *in vitro* that initiates near the 5′ boundary of the transformation-specific (*src*) region of the virus. Initiation by this promoter was insensitive to low levels of α-amanitin, and the RNA transcript could be terminated to yield a 340-nucleotide product.**

The retroviruses are a group of single-strand RNA viruses that integrate a double-strand DNA copy of their genome into host cell DNA as an obligatory intermediate of their replication cycle (1). Although it has been established that cellular RNA polymerase II (pol II) is responsible for transcription of the integrated viral genome (2, 3), it remains uncertain whether information necessary for the initiation of viral transcription is encoded by the viral genome or by host DNA adjacent to the site of integration. The discovery that proviruses contain a directly repeated sequence located at either end of the molecule (4–8) raised the possibility that proviral DNA contains the control elements necessary to function as an independent transcriptional unit. Sequence analysis of these DNA repeats (the long terminal repeat or LTR) for several retroviruses (9–12) confirmed that they contain sequences associated with both initiation and termination of eukaryotic pol II-dependent transcription.

Recently, cell-free systems that have the ability to accurately initiate transcription on cloned promoters *in vitro* have been described (13, 14). Subsequently, transcription initiation on various genes (15–17), including a cloned cDNA fragment of an avian retrovirus (18), have been reported. We have studied the initiation of transcription *in vitro* at promoters encoded by murine type C retroviral DNA. In agreement with the predictions suggested by integrated viral DNA structures, we find that initiation of transcription occurs within the LTR at both ends of the integrated genome and with equal efficiency. We also find an unexpected site for initiation by an α-amanitin-resistant RNA polymerase encoded near the left end of the p21 *src* gene of Harvey murine sarcoma virus (Ha-MuSV).

MATERIALS AND METHODS

Preparation of Cell-Free Extracts. HeLa cell extracts were prepared essentially as described (14). The dialysis buffer was 20 mM Hepes, pH 7.9/100 mM KCl/12.5 mM $MgCl_2$/0.1 mM EDTA/2 mM dithiothreitol/17% glycerol (vol/vol). Extracts that supported accurate initiation had a protein concentration of 25–30 mg/ml. The optimum DNA concentration for *in vitro* transcription experiments was 40 μg/ml of these extracts, regardless of the promoter being studied (data not shown). The extracts were stored in liquid N_2 and retained their activity for ≈2 months.

***In Vitro* Incubation and Analysis of RNA.** Standard reaction mixtures were 50 μl and contained 12 mM Hepes, pH 7.9; 60 mM KCl; 7 mM $MgCl_2$; 0.06 mM EDTA; 1.2 mM dithiothreitol; 10.20% glycerol (vol/vol); 500 μM each ATP, CTP, and GTP; 50 μM UTP containing 20 μCi of [α-^{32}P]UTP (1 Ci = 3.7×10^{10} becquerels; Amersham); 5 mM phosphocreatine; 30 μl of extract; and DNA at 40 μg/ml. Reaction mixtures were incubated at 30°C for 60 min. Reactions were terminated by the addition of 400 μl of 0.2 M NaOAc, pH 5/0.2% $NaDodSO_4$ containing yeast tRNA at 30 μg/ml. The mixtures were extracted once with phenol/$CHCl_3$/isoamyl alcohol (1:1:0.05) and twice with $CHCl_3$. RNA in the aqueous phase was precipitated by addition of 2 vol of EtOH. EtOH precipitation was repeated two more times.

The final RNA pellet was incubated in 100 μl of 1 M glyoxal/0.1 M Hepes, pH 7.9/50% dimethyl sulfoxide (vol/vol) at 37°C for 1 hr (13) and, after 1 final EtOH precipitation, analyzed on 4.5% polyacrylamide gels as described (13).

Preparation of DNA Templates. Two type C retroviral DNAs were used as transcription templates. One was clone 623, a molecular clone of AKR ecotropic murine leukemia virus derived by insertion of the integrated provirus into λ Charon 4A (19). Subgenomic clones consisting of the 5′ and 3′ halves of the molecule, including adjacent host sequences, were obtained by transfer to pBR322 using the single *Sal* I site located 4200 base pairs (bp) from the 5′ end of viral information and either the *Hin*dIII site 1000 bp upstream of the 5′ end or the *Eco*RI site located 600 bp downstream of the 3′ end of the provirus.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U. S. C. §1734 solely to indicate this fact.

Abbreviations: pol II, RNA polymerase II; repeat; LTR; long terminal Ha-MuSV; Harvey murine sarcoma virus; bp, base pair(s); AKR-MuLV; AKR murine leukemia virus; Ad2, adenovirus serotype 2; Mo-MuLV; Moloney murine leukemia virus.

The second viral template was derived from the permuted molecular clone of unintegrated Ha-MuSV (20). The large *Pst* I restriction endonuclease fragment was subcloned in pBR322 as described (21). Cloning of the 2200-bp *Eco*RI/*Bam*HI fragment has also been described (22). It should be noted that both the AKR murine leukemia virus (AKR-MuLV) and Ha-MuSV parental clones are transfectious (19, 20).

The *Bal* I E restriction endonuclease fragment E of adenovirus serotype 2 (Ad2) virus, which contains the major late Ad2 promoter and was cloned at the *Bam*HI site of pBR322 (14), was the gift of J. R. Manley.

Plasmids were extracted and purified by standard procedures (12). In preparation for use as *in vitro* templates, plasmid was restricted, phenol and $CHCl_3$ extracted, EtOH precipitated, and stored in 10 mM Tris chloride, pH 8/1 mM EDTA at −20°C.

Determination of 5′ Ends of *In Vitro*-Synthesized RNAs. For mapping of *in vitro* transcripts, RNA from 400-μl reaction mixtures containing no ^{32}P label was isolated as described above. The RNA pellet after the first EtOH precipitation was suspended in 125 μl of 50 mM Tris chloride, pH 7.9/10 mM $MgCl_2$ and incubated at 37°C for 10 min with RNase-free DNase I at 50 μg/ml (23). After extraction with phenol and $ChCl_3$, the RNA was coprecipitated with the appropriate viral DNA fragments ^{32}P labeled at their 5′ ends. The pellet was dissolved in 25 μl of 80% formamide/40 mM 1, 4-piperazinediethane sulfonic acid, pH 6.4/0.4 M NaCl/1 mM EDTA; the mixture was heated to 85°C for 5 min and then incubated at 45°C for 12 hr. The reactions were diluted with 250 μl of 0.03 M NaOAc, pH 4.6/0.25 M NaCl/2 mM $ZnCl_2$ containing 500 units of mung bean nuclease (P-L Biochemicals) and incubated for 1 hr at 37°C (17). After extraction with phenol and $CHCl_3$ and precipitations with EtOH, the samples were analyzed on 20% acrylamide/urea thin sequencing gels side by side with sequencing ladders obtained by chemical cleavage of the ^{32}P-labeled probe (24).

RESULTS

Specific Initiation of Transcription on AKR-MuLV DNA *in Vitro*. We set out to determine whether specific RNA pol II transcripts are initiated *in vitro* from retroviral templates by using the cell-free system described by Manley *et al.* (14). DNA sequence analysis of several cloned retroviruses has shown that the LTRs contain a T-A-T-A-A-A consensus sequence, or "Hogness box," ≈170 bp upstream of the 3′ end of the LTR (9–12). This structure is frequently associated with eukaryotic RNA pol II promoters.

Fig. 1*A* shows the restriction map for the 5′ end of clone 623 AKR-MuLV proviral DNA (19) as subcloned in pBR322. Use of this plasmid, cleaved with *Pvu* I, *Bgl* II, or *Bst* EII, as a template for *in vitro* initiation of transcription should give RNA "run-off" products 240, 340, or 740 nucleotides long, respectively, if transcription begins at a site associated with the canonical Hogness box. Analysis of labeled RNAs from such an experiment on glyoxal gels shows discrete RNA bands at 230, 330, and 710 nucleotides, in excellent agreement with the predicted result (Fig. 2*A*, lanes 1, 2, and 3). Further, addition of α-amanitin to the reaction mixtures results in the loss of these bands, indicating that these transcripts are RNA pol-II dependent (Fig. 2*A*, lanes 4, 5, and 6). On these and subsequent glyoxal gels, the transcription run-off products of the cloned late Ad2 promoter generated by cleavage of the plasmid with *Sma* I (530 nucleotides), *Hin*cII (405 nucleotides), or *Hin*dIII (190 nucleotides) were used as markers to determine the sizes of the retroviral transcripts (16). If either no DNA or pBR322 cleaved with various single-cut restriction enzymes is added to the reaction mixture, no specific transcripts are detected (data not shown).

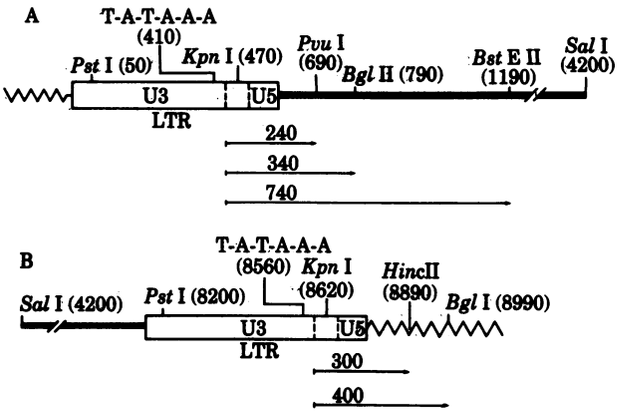

FIG. 1. Restriction endonuclease maps of the 5′ (*A*) and 3′ (*B*) regions of integrated AKR-MuLV. The T-A-T-A-A-A structures are positioned based on sequencing information (C. Van Beveren, D. Lowy, and I. M. Verma, personal communication). Numbers in parentheses indicate the distance in nucleotides from the 5′ end of the viral genome. Arrows under the maps indicate the length of RNA chains that should be generated *in vitro* if synthesis is initiated 25 bp downstream of the T-A-T-A-A-A sequence and the DNA template is truncated with the appropriate restriction enzyme so that RNA pol II transcription complexes run-off the template when they reach the cleavage site.

As the AKR-MuLV provirus also has an LTR at its 3′ end, we proceeded to test whether this LTR contains an active promoter *in vitro*. Fig. 1*B* shows the restriction map of the 3′ end of clone 623 AKR-MuLV and the sizes of the expected run-off products generated by cleaving with either *Hin*cII or *Bgl* I, whose restriction sites are located in the adjacent cellular information. Once again, the RNAs actually synthesized *in vitro*

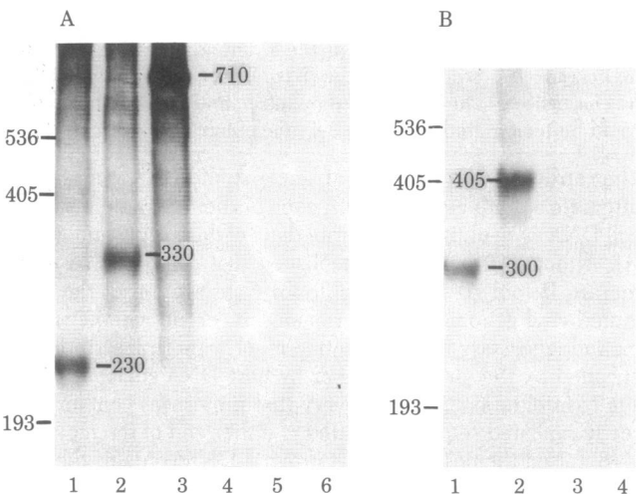

FIG. 2. Autoradiographs of polyacrylamide gels containing glyoxalated RNAs initiated *in vitro* in standard reaction mixtures by truncated AKR-MuLV templates. (*A*) Run-off products initiated from 5′ AKR-MuLV sequences truncated at the *Pvu* I, *Bgl* II, or *Bst*EII sites in the absence (230, 330, or 710 bases, respectively, lanes 1, 2, and 3) or presence (lanes 4, 5, and 6) of α-amanitin (1 μg/ml). (*B*) Transcripts initiated by the 3′ MuLV LTR, which run-off at either the *Hin*dII (300 bases, lane 1) or *Bgl* I (405 bases, lane 2) sites in adjacent host sequences. The presence of α-amanitin (1μg/ml) in the reaction mixture eliminates synthesis of these species (lanes 3 and 4, respectively). Numbers on the left indicate run-off RNAs generated by cleavage of the late Ad2 promoter with *Sma* I (536 bases), *Hin*cII (405 bases), or *Hin*dIII (193 bases) and used to obtain size estimates of retroviral products. Samples contained 5000 cpm (Cerenkov) per lane. Exposure was for 12 hr.

(300 or 405 bases, respectively) show good correspondence with the predicted results (Fig. 2*B*, lane 1 and 2). These bands are also sensitive to α-amanitin (Fig. 2*B*, lanes 3 and 4).

Quantitatively, an equal amount of label is incorporated into specific transcripts initiated from either AKR-MuLV LTR (data not shown), indicating that these promoters are used with equal efficiency *in vitro*.

Specific Initiation of Transcription on Ha-MuSV DNA *in Vitro*. To conclusively show that retroviral promoters could function *in vitro* independent of any host information, we studied the initiation of transcription on cloned Ha-MuSV. The clone used was derived from a non-integrated circular proviral molecule that contains only one LTR and was permuted at its single internal *Eco*RI restriction site (20). The restriction map of the LTR and adjacent 5′ unique viral sequences of this molecule is shown in Fig. 3. By using *Pvu* I-restricted material as template, a 230-base product would be predicted if initiation of transcription occurs ≈25 bases downstream of the Hogness box in Ha-MuSV (Fig. 3*A*). As seen in Fig. 4*A*, such a 230-base-long product is actually observed. Fig. 4*B* illustrates the run-off product obtained by using the Ha-MuSV template truncated with three different enzymes in combination with *Sst* II. The sizes of the run-off products agree with the predicted results (Fig. 3*A*). Synthesis of these RNAs is also sensitive to α-amanitin at 1 μg/ml (data not shown).

To determine the involvement of sequences proximal to the Hogness box in promoter recognition, a 950-bp agarose gel-purified Ha-MuSV fragment, obtained after *Cla* I/*Bam*HI digestion of the recombinant plasmid and containing the LTR, was further cleaved with *Xba* I, *Tha* I, or *Sac* I. The resulting fragments contain double-stranded regions to the 5′ side of the Ha-MuSV Hogness box of 118 bases, 19 bases, and 1 base, respectively (Fig. 3). The results of using these 5′-truncated fragments as templates for *in vitro* transcription are shown in Fig. 4*C*. The 950-bp fragment or the fragment cut with *Xba* I serve as efficient templates for the expected product, but the *Tha* I-and *Sac* I-cut templates do not. In contrast to these results, Wasylyk *et al.* (15), in an analogous experiment using a restriction fragment containing the conalbumin promoter, find that templates truncated ≈14 bp upstream of the Hogness box are capable of initiating transcription. We tentatively conclude that the failure of the *Tha* I fragment to support efficient initiation *in vitro* is not the result of an end-effect but may reflect the involvement of sequences between *Xba* I and *Tha* I in promoter recognition. In this regard, a second concensus sequence found 75 bases

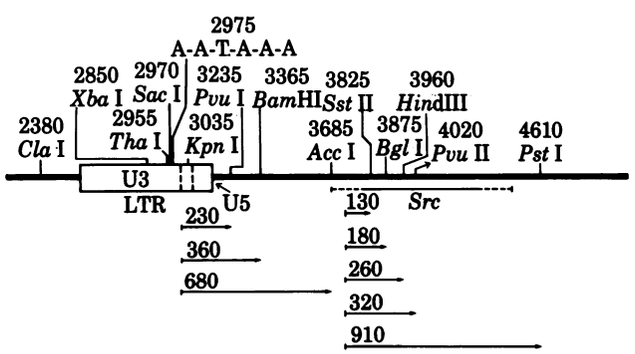

FIG. 3. Restriction endonuclease map of a part of the unintegrated Ha-MuSV genome that includes the single LTR and the transforming (*src*) region (21). The A-A-T-A-A-A sequence is the consensus promoter signal for this LTR. Arrows beginning in the LTR show the sizes of RNA run-off products expected from templates truncated with *Pvu* I, *Bam*HI, or *Acc* I. Arrows beginning near the boundary of the *src* region indicate the sizes of run-off RNAs expected for the α-amanitin-resistant (1 μg/ml) start observed *in vitro* (see Fig. 4). Dashed portions of the *src* region brackets indicate uncertainty in its boundaries (21).

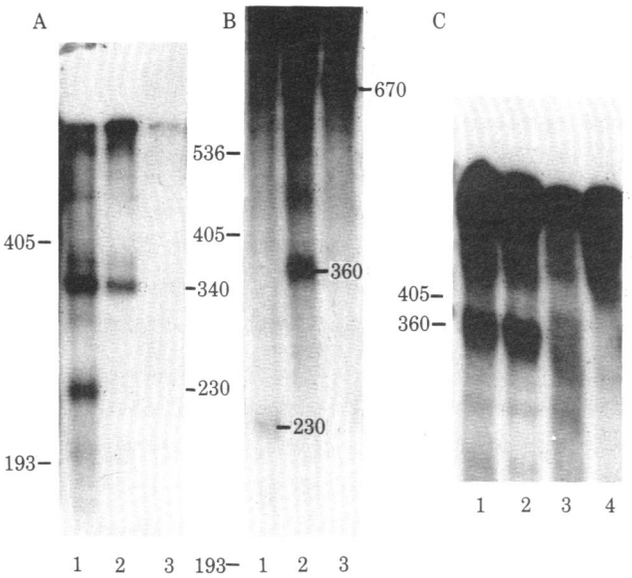

FIG. 4. Autoradiographs of RNAs initiated *in vitro* in standard reaction mixtures by Ha-MuSV templates. (*A*) Transcripts from Ha-MuSV clone Pst 8 (21) cleaved with *Pvu* I. Reaction mixtures contained α-amanitin at 0, 1, or 50 μg/ml (lanes 1, 2, and 3, respectively). (*B*) Run-off RNAs from clone Pst 8 cleaved with *Sst* II and either *Pvu* I, *Bam*HI, or *Acc* I (230, 360, or 670 bases; lanes 1, 2, and 3, respectively). (*C*) RNA initiated by gel-purified 950-bp restriction fragment generated by digestion of clone Pst 8 with *Cla* I/*Bam*HI (340 bases, lane 1). Lanes 2–4: specific transcripts initiated by this 950-bp fragment after further cleavage with *Xba* I, *Tha* I, or *Sac* I. Numbers indicate run-off products from the Ad2 late promoter used as size standards. Samples contained 5000 cpm (Cerenkov) per lane. Exposure was for 12 hr for *A* and *B* and 48 hr for *C*.

proximal to many pol II transcription starts (25) is present in the Ha-MuSV and AKR-MuLV sequences at the corresponding position, and this site is between the *Xba* I and *Tha* I restriction sites.

α-Amanitin-Resistant Initiation Site Encoded by the Ha-MuSV Genome. An unexpected 340-nucleotide transcript is also seen with the *Pvu* I-restricted Ha-MuSV template (Fig. 4*A*). Although the predicted Ha-MuSV LTR promoter is sensitive to α-amanitin at 1 μg/ml, this additional band is not (Fig. 4*A*, lane 2). The 340-base band is, however, sensitive to α-amanitin at 50 μg/ml (Fig. 4*A*, lane 3).

Preliminary nuclease SI mapping experiments indicated that the 340-base band (Fig. 4*A*) mapped between the *Bam*HI and the *Pst* I sites on the Ha-MuSV genome (see Fig. 3). This is also the region of the viral genome that contains the p21 (*src*) gene (21). To further characterize this α-amanitin (50 μg/ml)-sensitive transcript, a subgenomic clone of Ha-MuSV that contains no viral information to the 5′ side of the *Bam*HI site, and therefore lacks the LTR pol II promoter, was used. When this plasmid is cleaved with *Pst* I and added to the *in vitro* transcription reaction mixture, two predominant species of 340 and 950 bp are seen (Fig. 5*A*, lane 1). These bands are not sensitive to α-amanitin at 1 μg/ml but are sensitive to 50 μg/ml (Fig. 5*A*, lanes 2 and 3). If the plasmid is cut with *Pvu* II, *Hin*dIII, or *Bgl* I, RNAs (insensitive to α-amanitin at 1 μg/ml) 340, 270, or 180 bp long are observed (Fig. 5*B*, lanes 1–3). Taken together, these data indicate that an α-amanitin-resistant transcript is initiated near map coordinate 3700 (see Fig. 3). This transcript can terminate *in vitro* near the *Pvu* II site, yielding a species 340 bp long. Comparison of the ratio of radioactivity in the 950-bp run-off product with that in the 340-bp terminated product (Fig. 5*A*) indicates that the *in vitro* RNA is terminated ≈50% of the time. When *Sst* II-truncated templates are used, a 130 bp run-off

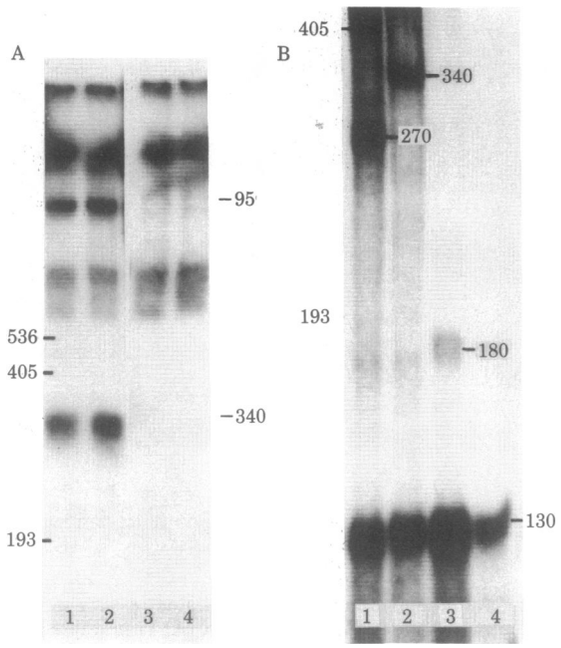

FIG. 5. Autoradiographs of transcripts initiated *in vitro* by Ha-MuSV clone 514 (18), which contains unique viral sequences including *src*, but no LTR. (*A*) Products obtained by addition of *Pst* I-cleaved clone 514 to standard reaction mixtures Reaction mixtures contained α-amanitin at 0, 1, and 50 μg/ml (lanes 1, 2, and 3, respectively. Reaction analyzed in lane 4 had no exogenous DNA. (*B*) Products obtained by addition of clone 514 cleaved with *Hin*dIII (lane 1), *Pvu* II (lane 2), *Bgl* I (lane 3), or *Sst* II (lane 4) in the presence of α-amanitin at 1 μg/ml. Late Ad2 promoter run-off RNAs were used as size standards. Samples contained 5000 cpm (Cerenkov) per lane. Exposure was for 12 hr.

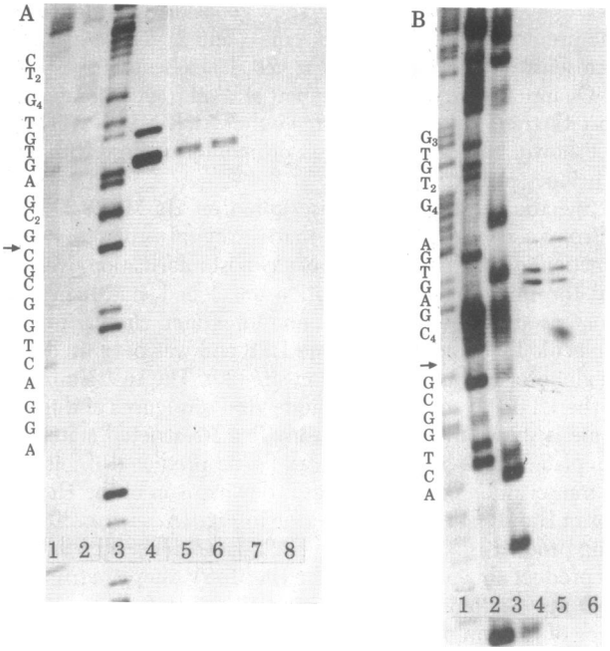

FIG. 6. Mapping of the 5′ ends of viral RNAs. The *Pst* I/*Kpn* I restriction fragment containing the AKR-MuLV LTR (Fig. 1) and the *Xba* I/*Kpn* I restriction fragment containing the Ha-MuSV LTR (Fig. 3) were labeled with ^{32}P at the 5′ end of the *Kpn* I site (specific activities 150,000 and 50,000 dpm/pmol, respectively; 1 dpm = 16.7 mBq). RNA was hybridized to 0.5 pmol of the appropriate labeled fragment, and hybrids were treated with mung bean nuclease. (*A*) Lanes 1–3: sequencing ladders produced by base-specific cleavage of the AKR-MuLV *Pst* I/*Kpn* I fragment (24). They represent the A < C, G + A, or C + T reactions, respectively, for the (minus) DNA strand. The other lanes contain the fraction of this restriction fragment protected from nuclease digestion by different RNAs; lane 4, 100 ng of AKR-MuLV 70S virion RNA; lanes 5 and 6, *in vitro*-synthesized RNA initiated by either the 5′ LTR (*Bgl* II truncated) or the 3′ LTR (*Hin*cII truncated), respectively; lane 7, *in vitro*-synthesized RNA initiated by the Ad2 late promoter (*Sma* I truncated); lane 8, RNA from *in vitro* reaction mixture that had no DNA template added. Exposure for lanes 1–4 was 12 hr and for lanes 5–8 was 72 hr. (*B*) Lanes 1–3: sequencing ladders representing G + A, C + T or A > C chemical cleavage of the Ha-MuSV *Xba* I/*Kpn* I fragment (minus) DNA strand, respectively. Lanes 4–6 represent the fraction of probe protected by Ha-MuSV RNAs; lane 4, 25 μg of poly(A)-selected RNA from a Ha-MuSV nonproducer cell line (26); lane 5, *in vitro* RNA initiated by *Bam*HI-truncated Ha-MuSV DNA; lane 6, RNA from *in vitro* reaction mixture that had no DNA template added. Exposure was for 96 hr. Arrows indicate the base on the minus strand complementary to the initiating nucleotide.

product would be predicted. This transcript would migrate in the same area of the gel in which a heavily labeled species is found in all samples (Fig. 5*B*). This 130-bp band appears in reactions when no exogenous DNA is added. It probably represents the end labeling of 5S RNA present in the HeLa extracts (14).

Determination of 5′ End of pol II-Dependent *in Vitro*-synthesized RNA. The results described above for *in vitro* transcripts initiated from either AKR-MuLV or Ha-MuSV LTRs indicate that transcription can be initiated near the region predicted by the proviral DNA sequence. To more precisely position the 5′ ends of these RNAs, we used a mung bean nuclease mapping procedure (17). The ^{32}P-5′-end-labeled probes used were either (*i*) a *Pst* I/*Kpn* I 400-bp fragment from the AKR-MuLV LTR or (*ii*) an *Xba* I/*Kpn* I 175-bp fragment from the Ha-MuSV LTR. These probes were hybridized to the appropriate *in vitro* RNAs and to either AKR 70S viral RNA or poly(A)-selected RNA obtained from Ha-MuSV-infected nonproducer cells (26). Inspection of the sequencing gels on which protected probe is analyzed allows several observations (Fig. 6). Significantly, the *in vitro* and *in vivo* RNAs yield the same pattern of protected bands. This indicates that *in vitro* and *in vivo* initiation appear to occur at precisely the same site. The identification of the initiating nucleotide in the DNA sequence is made somewhat difficult by the fact that, in our hands, the samples containing the protected probe migrate more slowly than the sequencing ladders with which they are compared. Based on the migration of xylene cylanol FF tracking dye (which fortuitously comigrates with the major protected band) present in each sample, this differential was determined to be three bases. The reason for this reproducibly slower migration is unknown but probably derives from differences in the sample loading buffers (17). The redundancy in protected bands can be interpreted either as incomplete nuclease digestion or heterogeneous 5′ ends (27). We feel that they represent partial digestion, as the major protection products map to the same position in the two viruses and the minor bands are all larger. Finally, a one and one-half base correction must be added to account for the difference in chemical versus nuclease cleavage (29).

The AKR-MuLV major band and the Ha-MuSV major band then both correspond to an uncapped 5′ end of G-C-G-C-C-A-G-T. . . (arrows in Fig. 6). The DNA sequences in this region of the LTRs for Mo-MuLV (. . .C-G-G-G-G-C-G-C-C-A-G-T-C-C-T. . .) (11) and AKR-MuLV (. . .C-G-G-C-G-C-G-C-C-A-G-T-C-C-T. . .) (C. Van Beveren, D. Lowy, and I. M. Verma, personal communication) have been determined. In addition, the first two nucleotides for uncapped Mo-MuLV RNA (the MuLV parent of the Ha-MuSV recombinant) have been determined as GpC (28). This dinucleotide is found in three positions in this region of the AKR-MuLV sequence and two positions in the Mo-MuLV sequence. Our data are consistent with the initiation dinucleotide being the first GpC of

the Mo-MuLV sequence and the corresponding second GpC of the AKR-MuLV sequence. Although nuclease digestion experiments cannot unambiguously identify the 5′ end of transcripts without corraborating RNA sequence information, the data presented here support the hypothesis that the transcripts are initiated at the same position *in vitro* as *in vivo*.

DISCUSSION

Our results show that type C proviruses contain sufficient information to initiate transcription of RNA pol II-dependent products *in vitro* independent of host information. Further, the prediction based on structural considerations that integrated proviruses should contain two promoters was confirmed *in vitro*. These results agree with the report of initiation *in vitro* from cDNA clones containing the Rous sarcoma virus 3′ LTR (18). Transcription of host DNA sequences from the 3′ LTR represents a novel mechanism for the activation of cellular genes by retroviruses (30). Indeed, emerging evidence (31, 32) suggests that host genetic information may be activated by transcription from either the 5′ or the 3′ LTR.

We find that removing sequences 5′ proximal to the Hogness box by restriction with *Tha* I (18 bases upstream) or *Sac* I (1 base upstream) eliminates promotion of RNA from the Ha-MuSV LTR. As removing sequences 14 bases proximal to the conalbumin Hogness box in an experiment directly comparable with that reported here has no effect on *in vitro* transcription by the conalbumin promoter (15), end effects are probably not an explanation of our results. Apparently, the T-A-T-A-A-A concensus sequence alone is not sufficient for correct initiation in this system. Sequences further proximal in the U3 region of the LTR apparently are necessary, in contrast to results reported for the conalbumin promoter (15); however, the cell-free systems used in the two studies are not identical. Our finding that transcripts are initiated with equal efficienty *in vitro* from the 5′ and 3′ LTRs of the AKR provirus does not exclude the possibility that host sequences may exert a *cis*-dominant effect *in vivo*.

We estimate that the efficiency of specific initiation in 1-hr incubations on the three retroviral pol II promoters is approximately the same, but the accumulation of transcripts from these promoters is about one-eighth that observed for the Ad2 late promoter. These retroviral promoters could therefore be classified with the "weak" *in vitro* promoters tested so far, including those for ovalbumin (15) and mouse globin (16).

Nuclease mapping of the 5′ ends of *in vitro* RNAs shows they are the same as those of the corresponding mature viral RNAs. This result argues that the promoters active in the 5′ LTRs *in vitro* are also active intracellularly. This conclusion must be qualified to include the possibility that larger transcripts are subject to processing in the *in vitro* system.

The most unusual of our results is the α-amanitin-resistant promoter located in unique Ha-MuSV DNA sequences. The properties of the enzyme active at this site are consistent with those of RNA polymerase III. In addition to being resistant to low levels and sensitive to high levels of α-amanitin, this transcript can be specifically terminated near map coordinate 4000. To date, pol III is the only enzyme that has been shown to specifically terminate *in vitro* (33, 34). Furthermore, the cell-free extract used in these experiments is known to contain pol III activity (14). However, our data are indirect and further characterization with purified components is needed to clarify this issue.

Whether this promoter is active *in vivo* and whether it has any biological relevance remains unknown at present. However, it is interesting that the 5′ end of this *in vitro* transcript is approximately coincident with the 5′ end of the Ha-MuSV *src* gene (21).

We wish to thank Edward Scolnick for helpful suggestions and encouragement throughout the course of this work. Esther Chang contributed the *Eco*RI/*Bam*HI subclone of Ha-MuSV, Ron Ellis contributed the *Pst* I subclone, and Jim Manley contributed the *Bal* I E fragment subclone of Ad2. Drs. Chang, Ellis, and D. Lowy also provided valuable mapping information for the retroviral genomes. We thank Lorraine Shaughnessy for typing the manuscript.

1. Bishop, J. M. (1978) *Annu. Rev. Biochem.* **47**, 35–88.
2. Jacquet, M., Groner, Y. & Hurwitz, J. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3045–2049.
3. Rymo, L., Parsons, J. T., Coffin, J. M & Weissman, C. (1974) *Proc. Natl. Acad. Sci.* **71**, 2782–2786.
4. Hsu, T. W., Sabran, J. L., Mark, G. E., Guntaka, R. V. & Taylor, J. M. (1978) *J. Virol.* **28**, 810–818.
5. Shank, P. R., Hughes, S. H., Kung, H.-J., Majors, J. E., Quintrell, N., Guntaka, R. V., Bishop, J. M. & Varmus, H. E. (1978) *Cell* **15**, 1383–1395.
6. Coffin, J. M. (1979) *J. Gen. Virol.* **42**, 1–26.
7. Sabran, J. L., Hsu, T. W., Yeater, C., Kaji, A., Mason, S. W. & Taylor, J. M. (1979) *J. Virol.* **29**, 170–178.
8. Hughes, S. H., Shank, P. R., Spector, D. H., Kung, H.-J., Bishop, J. M., Varmus, H. E., Bogt, P. K. & Breitman, M. L. (1978) *Cell* **15**, 1397–1410.
9. Majors, J. E. & Varmus, H. E. (1980) in *Animal Virus Genetics*, ICN–UCLA Symposia on Molecular and Cellular Biology, eds. xxxx (Academic, New York), pp. 241–253.
10. Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3937–3941.
11. Sutcliffe, J. G., Schinnick, T. M., Verma, I. M. & Lerner, R. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3302–3306.
12. Donehower, L. A., Huang, A. L. & Hager, G. L. (1981) *J. Virol.* **37**, 226–238.
13. Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. (1979) *Cell* **18**, 469–484.
14. Manley, J. L., Fire, A., Cano, A., Charp, P. A. & Gefter, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3855–3859.
15. Wasylyk, B., Kedinger, C., Corden, J., Brison, O. & Chambon, P. (1980) *Nature (London)* **285**, 367–373.
16. Luse, D. S. & Roeder, R. G. (1980) *Cell* **20**, 691–699.
17. Green, M. R. & Roeder, R. G. (1980) *Cell* **22**, 231–242.
18. Yamamoto, T., deCrombrugghe, B. & Pastan, I. (1980) *Cell* **22**, 787–797.
19. Lowy, D. R., Rands, E., Chattopadhyay, S. K., Garon, C. F. & Hager, G. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 614–618.
20. Hager, G. L., Chang, E. H., Chan, H. W., Garon, C. F., Israel, M. A., Martin, M. A., Scolnick, E. M. & Lowy, D. R. (1979) *J. Virol.* **31**, 795–809.
21. Ellis, R. W., DeFeo, D., Maryak, J. M., Young, H. A., Shih, T. Y., Chang, E. H., Lowy, D. R. & Scolnick, E. M. (1980) *J. Virol.* **36**, 408–420.
22. DeFeo, D., Gonda, M. A., Young, H. A., Chang, E., Lowy, D. R. Scolnick, E. M. & Ellis, R. W. (1981) *Proc. Natl. Acad. Sci. USA* **78**, 3332–3328.
23. Zimmerman, S. B. & Sandeen, G. (1966) *Anal. Biochem.* **14**, 269–277.
24. Maxam, A. & Gilbert, W. (1980) *Proc. Natl. Acad. Sci.* **77**, 560–564.
25. Benoist, C., O'Hare, K., Breathnach, R. & Chambon, P. (1980) *Nucleic Acids Res.* **8**, 127–142.
26. Scolnick, E. M. & Parks, W. P. (1974) *J. Virol.* **13**, 1211–1219.
27. Baker, C. C., Herisse, J., Coeirtais, G., Galibert, F. & Ziff, E. (1979) *Cell* **18**, 569–580.
28. Rose, J. K., Haseltine, W. A. & Baltimore, D. (1976) *J. Virol.* **20**, 324–329.
29. Sollner-Webb, B. & Reeder, R. H. (1979) *Cell* **18**, 485–499.
30. Robinson, H. L., Pearson, M. N., DeSimone, D. W., Tsichlis, P. N. & Coffin, J. M. (1980) *Cold Spring Harbor Symp. Quant. Biol.* **44**, 1133–1140.
31. Neel, B. G., Hayward, W. S., Robinson, H. L., Fang, J. F. & Astrin, S. M. (1981) *Cell* **23**, 323–334.
32. Payne, G. S., Courtneidge, S. A., Crittenden, L. B., Fadly, A. M., Bishop, J. M. & Varmus, H. E. (1981) *Cell* **23**, 311–322.
33. Weil, P. A., Segall, J., Harris, B., Ng, S. Y. & Roeder, R. G. (1979) *J. Biol. Chem.* **254**, 6163–6173.
34. Bogenhagen, D. R., Sakonju, S. & Brown, D. D. (1980) *Cell* **19**, 27–35.