JOURNAL OF VIROLOGY, Dec. 1978, p. 810-818
0022-538X/78/0028-0810$02.00/0

Vol. 28, No. 3
Printed in U.S.A.

Analysis of Unintegrated Avian RNA Tumor Virus Double-Stranded DNA Intermediates

TERESA W. HSU,[1] JAN L. SABRAN,[2] GEORGE E. MARK,[1] RAMAREDDY V. GUNTAKA,[3] AND JOHN M. TAYLOR[1]*

The Institute for Cancer Research, Fox Chase Center, Philadelphia, Pennsylvania 19111[1]; Department of Microbiology, University of Pennsylvania, Philadelphia, Pennsylvania 19174[2]; and Department of Microbiology, College of Physicians and Surgeons, Columbia University, New York, New York 10032[3]

Received for publication 26 May 1978

Previous studies by Guntaka et al. have shown that the unintegrated DNA intermediates of avian RNA tumor virus replication can be readily isolated from cultures of the quail tumor line QT-6 at 1 day after infection. The intermediates include double-stranded linear and covalently closed circular DNA species. Using the analysis procedure of Southern together with previously obtained information regarding the sites of action of certain restriction endonucleases on avian sarcoma virus DNA, we have further characterized the viral DNA intermediates. Evidence is presented that, relative to the RNA genome, most of the linear species possess a direct terminal sequence redundancy equivalent to $0.5 \times 10^6 \pm 0.3 \times 10^6$ daltons of double-stranded DNA. Some of the circular forms also possess a sequence redundancy of $0.21 \times 10^6 \pm 0.03 \times 10^6$ daltons.

The intracellular replication of an RNA tumor virus genome involves transcription by the viral reverse transcriptase of the viral RNA into DNA, integration of such DNA into the host DNA, and the subsequent transcription, by the cellular DNA-dependent RNA polymerase II, of the integrated viral DNA to yield new viral RNA (as reviewed in reference 24). Guntaka et al. (6) have shown that unintegrated avian sarcoma virus (ASV) DNA intermediates can be readily isolated from monolayer cultures of the quail tumor line QT-6 (12) at 1 day after infection. They have isolated double-stranded DNA intermediates of two conformations: linear and covalently closed circular. A third conformation, open-circular, arises from nicking of the covalently closed form. Shank and Varmus (16) have presented evidence that some of the linear species detected in the cytoplasm are precursors to the circular species which are detected only in the nucleus. Both the linear and circular species are capable of transfecting avian cells (6), and Guntaka et al. (5) found that addition of ethidium bromide to infected avian cells could inhibit both the formation of circular intermediates and the appearance of new virus. Such findings allow but nevertheless do not establish that the immediate precursor to integrated viral DNA is the unintegrated circular species.

This manuscript reports a characterization of those unintegrated double-stranded ASV DNA intermediates detected in QT-6 cells at 1 day after infection. The novel feature of this study relative to the initial study of Guntaka et al. (6) is the use of the analysis procedure of Southern (17) as applied to the agarose gel electrophoresis of in vivo viral DNA intermediates before and after digestion with restriction enzymes. The results are compared and contrasted with information we have recently obtained on the sites of action of certain restriction enzymes on viral DNA synthesized in vitro, that is, synthesized outside the cell in the endogenous reaction of detergent-disrupted virions (19). Relevant information concerning in vitro DNA is summarized in Fig. 1.

MATERIALS AND METHODS

Isolation of unintegrated viral DNA intermediates. QT-6 cells were derived from a fibrosarcoma induced in Japanese quail with methylcholanthrene by Moscovici et al. (12). These cells were seeded at a density of 5×10^6 cells per 10-cm-diameter petri dish and infected after 2 days with either the PrC or the B77 strain of ASV at a multiplicity of infection of about 1 focus-forming unit per cell. At 1 day after infection, cells were harvested and fractionated into nuclei and cytoplasm as previously described (6).

Covalently closed circular double-stranded viral DNA was isolated from the nuclei as previously described (6) or by an improved method that involves chromatin removal (Guntaka, Anal. Biochem., in press). Additional purification was achieved by equilibrium centrifugation in cesium chloride gradients containing propidium diiodide and by the use of a benzoylated-naphthoylated DEAE-cellulose column (6).

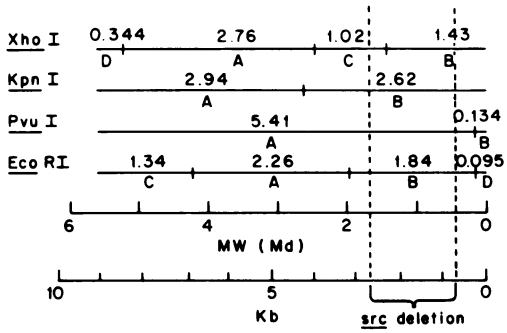

FIG. 1. *Ordering of restriction fragments obtained with ^{32}P-labeled double-stranded ASV DNA synthesized in vitro. This information is extracted from Taylor et al. (19). The right side represents the 5' terminus of that DNA transcript complementary to the viral RNA. The location of the src deletion in the td genome is based on heteroduplex studies of Lai et al. (9). The numbers indicate the mass (in Md) of the fragments.*

Linear double-stranded viral DNA was obtained as part of the total cytoplasmic DNA and was used without further purification. This cytoplasmic DNA contains a small amount of circular viral DNA intermediate relative to linear DNA.

Gel electrophoresis. The horizontal slab gel apparatus of McDonell et al. (10) was obtained from the Aquebogue Machine and Repair Shop, Long Island, N. Y. Samples of not more than 40 μl were applied to horizontal gels (13 by 13 by 0.6 cm) of 0.7% agarose (SeaKem), and electrophoresis was carried out as described in the figure legends. Where appropriate, electroelution of DNA from gels was carried out as previously described (19), by a method which was basically that of McDonell et al. (10). For molecular weight markers, *Hin*dIII fragments of ^{32}P-labeled phage lambda DNA were prepared as previously described (19).

Restriction endonuclease digestion. Enzymes were obtained commercially; *Pvu* I, *Xho* I and *Kpn* I were obtained from New England Biolabs, and *Eco*RI was obtained from Miles Laboratories, Inc. Digestions were carried out for 1 h at 37°C in not more than 40 μl with 4 U of enzyme and not more than 2 μg of DNA. For *Xho* I and *Kpn* I, the digestion buffer contained 0.006 M Tris-hydrochloride (pH 7.4), 0.05 M NaCl, 0.006 M $MgCl_2$, and 0.006 M mercaptoethanol. For *Pvu* I the NaCl was increased to 0.15 M and for *Eco*RI the Tris-hydrochloride was increased to 0.05 M, as suggested by Polisky et al. (13). Digestions were stopped by the addition of 0.1% sodium dodecyl sarcosinate, 7.5% sucrose, and 0.01% bromophenol blue.

Detection of viral sequences. After electrophoresis of samples into agarose gels, the nucleic acids were transferred to sheets of cellulose nitrate filter paper (B6; Schleicher & Schuell Co.) by the method of Southern (17). Filter papers were dried at room temperature, baked for 2 h at 80°C in a vacuum oven, and then pretreated at 68°C for at least 1 h in a solution containing 0.9 M NaCl, 0.09 M sodium citrate, 0.02% polyvinylpyrrolidone, 0.02% Ficoll, 0.02% bovine serum albumin, 1% glycine, and 500 μg of calf thymus DNA per ml (2).

A ^{32}P-labeled DNA probe complementary to the viral RNA was synthesized in the following reaction mixture: 0.1 M Tris-hydrochloride (pH 8.1), 0.006 M $MgCl_2$, 0.1 mM dATP, 0.1 mM dGTP, 0.1 mM TTP, 5 μM α-^{32}P-labeled dCTP (350 Ci/mmol; New England Nuclear Corp.), 1% mercaptoethanol, 500 μg of calf thymus DNA primers per ml (21), 5 μg of 70S viral RNA per ml, 33 μg of actinomycin D (Calbiochem) per ml and 100 U of purified reverse transcriptase per ml. After 1 h at 37°C, synthesis was stopped with 0.1 M EDTA. The sample was subjected to ethanol precipitation, alkali treatment (0.6 N NaOH for 1 h at 37°C), neutralization with Tris-hydrochloride, and further ethanol precipitation. In this way a reaction of 0.1 ml yielded about 25 × 10^6 cpm of probe at about 500 cpm/pg.

In some cases a probe complementary to the 3'-terminus of the viral RNA was made by a slightly different protocol. Viral 70S RNA was heat denatured and passed twice through a column of oligodeoxythymidylic acid-cellulose (11). The polyadenylic acid [poly(A)]-containing RNA thus isolated was used as template in the presence of 1 μg of oligo$(dT)_{10}$ (Miles Laboratories) per ml as primer rather than the calf thymus DNA primers. For hybridization of the labeled probe, the cellulose nitrate filter pretreated as described above was sealed in a plastic bag containing 18 μl of the following solution per cm^2: 0.9 M NaCl, 0.09 M sodium citrate, 0.02% polyvinylpyrrolidone, 0.02% Ficoll, 0.02% bovine serum albumin, 0.5% sodium dodecyl sulfate, 500 μg of calf thymus DNA per ml and 5 ng of ^{32}P-labeled probe per ml. After annealing for about 20 h at 68°C, the filter was washed three times for 20 min at 68°C in a plastic bag containing 0.3 M NaCl–0.03 M sodium citrate–0.5% sodium dodecyl sulfate. A final wash, with rocking, was carried out at room temperature in the same buffer, except that sodium dodecyl sulfate was omitted. After drying at room temperature and baking at 80°C in a vacuum oven, the filter was subjected to autoradiography at −70°C with Kodak RP-5 film and a Dupont Lightning Plus intensifying screen. The films obtained were scanned for optical density at 620 nm with a densitometer (E-C Apparatus Corp.).

Exonuclease digestion. Samples of gel-purified in vivo linears were digested with 2 μg of purified 3'-exonuclease per ml for 20 min at 37°C in 20 μl of a solution containing 0.05 M NaCl, 0.05 M Tris-hydrochloride (pH 7.4), 0.01 M $MgCl_2$, 0.1 M mercaptoethanol, and 500 μg of calf thymus DNA per ml. Sodium dodecyl sulfate was added to a final concentration of 0.5% to inactivate the enzyme. To allow circularization, a subsequent annealing at 68°C for 10 min was carried out after the addition of NaCl to a final concentration of 0.3 M. The sample was then collected by ethanol precipitation and resuspended in 20 μl of electrophoresis buffer containing sodium dodecyl sarcosinate, sucrose, and bromophenol blue as described above. Other exonuclease studies were carried out with 300 U of the 5'-exonuclease of phage lambda per ml for 10 min at 16°C in 20 μl of a solution containing 0.07 M glycine-KOH (pH 9.4) and 0.003 M $MgCl_2$.

RESULTS

The following experiment was undertaken to characterize the electrophoretic mobility of those in vivo linear DNA intermediates present in the cytoplasm of QT-6 cells at 1 day after infection by PrC strain ASV (6). This viral DNA, grossly contaminated with cellular DNA, was subjected to electrophoresis on a slab gel of 0.7% agarose. The viral DNA in the gel was assayed by the procedure of Southern (17) in which the DNA is denatured in the gel, transferred to a cellulose nitrate filter, and then detected by its ability to hybridize to a labeled DNA probe complementary to the viral genome. In this and subsequent experiments, the filter after hybridization was subjected to autoradiography and subsequent optical density scanning of the exposed X-ray film (Fig. 2). Molecular weight values were deduced relative to *Hin*dIII fragments of ^{32}P-labeled phage lambda DNA (19). Two flanking molecular weight markers were used in each electrophoresis. However, in Fig. 2 and subsequent figures we show only one marker. Also, we have reassorted the data for logic of presentation and in some cases used a composite of data from separate experiments. The deduced mass values (in megadaltons [Md]; 1 Md = 10^6 daltons) indicated on the panels are estimated to have an absolute error of not more than 5%. The in vivo linear viral DNA (Fig. 2a) migrates as two major bands, with modal equivalent masses of 6.36 and 5.10 Md. As discussed later, there is implicit in these estimations the assumption that the linear DNA is completely double stranded. From the earlier work of others, the smaller species is recognizable as the DNA derived from transformation-defective (*td*) genomes that possess a deletion of about 1.2 Md in the *src* region of the nondefective (*nd*) genomes (9). The band of *nd* linears centered at 6.36 Md is larger than expected in that a uniform double-stranded DNA transcript of the viral RNA, excluding the 3'-terminal poly(A), would have a mass of 5.8 to 6.0 Md (6, 9). We will refer to such a hypothetical transcript as a unit-length molecule, and its mass is subsequently identified as 5.85 Md. Relative to this, the linear DNA is 0.5 ± 0.3 Md larger.

We have attempted to compare and contrast the sites of action of certain restriction endonucleases on the in vivo linear DNA to the results of an earlier study in which we determined the sites of action of those enzymes on ASV DNA synthesized in vitro (19). Some of those data for an *nd* genome are summarized in Fig. 1. The site of the *src* deletion that generated *td* genomes is indicated. Note the convention that the right side of the map corresponds to the origin, or 5'-terminus, of that DNA transcript initiated on

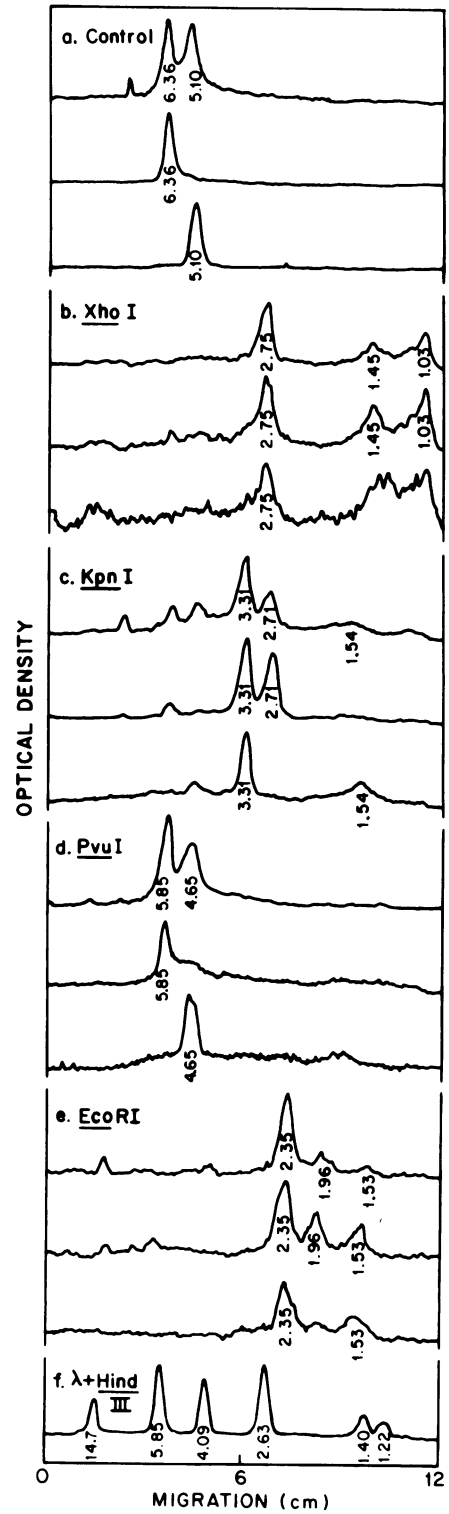

FIG. 2. *Agarose gel electrophoresis of unintegrated linear viral DNA intermediates before and after digestion with restriction enzymes. Samples*

the viral RNA and, therefore, complementary to the viral RNA. As explained below, our studies indicate that for both in vivo and in vitro DNA the right side of the genome is the same. However, at the left side the in vivo DNA has additional sequences.

The two main size classes of in vivo unintegrated linear DNA, *nd* and *td*, were purified by agarose gel electrophoresis and subsequent electroelution (Fig. 2a). When the *nd* in vivo linears are digested with *Xho* I before electrophoresis, we detect the three major fragments having masses of 2.75, 1.45, and 1.03 Md (Fig. 2b), comparable in size to fragments A, B, and C previously mapped at the right side of the in vitro DNA (Fig. 1). The combined mass of these is 5.23 ± 0.26 Md. If, as indicated earlier, the intact linear *nd* DNA is 6.36 ± 0.32 Md, then we would expect a fragment at the left side of 1.13 ± 0.59 Md. This fragment could be migrating in the vicinity of the observed 1.03-Md fragment.

Digestion of total in vivo linears with *Kpn* I yields three main fragments: 3.31, 2.71, and 1.54 Md (Fig. 2c). (Three bands of masses higher than 3.31 Md are also seen. The largest is a quail band containing virus-related sequences. The other two represent the opened form of *nd* and *td* circles, because the cytoplasmic fraction used as the source of linear DNA intermediates contains a small amount of circular DNA intermediates. However, there could be an additional component of incompletely digested linear DNA.) The 2.71-Md fragment is identified as fragment B on the right side of in vitro DNA (Fig. 1). The 1.53-Md fragment present in the *td* genome presumably arises from the *src* deletion in fragment B. The fragment having a mass of 3.31 ± 0.16 Md is considered to be significantly larger than fragment A which is located at the left side of in vitro DNA and has a mass of 2.94 ± 0.15 Md.

The digestion of in vivo linears with *Pvu* I (Fig. 2d) provides better evidence that additional sequences are present at the left side of the map and at the same time provides a clue as to the nature of these additional sequences. On in vitro DNA, *Pvu* I makes one cut at a site only 0.134 Md from the right side (Fig. 1). However, *Pvu* I digestion of in vivo DNA reduces the 6.36-Md band by 0.5 ± 0.3 Md to a fragment of 5.85 Md. Thus, *Pvu* I has more than one site of action on in vivo DNA. Our interpretation is that the new *Pvu* I site is on the left and is actually a second copy of that on the right. In other words, the unit-length molecule defined earlier is the 5.85-Md fragment released with *Pvu* I. The additional sequences of 0.5 ± 0.3 Md are located on the left end of the in vivo linear and represent a direct terminal repeat of the right side sequences. An analogous interpretation is given to the *Pvu* I digestion of in vivo *td* linears.

The results of the digestion of in vivo linears with *Eco*RI are also consistent with the above interpretation. The *nd* linear DNA yields three main fragments: 2.35, 1.96, and 1.53 Md (Fig. 2e). The two larger fragments correspond in size to the fragments A and B located toward the right side of the in vitro DNA (Fig. 1). The third fragment, 1.53 Md, is somewhat larger (0.19 ± 0.14 Md) than the fragment C located at the left side of in vitro DNA. However, in vitro DNA is cut very close to the right side by *Eco*RI. Thus, if the extra sequences at the left side of in vivo DNA are as proposed above, a direct redundancy of left-side sequences, then the *Eco*RI site at the left would be repeated and in vivo DNA would contain, as observed, a band somewhat larger than that observed in vitro. More importantly, the total mass of the three main bands, 5.84 ± 0.19 Md, should be, and is, the same size as the unit-length molecule of 5.85 ± 0.19 Md. Incidentally, in the *Eco*RI digest of in vivo *td* linears, the mass of fragment B is reduced to about 0.68 Md as a consequence of the *src* deletion (not included in Fig. 2e, but see Fig. 3f).

The data described above indicate that the average in vivo linear *nd* DNA migrates as if it were 0.5 ± 0.3 Md larger than the hypothetical 5.85-Md unit-length species. Restriction enzyme studies have been interpreted as evidence that the extra sequences are located on the left side and represent a direct terminal repeat of right-side sequences. Two additional independent lines of evidence have been obtained which are consistent with this interpretation.

If the in vivo linear double-stranded DNA species contain a direct terminal repeat, then pretreatment with either the 3'-exonuclease of *Escherichia coli* or the 5'-exonuclease of phage lambda should expose cohesive ends on the DNA and allow circularization. For such an experiment the in vivo *nd* linears were purified by agarose gel electrophoresis and subsequent electroelution. Relative to the *Hin*dIII fragments of phage lambda (Fig. 4g), the purified linear DNA migrates as a band with a modal mass equivalent to about 6.36 Md (Fig. 4c). Annealing of this DNA has no major effect on migration (Fig. 4d),

were subjected to electrophoresis for 20 h at 0.8 V/cm on a gel of 0.7% agarose, after which viral sequences were detected by the method of Southern (17). (a) The upper lane is of the total linears; the second and third lanes represent purified nd and td linear species, respectively; (b) as in (a), digested with Xho I; (c) as in (a), digested with Kpn I; (d) as in (a), digested with Pvu I; (e) as in (a), digested with EcoRI; (f) molecular weight markers provided by HindIII fragments of ^{32}P-labeled phage lambda DNA.

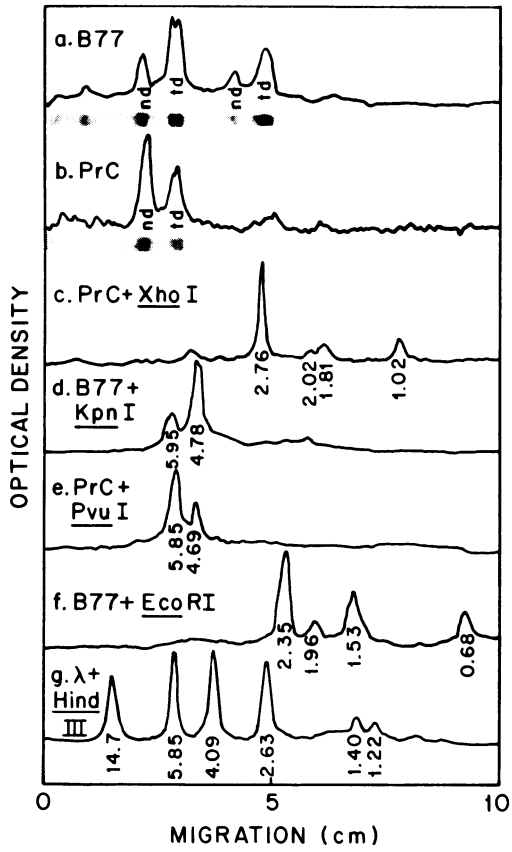

FIG. 3. *Agarose gel electrophoresis of unintegrated circular viral DNA intermediates before and after digestion with restriction enzymes. Samples were subjected to electrophoresis for 16 h at 0.8 V/cm on a gel of 0.7% agarose after which viral sequences were detected by the method of Southern (17). (a) Circles of B77 strain ASV; (b) circles of PrC strain ASV; (c) as in (b), digested with Xho I; (d) as in (a), digested with Kpn I; (e) as in (b), digested with Pvu I; (f) as in (a), digested with EcoRI; (g) molecular weight provided by HindIII digest of ^{32}P-labeled phage lambda DNA.*

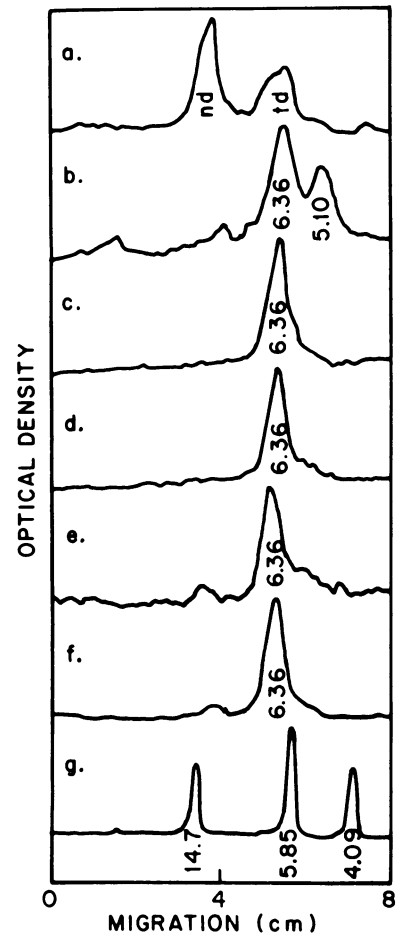

FIG. 4. *Conversion of linear unintegrated viral DNA to a circular form after digestion with exonuclease III. Samples were subjected to electrophoresis for 15 h at 1.6 V/cm on a gel of 0.7% agarose, after which viral sequences were detected by the method of Southern (17). (a) Total circular viral DNA intermediates; (b) total viral linear DNA intermediates; (c) linear viral intermediates of modal mass equivalent to about 6.36 Md, purified by electrophoresis and subsequent electroelution; (d) as in (c), but subjected to annealing; (e) as in (c), but subjected to digestion with 3′-exonuclease and then subjected to annealing; (f) as in (e), but subjected to digestion with 5′-exonuclease; (g) molecular weight markers provided by HindIII fragments of ^{32}P-labeled phage lambda DNA.*

but pretreatment with 3′-exonuclease followed by annealing (Fig. 4e) allows conversion of approximately 14% of the linear DNA to a form that migrates as expected for a noncovalently closed circular form (Fig. 4a). Increasing the annealing time by eightfold did not increase the extent of conversion of linears to a circular form (data not shown). Exonuclease treatment did not significantly alter the migration of the linear DNA and in the absence of subsequent annealing did not produce circles (data not shown). Similar results have been obtained with linear DNA pretreated with 5′-exonuclease and then annealed (Fig. 4f). In summary, our results indicate that at least some, but not necessarily all, of the additional sequences on the linear DNA represent a direct terminal repeat.

A third line of evidence for a direct terminal repeat on the in vivo linear DNA has been obtained by the use of a specific radioactive probe in the Southern procedure. Samples of in vivo linears were digested with either *Eco*RI or

Xho I, subjected to agarose gel electrophoresis, and subsequently transferred by the procedure of Southern (17) to a cellulose nitrate filter. The filter was hybridized by using a labeled probe complementary to the 3'-terminus of the viral RNA. In the absence of a terminal redundancy on the linear DNA, it would be expected that the 3'-terminal probe would anneal only to those fragments from the right side of the map. As summarized in Table 1, the 3'-terminal probe anneals as expected primarily to the right side *Eco*RI fragment B. After this there is relatively more labeling to the left-side fragment C than to the middle fragment A. This result is consistent with the earlier interpretation of a direct terminal repeat. When the 3'-terminal probe is applied to the *Xho* I digest, labeling again occurs primarily to the right-side fragment B. Relative to this there is less labeling in the central fragments A and C. In the light of the labeling by the 3' probe of *Eco*RI fragments derived from both ends, the redundancy would have to be to the left side of *Xho* I fragment A (Fig. 1).

The results presented above have considered the in vivo linear DNA species and the evidence for a direct terminal repeat of an average of 0.5 ± 0.3 Md. An electrophoretic analysis has also been made of in vivo circular DNA species obtained from the nucleus of QT-6 cells infected with either the B77 or the PrC strain of ASV (Fig. 3a and b, respectively). The purification procedure, as previously described by Guntaka et al. (6), enriches for covalently closed circular forms. However, on storage, even at −70°C, these species readily acquire at least one single-stranded nick per circle because they are converted predominantly to an open-circular form. On the basis of the studies of Guntaka et al. (6) we identified the two slower migrating bands as open *nd* and *td* circles. The two faster bands are covalently closed *nd* and *td* circles. There are more *td* circles relative to *nd* circles for B77-infected cells than for PrC-infected cells. Upon closer examination, each of these four circular species has its own particular heterogeneity. High-resolution electrophoresis resolves each as two closely migrating species. That is, for both open and covalently closed circles and both for *nd* and *td* genomes, there is a size heterogeneity within which two distinct size classes are more frequent.

TABLE 1. *Relative labeling of restriction fragments of linear viral DNA intermediates*[a]

Restriction enzyme used to derive fragments	Fragment nomenclature	Fragment mass (Md)	Relative labeling with 3' probe	Relative molar labeling with 3' probe
*Eco*RI	B	1.96	1.00	1.00
	A	2.35	0.13	0.11
	C	1.53	0.23	0.29
Xho I	B	1.45	1.00	1.00
	C	1.03	0.07	0.10
	A	2.75	0.17	0.09

[a] Total linear *nd* DNA intermediates were digested with either *Eco*RI or *Xho* I, subjected to electrophoresis on a gel of 1% agarose, and transferred to a cellulose nitrate filter by the method of Southern (17). The filter was annealed with a ^{32}P-labeled DNA probe complementary to the 3' terminus, excluding the poly(A), of the RNA genome. After autoradiography, relative labeling of particular restriction fragments was quantitated with a transmission densitometer. The nomenclature, masses, and relative order of restriction fragments are as summarized in Fig. 1.

To characterize the size heterogeneity of the in vivo circular DNA intermediates, we used a preliminary restriction enzyme digestion. For example, digestion with *Xho* I, as shown in Fig. 3c, yields the 2.76- and 1.02-Md fragments present for both in vivo linear (Fig. 2b) and in vitro linear DNA (Fig. 1). After circularization of a linear molecule, digestion with *Xho* I should produce a junction fragment not present in the linear molecule. Actually, two new main fragments are produced at 2.02 and 1.81 Md. Our interpretation is that these junction fragments are equivalent to the joining of the 1.43-Md fragment B at the right side of the linear DNA and two fragments of different sizes at the left side. The difference in mass between the two new junction fragments is 0.21 ± 0.03 Md and probably represents the size difference between the two predominant forms of *nd* circles. There is a problem in Fig. 4 in that the two expected junction fragments from the *td* circle are not detected. If the *src* deletion takes out the *Xho* I site between fragments B (1.43 Md) and C (1.02 Md) as indicated in Fig. 1, then along the lines of our interpretation the 1.2-Md *src* deletion should replace the 2.02- and 1.81-Md junction fragments by 2.02 − 1.2 + 1.02 = 1.84 Md and 1.81 − 1.2 + 1.02 = 1.63 Md, respectively. Such fragments are not seen in Fig. 3, but in other experiments with PrC circles we have seen what may be these fragments but at a lower intensity relative to the 2.02- and 1.81-Md junction fragments derived from the *nd* circles. The lower intensity is reasonable because our preparations of PrC circles have less *td* relative to *nd*. However, B77 circles have more *nd* than *td*, and yet we do not see the expected 1.84- and 1.63-Md fragments. We propose that in B77 circles there can be heterogeneity in the size and/or location of the *src* deletion. In some cases this deletion does not seem to include the *Xho* I site between

fragments B and C (Fig. 1) because we have observed fragments at about 0.8 and 0.6 Md of the size expected for a 1.2-Md deletion in the 2.02- and 1.81-Md junction bands of *nd* circles.

Digestion of *nd* circles with *Eco*RI (Fig. 3) yields the same three fragments, A, B and C, as detected after digestion of in vivo linears (Fig. 2e). This means that the extra sequences on the larger *nd* circle with respect to the smaller *nd* circle are located outside these three fragments. In this case expected new junction fragments would be too small to be detected by our Southern analysis. Digestion of circles with *Kpn* I (Fig. 3d), which makes only one cut on the in vivo linears, opens the circles and allows an estimate of the molecular weight relative to the *Hin*dIII fragments of phage lambda (Fig. 3g). The two predominant species of opened *nd* circles almost comigrate at about 5.95 Md, and the two species of opened *td* circle almost comigrate at 4.78 Md. More convincing data have been obtained than are represented in Fig. 4. Digestion of circles with *Pvu* I (Fig. 3e) yields an answer different than that obtained with *Kpn* I. Now one, rather than two, species seems to be released from both the *nd* and *td* intact circles. Our interpretation is that the circle that is 0.21 Md larger has a second copy of the *Pvu* I site present at the right side of the linears. Any circle whose structure is equivalent to joining the ends of linears with even variable amounts of direct terminal repeats of at least 0.134 Md from the right side (Fig. 5) will, after digestion with *Pvu* I, yield unit-length molecules. If a repeat is present but less than 0.134 Md, then the circles will be cut only once, and circles of as much as 0.134 Md larger than unit length will be obtained. Our data are not sufficient to test the acceptability of this latter prediction.

DISCUSSION

Evidence has been presented above that the unintegrated linear viral DNA intermediates synthesized in vivo, that is, in infected QT-6 cells, are 0.5 ± 0.3 Md larger than expected for a hypothetical unit-length double-stranded transcript of the viral RNA, exluding poly(A). Relative to the map of restriction enzyme sites on viral DNA synthesized in vitro (Fig. 1), the additional sequences on the in vivo DNA seem to be located at the left end of the map and represent a direct terminal redundancy of sequences located at the right side. This interpretation is summarized in Fig. 5. Assuming this interpretation, the unit-length *nd* DNA has been identified as having a mass of 5.85 Md, because this is the fragment released from in vivo linears with the enzyme *Pvu* I. This enzyme makes only one cut at the right side of in vitro DNA, but it

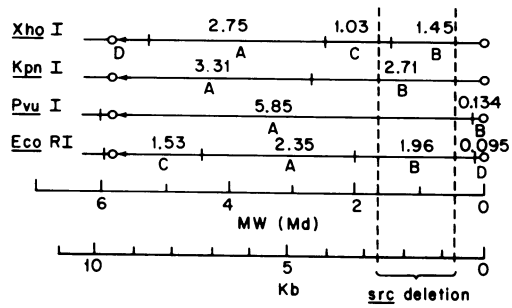

FIG. 5. *Ordering of restriction fragments obtained with in vivo unintegrated linear DNA intermediates. The numbers indicate the mass (in Md) of the fragments as deduced in this manuscript. The molecular weights of the smallest Pvu I and EcoRI fragments are as deduced from previous in vitro studies (Fig. 1 and reference 19). The right circle corresponds to the origin of transcription of viral RNA into complementary DNA. Moving toward the left, you reach the arrowhead, which is interpreted as a hypothetical unit-length transcript of 5.85 Md. The in vivo linear DNA has a right side identical to that obtained for in vitro DNA, but its left side is 0.5 ± 0.3 Md of double-stranded DNA larger than the unit-length molecule. This extra DNA is interpreted as a redundancy of sequences present at the right side of unit-length species. The Pvu I and EcoRI sites at the left are thus repeats of those on the right.*

makes two cuts on the redundant in vivo DNA. In our previous study potentially full-size in vitro DNA had an equivalent mass of 5.6 Md (19). This is shorter than the unit-length DNA at the left side, presumably due to premature termination of complementary DNA synthesis and/or inadequate copying of the 3′ terminus of that complementary DNA into double-stranded DNA.

The redundant sequences present at the left end of the in vivo linear DNA increase the apparent molecular weight with respect to that of unit-length DNA intermediates. In our analysis we have implied but not shown that the redundant sequences are exclusively double stranded. Analysis of linears digested with single-strand-specific nuclease S1 has been unsuccessful because of multiple single-stranded sites distributed along the linears (data not shown). If some of the redundant sequences are single stranded, we would expect that those single-stranded regions would have to be of the DNA strand complementary to the viral genome.

Canaani et al. (3) have reported studies on the sites of action of certain restriction enzymes on the linear unintegrated intermediates of murine sarcoma virus. They found that a probe to the 3′ terminus of the viral RNA was able to anneal to restriction fragments from not just one but from both ends of linear DNA. They did not

establish which sequences were repeated. We have presented here a similar result for ASV (Table 1) together with two other lines of evidence for a terminal redundancy on the DNA. Our ability to circularize the linears after nuclease digestion (Fig. 4) is evidence that the repeat is direct rather than inverted. Our evidence as to the amount of sequences repeated on the linears relative to the hypothetical unit-length genome is primarily the digestion of *Pvu* I. Our previous experience with in vitro DNA was that *Pvu* I made one cut, at 0.134 Md from the right side (Fig. 1) (19). However, on the in vivo linear DNA *Pvu* I makes two cuts, and our interpretation of the extra cut is that the repeat on the linear DNA relative to the unit-length genome is at least 0.134 Md. In the absence of enzyme digestion, the in vivo linears migrate with an average migration equivalent to double-stranded DNA that is 0.5 ± 0.3 Md larger than the unit-length genome, but, as discussed in the previous paragraph, this analysis implies that the redundancy at the left side of the linear is strictly double stranded.

It is relevant to note that Rothenberg et al. (14) have recently studied an in vitro single-stranded DNA transcript of murine leukemia virus. They characterized by electron microscopy those heteroduplexes that could be made between this transcript of 8.8 kilobases and a particular subgenomic viral RNA detected in infected cells. An interpretation not considered by them but possibly more consistent with their data and certainly in agreement with the data presented in this manuscript is that their 8.8-kilobase transcript contained a direct terminal repeat of, on average, 0.6 kilobase (equivalent to 0.37 Md of double-stranded DNA).

How the direct terminal sequence redundancy on the linears is acquired can only be speculated upon at this time. One hypothesis based upon in vitro studies of reverse transcription (18) is as follows. DNA synthesis is initiated on the 3′ terminus of the $tRNA^{trp}$ primer bound near the 5′ terminus of the viral RNA (20). DNA is elongated by transcription of 101 nucleotides after the 5′ terminus of the viral 35S RNA (7). After digestion by the RNase H activity of the reverse transcriptase of that RNA already transcribed (4, 22), there can be a base pairing interaction between at least 15 nucleotides at the 3′ terminus of the DNA and a complementary sequence near the poly(A) at the 3′ terminus of the viral 35S RNA (15). This jump to the 3′-terminal region allows further transcription of the viral RNA. However, if the jump occurs to the 3′-terminal region of that 35S RNA on which DNA synthesis was initiated, then DNA synthesis will not be able to proceed beyond the primer binding site to generate the observed direct terminal repeat because the early RNase H action has digested away the relevant template. To explain the direct terminal repeat on the majority of the unintegrated linear DNA, we propose that the initial jump is almost exclusively from the 5′ terminus of one subunit to the 3′ terminus of another subunit. This could be a practical reason why the RNA tumor viral genomes exist as exactly two hydrogen-bonded 35S subunits (1, 8). After the jump to the second subunit, transcription can proceed to the 5′ terminus of that subunit and carry out by the same mechanism a second jump, this time to the 3′ terminus of the first 35S RNA. Elongation after this jump would explain the observed terminal redundancy. Haseltine et al. (7) have proposed a very credible model of how the two 35S RNAs and two associated tRNA primers are held together in a 70S complex. A consequence of their structure is that the initial transcription of DNA beginning from one primer to the 5′ terminus of the subunit to which that primer is bound could facilitate the displacement of the other primer from the subunit to which it is bound. This displacement and the utilization of just one primer are essential features of the model proposed here. After we had considered this model, we became aware of studies of Varmus et al. (23) who find in infected cells a unique species of plus-strand DNA, approximately 250 to 300 nucleotides long, that contains sequences specific for both the 3′ and 5′ termini of the viral RNA. Rather than the viral RNA, this unique DNA fragment is a more likely candidate as template for the redundant sequences present at the left side of in vivo DNA intermediates.

We have also presented evidence in this manuscript that the circular DNA intermediates present in infected cells are heterogeneous in size. There are two predominant species of circle, one being about 0.21 ± 0.03 Md larger than the other. The enzyme *Pvu* I converts the larger of the *nd* circles to a linear of 5.85 Md, which we interpret as the unit-length genome. This circle has two copies of the *Pvu* I site, whereas we would expect a unit genome to have only one copy (19). The smaller *nd* circle has a mobility comparable to that of a unit genome. It does not have redundancy of the *Pvu* I site and may in fact have little, if any, redundancy. Unpublished studies by other laboratories also indicate circular intermediates of two predominant size classes: ASV as studied by P. Shank, H. Varmus, and M. Bishop; mouse mammary tumor virus as studied by P. Shank, J. Cohen, H. Varmus, K. Yamamoto and G. Ringold; and murine leukemia virus as studied by R. Weinberg. The data presented here suggest that at least the larger

ASV circle contains 0.21 ± 0.03 Md of redundant sequences. At this stage we do not know whether the larger circle is of the same size or smaller than the linear intermediate. However, under our conditions of electrophoresis, the linear DNA does migrate more slowly.

As mentioned above, there is as yet no firm evidence as to whether circular species are immediate precursors to integrated viral DNA. Nevertheless, this manuscript has presented evidence that both linear and some circular viral DNA species have a direct sequence redundancy. And, in a subsequent manuscript (J. Sabran, T. Hsu, C. Yeater, A. Kaji, W. Mason, and J. Taylor, J. Virol., in press) evidence is presented that even after integration of the ASV DNA direct terminal repeated sequences are maintained on the viral DNA.

ACKNOWLEDGMENTS

The technical assistance of Amy Weiner in the preparation of viral DNA intermediates is gratefully acknowledged. Purified reverse transcriptase from avian myeloblastosis virus was kindly provided by the Office of Program Resources and Logistics, Viral Cancer Program, Bethesda, Md. Purified 3'-exonuclease of *E. coli* was a gift of Larry Loeb, and purified 5'-exonuclease of phage lambda-infected *E. coli* was a gift of Charles Hill.

J.M.T. is supported by grant VC-155B from the American Cancer Society, by Public Health Service institutional grants CA-06927 and RR-05539 from the National Cancer Institute and the Division of Research Resources, National Institues of Health, respectively, and by an appropriation from the Commonwealth of Pennsylvania. R.V.G. is supported by Public Health Service contract CP-71055 and grant CA-19152 from the National Cancer Institute.

LITERATURE CITED

1. **Bender, W., Y.-H Chien, S. Chattopadhyay, P. K. Vogt, M. B. Gardner, and N. Davidson.** 1978. High-molecular-weight RNAs of AKR, NZB, and wild mouse viruses and avian reticuloendotheliosis virus all have similar dimer structures. J. Virol. **25**:888-896.
2. **Botchan, M., W. Topp, and J. Sambrook.** 1976. The arrangement of simian virus 40 sequences in the DNA of transformed cells. Cell **9**:269-287.
3. **Canaani, E., P. Duesberg, and D. Dina.** 1977. Cleavage map of linear mouse sarcoma virus DNA. Proc. Natl. Acad. Sci. U. S. A. **74**:29-33.
4. **Collett, M. S., P. Dierks, J. T. Parsons, and A. J. Faras.** 1978. Specific removal of the 5'-terminus of the avian sarcoma virus genome by viral RNase H during reverse transcription *in vitro*. Nature (London) **272**:181-184.
5. **Guntaka, R. V., B. W. Mahy, J. M. Bishop, and H. E. Varmus.** 1975. Ethidium bromide inhibits the appearance of closed circular viral DNA and integration of virus-specific DNA in duck cells infected by avian sarcoma virus. Nature (London) **253**:507-511.
6. **Guntaka, R. V., O. C. Richards, P. R. Shank, H.-J. Kung, N. Davidson, E. Fritsch, J. M. Bishop, and H. E. Varmus.** 1976. Covalently closed circular DNA of avian sarcoma virus: purification from nuclei of infected quail tumor cells and measurement by electron microscopy and gel electrophoresis. J. Mol. Biol. **106**:337-357.
7. **Haseltine, W. A., A. M. Maxam, and W. Gilbert.** 1977. Rous sarcoma virus is terminally redundant: the 5'-sequence. Proc. Natl. Acad. Sci. U. S. A. **74**:989-993.
8. **Kung, H.-J., J. M. Bailey, N. Davidson, M. O. Nicholson, and R. M. McAllister.** 1975. Structure, subunit composition, and molecular weight of RD-114 RNA. J. Virol. **16**:397-411.
9. **Lai, M. M. C., S. S. A. Hu, and P. K. Vogt.** 1977. The occurrence of partial deletion and substitution of the *src* gene in the RNA genome of avian sarcoma viruses. Proc. Natl. Acad. Sci. U. S. A. **74**:4781-4785.
10. **McDonell, M. W., N. Simon, and F. W. Studier.** 1977. Analysis of restriction fragments of T7 DNA and determination of molecular weights by electrophoresis in neutral and alkaline gels. J. Mol. Biol. **110**:119-146.
11. **McLaughlin, C. S., J. R. Warner, M. Edmonds, H. Nakazato, and M. H. Vaughan.** 1973. Polyadenylic acid sequences in yeast messenger RNA. J. Biol. Chem. **248**:1466-1471.
12. **Moscovici, C., M. G. Moscovici, H. Jiminez, M. M. C. Lai, M. J. Hayman, and P. K. Vogt.** 1977. Continuous tissue culture cell lines derived from chemically induced tumors of Japanese quail. Cell **11**:95-103.
13. **Polisky, B., P. Greene, D. E. Garfin, B. J. McCarthy, H. M. Goodman, and H. W. Boyer.** 1975. Specificity of substrate recognition by the *Eco*RI restriction endonuclease. Proc. Natl. Acad. Sci. U. S. A. **72**:3310-3314.
14. **Rothenberg, E., D. J. Donoghue, and D. Baltimore.** 1978. Analysis of a 5' leader sequence on murine leukemia virus 21S RNA: heteroduplex mapping with long reverse transcriptase products. Cell **13**:435-451.
15. **Schwartz, D. E., P. C. Zamecnik, and H. L. Weith.** 1977. Rous sarcoma virus is terminally redundant: the 3' sequence. Proc. Natl. Acad. Sci. U. S. A. **74**:994-998.
16. **Shank, P. R., and H. E. Varmus.** 1978. Virus-specific DNA in the cytoplasm of avian sarcoma virus-infected cells is a precursor to covalently closed circular viral DNA in the nucleus. J. Virol. **5**:104-114.
17. **Southern, E. M.** 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. J. Mol. Biol. **98**:503-517.
18. **Taylor, J. M.** 1977. An analysis of the role of tRNA species as primers for the transcription into DNA of RNA tumor virus genomes. Biochim. Biophys. Acta **473**:57-71.
19. **Taylor, J. M., T. W. Hsu, and M. M. C. Lai.** 1978. Restriction enzyme sites on the avian RNA tumor virus genome. J. Virol. **6**:479-484.
20. **Taylor, J. M., and R. Illmensee.** 1975. Site on the RNA of an avian sarcoma virus at which primer is bound. J. Virol. **16**:553-558.
21. **Taylor, J. M., R. Illmensee, and J. Summers.** 1976. Efficient transcription of RNA into DNA by avian sarcoma virus polymerase. Biochim. Biophys. Acta **442**:324-330.
22. **Taylor, J. M., R. Illmensee, L. Trusal, and J. Summers.** 1976. Transcription of the avian sarcoma virus RNA, p. 172-213. *In* D. Baltimore, A. H. Huang, and C. F. Fox (ed.), Animal virology, vol. 4. Academic Press Inc., New York.
23. **Varmus, H. E., S. Heasly, H.-J. Kung, H. Opperman, V. C. Smith, J. M. Bishop, and P. R. Shank.** 1978. Kinetics of synthesis, structure and purification of avian sarcoma virus-specific DNA made in the cytoplasm of acutely infected cells. J. Mol. Biol. **120**:55-82.
24. **Weinberg, R.** 1977. Synthesis of the intermediates leading to the integrated provirus. Biochim. Biophys. Acta **473**:39-45.