Proc. Natl. Acad. Sci. USA
Vol. 78, No. 9, pp. 5411–5415, September 1981
Biochemistry

Identification of a RNA polymerase II initiation site in the long terminal repeat of Moloney murine leukemia viral DNA

(*in vitro* transcription/RNase T1/nuclease S1/35S virion RNA)

SHELLA A. FUHRMAN*, CHARLES VAN BEVEREN†, AND INDER M. VERMA†

*Department of Biology, University of California at San Diego, La Jolla, California 92093; and †Tumor Virology Laboratory, The Salk Institute, San Diego, California 92138

Communicated by E. Peter Geiduschek, May 26, 1981

ABSTRACT **We have used a soluble *in vitro* RNA polymerase II transcription system to define the site of initiation of Moloney murine leukemia viral RNA synthesis. Molecularly cloned integrated and unintegrated Moloney murine leukemia virus DNAs were used as templates. The 5′ ends of *in vitro* transcripts and virion RNA of Moloney murine leukemia virus were compared by nuclease S1 protection experiments. Our results indicate that viral sequences upstream of the *in vivo* cap site are implicated in the transcription of viral RNA and that the 5′ end of an *in vitro* transcript derived from an integrated Moloney murine leukemia virus clone corresponds to the 5′ end of viral genomic RNA.**

During the life cycle of retroviruses, viral genomic RNA is transcribed into double-stranded RNA by reverse transcriptase (RNA-dependent DNA nucleotidyltransferase; refs. 1 and 2). After integration into the host chromosome, viral DNA is transcribed by cellular DNA-dependent RNA polymerase II (3, 4). No large precursor of viral genomic RNA has been observed in infected cells (5, 6). Furthermore, the unintegrated forms of viral DNA (whether synthesized *in vivo* or *in vitro*) are infectious, regardless of their subsequent integration sites in the chromosomal DNA (7–11). It thus appears plausible that proviral DNA might contain its own promoter for initiation of transcription, although the possibility that the transcription of an integrated provirus is under the control of closely juxtaposed cellular promoters cannot be excluded. Structural analysis of the long terminal repeat (LTR) from several murine and avian retroviruses shows a sequence similar to a Hogness–Goldberg (T-A-T-A) box, upstream from the 5′-cap nucleotide (12–18). The Rous sarcoma virus LTR contains an *in vitro* RNA polymerase II initiation site 23 base pairs (bp) downstream from a T-A-T-A-box-like sequence (18). The complete nucleotide sequences of the 5′ LTR of an integrated Moloney murine leukemia virus (M-MuLV) DNA and of an LTR of an unintegrated form of M-MuLV DNA have been determined (ref. 13; unpublished results); the sequences include the 5′-cap site of viral RNA and two possible T-A-T-A-like boxes (13). One of the T-A-T-A-like sequences is located immediately to the right of the *Sac* I site, at position −25 to −31 of the LTR (Sac-T-A-T-A-like), and is likely to be part of the initiation site used for viral RNA synthesis. However, a C-A-A-T-A-A-A sequence that is 15 nucleotides (nt) to the right of the *Kpn* I site (Kpn-T-A-T-A-like) is identical to the Sac-T-A-T-A-like box and should therefore be regarded as part of a second candidate RNA synthesis initiation site.

Because we had molecularly cloned integrated M-MuLV DNA fragments that contain the 5′ LTR together with adjacent cellular sequences and also unintegrated M-MuLV (Fig. 1), we designed experiments to determine: (*i*) whether any part of the M-MuLV LTR, particularly either of the T-A-T-A-like sequences, can function as sites for initiation of transcription *in vitro* by RNA polymerase II and (*ii*) whether there are signals in the cloned cellular flanking sequences that could give rise to the initiation of viral transcripts. We used the *in vitro* soluble transcription system described by Manley *et al.* (21) as the source of RNA polymerase II and factors.

In this report, we show that (*i*) the Sac-T-A-T-A-like sequence is implicated in the initiation of transcription of viral RNA, (*ii*) an *in vitro* transcript of an integrated M-MuLV clone and viral genomic RNA have the same 5′ ends, and (*iii*) there is an *in vitro* RNA polymerase II initiation site in the plasmid pBR322.

METHODS

Recombinant DNA Materials. All recombinant DNA materials were handled according to the National Institutes of Health guidelines for recombinant DNA research under the supervision of the local Institutional Biosafety Committee. Specific plasmids are described in Fig. 1 (M-MuLV clones) and in ref. 20 (pSmaF). Plasmid DNA was isolated as described (22) and purified by centrifugation in CsCl/ethidium bromide. Ethidium bromide was removed with isopropyl alcohol; the DNA was then gel filtered through a Bio-Gel A5M column (Bio-Rad). Restriction endonuclease-digested DNA was extracted with phenol and concentrated by ethanol precipitation prior to transcription. DNA fragments were eluted from preparative agarose gels as described in ref. 23.

RNA Polymerase II Extracts and *in Vitro* RNA Synthesis. Extracts from HeLa cells (21) were used as the source of RNA polymerase II and factors for *in vitro* transcription. Reaction mixtures (30 μl) contained 12 μl of extract and 1–1.5 μg of restriction endonuclease-digested DNA or 0.4–0.6 μg of purified DNA fragments. Other components were as described (19), except for the addition of 3.3 mM phospho*enol*pyruvate. RNA synthesis was started by adding extract to the other components, continued for 60 min at 30°C, and then stopped by adding 15 μl of 30 mM EDTA/0.6% Sarkosyl/tRNA (1.5 mg/ml)/0.4 mM aurintricarboxylic acid. After phenol extraction and ethanol precipitation, RNA was subjected to electrophoresis on 4% acrylamide/7 M urea/0.05 M Tris borate, pH 8.3/1 mM EDTA gels (24), and the gels were autoradiographed at −70°C using unflashed Cronex 4 film and Quanta III intensifying screens (DuPont).

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: M-MuLV, Moloney murine leukemia virus; LTR, long terminal repeat; nt, nucleotides; bp, base pair(s); kbp, kilobase pair(s).

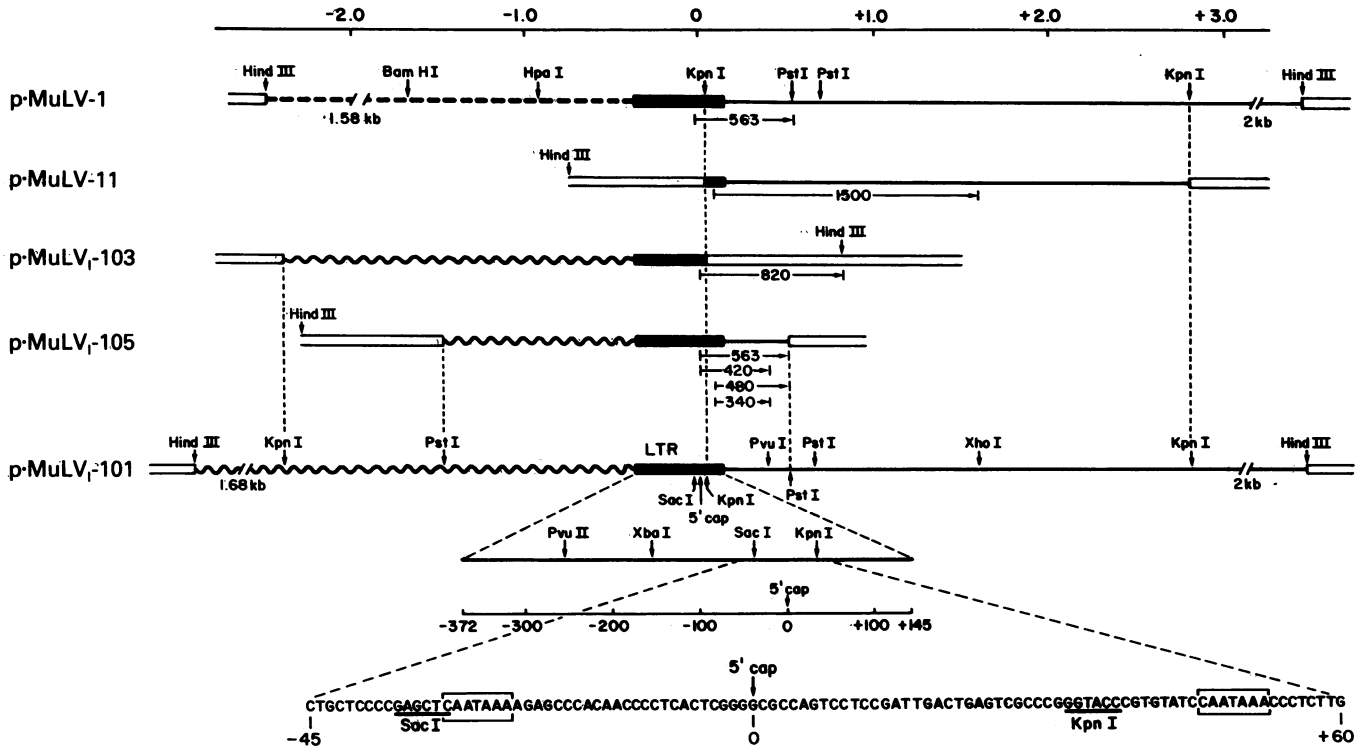

FIG. 1. Physical map of cloned M-MuLV DNAs. The construction of p·MuLV-101 has been described (19); it contains the 5′ LTR of M-MuLV DNA, 4.9 kilobase pairs (kbp) of proviral DNA, and 3.5 kbp of adjacent cellular DNA inserted at the *Hin*dIII site of pBR322. p·MuLV$_I$-103 and p·MuLV-11 have the indicated *Kpn* I fragments of p·MuLV$_I$-101 inserted into the *Pst* I site of pBR322 by (dC)·(dG) tailing (13). The inserted sequences can be separated from vector sequences with *Pst* I or *Kpn* I. In both plasmids, the M-MuLV LTR/pBR322 junction is ≈780 bp away from the single pBR322 *Hin*dIII site that is shown on each line. [A measure of uncertainty is imposed by the unknown lengths of the (dC)·(dG) tails.] p·MuLV$_I$-105 contains the indicated *Pst* I fragment of p·MuLV$_I$-101 cloned directly into the *Pst* I site of pBR322. In this plasmid, the cellular DNA-pBR322 junction is 779 bp from the pBR322 *Hin*dIII site. p·MuLV-1 contains the entire genome of M-MuLV inserted in the *Hin*dIII site of pBR322 and is derived from the unintegrated M-MuLV insert in λ·MuLV-1 (19, 20). An expanded map of the M-MuLV LTR, including the positions of relevant restriction endonuclease sites, is shown below the p·MuLV$_I$-101 map; the nucleotide sequence between −45 (left of the *Sac* I site) and +60 (right of the *Kpn* I site) is shown below the expanded LTR map. The Sac-T-A-T-A-like box, the Kpn-T-A-T-A-like box, and the 5′-cap nucleotide are indicated. Potential transcripts originating downstream from either the Sac-T-A-T-A-like box (563 nt, 420 nt, and 820 nt) or the Kpn-T-A-T-A-like box (480 nt, 340 nt, and 1500 nt) are indicated below the maps of p·MuLV-1, p·MuLV-11, p·MuLV$_I$-103, and p·MuLV$_I$-105. ▬, LTR of M-MuLV; ——, 5′ viral sequences; - - -, 3′ viral sequences; ∿∿, cellular sequences; □, pBR322 sequences; →, potential transcripts.

RNase T1 Oligonucleotide Analysis. RNA was eluted from unfixed gels as described in ref. 25. *In vivo* ^{32}P-labeled yeast 5S and 5.8S RNAs were prepared as described (26). Labeled RNA was digested at 37°C with 0.2 μg of RNase T1 (Calbiochem) in the presence of 5 μg of unlabeled yeast tRNA in 5 μl of H_2O; the digests were then diluted with 10 μl of 95% formamide/0.5% xylene cyanol/0.5% bromphenol blue, heated 5 min at 65°C, and subjected to electrophoresis on a 20% acrylamide sequencing gel (24).

Nuclease S1 Mapping. The p·MuLV$_I$-105 DNA was digested with *Sau* 3A and treated first with calf intestine alkaline phosphatase and then with terminally labeled T4 polynucleotide kinase to a specific activity of 5–6 × 10^6 cpm/pmol of 5′ ends; the products were fractionated by electrophoresis on a 3.5% acrylamide gel so that the 442-bp fragment from the LTR could be isolated. Single strands of that fragment were obtained by electrophoresis on an 8% acrylamide/(bis)acrylamide (60:1) strand separation gel; the identity of the single strands was determined by standard sequence analysis (24).

The 540-nt transcript from *Pst* I-digested p·MuLV$_I$-105 or virion M-MuLV 35S RNA (a gift of R. Bosselman) was hybridized to the labeled 442-nt probe (2 to 3 fmol) in 25 μl of 0.3 M NaCl/10 mM Tris·HCl, pH 8/0.1 mM EDTA for 4 hr at 65°C. The mixture was cooled; 200 μl of 250 mM NaCl/1 mM $ZnCl_2$/30 mM NaOAc, pH 5.2, containing denatured calf thymus DNA (20 μg/ml) was added; and this mixture was incubated at 37°C with 60 units of nuclease S1 (Sigma). Digestion was stopped after 1 hr and EDTA and 100 ng of carrier DNA was added. After phenol extraction and ethanol precipitation, samples were subjected to electrophoresis on 6% acrylamide sequencing gels.

RESULTS

Characterization of the *in Vitro* Transcription System. First, by way of control, the *in vitro* system was programmed with *Sma* I-digested pSmaF DNA. The SmaF fragment of adenovirus-2 DNA contains the major late adeno promoter (27); the 536-nt transcript from *Sma* I-digested pSmaF (Fig. 2*A*, lane 7) has been well characterized in other laboratories (21, 27). The minor α-amanitin-sensitive RNAs that are smaller than the major 536-nt runoff transcript have not been characterized as their appearance varies from experiment to experiment. Such minor bands could result from (*i*) permature termination by the RNA polymerase II, (*ii*) "pausing" of the RNA polymerase during transcription, (*iii*) breakdown of the major transcripts, or (*iv*) low-level utilization of other initiation sites in the *in vitro* system. We have not investigated these alternatives further. The 600- and 2000-nt RNAs that are seen in all the even-numbered lanes are also observed when the extract is incubated without added DNA (Fig. 2*A*, lane 13), and their synthesis is not sensitive to actinomycin D (F. Keppel, personal communication).

Transcription of M-MuLV-Containing DNA Fragments. The sizes of the transcripts (Fig. 2*A*) that are marked by arrows

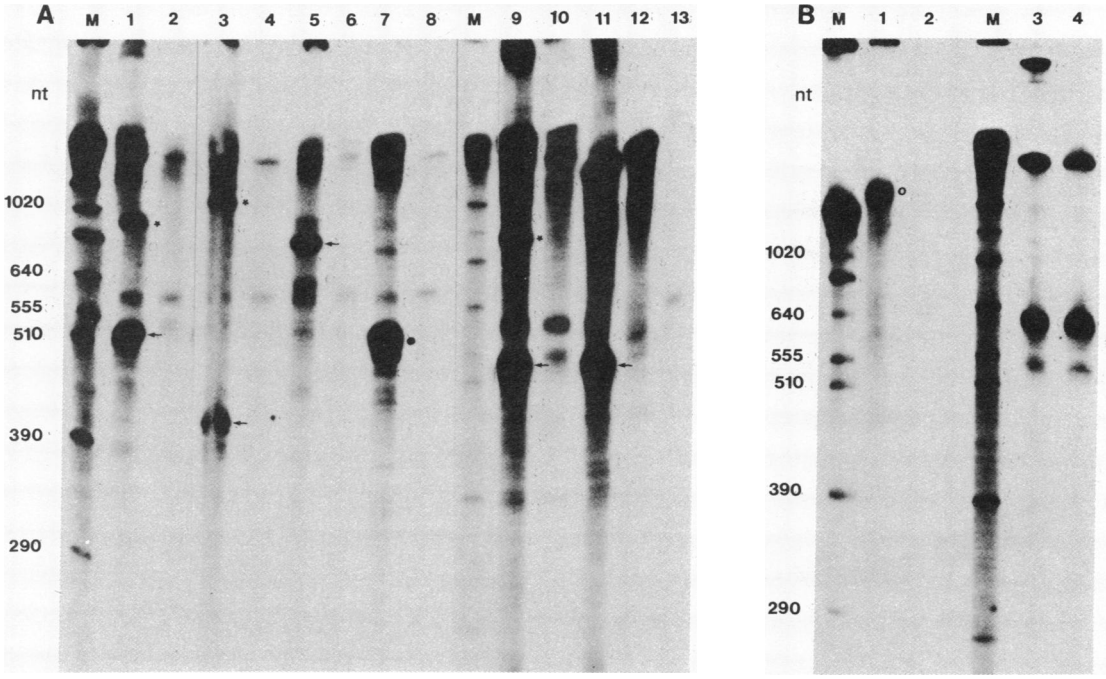

FIG. 2. Electrophoretic analysis of *in vitro* RNA transcription with restriction endonuclease-cleaved DNA templates. Odd-numbered lanes, RNA synthesis without α-amanitin; even-numbered lanes, controls to which α-amanitin (1 μg/ml) was added to inhibit RNA polymerase II. Marker lanes (M) are *in vitro* RNA synthesized by *Bacillus subtilis* holoenzyme on intact phage SP01 DNA as template (28). (*A*) Templates are *Pst* I-digested p·MuLV$_I$-105 DNA (lanes 1, 2, 9, and 10), *Pvu* I-digested p·MuLV$_I$-105 (lanes 3 and 4), *Hin*dIII-digested p·MuLV$_I$-103 (lanes 5 and 6), *Sma* I-digested pSmaF (major late adeno promoter; ref. 20) (lanes 7 and 8), 1.65-kbp *Hpa* I/*Pst* I fragment of p·MuLV-1 (lanes 11 and 12), and none (lane 13). (*B*) Templates are *Xho* I-digested p·MuLV-11 (lanes 1 and 2) and 1.3-kbp *Pst* I/*Xba* I fragment of p·MuLV$_I$-105 (lanes 3 and 4). α-Amanitin-sensitive transcripts: →, M-MuLV clones; ●, pSmaF; *, derived from pBR322 DNA; ○, derived from *Xho* I-digested p·MuLV-11 (this is not mapped but probably is also derived from pBR322 DNA).

in lanes 1 and 9 (540 nt, *Pst* I-cut p·MuLV$_I$-105), lane 3 (400 nt, *Pvu* I-digested p·MuLV$_I$-105), lane 5 (820 nt, *Hin*dIII-cut p·MuLV$_I$-103), and lane 11 (540 nt, 1.65-kbp *Hpa* I/*Pst* I fragment of p·MuLV-1) are consistent with transcription initated at a site near the Sac-T-A-T-A-like box. As 480-nt RNA from *Pst* I-cleaved p·MuLV$_I$-105 DNA and 340-nt RNA from *Pvu* I-cut p·MuLV$_I$-105 are not seen, it is unlikely that transcription initiates at a site near the Kpn-T-A-T-A-like box. Furthermore, there is no α-amanitin-sensitive 1500-nt transcript from *Xho* I-digested p·MuLV-11 DNA (however, see below). We have not characterized the minor α-amanitin-sensitive transcripts seen in lanes 1, 3, 5, 9, and 11 of Fig. 2*A*, again because their appearance varies from experiment to experiment.

No transcripts are observed that are both longer than the 540-nt runoff RNA of *Pst* I-digested p·MuLV$_I$-105 and shorter by 140 nt when *Pvu* I-digested p·MuLV$_I$-105 is used as template. In addition, there are no α-amanitin-sensitive transcripts from the 1.3-kbp *Pst* I/*Xba* I fragment of p·MuLV$_I$-105 (Fig. 2*B*, lanes 3 and 4). This fragment contains LTR sequences to the left of the *Xba* I site, as well as 1.1 kbp of adjacent cellular sequences. These data suggest that there are no *in vitro* RNA polymerase II initiation sites to the left of the Sac-T-A-T-A-like box or in the flanking cellular sequence.

We have compared the RNase T1 oligonucleotide pattern of the 540-nt transcripts from *Pst* I-digested p·MuLV$_I$-105 and from the 1.65-kbp *Hpa* I/*Pst* I fragment of p·MuLV-1. The oligonucleotide patterns of the two transcripts are essentially identical (Fig. 3, lanes 3 and 4). The DNA sequence of M-MuLV from the *Sac* I site in the LTR to the *Pst* I site in viral coding sequence (unpublished data) predicts RNase T1 oligonucleotides of length 19, 13, 12, 11, and 9 nt (when [α-^{32}P]UTP is used for RNA synthesis, the relative labeling of these is 5, 1, 4, 11, and 3). We observe oligonucleotides of length 19, 12, 11, and 9 (marked A, B, C, and D, respectively, in Fig. 3); their labeling is consistent with that predicted by the DNA sequence. The single predicted 19-base oligonucleotide spans the Kpn-T-A-T-A-like box (Fig. 1, +41 to +60 of the LTR sequence shown); thus, it is likely that the 540-nt transcript has its 5′ end to the left of the Kpn-T-A-T-A-like box. The predicted 13-mer is barely detectable, presumably because of a combination of its low relative labeling and the fact that unflashed film is nonlinear in its response to low amounts of radioactivity (31). We observe other very faint labeled oligonucleotides not predicted by the DNA sequence. These could be derived from the background RNA that contaminante the excised 540-nt transcripts.

From these observations, we infer that (*i*) the Sac-T-A-T-A-like box is functional *in vitro*; (*ii*) the 540-nt runoff transcripts of p·MuLV$_I$-105 and p·MuLV-1 are identical; and (*iii*) if either the Kpn-T-A-T-A-like box or the left half of the LTR and adjacent cellular sequences have promoter function *in vitro*, it must be at an undetectable level.

Nuclease S1 Mapping of the 5′ End of the 540-nt Runoff Transcript of *Pst* I-Digested p·MuLV$_I$-105 and M-MuLV Virion RNA. To define more precisely the 5′ end of the *in vitro* transcript, we chose a nuclease S1 protection experiment (32). We hybridized the 540-nt *in vitro* transcript of p·MuLV$_I$-105 and M-MuLV virion RNA to a labeled single-strand DNA probe that extends from +164 of the viral sequence (^{32}P-labeled 5′ end) to −275 of the LTR sequence (3′ end). The 5′ terminus of M-MuLV 35S RNA is m^{7}G$^{5'}$ppp$^{5'}$GmpCp (33), and the 3′ end of strong-stop DNA is 5′...G-G-C-G-C 3′ (13); thus, a 164-nt duplex should form between M-MuLV virion DNA and the labeled DNA probe. We observe 165-nt DNA after nuclease S1 digestion, as well as other bands at 166 and 167 nt (Fig. 4, lane

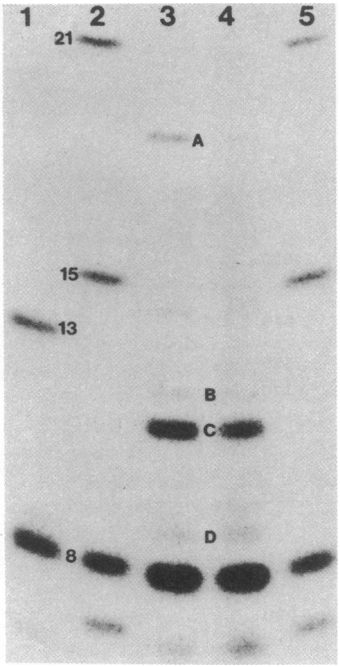

FIG. 3. RNase T1-generated oligonucleotides of *in vitro* transcripts. RNA samples were digested with RNase T1 and subjected to electrophoresis. Lanes: 1, yeast 5S RNA; 2 and 5, yeast 5.8S RNA; 3, 540-nt runoff transcript of *Pst* I-digested p·MuLV$_I$-105; 4, 540-nt runoff transcript of the 1.65-kbp *Hpa* I/*Pst* I fragment of p·MuLV-1. Lengths of the large RNase T1 oligonucleotides of yeast 5S and 5.8S RNAs are indicated between the lanes (29, 30). Estimated lengths of the oligonucleotides in lanes 3 and 4 are A, 19 nt; B, 12 nt; C, 11 nt; and D, 9 nt. Oligonucleotides of <8 bases are not shown. In this size range, there is no detectable difference in the pattern between lanes 3 and 4.

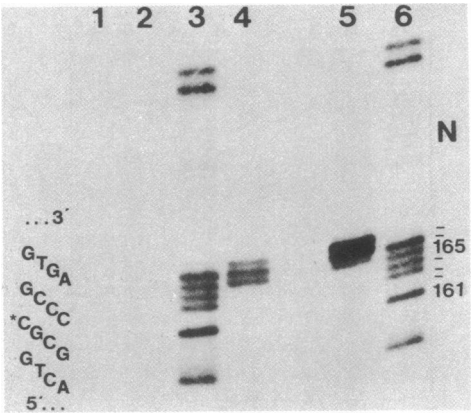

FIG. 4. Nuclease S1 mapping of the 5' end of *in vitro* and *in vivo* M-MuLV RNAs. RNA samples were hybridized to the labeled 442-nt DNA probe and treated with nuclease S1 unless otherwise noted. RNA in hybrids was degraded by alkali treatment and the 5'-end-labeled DNA probe was subjected to electrophoresis. Lanes: 1, 442-nt probe, no nuclease S1; 2, 442-nt probe; 3 and 6, markers [a sequence ladder obtained with cytidine cleavage of probe (24)]; 4, probe hybridized to 540-nt runoff transcript of *Pst*-I digested p·MuLV$_I$-105; 5, probe hybridized to M-MuLV virion RNA (1.7 fmol). A band in the sequencing ladder representing the *n*th nucleotide from the labeled 5' end of the probe is actually of length $N = n - 1$ because the sequencing reaction removes the *n*th nucleotide (24, 32). The DNA sequence of the relevant portion of the 442-nt probe is shown on the left. The cytidine that represents the 3' end of strong-stop DNA is marked by an asterisk.

5). This heterogeneity probably results from partial protection of the DNA by the cap structure at the 5' end of virion RNA and has been observed by others (32, 34–36). More importantly, the same three DNA fragments are observed when *in vitro* p·MuLV$_I$-105 RNA is used to protect the labeled DNA probe (Fig. 4, lane 4). The relative intensities of the three bands differ slightly from those protected by virion RNA. This, however, could be a reflection of incomplete methylation *in vitro* (18), as the 165-nt fragment is (relatively) more intense with *in vitro* RNA than with virion RNA. From these data, we conclude that the 5' end of *in vitro*-synthesized M-MuLV RNA corresponds to the 5' end of *in vivo* virion RNA.

Transcription from pBR322 Sequences. We do observe α-amanitin-sensitive transcripts that are longer than 540 and 400 nt; these are indicated by asterisks in Fig. 2*A*, lanes 1, 3, and 9. The transcript from *Pst* I-digested p·MuLV$_I$-105 is ≈930 nt long and the transcript from *Pvu* I-digested p·MuLV$_I$-105 is ≈1100 nt long. The relationship between these sizes is inconsistent with transcription starting at a point to the left of the Sac-T-A-T-A-like box and reading to the right (because the *Pvu* I transcript is longer, rather than shorter, than the *Pst* I transcript). Southern mapping (37) of these transcripts (data not shown) demonstrates that they hybridize to pBR322 sequences that are located to the right of the viral *Pst* I site of p·MuLV$_I$-105 and not to the cloned cellular/viral sequences. There is a sequence, C-T-A-T-A-A-A-G, 1 kbp to the right of the proviral DNA/pBR322 junction in p·MuLV$_I$-105 at position 2611–2618 in pBR322 DNA [in the numbering system of Sutcliffe (38)], that may be associated with RNA polymerase II-dependent initiation in this *in vitro* system. We have not mapped the >2000-nt α-amanitin-sensitive transcript from *Xho* I-digested p·MuLV-11 (Fig. 2*B*, lane 1); however, the length of that RNA is consistent with initiation within pBR322 DNA and elongation leftward to the *Xho* I site in proviral DNA.

DISCUSSION

We have shown that a site in the M-MuLV LTR used for initiation of RNA synthesis *in vitro* corresponds to a site used *in vivo* for synthesis of M-MuLV virion RNA. That this site is 25 nt downstream from the Sac-T-A-T-A-like box provides evidence that the Sac-T-A-T-A-like box is important for initiation of transcription at the known cap site both *in vivo* and *in vitro*.

We find no evidence that the Kpn-T-A-T-A-like box is functional as an *in vitro* promoter; the differences that create this distinction must reside in sequences outside the obvious T-A-T-A boxes. Corden *et al.* (36) have shown that sequences from −10 to −44 nt upstream of the cap site are required for correct *in vitro* transcription of the conalbumin gene; sequences from −12 to −32 are required for transcription of the major late adenovirus promoter. Construction of appropriate deletion/fusion mutants of the two M-MuLV LTR T-A-T-A-like boxes will be instructive. As the Sac-T-A-T-T-like and Kpn-T-A-T-A-like sequences are in the 5' LTR, they must be present in the 3' LTR. It will be of interest to determine whether the T-A-T-A-like sequences in the 3' LTR can also be used as sites for initiation of RNA transcription. A transcript initiated in the 3' LTR can presumably elongate into adjacent cellular sequences. Such a mechanism has been postulated for oncogenesis by avian leukosis virus (39, 40).

Although we are unable to detect *in vitro* promoter sites to the left of the Sac-T-A-T-A-like box (in the LTR or cellular flanking sequences) that could give rise to viral transcripts of the correct orientation, it seems quite likely that these sequences influence viral expression *in vivo*: there is evidence of different phenotypes of viral expression in four BALB/Mo mouse substrains that carry a single M-MuLV genome at different chromosomal integration sites (41). That LTR sequences upstream

of the Sac-T-A-T-A-like box may be important for viral expression *in vivo* is suggested by experiments showing that the T-A-T-A box is not an absolute requirement for gene expression *in vivo* and that sequences further upstream from the 5′ terminal nucleotide are more important in directing the reading of RNA polymerase II transcription units in intact cells (34, 36). It will be of interest to construct deletion mutants in the M-MuLV LTR to the left of the Sac-T-A-T-A-like box and to test their infectivity *in vivo*.

Note Added in Proof. Nuclease S1 mapping of the 5′ end of the 540-nt transcript from *Pst* I-digested p·MuLV-1 results in a pattern identical to that shown in Fig. 4 for M-MuLV virion RNA and the 540-nt transcript from *Pst* I-digested p·MuLV$_I$-105.

We thank D. Engelke and A. Berns for discussions; A. Esty and T. Friedmann for HeLa cells; and M. Brennan, J. T. Simon, and S. Burns for help in the preparation of the manuscript. S.A.F. was supported by a postdoctoral fellowship from the National Institute of General Medical Sciences. This work was supported by grants from the National Institute of General Medical Sciences, the National Cancer Institute, and the American Cancer Society (to I.M.V. and E. P. Geiduschek).

1. Temin, H. M. & Baltimore, D. (1972) in *Advances in Virus Research*, eds. Smith, K. M. & Lauffer, M. A. (Academic, New York), Vol. 17, pp. 129–186.
2. Verma, I. M. (1977) *Biochim. Biophys. Acta* **473**, 1–38.
3. Jacquet, M., Groner, Y., Monroy, G. & Hurwitz, J. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 3045–3049.
4. Rymo, L., Parsons, J. T., Coffin, J. M. & Weissmann, C. (1974) *Proc. Natl. Acad. Sci. USA* **71**, 2782–2786.
5. Haseltine, W. A. & Baltimore, D. (1976) *J. Virol* **19**, 331–337.
6. Fan, H. (1977) *Cell* **11**, 297–305.
7. Rothenberg, E., Smotkin, D., Baltimore, D. & Weinberg, R. A. (1977) *Nature (London)* **269**, 122–126.
8. Andersson, P., Goldfarb, M. P. & Weinberg, R. A. (1979) *Cell* **16**, 63–75.
9. Lai, M. H. T. & Verma, I. M. (1980) *Virology* **100**, 194–198.
10. Clayman, C. H., Mosharrafa, E. & Faras, A. J. (1979) *J. Virol.* **29**, 242–249.
11. Lai, M. H. T. & Verma, I. M. (1980) *Virology* **104**, 407–417.
12. Kunitada, S., Mizutani, S. & Temin, H. M. (1980) *Nature (London)* **285**, 550–554.
13. Van Beveren, C., Goddard, J. G., Berns, A. & Verma, I. M. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3307–3311.
14. Sutcliffe, J. G., Shinnick, T. M., Verma, I. M. & Lerner, R. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3302–3306.
15. Dhar, R., McClements, W. L., Enquist, L. W. & Vande Woude, G. F. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3937–3941.
16. Premkumar Reddy, E., Smith, M. J., Canaani, E., Robbins, K. C., Tronick, S. R., Zain, S. & Aaronson, S. A. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 5234–5238.
17. Czernilofsky, A. P., Levinson, A. D., Varmus, H. E., Bishop, J. M., Tischer, E. & Goodman, H. M. (1980) *Nature (London)* **287**, 198–203.
18. Yamamoto, T., de Crombrugghe, B. & Pastan, I. (1980) *Cell* **22**, 787–797.
19. Berns, A. J. M., Lai, M. H. T., Bosselman, R. A., McKennett, M. A., Bacheler, L. T., Fan, H., Robanus Maandag, E. C., van den Putten, H. & Verma, I. M. (1980) *J. Virol.* **36**, 254–263.
20. Van Beveren, C., Galleshaw, J. A., Jonas, V., Berns, A. J. M., Doolittle, R., Donoghue, D. J. & Verma, I. M. (1981) *Nature (London)* **289**, 258–262.
21. Manley, J. L., Fire, A., Cano, A., Sharp, P. A. & Gefter, M. L. (1980) *Proc. Natl. Acad. Sci. USA* **77**, 3855–3859.
22. Clewell, D. B. & Helinski, D. (1970) *Biochemistry* **9**, 4428–4440.
23. Klemenz, R. & Geiduschek, E. P. (1980) *Nucleic Acids Res.* **8**, 2679–2689.
24. Maxam, A. M. & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. USA* **74**, 560–565.
25. Hunter, T. & Garrels, J. I. (1977) *Cell* **12**, 767–781.
26. Knapp, G., Beckmann, J., Johnson, P. F., Fuhrman, S. A. & Abelson, J. (1978) *Cell* **14**, 221–236.
27. Weil, P. A., Luse, D. S., Segall, J. & Roeder, R. G. (1979) *Cell* **18**, 469–484.
28. Brennan, S., Romeo, J., Chelm, B. K. & Geiduschek, E. P. (1981) *Virology* **111**, 604–628.
29. Bell, G. I., DeGennaro, L. J., Gelfand, D. H., Bishop, R. J., Valenzuela, P. & Rutter, W. J. (1977) *J. Biol. Chem.* **252**, 8118–8125.
30. Valenzuela, P., Bell, G. I., Venegas, A., Sewell, E. T., Masiarz, F. R., DeGennaro, L. J., Weinberg, F. & Rutter, W. J. (1977) *J. Biol. Chem.* **252**, 8126–8135.
31. Laskey, R. A. (1980) *Methods Enzymol.* **65**, 363–371.
32. Weaver, R. F. & Weissmann, C. (1979) *Nucleic Acids Res.* **7**, 1175–1193.
33. Rose, J. K., Haseltine, W. A. & Baltimore, D. (1976) *J. Virol.* **20**, 324–329.
34. Benoist, C. & Chambon, P. (1981) *Nature (London)* **290**, 304–310.
35. Wasylyk, B., Kédinger, C., Corden, J., Brison, O. & Chambon, P. (1980) *Nature (London)* **285**, 367–373.
36. Corden, J., Wasylyk, B., Buchwalder, A., Sassone-Corsi, P., Kédinger, C. & Chambon, P. (1980) *Science* **209**, 1406–1414.
37. Southern, E. M. (1975) *J. Mol. Biol.* **98**, 503–517.
38. Sutcliffe, J. G. (1978) *Nucleic Acids Res.* **5**, 2721–2728.
39. Payne, G. S., Courtneidge, S. A., Crittenden, L. B., Fadly, A. M., Bishop, J. M. & Varmus, H. E. (1981) *Cell* **23**, 311–322.
40. Neel, B. G., Hayward, W. S., Robinson, H. L., Fang, J. & Astrin, S. M. (1981) *Cell* **23**, 323–334.
41. Jähner, D. & Jaenisch, R. (1980) *Nature (London)* **287**, 456–458.