JOURNAL OF VIROLOGY, Oct. 1980, p. 79-92
0022-538X/80/10-0079/14$02.00/0

Vol. 36, No. 1

Transcripts of the Adeno-Associated Virus Genome: Mapping of the Major RNAs

MICHAEL R. GREEN AND ROBERT G. ROEDER*

Departments of Biological Chemistry and Genetics, Division of Biology and Biomedical Sciences, Washington University School of Medicine, St. Louis, Missouri 63110

The four major adeno-associated virus type 2 (AAV2)-specific RNAs were mapped on the linear viral genome by a variety of biochemical techniques, including S1 nuclease and exonuclease VII mapping, RNA gel-transfer hybridization, and analysis of reverse transcriptase extension products. All the major AAV2 RNAs were derived from the minus DNA strand and had 3' termini at position 96. The nucleus-specific 4.3- and 3.6-kilobase (kb) RNAs had 5' termini at positions 6 and 19, respectively. The 5' terminus of the 2.6-kb RNA mapped to position 38.5. The predominant 2.3-kb AAV2 mRNA was spliced and contained a short leader sequence (approximately 50 nucleotides) which mapped to position 38.5, coincident with the 5' terminus of the 2.6-kb RNA. The 5' end of the body of the 2.3-kb RNA mapped to position 46.5. These results are discussed in terms of the involvement of single versus multiple promoters (for transcription) and RNA splicing mechanisms in the generation of the AAV2 RNAs.

Mammalian DNA viruses have provided powerful models for the analysis and formulation of mechanisms of gene expression in eucaryotic cells. The well-studied DNA tumor viruses, papovaviruses and adenoviruses, have been particularly useful in this regard. An enormous amount of insight into the structure and function of eucaryotic genes has resulted directly from the intensive efforts to elucidate the structural relationship between these viral genomes and their products (i.e., RNAs and proteins).

In contrast, parvoviruses have been investigated much less intensely than other DNA viruses. However, these viruses also provide potentially useful systems for study since their programs of gene expression may be much less complex than those of the tumor viruses. For example, in contrast to papovaviruses and adenoviruses, which display at least two phases of gene expression (early and late) and require viral proteins to express late genes (17), the autonomous parvovirus H1 appears to produce the same viral RNAs at early and late times after infection and apparently utilizes host cell mechanisms and components which are independent of viral protein synthesis (18). Although the defective parvovirus AAV2 (adeno-associated virus type 2) does require heterologous viral functions for replication (8), there is as yet no evidence that the AAV2 transcripts vary qualitatively during the lytic cycle. The small size of the parvovirus single-stranded linear DNA genome, 1.4×10^6 to 1.6×10^6 daltons, also makes these viruses among the least complex viruses that infect eucaryotic cells (32, 39).

In our earlier studies of AAV2 (19), we defined and partially characterized four predominant AAV2 RNAs in virus-infected cells, indicating that the previous mapping studies of AAV2 transcripts were incomplete (for a discussion of these early mapping studies, see reference 7). In light of the small size of the viral genome, the number and sizes of the AAV2 RNAs we detected (19) initially suggested a transcriptional map with considerable overlaps in the viral RNAs; it therefore became necessary to map the AAV2 transcripts in more detail. For this purpose, we have used high-resolution techniques which have, in addition, the sensitivity necessary to detect the less abundant RNAs. In this analysis, we were aided by the very complete physical mapping studies that have been published for the AAV2 genome (4, 13). In this report, we describe the orientation of the four major AAV2 RNAs with respect to the viral genome. Moreover, these mapping data reveal a spliced structure for the major AAV2 mRNA, which has a short 5' leader sequence (about 50 nucleotides) ligated to a long (2.3-kilobase) body. Similar conclusions regarding map positions of the AAV2 transcripts and the spliced structure of the 2.3-kb RNA were reported while this manuscript was in preparation (27).

MATERIALS AND METHODS

Cells and virus. AAV2 was propagated in and purified from adenovirus type 2 (Ad2)-infected KB cells according to standard procedures (5). For preparation of ^{32}P-labeled AAV2, 1 liter of KB cells was infected with AAV2 (10 PFU/cell) and Ad2 (5 PFU/

cell). At 10 h postinfection, 0.25 mM ^{32}P-labeled inorganic phosphate (25 mCi) was added, and virus was purified at 40 h postinfection. AAV2 virions were banded two or three times in CsCl, followed by dialysis against 10 mM Tris-hydrochloride (pH 7.5).

Nucleic acids. To prepare AAV2 DNA, virions were digested with 300 μg of proteinase K (Beckman Instruments, Inc.) per ml in 0.5% Sarkosyl (2 h, 37°C) and extracted three times with phenol-chloroform-isoamyl alcohol (1:1:0.05) and once with ether. After removal of the ether with N_2 gas, the DNA was dialyzed against 0.15 mM sodium citrate–1.5 mM sodium chloride. RNA was extracted at 24 h after infection, and polyadenylated RNA was selected by oligodeoxythymidylate-cellulose chromatography as described (18).

AAV2 DNA preparations were routinely self-annealed overnight at 68°C in 0.72 M NaCl–10 mM piperazine-*N*,*N*′-bis(2-ethanesulfonic acid) (PIPES; pH 6.7)–1 mM EDTA–0.05% sodium dodecyl sulfate to ensure that all DNA was double stranded. After ethanol precipitation to remove reagents, the DNA was resuspended in a buffer appropriate for restriction endonuclease digestion. The enzymes *Bam*HI, *Hin*dIII, *Eco*RI, and *Hae*III were prepared in this laboratory according to published procedures. *Hpa*II was purchased from New England Biolabs. DNA fragments were resolved in 1.4% agarose gels or 5 to 20% gradient acrylamide gels modified from Jeppesen (23). After detection of DNA bands by ethidium bromide staining, DNA was purified from gel slices by overnight electroelution into dialysis bags. The DNA was passed through glass wool, concentrated by repeated 2-butanol extractions, and ethanol precipitated.

AAV2 DNA restriction fragments were selectively labeled at the 3′ ends by filling in recessed 3′ ends with Klenow DNA polymerase in the presence of an α-^{32}P labeled-deoxyribonucleoside triphosphate (dNTP). Reaction mixtures (40 μl) contained DNA, 10 mM Tris-hydrochloride (pH 7.9), 10 mM $MgCl_2$, 50 mM NaCl, 10 mM β-mercaptoethanol, 50 μM of each unlabeled dNTP, and 30 μCi of the radioactive dNTP (Amersham; 350 Ci/mmol; added after evaporation to dryness). After the addition of 2 U of Klenow polymerase (a gift of Wayne Barnes), the fill-in reaction was carried out at 5°C for 20 min. After inactivation of the polymerase at 70°C for 5 min, labeled DNA was purified from unincorporated triphosphates on a 1-ml Sephadex G-50 column.

DNA restriction fragments to be labeled at their 5′ ends were dephosphorylated immediately after the restriction endonuclease digestion by the addition of bacterial alkaline phosphatase (Bethesda Research Laboratories), and sodium dodecyl sulfate (to 0.5%; 34) and incubation at 37 to 45°C for 45 min. After phenol and ether extractions, the DNA was subjected to gel electrophoresis and the fragments were purified as described above. The 5′ labeling reactions contained (in 15 μl) 50 mM Tris-hydrochloride (7.5), 10 mM $MgCl_2$, 0.1 mM EDTA, 5 mM dithiothreitol, 0.5 mM spermidine, and at least 1 μM [γ-^{32}P]ATP, which was prepared as described (38). After the addition of 5 U of T4 polynucleotide kinase (P-L Biochemicals), the reaction was incubated at 37°C for 30 min. After kinase inactivation by heating at 70°C for 5 min, labeled DNA was purified from unincorporated [γ-^{32}P]ATP on a 1-ml Sephadex G-50 column.

S1 nuclease and exonuclease VII mapping. Hybridization reactions (20 μl) containing AAV2 [^{32}P]-DNA and RNA in an 80% formamide buffer (9, 37) were incubated at 70°C for 4 min and then at 52 to 60°C for 0.25 to 3 h. Reactions were terminated by adding 250 μl of S1 nuclease buffer, S1 nuclease (prepared as described in reference 18 or purchased from Sigma Chemical Corp.), and 6 μg of denatured salmon-sperm DNA and continuing the incubation at 37 or 50°C for 45 min (18). Alternatively, hybridization reactions were terminated by adding 250 μl of exonuclease VII buffer (15 mM Tris-hydrochloride, pH 7.9, 50 mM potassium phosphate, 10 mM β-mercaptoethanol, and 8 mM EDTA) and excess exonuclease VII (a gift of J. Chase) and continuing the incubation at 37°C for 1 h. The extent of DNA excess (10- to 100-fold in the experiments reported) was determined by assaying a small portion of the hybridization reaction for single-stranded [^{32}P]DNA (18). Decreasing the DNA/RNA ratio 10-fold had no effect on the pattern of S1 nuclease-generated [^{32}P]DNA-RNA hybrids (data not shown; see reference 18).

Electrophoretic resolution, transfer, and detection of AAV2 RNAs on diazobenzylmethyl-paper. RNA and DNA samples were glyoxylated by incubation for 2 h at 37°C in reactions (40 μl) containing 50% dimethyl sulfoxide, 0.25 M HEPES (*N*-2-hydroxyethylpiperazine-*N*′-2-ethanesulfonic acid) (pH 7.5), and 1 M glyoxal. After electrophoresis of the samples in 1.4% agarose gels (18), glyoxal adducts were removed and the RNAs were transferred to diazobenzylmethyl-paper according to published methods (1). Prehybridization (6 h) and hybridization (36 h) reactions were conducted at 45°C as described (1).

Reverse transcriptase extension products. AAV2 5′ [^{32}P]DNA restriction fragments (10^5 cpm) were hybridized at 52°C for 2 h to 30 μg of total cellular RNA from AAV2/Ad2-infected cells in 20 μl of the same 80% formamide buffer (9, 37) used in nuclease mapping studies. Nucleic acids were ethanol precipitated and redissolved in 40 mM Tris-hydrochloride (pH 8.0), 5 mM $MgCl_2$, 25 mM NaCl, 5 mM dithiothreitol, and 1 mM of all four dNTP's. Eighteen units of avian myeloblastosis virus reverse transcriptase (J. W. Beard, Life Sciences, Inc., Petersburg, Fla.) was added, and the reaction was incubated at 41°C for 3 h. After phenol and ether extractions, nucleic acids were precipitated with ethanol, collected and redissolved in alkaline-agarose electrophoresis buffer, and subjected to electrophoresis on 2% alkaline agarose gels.

RESULTS

Mapping of AAV2 RNAs by using large uniformly labeled restriction fragments. To initially assign the four major AAV2 RNAs to specific genomic regions, we used large uniformly labeled DNA restriction fragments as probes in S1 nuclease mapping studies. If an RNA maps entirely within the fragment, hybridization and S1 nuclease digestion produce a DNA-RNA hybrid whose length is identical to that of the original RNA. In contrast, if a portion of the RNA maps outside the boundaries of the

restriction fragment, the hybrid is truncated by the same extent that the RNA maps outside of the probe. A partial restriction map of AAV2, taken from de la Maza and Carter (13) is shown in Fig. 1.

*Bam*HI cleaves AAV2 DNA at position 22 to produce a 3.7-kb A fragment (22 to 100) and a 1.1-kb B fragment (0 to 22) (Fig. 2, lane 2). In an S1 nuclease mapping experiment with *Bam*HI-A and RNA from virus-infected cells, four bands were observed (lane 5). The uppermost faint band (unlabeled) corresponded to a DNA-DNA hybrid which was also visualized when RNA from mock-infected cells was used (lane 4). The other three bands in lane 5 corresponded to viral RNA-DNA hybrids and consisted of a predominant 2.3-kb hybrid and less abundant 2.6-kb and 3.6-kb hybrids. For comparison, the 4.3-, 3.6-, 2.6-, and 2.3-kb hybrids generated with an intact DNA probe are shown in lane 3 (hybrid positions indicated by small arrows; see also reference 19). The 2.3- and 2.6-kb hybrids formed with *Bam*HI-A were equivalent in length to the corresponding hybrids formed with full-length DNA, whereas the 3.6-kb *Bam*HI-A hybrid was slightly shorter than the 3.6-kb hybrid formed with full-length DNA. With *Bam*HI-A no 4.3-kb hybrid was observed, as expected from the length of the DNA fragment. From these data we conclude that, within the resolution of the gel system (50 to 100 nucleotides), all sequences encoding the 2.3- and 2.6-kb RNAs are within map positions 22 to 100. Additionally, the 4.3-kb, and probably the 3.6-kb, RNA contains sequences from positions 0 to 22.

*Hin*dIII cleaved AAV2 DNA once at position 40, producing a 2.9-kb A fragment (40 to 100) and a 1.9-kb B fragment (0 to 40) (Fig. 3, lane 1). With RNA from AAV2-infected cells, the *Hin*dIII A fragment produced two specific DNA-RNA hybrid bands in S1 nuclease mapping experiments (Fig. 3, lane 4). For comparison, the four hybrids formed with full-length DNA are shown in lane 2 (small arrows). A 2.3-kb band (lane 4) comigrated with and had a relative abundance equal to the control 2.3-kb DNA-RNA hybrid formed with full-length DNA (lane 2). However, the apparent 2.6-kb band generated with *Hin*dIII-A (lane 4) appeared slightly truncated relative to the control 2.6-kb hybrid band (lane 2). Neither a 4.3- nor a 3.6-kb band was

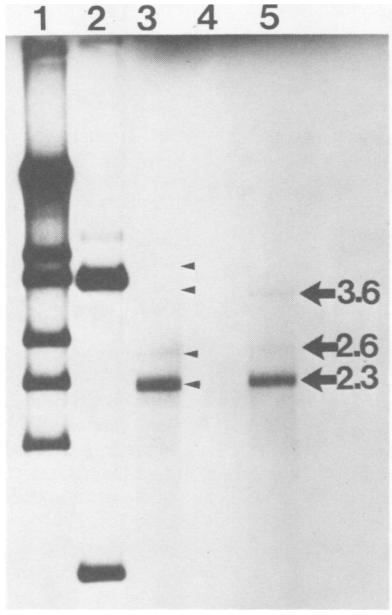

FIG. 2. *S1 nuclease-generated AAV2 BamHI-A [^{32}P]DNA-RNA hybrids displayed on 1.4% neutral agarose gels. Lane 1 contains an EcoRI digest of Ad2 [^{32}P]DNA (for fragment sizes, see reference 18). Lane 2 contains a BamHI digest of AAV2 [^{32}P]DNA. Lane 3 contains S1 nuclease digestion products of AAV2 [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells. Lane 4 contains S1 nuclease digestion products of AAV2 BamHI-A [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 5 contains S1 nuclease digestion products of AAV2 BamHI-A [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells.*

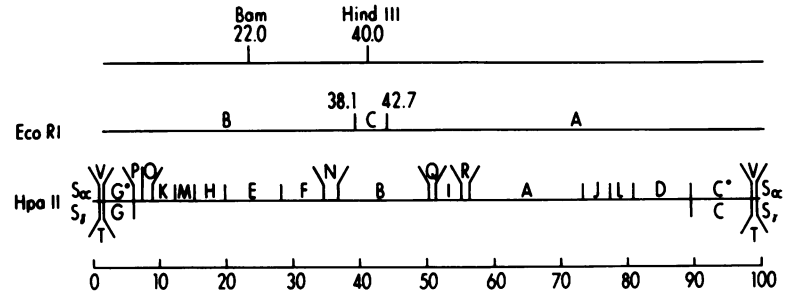

FIG. 1. *Restriction endonuclease cleavage map of AAV2 DNA. This map was redrawn from data presented in reference 13. However, also see reference 4 particularly in regard to map assignments and discussion of the terminal DNA fragments. One map unit is approximately 48 nucleotides.*

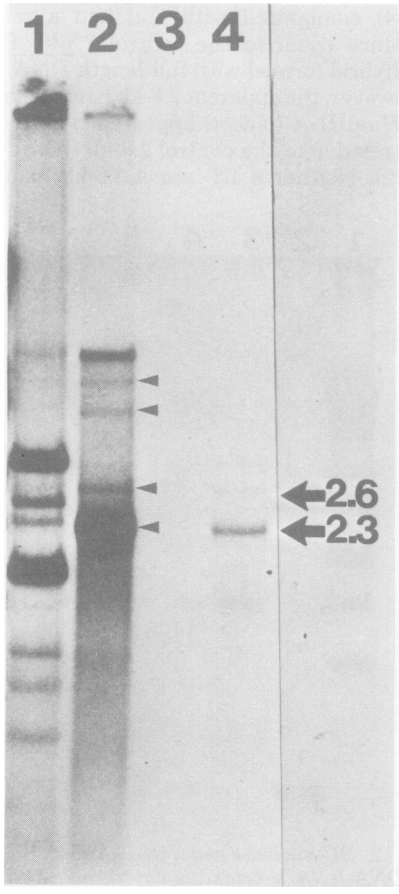

FIG. 3. *S1 nuclease-generated AAV2 HindIII-A [^{32}P]DNA-RNA hybrids displayed on 1.4% neutral agarose gels. Lane 1 contains a HindIII digest of AAV2 [^{32}P]DNA (this particular enzyme preparation was contaminated by HindII activity resulting in the minor bands). Lane 2 contains S1 nuclease digestion products of AAV2 [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells. Lane 3 contains S1 nuclease digestion products of AAV2 HindIII-A [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 4 contains S1 nuclease digestion products of AAV2 HindIII-A [^{32}P]-DNA annealed with whole-cell RNA from AAV2/ Ad2-infected KB cells.*

observed with the *Hin*dIII A fragment since this probe was itself shorter than 3.6 kb. Thus, within the resolution of the gel system, we conclude that the 2.3-kb RNA maps between positions 40 and 100 and that the 2.6-kb RNA contains some sequences from positions 22 to 40 (in addition to those derived from positions 40 to 100).

*Eco*RI cleaved AAV2 DNA at positions 38.1 and 42.7, producing a 2.7-kb A fragment (42.7 to 100), a 1.8-kb B fragment (0 to 38.1), and a 0.2-kb C fragment (38.1 to 42.7), the first two of which are seen as the major bands in lane 2 of Fig. 4A (the small C fragment has been run off the gel in this experiment). In an S1 nuclease mapping experiment with AAV2 RNA (Fig. 4A), the *Eco*RI A fragment produced a 2.3-kb DNA-RNA hybrid (lane 5) indistinguishable in length and relative abundance from the control 2.3-kb hybrid band seen with full-length DNA (lane 3). Thus, within the resolution of the gel system, the 2.3-kb RNA must map between positions 42.7 and 100. The *Eco*RI A fragment did not generate a hybrid that comigrated with the control 2.6-kb DNA-RNA hybrid (lane 3), but did produce a band which migrated slightly faster

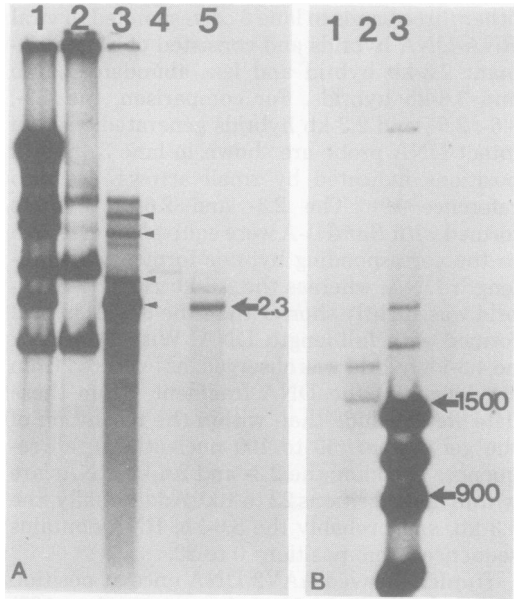

FIG. 4. *S1 nuclease-generated AAV2 EcoRI-A and -B [^{32}P]DNA-RNA hybrids displayed on 1.4% neutral agarose gels. (A) Lane 1 contains an EcoRI digest of Ad2 [^{32}P]DNA. Lane 2 contains an EcoRI digest of AAV2 [^{32}P]DNA. Lane 3 contains S1 nuclease digestion products of AAV2[^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells. Lane 4 contains S1 nuclease digestion products of AAV2 EcoRI-A [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 5 contains S1 nuclease digestion products of AAV2 EcoRI-A [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells. (B) Lane 1 contains S1 nuclease digestion products of AAV2 EcoRI-B [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 2 contains S1 nuclease digestion products of AAV2 EcoRI-B [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells. Lane 3 contains an HaeIII digest of ϕX174 replicative form [^{32}P]DNA. The sizes of the ϕX174 HaeIII DNA fragments in this and subsequent figures are based on nucleotide sequence analysis of ϕX174 DNA (33) and are (in nucleotide number): A, 1,353; B, 1,078; C, 872; D, 603; E, 310; F, 281; G, 271; H, 234; I, 194; J, 118; K, 72.*

(lane 5). Thus, in combination with the results obtained with *Hin*dIII-A and *Bam*HI-A, we conclude that the 2.6-kb RNA contains sequences between 22 and 42.7, as well as those derived from within the region of 42.7 to 100.

In an S1 nuclease mapping experiment with *Eco*RI-B (Fig. 4B), two specific DNA-RNA bands were observed (lane 2) in addition to the DNA-DNA renaturation band (compare lanes 1 and 2). The sizes of the DNA-RNA hybrids were 1,500 and 900 nucleotides, based on migration positions relative to those of radioactive ϕX174 (double-stranded) *Hae*III DNA fragments (lane 3). Since the previous experiments showed that the 4.3- and 3.6-kb RNAs contain sequences from the 0 to 40 region of the genome, it is likely that the 1,500- and 900-nucleotide hybrids result from hybridization of *Eco*RI-B with the left ends of the 4.3- and 3.6-kb RNAs, respectively. Experiments presented below prove this hypothesis.

Strand and region specificity of the large AAV2 nuclear specific RNAs. AAV2 RNA is reported to be encoded only by the minus DNA strand (7). However, since these early studies were based on results from RNA excess saturation-hybridization experiments, it is possible that only the very abundant 2.3-kb RNA was detected. To determine whether the much less abundant 4.3- and 3.6-kb RNAs were also derived from the minus DNA strand and to gain further mapping information, we constructed a region- and strand-specific probe for the large nuclear AAV2 RNAs.

After AAV2 DNA was cleaved with *Bam*HI, the larger A fragment was isolated and the minus strand was selectively labeled by "filling in" the staggered *Bam*HI site with Klenow polymerase in the presence of [α-^{32}P]dATP (see Materials and Methods). This radioactive *Bam*HI A fragment was then cleaved with *Eco*RI (position 38.1), and the resulting radioactive 22 to 38.1 DNA fragment was isolated and used as the probe in an RNA gel-transfer hybridization experiment (see Materials and Methods).

The probe detected only RNAs corresponding to the scarce 4.3-, 3.6-, and 3.4-kb RNAs and none corresponding to the more abundant 2.3- and 2.6-kb species (positions indicated; Fig. 5, lane 3). Thus, the 4.3-, 3.6-, and 3.4-kb RNAs were transcribed from the same DNA strand (minus) as was the more abundant species. Moreover, genomic sequences from positions 22 to 38.1 encoded only the larger AAV2 RNAs and not the 2.3- and 2.6-kb transcripts, as concluded in the previous section. These deductions about strand origin and map positions for all viral transcripts are corroborated by data presented above and below.

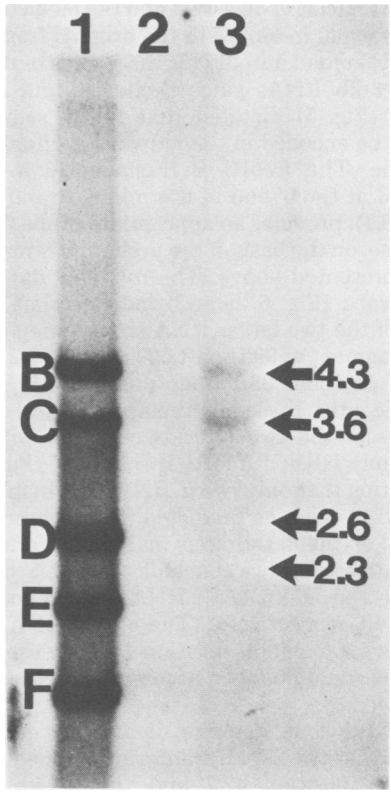

FIG. 5. *Detection of AAV2 RNAs with a position 22 to 38.1 minus-strand [^{32}P]DNA probe by RNA gel-transfer hybridization. RNAs were glyoxylated, subjected to electrophoresis on 1.4% neutral agarose gels, transferred to diazobenzylmethyl-paper, and analyzed by hybridization with the labeled DNA conducted as described (19). Construction of the 22 to 38.1 minus-strand [^{32}P]DNA probe is described in the text. Lane 1, glyoxylated [^{32}P]labeled Ad2 EcoRI fragments B to D run as markers. Lane 2, 2 μg of poly(A)(+), whole-cell RNA from Ad2-infected KB cells. Lane 3, 2 μg of poly(A)(+) whole-cell RNA from AAV2/Ad2-infected KB cells.*

Positioning the 5′ ends of the AAV2 RNAs. The 5′ termini of RNAs can be identified and positioned in S1 nuclease mapping experiments, using 5′ end-labeled DNA probes (2, 36, 40). If a 5′ end-labeled DNA fragment encodes only the 5′, and not the 3′, terminus of an RNA, hybridization results in an overlapping hybrid in which the 5′ label is protected from S1 nuclease digestion. S1 nuclease digestion generates a shortened DNA fragment, the length of which is equivalent to the distance between the 5′ end of the RNA and the labeled DNA end. Alternatively, if the DNA encodes only an internal portion of the RNA, the resulting hybrid DNA will be flanked by RNA at both the 5′ and 3′ ends. The 5′ label will, therefore, be protected

from S1 nuclease digestion and the labeled DNA will be equal in length to the original fragment.

We searched initially for the 5′ ends of the 4.3- and 3.6-kb RNAs since their size and strand origin (Fig. 5) dictated that the 5′ sequences would be encoded in the leftward portion of the genome. The *Eco*RI B fragment (0 to 38.1), labeled at the 5′ end of the minus strand (position 38.1), provides an appropriate probe for this purpose, on the basis of the preliminary mapping data presented above. The mapping data with this probe (Fig. 6, lane 3) indicate that the 5′ ends of the two larger RNA species map within this fragment at 900 and 1,500 nucleotides to the left of positions 38.1 (map positions 19 and 6). Significantly, these lengths are identical to the lengths of the two hybrids observed with the uniformly labeled *Eco*RI B fragment (Fig. 4B), indicating that only two 5′ RNA ends map in the 0 to 38.1 region as predicted. From these data, as well as the 3′ terminus mapping information presented below, we conclude that the 5′ ends of the 4.3- and 3.6-kb AAV2 RNAs map at positions 6 and 19, respectively. (The 4.3-kb RNA could not have a 5′ end at position 19 since the small genome could not then accommodate the molecule.)

The previous mapping data (Fig. 4A) suggested that the 2.3-kb transcript mapped to the right of position 42.7 and the 2.6-kb RNA mapped predominantly to the right of position 40 (possibly with some sequences to the left of position 40). Based on early AAV2 transcription studies (7), we felt it highly probable that these RNAs were also encoded on the minus DNA strand. Therefore, we chose the 5′ end-labeled 0.7-kb *Hpa*II B fragment (≃36 to 50) as a probe to map the 5′ ends of these RNAs. This DNA fragment produced two strong AAV2 RNA-dependent bands in S1 nuclease mapping experiments (Fig. 7, lane 3). (The 700-nucleotide band in lanes 2 and 3 was due to reannealing of the [^{32}P]DNA probe.) The more intense of these hybrid bands was 180 nucleotides and the second was 550 nucleotides in length (based on calibration with 5′ end-labeled ϕX174 *Hae*III DNA fragments; lane 1). Based on the size and map positions of the bodies of the 2.3- and 2.6-kb RNAs (Fig. 1 to 3, above), the relative intensities of the 180- and 550-nucleotide bands, and 3′ terminus mapping results (below), we interpret the 180- and 550-nucleotide bands as the 5′ ends of the 2.3- and 2.6-kb RNAs, respectively. These results also provide strong and consistent evidence that the 2.3- and 2.6-kb transcripts are encoded by the minus DNA strand, since positive strand transcripts of this length could not map entirely to the right of position 22 (Fig. 2) and still have 5′ ends mapping in the *Hpa*II B fragment (Fig. 7). Therefore, we position the 5′ ends of the 2.3- and 2.6-kb RNAs (as detected by this methodology) at 180 and 550 nucleotides to the left of position 50, at positions 46.5 and 38.5, respectively.

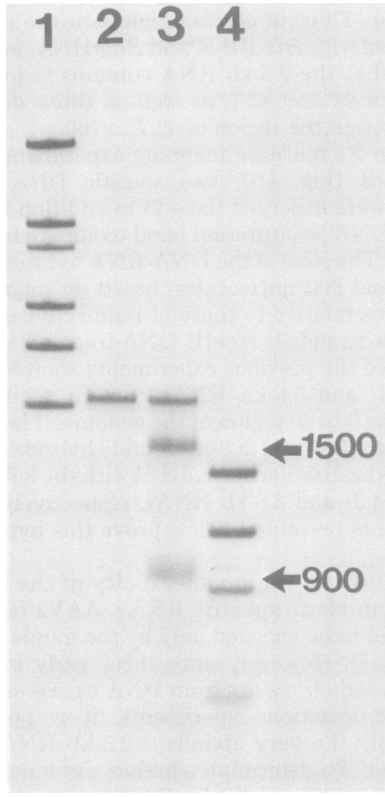

FIG. 6. *S1 nuclease-generated 5′-labeled AAV2 EcoRI-B [^{32}P]DNA-RNA hybrids displayed on 1.4% neutral agarose gels. Lane 1 contains an EcoRI digest of Ad2 [^{32}P]DNA. Lane 2 contains S1 nuclease digestion products of 5′-labeled AAV2 EcoRI-B [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 3 contains S1 nuclease digestion products of 5′-labeled AAV2 EcoRI-B [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells. Lane 4 contains an HaeIII digest of ϕX174 replicative form [^{32}P]DNA.*

As a check on our methodology and to investigate the possibility that other viral RNA 5′ termini existed, we also performed S1 nuclease mapping experiments with other 5′ end-labeled DNA fragments. Significantly, the *Hpa*II A, D, E, and I fragments, *Hin*dIII A fragment, and *Eco*RI A fragment all failed to result in the production of any labeled hybrids shorter than the [^{32}P]DNA probe (not shown).

Major AAV2 transcripts form a 3′ coterminal family. The sizes, relative mapping assignments, strandedness, and location of the 5′ ends of the four AAV2 RNAs suggested that all

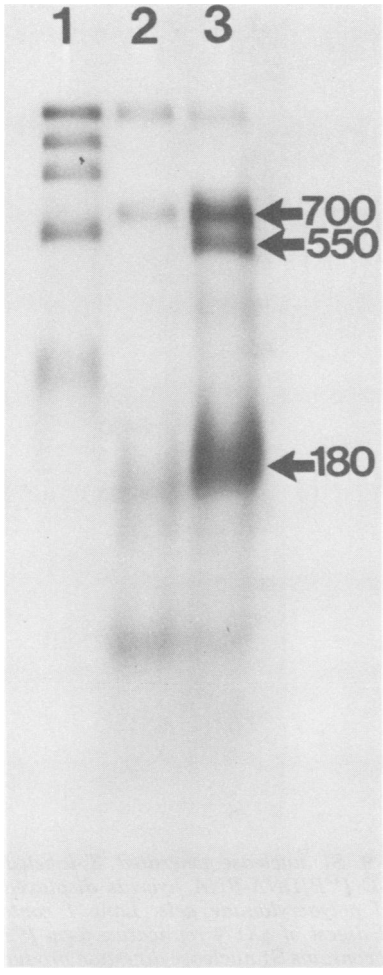

FIG. 7. *S1 nuclease-generated 5'-labeled AAV2 HpaII-B [^{32}P]DNA-RNA hybrids displayed on 2% alkaline agarose gels. Lane 1 contains 5'-labeled φX174 replicative form HaeIII [^{32}P]DNA fragments. Lane 2 contains S1 nuclease digestion products of 5'-labeled AAV2 HpaII-B [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 3 contains S1 nuclease digestion products of 5'-labeled AAV2 HpaII-B [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells.*

the AAV2 transcripts might coterminate at about position 95 on the genome. This contention was directly tested in S1 nuclease mapping experiments using 3'-labeled DNA fragments, which are specific for the 3' ends of RNAs. To initially search for 3' termini of the less abundant 4.3-, 3.6-, and 2.6-kb RNAs, we constructed an *Eco*RI-A probe (42.7 to 100), labeled at the 3' end of the minus DNA strand (42.7). In S1 nuclease experiments, this [^{32}P]DNA fragment produced only one DNA-RNA hybrid band (2.5 kb), about 200 nucleotides shorter than the original *Eco*RI A fragment (Fig. 8, lane 3). This result indicates that 3' ends of the AAV2 RNAs detected mapped to position 96 and were not distributed heterogeneously. Also shown in Fig. 8 (lanes 4 and 5) is an S1 nuclease mapping experiment using a 3'-labeled *Eco*RI B fragment. Only the DNA-DNA renaturation band, and no DNA-RNA hybrid bands, was evident. Therefore, no AAV2 RNA had a 3' end mapping between positions 0 and 38.1 on the plus DNA strand, a result consistent with the other mapping data and strand assignments presented above.

The experiment with the *Eco*RI DNA fragment A could not detect the 3' terminus of the predominant 2.3-kb RNA, since the body of this transcript mapped to the right of the labeled

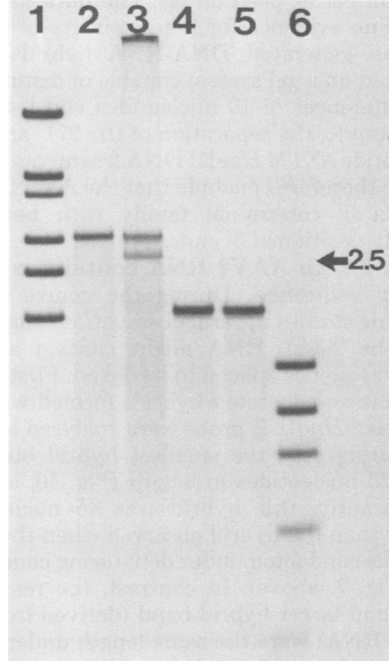

FIG. 8. *S1 nuclease-generated 3'-labeled AAV2 EcoRI-A and -B [^{32}P]DNA-RNA hybrids displayed on 1.4% neutral agarose gels. Lane 1 contains an EcoRI digest of Ad2 [^{32}P]DNA. Lane 2 contains S1 nuclease digestion products of 3'-labeled AAV2 EcoRI-A [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 3 contains S1 nuclease digestion products of 3'-labeled AAV2 EcoRI-A [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells. Lane 4 contains S1 nuclease digestion products of 3'-labeled AAV2 EcoRI-B [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 5 contains S1 nuclease digestion products of 3'-labeled AAV2 EcoRI-B [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells. Lane 6 contains an HaeIII digest of φX174 replicative form [^{32}P]DNA.*

DNA terminus at position 42.7. To map the 3′ end of this transcript, to further analyze for heterogeneity of the 3′ ends of all RNAs, and to more precisely delineate the 3′ map position(s), we used a 3′-labeled *Hpa*II C fragment (≃89 to 99) in S1 nuclease mapping experiments. (The size, strandedness, and general location of the 2.3-kb RNA imply that the 3′ end of this transcript maps within the *Hpa*II C fragment.) The results are shown in Fig. 9 (lane 3), in this case displayed on a nondenaturing polyacrylamide gel. The faint band (indicated by upper arrow) represents the renatured full-length DNA (480 nucleotides). The strong single DNA-RNA hybrid band (lower arrow) was 320 nucleotides in length, as estimated from the length standards run in lane 1. Thus, the 3′ ends of all the AAV2 RNAs mapped 320 nucleotides to the right of position 89, at position 96. The data in Fig. 9 reveal no evidence for heterogeneity of the S1 nuclease-generated DNA-RNA hybrids when analyzed on a gel system capable of distinguishing differences of 10 nucleotides and less (see, for example, the separation of the 271- and 281-nucleotide φX174 *Hae*III DNA fragments in lane 1). We therefore conclude that the AAV2 RNAs form a 3′ coterminal family with heterogeneously positioned 5′ ends.

The 2.3-kb AAV2 RNA contains a short leader sequence. During the course of our mapping studies, several observations suggested that the 2.3-kb RNA might contain a short leader sequence spliced to its 5′ end. First, when S1 nuclease-generated hybrids formed with the 5′-labeled *Hpa*II-B probe were analyzed on nondenaturing gels, the smallest hybrid observed was 225 nucleotides in length (Fig. 10, lane 3). Significantly, this hybrid was 45 nucleotides longer than the hybrid observed when the analysis was conducted under denaturing conditions (see Fig. 7, above). In contrast, the renatured DNA and larger hybrid band (derived from the 2.6-kb RNA) were the same length under denaturing and nondenaturing conditions (approximately 700 and 550 nucleotides, respectively). A plausible interpretation of this result is that the 2.3-kb RNA contains a 5′ leader sequence, approximately 45 nucleotides long, encoded in the *Hpa*II B fragment. The leader would be released from the hybrid under denaturing (but not nondenaturing) conditions, resulting in the disparate length measurements. No evidence for this leader was obvious when full-length hybrids were analyzed under denaturing versus nondenaturing conditions (19) because the gel system used could not resolve 45-nucleotide differences in this size range (2,000 to 4,000 nucleotides). However, differences of 45 nucleotides can be easily resolved when hybrids of several hundred nucleotides are analyzed.

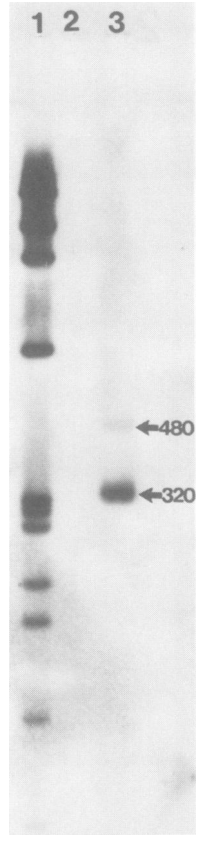

FIG. 9. *S1 nuclease-generated 3′-labeled AAV2 HpaII-C [^{32}P]DNA-RNA hybrids displayed on 7% neutral polyacrylamide gels. Lane 1 contains an HaeIII digest of φX174 replicative form [^{32}P]DNA. Lane 2 contains S1 nuclease digestion products of 3′-labeled AAV2 HpaII-C [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 3 contains S1 nuclease digestion products of 3′-labeled AAV2 HpaII-C [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells.*

The second result suggesting that the 2.3-kb RNA contained a 5′ leader sequence was from experiments localizing the 5′ ends of viral RNAs by analysis of reverse transcriptase extension products. 5′-labeled DNA fragments are hybridized to RNA (DNA excess, high formamide conditions) followed by extension of the labeled DNA primer by reverse transcriptase. The length of these extension products provides information about the location of the 5′ ends of the RNAs relative to the restriction fragment. We hose the *Hpa*II Q fragment (50 to 51.1) as a primer since it was small and located close to the presumptive 5′ end of the 2.3-kb RNA. The

FIG. 10. *S1 nuclease-generated 5′-labeled AAV2 HpaII-B [^{32}P]DNA-RNA hybrids displayed on 7% neutral polyacrylamide gels. Lane 1 contains 5′-labeled φX174 replicative form HaeIII [^{32}P]DNA fragments. Lane 2 contains S1 nuclease digestion products of 5′-labeled AAV2 HpaII-B [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells.*

reverse transcriptase extension products of this primer are shown in Fig. 11. If no reverse transcriptase was added to the hybrids, only a single band equivalent in length (50 nucleotides) to the original *Hpa*II Q fragment was observed (lane 2). However, if the hybridized primer was extended with reverse transcriptase two discrete products, in addition to the *Hpa*II-Q band, were observed (lane 3). The shorter, more intense band was 280 nucleotides and resulted from extension of the primer in hybrid with the 2.3-kb RNA. The longer extension product was 600 nucleotides and resulted from extension of the primer in hybrid with the 2.6-kb RNA.

The S1 nuclease mapping experiment using the 5′ *Hpa*II-B probe detected a very abundant 180-nucleotide DNA-RNA hybrid (Fig. 7). If this represented the real 5′ terminus of the 2.3-kb

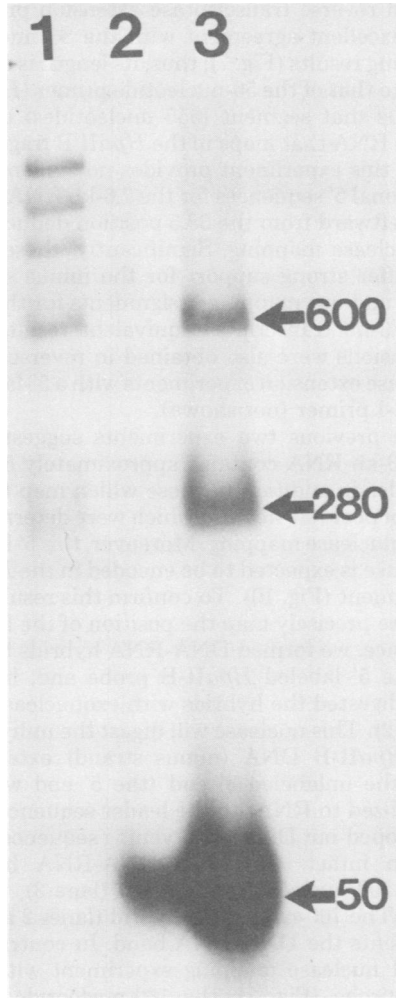

FIG. 11. *Reverse transcriptase extension products of 5′-labeled AAV2 HpaII-Q [^{32}P]DNA-RNA hybrids displayed on 2% alkaline agarose gels. Lane 1 contains 5′-labeled φX174 replicative form HaeIII [^{32}P]DNA fragments. Lane 2 contains 5′-labeled AAV2 HpaII-Q [^{32}P]DNA-RNA hybrids after an extension reaction in which the reverse transcriptase was omitted. Lane 3 contains 5′-labeled AAV2 HpaII-Q [^{32}P]DNA-RNA hybrids after extension with reverse transcriptase.*

RNA, then the major reverse transcriptase extension product of the *Hpa*II-Q fragment should be equal in length to *Hpa*II-Q (≃50 nucleotides) plus the 180 nucleotides of RNA that maps into the *Hpa*II B fragment (230 nucleotides). Significantly, the 280-nucleotide length of the major reverse transcriptase extension product is 50 nucleotides longer than this 230-nucleotide prediction, suggesting that the 2.3-kb RNA contains approximately 50 additional nucleotides at its 5′ end. Moreover, the 600-nucleotide length of the second reverse transcriptase extension product is in excellent agreement with the S1 nuclease mapping results (Fig. 7); thus, its length is equivalent to that of the 50-nucleotide primer (*Hpa*II-Q) plus that segment (550 nucleotides) of the 2.6-kb RNA that maps in the *Hpa*II B fragment. Thus, this experiment provides no evidence for additional 5′ sequences for the 2.6-kb RNA mapping leftward from the 38.5 position deduced by S1 nuclease mapping. Significantly, these data also offer strong support for the minus strand origin and our mapping assignments for the 2.6- and 2.3-kb transcripts. Equivalent results and conclusions were also obtained in reverse transcriptase extension experiments with a 5′-labeled *Hpa*II-I primer (not shown).

The previous two experiments suggest that the 2.3-kb RNA contains approximately 50 nucleotides in addition to those which map to the right of position 46.5 and which were determined by S1 nuclease mapping. Moreover, this 5′ leader sequence is expected to be encoded in the *Hpa*II B fragment (Fig. 10). To confirm this result and to more precisely map the position of the leader sequence, we formed DNA-RNA hybrids by using the 5′ labeled *Hpa*II-B probe and, in this case, digested the hybrids with exonuclease VII (Fig. 12). This nuclease will digest the unhybridized *Hpa*II-B DNA (minus strand) extending from the unlabeled 3′ end (the 5′ end will be hybridized to RNA) to the leader sequence, but any looped-out DNA (intervening sequence) will remain intact. Only one DNA-RNA hybrid band, 550 nucleotides in length (lane 3), is obvious. The 700-nucleotide hybrid (lanes 2 and 3) represents the DNA-DNA band. In contrast to the S1 nuclease mapping experiment with the same probe (Fig. 7), the 180-nucleotide band (due to the 2.3-kb RNA) was totally absent, but the 550-nucleotide band (due to the 2.6-kb RNA) was identical in both experiments. These results demonstrate that the 2.3-kb RNA is spliced and contains a 5′ leader sequence which maps coincident with the 5′ end of the unspliced 2.6-kb RNA. This result was suggested by an earlier mapping experiment using exonuclease VII and full-length AAV2 [^{32}P]DNA (19). In that case, we observed a significantly increased ratio of the

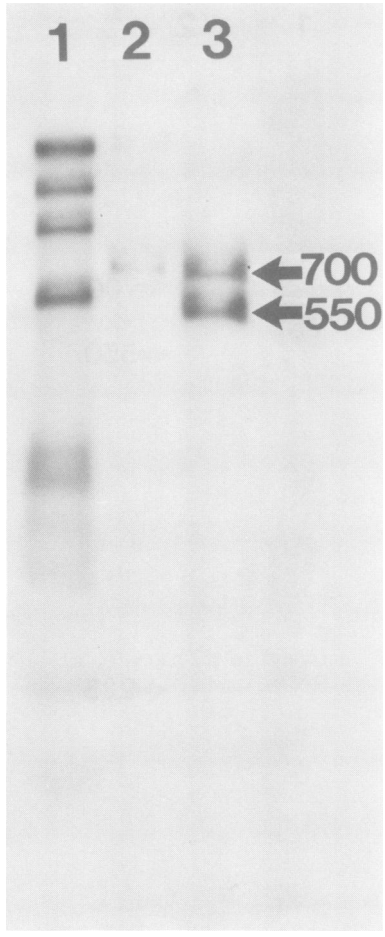

FIG. 12. *Exonuclease VII-generated 5′-labeled AAV2 HpaII-B [^{32}P]DNA-RNA hybrids displayed on 2% alkaline agarose gels. Lane 1 contains 5′-labeled φX174 replicative form HaeIII [^{32}P]DNA fragments. Lane 2 contains exonuclease VII digestion products of 5′-labeled AAV2 HpaII-B [^{32}P]DNA annealed with whole-cell RNA from Ad2-infected KB cells. Lane 3 contains exonuclease VII digestion products of 5′-labeled AAV2 HpaII-B [^{32}P]DNA annealed with whole-cell RNA from AAV2/Ad2-infected KB cells.*

2.6-kb to the 2.3-kb bands compared with S1 nuclease digestion products, although a 2.3-kb band was also present. It is likely that in the earlier experiment, the higher hybridization temperatures (60 versus 52°C) or the size of the probe (full-length DNA versus the *Hpa*II B fragment) resulted in some hybrids containing leaders not hybridized to DNA, thus allowing digestion of the intervening sequence by exonuclease VII. Significantly, both exonuclease VII experiments, interpreted in conjunction with our other mapping data, indicate that 5′ ends of viral RNAs map to positions 6, 19, and 38.5. Of course,

our results do not exclude the possibility that a small number of nucleotides, derived from other genome regions, are spliced to either end of these RNAs.

Our mapping data are summarized in Fig. 13. We believe that this is the only model consistent with all of the data. The length of the transcripts deduced from the map positions agrees well with the lengths we originally reported based on sizing full-length hybrids in S1 nuclease mapping experiments (19).

DISCUSSION

Elsewhere (19) we demonstrated that cells productively infected with AAV2 accumulate multiple viral RNAs, which were also analyzed with respect to size, cellular location, and specific structural features. In the present communication, we have extended these studies by determining the positional relationships between the four major AAV2 transcripts and the viral genome (Fig. 13). As indicated, the four predominant AAV2 RNAs coterminate at position 96. The 5′ ends of the 4.3-, 3.6-, and 2.6-kb RNAs map, respectively, to positions 6, 19, and 38.5. The 2.3-kb RNA is spliced and contains a 50-nucleotide leader sequence (5′ end at 38.5) and a 2.3-kb body (5′ end at 46.5). As indicated, the arrangement of the AAV2 RNAs is similar in several respects to the arrangements of certain classes of transcripts observed in other eucaryotic viruses. Thus, in both early and late classes of Ad2 transcripts, there are examples of families of (overlapping) mRNAs whose main bodies have common 3′ termini but heterogeneously positioned 5′ terminal regions (10, 29, 30). Moreover, long polyadenylated nuclear RNAs with 3′ ends identical to those of shorter cytoplasmic mRNA's have been reported at late times of infection (2), and these could be analogous to the AAV2 4.3- and 3.6-kb nuclear RNAs. Papovaviruses (3), alphaviruses (26), and RNA tumor viruses (22) also encode overlapping RNAs with common 3′ sequences.

In contrast, the pattern and organization of the AAV2 RNAs appear to differ from the transcript patterns found for certain regions of several other DNA viruses. Thus, 5′ termini of the AAV2 RNAs map to distinct regions on the AAV2 genome, but in the case of Ad2 the major late viral mRNA's (11, 15) and long nuclear RNAs (2) contain identical 5′-terminal leader sequences which are encoded at the beginning of the late transcription unit (42). In this regard, however, we emphasize that the methodology used in this report may not have detected, on the various AAV2 RNAs, a short terminal sequence of nucleotides encoded at an "upstream" site(s) on the AAV2 genome.

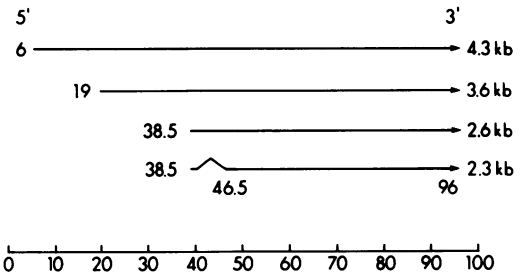

FIG. 13. *AAV2 transcription map. The four major AAV2 transcripts are referred to by size (in kilobases) in the right margin. All four viral RNAs are transcribed from the minus DNA strand, with 3′ termini at position 96 and 5′ termini at 6, 19, and 38.5 as indicated for each RNA. This map does not exclude the presence of 5′ RNA sequences encoded at other (upstream) regions (see text). The sequence spliced out of the 2.3-kb RNA is indicated by the caret.*

The relationship between the major 2.3-kb AAV2 RNA (a presumptive mRNA [19]) and the viral genome also shows a similarity to the relationship between autonomous parvovirus RNAs and their DNA templates. In the case of both minute virus of mice (MVM) (35) and H1 (18; M. R. Green and R. G. Roeder, unpublished data), the bodies of the abundant viral RNAs map to the right end of the genome; in all three cases (AAV2, MVM, and H1) transcription proceeds from left to right. However, there is a major difference between the major AAV2 RNA and the predominant H1 and MVM transcripts. The latter contain long leader sequences (200 to 400 nucleotides [18]) which in the case of MVM (35) map near the extreme left end of the genome; in contrast, the predominant AAV2 RNA contains a short leader which maps near the center of the genome.

On the basis of the relative concentrations and sequence overlaps (homologies) between the H1 virion polypeptides and on the basis of the sizes, relative concentrations, and splicing patterns (which indicated a large common sequence element) of the H1 RNAs, we speculated that the translation products of the 3.0- and 2.8-kb RNAs were, respectively, the minor (92,000-dalton) and major (72,000-dalton) H1 virion proteins. On the basis of similar reasoning, it is plausible that the predominant 2.3-kb AAV2 RNA, which is both polyadenylated and present in polysomes, encodes the major AAV2 virion protein (66,000 daltons) and that the 2.6-kb AAV2 RNA encodes one of the two minor AAV2 virion polypeptides (80,000 and 90,000 daltons). (The third virion polypeptide might be derived by proteolytic cleavage as occurs in nondefective parvoviruses [39], but this has not yet been demonstrated for AAV2.) The apparently low concentration of the

presumptive 2.6-kb mRNA (or any other minor species) in the polysome population may simply be a reflection of the exceedingly low molar ratio of the minor virion polypeptide(s) relative to the major virion protein. This hypothesis regarding multiple AAV2 mRNAs would also explain the reported observation of several primary translation products in AAV2-infected cells and in cell-free translation systems (6). Additionally, the RNA mapping data presented here demonstrate extensive sequence overlap between the 2.3- and 2.6-kb AAV2 RNAs (Fig. 13), consistent with the extensive sequence homology between the AAV2 virion proteins (24) and in support of a model in which two mRNA's are translated in the same reading frame.

A noteworthy feature of the AAV2 RNAs is that, of the four major species detected, only the very abundant 2.3-kb mRNA appears to exist in a predominantly spliced form. Thus, the 4.3-, 3.6-, and 2.6-kb RNAs are predominantly or exclusively (by mass) present as unspliced structures. Our data do not exclude the existence of spliced counterparts for the 4.3- and 3.6-kb RNAs, which have been detected in a recent AAV2 transcription study (27); however, if large spliced RNAs are present, they represent a minor fraction (by mass) of the population of large viral nuclear RNAs. For example, in our previous study we detected a 3.4-kb AAV2 RNA which might be a spliced counterpart to the 3.6-kb RNA, but even in this case, the unspliced 3.6-kb RNA is significantly more abundant than the 3.4-kb RNA (19).

These results for the AAV2 RNAs clearly differ from those reported in our studies of the autonomous parvovirus H1 transcripts. In the case of H1, the three major viral RNAs exist predominantly (by mass) as spliced structures and contain bodies that are identical in length (2.6 kb) (18). Potential unspliced forms of the three major spliced H1 RNAs were detected only in nuclear RNA preparations in drastically reduced amounts compared with the spliced H1 RNAs (18).

Our cellular localization studies (19) demonstrated that the unspliced AAV2 RNAs are predominantly confined to the nucleus, whereas the spliced 2.3-kb mRNA is also found in significant amounts on polysomes (19). These results are very consistent with recent studies which suggest that splicing may in some cases be essential for or enhance mRNA formation and stabilization (20, 21).

The mapping data allow the formulation of two extreme models for the biogenesis of the AAV2 RNAs. First, in a single-promoter model, a primary transcription event involving transcription of 90% or more of the AAV2 genome, followed by polyadenylation, would generate the 4.3-kb RNA. The other AAV2 RNA species could then be derived from this primary transcript by a series of 5′ exonucleolytic or endonucleolytic cleavages and splicing (in the case of the 2.3-kb RNA) reactions. A strong precedent exists for this model in that families of related adenovirus and papovavirus RNAs are produced from single transcriptional units by a similar series of events (for review, see reference 12 and references therein). Second, in a multiple-promoter model, the various AAV2 RNAs would result from independent and separate initiation events at different regions along the viral genome, followed by posttranscriptional modifications, including splicing. This model is reminiscent of the manner in which late viral RNAs are synthesized from some bacteriophage genomes (25).

Significantly, since all the major AAV2 RNAs are polyadenylated (19), both models are compatible with recent data suggesting that polyadenylation precedes splicing (31). Moreover, in both models the 2.6-kb RNA would be the unspliced counterpart of, and possibly the immediate precursor to, the spliced 2.3-kb mRNA. Therefore, AAV2 may provide an advantageous system for studying RNA splicing. The 2.6-kb RNA is present in nuclear RNA preparations in amounts easily detectable by biochemical procedures such as S1 nuclease mapping and RNA gel-transfer hybridization (19). Moreover, the 2.6-kb transcript can be directly detected in polyadenylate [poly(A)](+) RNA preparations after agarose gel electrophoresis by ethidium bromide staining (Green and Roeder, unpublished data). In contrast, unspliced nuclear RNAs (potential precursors) are often difficult to detect in other systems, presumably because unspliced substrates are spliced rapidly after synthesis. For example, as mentioned above, the three major parvovirus H1 RNAs are present almost exclusively (by mass) in spliced forms, including the largest nucleus-specific transcript which is 95% of the H1 genome in length (18).

The two different models suggested for the biosynthesis of AAV2 RNAs can be tested by several distinct experimental approaches. First, sequence analysis of the 5′ ends of the AAV2 RNAs may provide insight into the location, and possibly identity, of their promoters on viral DNA. Second, the number and location of the promoters can be determined by techniques such as nascent nuclear RNA mapping (14), and the size of the transcription unit(s) can be measured by determining the sensitivity of viral RNAs to UV irradiation (16). Third, recently

developed in vitro transcription systems (28, 41) that faithfully initiate RNA synthesis when programed with DNA containing a bona fide eucaryotic RNA polymerase II promoter sequence may provide strong indications of those regions of the AAV2 genome at which RNA synthesis is initiated in vivo.

If the AAV2 genome, in fact, acts as a single transcription unit, then AAV2 provides a model system to study transcription and RNA processing, since the viral RNAs are relatively simple in terms of number and structure and are present in sufficient quantity for detailed study. On the other hand, if the AAV2 genome contains multiple promoters, then AAV2 provides an interesting system to study a somewhat novel situation (in eucaryotes) in which transcription units with different promoters utilize common termination signals.

ACKNOWLEDGMENTS

This work was supported by American Cancer Society grant NP-284 and Public Health Service grants CA16640 and CA23615 to R.G.R. and cancer center support grant CA16217 to Washington University, the latter three from the National Cancer Institute. M.R.G. was supported by an insurance medical scientist scholarship.

LITERATURE CITED

1. **Alwine, J. C., D. J. Kemp, B. A. Parker, J. Reiser, J. Renart, G. R. Stark, and G. M. Wahl.** 1980. Detection of specific RNAs or specific fragments of DNA by fractionation in gels and transfer to diazobenzyloxymethyl-paper. Methods Enzymol. **68**:220–242.
2. **Berget, S. M., and P. A. Sharp.** 1979. Structure of late adenovirus 2 heterogeneous nuclear RNA. J. Mol. Biol. **129**:547–565.
3. **Berk, A. J., and P. A. Sharp.** 1978. Spliced early mRNAs of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**: 1274–1278.
4. **Berns, K. I., and W. W. Hauswirth.** 1979. Adeno-associated viruses. Adv. Virus Res. **25**:407–449.
5. **Berns, K. I., and J. A. Rose.** 1970. Evidence for a single-stranded adenovirus-associated virus genome: isolation and separation of complementary single strands. J. Virol. **5**:693–699.
6. **Buller, R. M. C., and J. A. Rose.** 1978. Characterization of adeno-associated-virus polypeptides synthesized in vivo and in vitro, p. 399–410. *In* D. C. Ward and P. Tattersall (ed.), The replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
7. **Carter, B. J.** 1978. Parovovirus transcription, p. 33–52. *In* D. C. Ward and P. Tattersall (ed.), The replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
8. **Carter, B. J., F. J. Koczot, J. Garrison, J. A. Rose, and R. Dolin.** 1973. Separate helper functions provided by adenovirus for adenovirus-associated virus multiplication. Nature (London) New Biol. **244**:71–73.
9. **Casey, J., and N. Davidson.** 1977. Rates of formation and thermal stabilities of RNA:DNA and DNA:DNA duplexes at high concentrations of formamide. Nucleic Acids Res. **4**:1539–1552.
10. **Chow, L. T., T. R. Broker, and J. B. Lewis.** 1979. Complex splicing patterns of RNAs from the early regions of adenovirus-2. J. Mol. Biol. **134**:265–303.
11. **Chow, L. T., R. E. Gelinas, T. R. Broker, and R. J. Roberts.** 1977. An amazing sequence arrangement at the 5′ ends of adenovirus-2 messenger RNA. Cell **12**:1–8.
12. **Darnell, J. E.** 1979. Transcription units for mRNA production in eukaryotic cells and their DNA viruses. Prog. Nucleic Acids Res. Mol. Biol. **22**:327–353.
13. **de la Maza, L. M., and B. J. Carter.** 1977. Adeno-associated virus DNA structure. Restriction endonuclease maps and arrangement of terminal sequences. Virology **82**:409–430.
14. **Evans, R. M., N. Fraser, E. Ziff, J. Weber, M. Wilson, and J. E. Darnell.** 1977. The initiation sites for RNA transcription in Ad2 DNA. Cell **12**:733–739.
15. **Gelinas, R. E., and R. J. Roberts.** 1977. One predominant 5′-undecanucleotide in adenovirus 2 late messenger RNAs. Cell **11**:533–544.
16. **Goldberg, S., J. Weber, and J. E. Darnell.** 1977. The definition of a large viral transcription unit late in Ad2 infection of Hela cells: mapping by effects of ultraviolet irradiation. Cell **10**:617–621.
17. **Green, M.** 1970. Oncogenic viruses. Annu. Rev. Biochem. **39**:701–756.
18. **Green, M. R., R. M. Lebovitz, and R. G. Roeder.** 1979. Expression of the autonomous parvovirus H1 genome: evidence for a single transcriptional unit and multiple spliced polyadenylated transcripts. Cell **17**:967–977.
19. **Green, M. R., S. E. Straus, and R. G. Roeder.** 1980. Transcripts of the adenovirus-associated virus genome: multiple polyadenylated RNAs including a potential primary transcript. J. Virol. **35**:560–565.
20. **Gruss, P., C. Lai, R. Dhar, and G. Khoury.** 1979. Splicing as a requirement for biogenesis of functional 16S mRNA of simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **76**:4317–4321.
21. **Hamer, D. H., and P. Leder.** 1979. Splicing and the formation of stable RNA. Cell **18**:1299–1302.
22. **Hayward, W. S.** 1977. Size and genetic content of viral RNAs in avian oncovirus-infected cells. J. Virol. **24**:47–63.
23. **Jeppesen, P. G. N.** 1974. A method for separating DNA fragments by electrophoresis in polyacrylamide concentration gradient slab gels. Anal. Biochem. **58**:195–207.
24. **Johnson, F. B., D. A. Vlazny, T. A. Thomson, P. A. Taylor, and M. D. Lubek.** 1978. Adeno-associated-virus polypeptides: molecular similarities, p. 411–421. *In* D. C. Ward and P. Tattersall (ed.), The replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
25. **Kassavetis, G. A., and M. J. Chamberlin.** 1979. Mapping of class II promoter sites utilized in vitro by T7-specific RNA polymerase on bacteriophage T7 DNA. J. Virol. **29**:196–208.
26. **Kennedy, S. I. T.** 1976. Sequence relationships between the genome and the intracellular RNA species of standard and defective-interfering Semliki Forest virus. J. Mol. Biol. **108**:491–511.
27. **Laughlin, C. A., H. Westphal, and B. J. Carter.** 1979. Spliced adenovirus-associated virus RNA. Proc. Natl. Acad. Sci. U.S.A. **76**:5567–5571.
28. **Luse, D. S., and R. G. Roeder.** 1980. Accurate initiation of transcription on a purified mouse β-globin DNA in a soluble system. Cell **20**:691–699.
29. **McGrogan, M., and H. J. Raskas.** 1978. Two regions of the adenovirus 2 genome specify families of late polysomal RNAs containing common sequences. Proc. Natl. Acad. Sci. U.S.A. **75**:625–629.
30. **Nevins, J. R., and J. E. Darnell.** 1978. Groups of adenovirus type 2 mRNAs derived from a large primary transcript: probable nuclear origin and possible common 3′ ends. J. Virol. **25**:811–823.
31. **Nevins, J. R., and J. E. Darnell.** 1978. Steps in the processing of Ad2 mRNA:poly(A)+nuclear sequences

are conserved and poly(A) addition precedes splicing. Cell **15**:1477–1493.

32. **Rose, J. A.** 1974. Parvovirus reproduction, p. 1–61. *In* H. Frankel-Conrat (ed.), Comprehensive virology, vol. 3. Plenum Publishing Corp., New York.
33. **Sanger, F., A. R. Coulson, T. Friedmann, G. M. Air, B. G. Barrell, N. L. Brown, J. C. Fiddes, C. A. Hutchinson, P. M. Slocombe, and M. Smith.** 1978. The nucleotide sequence of bacteriophage ϕX174. J. Mol. Biol. **125**:225–246.
34. **Shinagawa, M., and R. Padmanabhan.** 1979. Inhibition of a nuclease contaminant in the commercial preparations of *Escherichia coli* alkaline phosphatase. Anal. Biochem. **95**:458–464.
35. **Tal, J., D. Ron, P. Tattersall, S. Bratosin, and Y. Aloni.** 1979. About 30% of minute virus of mice RNA is spliced out following polyadenylation. Nature (London) **279**:649–651.
36. **Tsujimoto, Y., and Y. Suzuki.** 1979. The DNA sequence of *Bombyx mori* fibroin gene including the 5′ flanking, mRNA coding, entire intervening and fibroin protein coding regions. Cell **18**:591–600.
37. **Vogelstein, B., and D. Gillespie.** 1977. RNA-DNA hybridization in solution without DNA reannealing. Biochem. Biophys. Res. Commun. **75**:1127–1132.
38. **Walseth, T. F., and R. A. Johnson.** 1979. The enzymatic preparation of [^{32}P]nucleoside triphosphate, cyclic [α^{32}P]AMP, and cyclic [^{32}P]GMP. Biochim. Biophys. Acta **526**:11–31.
39. **Ward, D. C., and P. Tattersall (ed.).** 1978. The replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
40. **Weaver, R. F., and C. Weissmann.** 1979. Mapping of RNA by a modification of the Berk-Sharp procedure: the 5′ termini of 15S β-globin mRNA precursor and mature 10S β-globin mRNA have identical map coordinates. Nucleic Acids Res. **7**:1175–1193.
41. **Weil, P. A., D. S. Luse, J. Segall, and R. G. Roeder.** 1979. Selective and accurate initiation of transcription of the Ad2 major late promotor in a soluble system dependent on purified RNA polymerase II and DNA. Cell **18**:469–484.
42. **Ziff, E. B., and R. M. Evans.** 1978. Coincidence of the promoter and capped 5′ terminus of RNA from the adenovirus 2 major late transcription unit. Cell **15**: 1463–1475.