The "sarcoma-specific" region of Moloney murine sarcoma virus 124
(RNase H/18S RNA fragment/proviral DNA/restriction endonuclease mapping)

DINO DINA*

Department of Biochemistry, University of California, Berkeley, California 94720; and Department of Genetics, Albert Einstein College of Medicine, Bronx, New York 10461

Communicated by Harry Eagle, March 24, 1978

ABSTRACT Labeled, purified 30S RNA from Moloney murine sarcoma virus was annealed to an excess of Moloney murine leukemia virus complementary DNA. Upon treatment of the resulting DNA-RNA hybrids with RNase H followed by sucrose gradient sedimentation, an undigested 18S RNA molecule was recovered. This RNA molecule was shown to represent the "sarcoma-specific" region of the virus. The unintegrated linear DNA provirus of murine sarcoma virus 124 was isolated from newly infected cells and a physical map of the sarcoma-specific region was obtained. First, unintegrated full-length linear proviral DNA molecules were cleaved by several restriction endonucleases. The reciprocal position and orientation with respect to the viral RNA of the resulting fragments were established. The location of the sarcoma-specific region was determined by competition-hybridization with 32P-labeled viral genomic RNAs and proviral DNA fragments. A 1500-base-pair fragment was obtained by cleavage with HindIII + Bgl II. This fragment mapped between 750 and 2500 base pairs from the right end of the proviral DNA (corresponding to the 3' terminus of the viral RNA) and contained the whole set of the sarcoma-specific information. This murine sarcoma virus proviral restriction fragment is approximately of the same size and map position as the isolated 18S sarcoma-specific RNA.

Moloney murine sarcoma virus 124 (MuSV) and Moloney murine leukemia virus cl.1 (MuLV) are structurally and biologically related (1-4). MuLV has been shown to act as a helper for MuSV 124 but does not by itself transform fibroblasts in vitro although it causes leukemia in vivo (5). MuSV is defective for pol and env functions but it contains the information for fibroblast transformation (src) (2-4, 6-14).

Cross-hybridization studies performed by annealing viral RNAs with cDNAs have shown that Moloney MuSV and MuLV are indeed similar (2, 8, 15). Seventy percent of the MuSV genome is shared by MuLV (common sequences) and 30% is MuSV-specific sequences (about 2000 nucleotides). The arrangement of these sequences on both MuLV and MuSV genomes has been investigated by cross annealing (4) and heteroduplex mapping (3). These experiments have shown that the MuSV genome is made of three main portions: a 3'-terminal region 1000 nucleotides long [including poly(a)] containing sequences shared by MuLV (3' common); a sarcoma-specific region (src) 1500 nucleotides long, located between 1000 and 2500 nucleotides from the 3' end of the viral genome; and 5' region 3500 nucleotides long containing sequences shared with MuLV (5' common).

Upon infection of NIH 3T3 cells with Moloney MuSV 124, a vast excess of MuSV 124 proviral DNA over MuLV proviral DNA is synthesized (unpublished data). This DNA has been characterized and shown to be cleaved by several restriction endonucleases. A preliminary physical map of the proviral MuSV 124 DNA has been reported recently (16).

The results mentioned above indicated that the sarcoma-specific region of MuSV 124 was located at a unique site as a fully contiguous sequence without any "common" RNA spacers (3, 4). However, it was not possible to exclude the presence of short intermingled common sequences within the sarcoma-specific region. To test this, one needs an experimental approach that readily detects short RNA-DNA hybrids.

Wang and Duesberg (17) reported that intact 3S viral RNA could be recovered after Escherichia coli RNase H treatment in the presence of oligo(dT). The poly(A) sequences on the viral RNA were digested during the treatment in a specific fashion. This result implied that E. coli RNase H, purified and used under rigorous conditions, would specifically recognize short (200 nucleotides at most) DNA-RNA hybrids (18).

I reasoned that, if I could obtain intact hybrids from full-length labeled MuSV 30S RNA and a large excess of unlabeled MuLV cDNA, only the regions paired with DNA should be degraded by digestion with RNase H and unhybridized regions should be left intact. This report describes the isolation and partial purification of a unique MuSV genomic RNA fragment containing the "MuSV-specific" information by this approach.

A physical map of the same MuSV-specific region, obtained by restriction endonuclease cleavage of unintegrated linear MuSV DNA provirus, is also described. Cleavage by a combination of two enzymes, Bgl II and HindIII, yields a 1500-base-pair-long DNA fragment containing most or all of the sarcoma-specific information.

MATERIALS AND METHODS

Cells and Viruses. Moloney MuSV 124 (1-4, 19) produced by thymus/bone marrow mouse cells (TB cells) was used for all the experiments. Moloney MuLV cl.1, produced by NIH 3T3 cells, was a gift of H. Fan.

Preparation of Viral RNAs. Labeled and unlabeled viral RNAs were prepared as described (2). Specific activity of 32P-labeled viral RNAs ranged between 1 and 2 X 10^6 cpm/μg in different experiments. Labeling of viral RNA with 125I was performed by the procedure of Commerford (20). The labeled RNA was further purified following a protocol suggested by P. Shank (personal communication). Briefly, after the initial incubation with 125I, the RNA was heated at 80° and sedimented through a 15-30% sucrose gradient (10 mM Tris, pH 7.0/1 mM EDTA/10 mM NaCl) for 4 hr at 40,000 rpm in a Spinco SW 41 rotor at 20°. 125I radioactivity at the top of the

Abbreviations: MuSV, murine sarcoma virus; MuLV, murine leukemia virus; NaDodSO4, sodium dodecyl sulfate.

* Present address: Department of Genetics, Albert Einstein College of Medicine, Bronx, NY 10461.
gradient was discarded (0–4 S), and the heterogeneous RNA of larger size was pooled and ethanol precipitated. The RNA was then dissolved in 1.8 ml of water and heated for 1 min at 100°C in the presence of 50 μg of unlabeled purified yeast tRNA and 50 μg of calf thymus DNA as carriers. The solution was cooled and layered over 2.2 ml of a saturated solution of cesium sulfate (Gallard and Schlesinger, optical grade). The sample was centrifuged for 60 hr at 33,000 rpm at 20°C in a Spinco SW 50.1 rotor. The narrow band of 125I-labeled RNA was pooled, diluted 1:5 with water, and ethanol precipitated twice.

Preparation of Total Proviral DNA. MuSV 124 unintegrated proviral DNA was prepared from newly infected NIH 3T3 cells as described (16) except that the CsCl fractionation step was omitted. This DNA contained mainly linear forms (~90%) and was shown to be almost completely double-stranded. Both strands were analyzed by hybridization to viral RNA and cDNA. More than 95% of the (–) strand DNA molecules were shown to be full length. About 20% of the (–) strand was full length and the remaining 80% showed an average molecular weight of 1–2 × 10^6. No specific regions of single-stranded DNA were observed (unpublished data).

Restriction Endonucleases and Enzyme Assay Conditions. EcoRI, Bgl II, and Dpn II were gifts of Tom Kadesch. BarnHI, HindIII, HincII, and Hae II were purchased from New England Biolabs and used under the suggested assay conditions. After digestion (for incubation time, see figure legends), the reactions were stopped with 50 mM EDTA/1% sodium dodecyl sulfate (NaDodSO4). Proviral DNA was extracted with buffered phenol, precipitated with ethanol, and resuspended in 10 mM Tris, pH 7.0/1 mM EDTA. In all cases, 1 enzyme unit was the amount of enzyme required to completely digest 1 μg of DNA in 1 hr at 37°C.

Gel Electrophoresis. Vertical slab gel electrophoresis of nucleic acids in 1.4% agarose was performed according to published procedures (21). Gels were usually run at 50 V for 6 hr at room temperature.

Blotting and Hybridization. Proviral DNA fragments were transferred from the gels onto Schleicher and Schuell nitrocellulose membrane sheets according to the technique described by Southern (22) and modified by Ketner and Kelley (23). After the transfer, the membrane sheets were cut into pieces two lanes wide and dried in a vacuum oven for 2 hr at 80°C. Hybridization with 125I-labeled viral RNA (1 × 10^6 cpm; specific activity, 50–100 × 10^6 cpm/μg) were performed in 0.9 M NaCl/0.09 M Na citrate at 70°C in Lucite cuvettes (110 × 25 × 1.5 mm; about 3 ml) sealed with a Teflon top. The sheets were rinsed twice at room temperature with 50 ml of 0.3 M NaCl/0.03 M Na citrate, incubated 1 hr at 60°C in the same buffer, treated for 1 hr at room temperature with pancreatic RNase A at 20 μg/ml, rinsed again with 0.3 M NaCl/0.03 M Na citrate, dried, and autoradiographed.

Autoradiography. Exposure of XR5 Kodak films for variable amounts of time was performed at ~70°C with DuPont Lightning Plus enhancing screens.

RNase H Digestions. RNase H (a gift of J. Hurwitz) treatment was for 15 min at 30°C in digestion buffer (10 mM Tris, pH 7.0/0.1 M NaCl/10 mM MgCl2/10 mM dithiothreitol/30% glycerol). Two units of enzyme was used for 0.2 μg of RNA. Reactions were stopped by the addition of NaDodSO4 to 1% and EDTA to 50 mM. Samples were boiled for 1 min, quenched in ice water, and loaded directly onto 15–30% sucrose gradients for size measurements.

Liquid Hybridization. Hybridizations were performed in 0.3 M Na acetate, pH 6.1/50% formamide at 43°C. The viral RNA was optimally stable under these conditions.

RESULTS

Isolation of 18S Sarcoma-Specific RNA. Fig. 1A shows that treatment of purified 30S MuSV RNA with RNase H in the absence of MuLV cDNA produced only a minor amount of degradation. No low molecular weight RNA was formed during this reaction, and the bulk of the MuSV RNA cosedimented with the 30S viral RNA marker run on the same gradient.

Fig. 1B compares the size of the 30S viral RNA at the end of the hybridization with MuLV cDNA, with and without addition of RNase H. In the profile of the control sample (without RNase H), the main peak cosedimented with 30S intact viral RNA. There was a broadening of the peak, but no shift in mean sedimentation value. The hybrid RNA treated with RNase H was extensively degraded. However, a large percentage of the RNA (~30%) sedimented as a peak into the gradient. In a separate experiment (Fig. 1C) an excess of unlabeled MuSV 124 cDNA was annealed to 32P-labeled MuSV RNA and digested with RNase H. All of the RNA radioactivity was found at the top of

![Figure 1](https://example.com/fig1.png)

**Fig. 1.** Isolation of the MuSV-specific RNA region by RNase H treatment. (A) 3P-Labeled purified 30S MuSV RNA (0.05 μg) was digested with RNase H for 15 min at 30°. 10 μg of unlabeled 4S yeast RNA was present as carrier. An identical amount of H-labeled 30S MuSV RNA was incubated for the same time in digestion buffer without any enzyme. At the end of the incubation NaDodSO4 to 1% and EDTA to 50 mM were added and the two samples were heated and quenched. The samples were then loaded on two 15–30% sucrose/NaDodSO4 gradients (10 mM Tris, pH 7.0/10 mM NaCl/1 M EDTA/0.2% NaDodSO4) and centrifuged for 6 hr at 40,000 rpm in the Spinco SW 41 rotor at 20°C. (B) Unlabeled MuLV cDNA (10 μg) was annealed to 0.2 μg of 32P-labeled purified 30S MuSV RNA for 2 hr. DNA concentration was 200 μg/ml. The hybridized sample was then treated with 2 units of RNase H for 15 min at 30°C in the presence of 10 μg of unlabeled 4S yeast RNA. An aliquot of the hybrid was kept undigested for size measurements. The two samples (digested and undigested) were heated, quenched, and run on a 15–30% sucrose gradient for 6 hr at 40,000 rpm in the Spinco SW 41 rotor at 20°C. O, 32P cpm, undigested; A, 32P cpm, RNase H-digested. (C) Unlabeled MuSV cDNA (10 μg) was annealed to 0.02 μg of 32P-labeled MuSV 30S RNA as above. After digestion with RNase H, the sample was heated, quenched, and run for 2 hr at 50,000 on a sucrose/NaDodSO4 gradient at 20°C in a Spinco SW 50.1 rotor. (D) The peak fractions (20–24) from the RNase H-digested sample fractionated on the first sucrose gradient (B) were pooled, extracted with phenol, and ethanol precipitated. The RNA was then loaded on a 15–30% sucrose/NaDodSO4 gradient and sedimented for 3 hr at 50,000 rpm in the Spinco SW 50.1 rotor at 20°C. 4S, 18S, and 28S ribosomal RNA markers were run on a parallel gradient (not shown).

cDNA Synthesis. cDNA was synthesized by using Triton X-100-disrupted virions as described (2).

![Graph](https://example.com/graph.png)
the gradient, indicating that complete degradation of the homologous cDNA-RNA hybrids had occurred.

The fractions containing the undigested RNA (Fig. 1B) were pooled, extracted with phenol to remove residual nucleases, ethanol precipitated, and rerun in a second sucrose gradient (Fig. 1D). The main peak showed an approximate sedimentation value of 18 S with a small shoulder of lower molecular weight material. These results indicated that the MuSV-specific region can be isolated as a fully contiguous RNA stretch of about 1800 nucleotides.

To test the purity and sequence specificity of the isolated 18S MuSV-specific RNA, hybridization experiments with purified MuSV-specific cDNA (2) and MuLV cDNA were performed. MuSV-specific cDNA and MuLV cDNA were prepared as described (2). An excess of both MuSV-specific DNA and MuLV cDNA was annealed to 32P-labeled 18S sarcoma-specific RNA. The purified sarcoma-specific 18S RNA completely hybridized with MuSV-specific cDNA but only 19% was homologous to MuLV sequences (Table 1). This 19% residual hybridization may represent either incomplete digestion by RNase H (unlikely) or contamination by degraded total MuSV RNA (most likely). The 18S RNA was poly(A)-free by oligo(dT)-cellulose chromatography (data not shown), a result in good agreement with its location on the MuSV genome (3, 4).

Thus, cross-hybridization of MuSV 30S RNA with excess MuLV cDNA followed by RNase H digestion allowed the isolation of an RNA molecule about 1800 nucleotides long and containing the sarcoma-specific sequence of MuSV 124. This RNA fraction was about 80% pure, as judged by hybridization to MuLV cDNA and presumably represented the majority of the genetic information by which MuSV differs from MuLV. In these experiments, as in previous ones (3, 4), shorter sarcoma-specific regions, located elsewhere on the RNA genome, would not have been detected and therefore cannot be completely excluded.

Physical Map of Moloney MSV 124. Different sets of results including hybridization to sarcoma-specific cDNA (4), heteroduplex mapping (3), and the RNase H digestion of cDNA-RNA hybrids support the proposition that the sarcoma-specific information maps at a unique site on the MuSV chromosome. It therefore should be possible to obtain a physical map of the sarcoma-specific region by restriction endonuclease digestion of MuSV double-stranded proviral DNA (16).

A cleavage map of MuSV 124 proviral DNA had been previously obtained by direct hybridization of labeled total MuSV 124 cDNA and RNA to agarose-gel-fractionated fragments (16). The levels of detection and reproducibility allowed by that method were clearly insufficient for further accurate mapping of the sarcoma-specific region. I therefore decided to obtain a new physical map by the agarose/nitrocellulose filter blotting technique of Southern (22). MuSV 124 proviral DNA was fractionated on 1.4% agarose gel as intact linear provirus and after cleavage with several enzymes and enzyme combinations. The gels were blotted onto nitrocellulose membrane strips and hybridized to 32P-labeled MuSV 30S RNA. The patterns obtained were similar to the ones originally reported (16) but some small important differences were revealed by the improved method. The intact linear proviral DNA had an apparent double-stranded molecular weight of 3.6-3.7 x 10^6 instead of 3.4 x 10^6 (Fig. 2). This molecular weight corresponds to a total length of about 5600 base pairs, in good agreement with the reported genome length of MuSV viral RNA (1-4). HincII produced three fragments of molecular weights about 1.9 x 10^6, 1.1 x 10^6, and 0.7 x 10^6. Bgl II produced two fragments (2.3 x 10^6 and 1.4 x 10^6). All other enzymes [HindIII and Hae II; Bam H1, Sma, and EcoRI (not shown)] and enzyme combination (Bgl II + HindIII, Bgl II + Hae II, Hae II + HindIII) gave fragments with molecular weights summing to 3.6-3.7 x 10^6. These results are summarized in Table 2.

Mapping the Sarcoma-Specific Sequences. Canaani et al. (16) reported that the linear proviral DNA was not circularly permuted with respect to the viral genomic RNA. Thus, by comparison of the data presented in Fig. 2 and Table 2 with RNA mapping data (3, 4), the sarcoma-specific sequence, an estimated 1500-1800 nucleotides long (see above and refs. 2-4), would be located on either the middle HincII fragment or the middle Bgl II + HindIII fragment.

To test these possibilities an experiment was designed based on the following rationale:
Table 2. Cleavage products of restriction endonucleases

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>No. of fragments</th>
<th>Fragment length, base pairs</th>
<th>Total M_r of fragments \times 10^{-6}</th>
</tr>
</thead>
<tbody>
<tr>
<td>EcoRI</td>
<td>1</td>
<td>5600</td>
<td>3.6-3.7</td>
</tr>
<tr>
<td>BamHI</td>
<td>1</td>
<td>5600</td>
<td>3.6-3.7</td>
</tr>
<tr>
<td>HincII</td>
<td>3</td>
<td>2900</td>
<td>3.7</td>
</tr>
<tr>
<td></td>
<td>1.9</td>
<td>1700</td>
<td>Middle</td>
</tr>
<tr>
<td></td>
<td>0.7</td>
<td>1050</td>
<td>3'</td>
</tr>
<tr>
<td>HindIII</td>
<td>2</td>
<td>4750</td>
<td>5'</td>
</tr>
<tr>
<td></td>
<td>3.1</td>
<td>750</td>
<td>3'</td>
</tr>
<tr>
<td>Bgl II</td>
<td>2</td>
<td>3550</td>
<td>5'</td>
</tr>
<tr>
<td></td>
<td>2.2</td>
<td>2250</td>
<td>3'</td>
</tr>
<tr>
<td>Hae II</td>
<td>2</td>
<td>3900</td>
<td>3'</td>
</tr>
<tr>
<td></td>
<td>2.6</td>
<td>1500</td>
<td>5'</td>
</tr>
<tr>
<td>HindIII + Bgl II</td>
<td>3</td>
<td>1200</td>
<td>Middle</td>
</tr>
<tr>
<td></td>
<td>2.2</td>
<td>750</td>
<td>3'</td>
</tr>
<tr>
<td>Hae II + HindIII</td>
<td>3</td>
<td>3100</td>
<td>Middle</td>
</tr>
<tr>
<td></td>
<td>2.0</td>
<td>1500</td>
<td>5'</td>
</tr>
<tr>
<td></td>
<td>0.5</td>
<td>750</td>
<td>3'</td>
</tr>
<tr>
<td>Hae II + Bgl II</td>
<td>3</td>
<td>2250</td>
<td>3'</td>
</tr>
<tr>
<td></td>
<td>1.5</td>
<td>1800</td>
<td>Middle</td>
</tr>
<tr>
<td></td>
<td>1.2</td>
<td>1500</td>
<td>5'</td>
</tr>
</tbody>
</table>

Data in this table are based on the experiment shown in Fig. 2.

* The reciprocal position of the double-stranded DNA fragments is based on the restriction patterns shown in Fig. 2. The 3' and 5' locations with respect to the viral RNA genome are based on the experiments of Canaani et al. (16).

† Several independent experiments have shown that these molecular weights are underestimated. The correct value for the 5' Hae II fragment is 1.1 × 10^6 or 1700 nucleotides.

‡ Molecular weights for these fragments, as calculated from the blot shown in Fig. 2, are lower than expected from the total map. The possibility that HindIII may cut more than once, yielding an undetected smaller fragment and therefore being responsible for the apparent discrepancy in molecular weights, was excluded by analyzing restriction patterns obtained by partial digestion with Bgl II and HindIII.

No “larger” middle fragments were observed under these conditions. Several independent experiments have conclusively shown that the correct molecular weight for the “middle” fragment is 1.0 × 10^6 or 1500 nucleotides.

Fig. 3 shows the results of such an approach using HincII and Bgl II + HindIII restriction fragments. In both experiments, annealing of the fragments to 32P-labeled MuSV RNA resulted in visualization of the whole restriction pattern. When a blot containing HincII-cleaved MuSV DNA was annealed to MuLV RNA (lane b) all three fragments were visualized to a large extent, indicating that all fragments contain common sequences. In the competition experiment (lane c), the major 5' fragment was almost completely erased by unlabeled MuLV RNA but the middle and 3' fragments were detected. This indicated that both these fragments contain MuSV-specific sequences. Fig. 3B shows the same kind of experiment performed by using blots of Bgl II + HindIII fragments. In this case, MuLV failed to visualize the middle fragment (lane b), which was in turn the only one visualized by the competition experiment (lane c). Thus, the majority of MuSV-specific sequences are located in the 1500-nucleotide-long middle fragment generated by Bgl II + HindIII. The faint band visualized by MuLV annealing in lane b, corresponding to the MuSV-specific fragment, has been consistently found in a series of similar experiments. It could be due either to a short common sequence at either end of the fragment or to a low level of cross-hybridization not detected by other methods (RNase H, heteroduplex analysis). Moreover, because short fragments (<1000 nucleotides) are blotted onto nitrocellulose with lower efficiency and are not visualized very well, the possibility that a small portion of the MuSV-specific sequences may be present on the 3' fragment cannot be excluded.

CONCLUSIONS

The physical map of the linear double-stranded proviral DNA of MuSV 124 is shown in Fig. 4. The various fragments obtained by restriction endonuclease digestion of MuSV 124 DNA are ordered according to the results shown in Fig. 2 and the mapping experiments reported earlier (16). The molecular weight of the intact linear MuSV 124 DNA provirus is about 3.7 × 10^6 or 5600 base pairs instead of 3.4 × 10^6 for 5000 base pairs as previously reported (16). This difference could be explained by the fact that the proviral DNA analyzed previously was partially single stranded. The proviral DNA preparation used in the experiments described in this report was shown to be almost completely double stranded. Molecular weight measurements by blotting of agarose gel-fractionated DNA were highly reproducible except when markers were run in the outside lanes. In these cases an obvious underestimation of the molecular weight of fragments in the central lanes occurred (see Table 2). The molecular weights of the fragments obtained with
hybridizations described earlier (4) in a heteroduplex nonhomology loop of 1540 bases (25) reported by Hu et al. (3) and here has been isolated as an 18S sarcoma-specific genomic RNA fragment. Furthermore, the sarcoma-specific region of the proviral DNA itself has been separated and mapped. This region appears to map at approximately the same location on both genomic DNA (4) and proviral DNA, implying that viral DNA made in vitro and viral RNA are colinear molecules. The finding of enzyme combinations dissecting specifically the MuSV-specific region from the rest of the genome should make possible DNA transfection experiments using subgenomic length fragments (25, 26) and should be helpful in establishing a transcriptional map of MuSV 124 in virus-transformed cells.

I thank Dr. E. E. Penhoet for the use of his laboratory space and facilities and Dr. E. Benz, Dr. F. Plata, and C. Falcì for critical reading of the manuscript and for their suggestions. This work was supported in part by Grant CA 20637 to E. E. Penhoet.