JOURNAL OF VIROLOGY, Oct. 1988, p. 3684–3694
0022-538X/88/103684-11$02.00/0
Copyright © 1988, American Society for Microbiology
Vol. 62, No. 10

Detailed Transcription Map of Aleutian Mink Disease Parvovirus

SOREN ALEXANDERSEN, MARSHALL E. BLOOM,* AND SYLVIA PERRYMAN

National Institute of Allergy and Infectious Diseases, Laboratory of Persistent Viral Diseases, Rocky Mountain Laboratories, Hamilton, Montana 59840

Received 28 March 1988/Accepted 10 June 1988

We studied the transcription program of Aleutian mink disease parvovirus (ADV) by using a combination of cDNA cloning and sequencing, primer extension, and Northern (RNA) blot hybridization with splice-specific oligonucleotides. The 4.8-kilobase ADV genome was transcribed in the rightward direction, yielding plus-sense polyadenylated transcripts of 4.3 (R1 RNA), 2.8 (R2), 2.8 (R3), 1.1 (RX), and 0.85 (R2′) kilobases. Each RNA transcript had potential translation initiation sites within open reading frames, suggesting protein translation, and a scheme encompassing ADV structural and nonstructural proteins is proposed. Each of the five RNA transcripts had a characteristic set of splices and originated from a promoter at nucleotide 152 (map unit 3 [R1, R2, R2′, and RX]) or at nucleotide 1729 (map unit 36 [R3]). The transcripts terminated with a poly(A) tail at one of two positions: either at map unit 53 (R2′ and RX) or at map unit 92 (R1, R2, and R3). Similarities with and differences from the transcription maps of other parvoviruses are discussed, and possible roles of the unique features found in ADV transcription are related to the special pathogenic features of this virus.

Aleutian mink disease parvovirus (ADV) is an autonomous parvovirus (14) which causes a number of interesting clinical and pathological syndromes in mink. These syndromes span from abortion (1, 25, 26, 39) through acute pneumonia in neonatal mink (1, 2, 4, 5) to chronic immune complex-mediated glomerulonephritis and arteritis in adult mink (1, 3, 27, 41–43). Previous studies have shown that acute disease is caused by a primary cytopathic injury of infected cells associated with permissive high-level replication and expression of the viral genome, whereas, in contrast, the chronic disease is characterized by restricted low-level viral replication and expression leading to persistent infection and immune-complex disease (3, 4). Since the input viral genome can be the same in each case, the differences in viral gene expression may reflect differential interaction of various viral and cellular gene products, and, furthermore, this regulation may determine the outcome of infection. Analysis of viral transcription is an obvious starting point for comparing viral gene expression in permissive and nonpermissive systems, and definition of the ADV transcription program in a permissive cell culture system will provide a framework on which in vivo comparisons may be made.

We recently sequenced the ADV genome and showed that ADV had the same general genetic organization as other autonomously replicating parvoviruses (11). This general parvovirus organization includes an approximately 5,000-base single-stranded DNA genome of minus polarity with terminal hairpin structures. The transcription program uses overlapping transcripts and multiple reading frames to increase the genetic information included in the short genome. Two or three viral structural proteins are encoded by the right side of the genome by spliced RNA transcripts starting from a promoter around map unit (mu) 39. Several nonstructural proteins are specified by spliced transcripts from the left side of the genome using a promoter around mu 4. Because these nonstructural proteins are presumably involved in viral replication and in regulation of viral gene expression, it may be especially interesting to study these proteins and the transcripts from which they are encoded during ADV infection.

In the present report we define the transcription scheme of ADV-G in cell culture and suggest a model for the translation of these RNA transcripts into viral proteins. This model is compared with the transcriptional and translational scheme of related parvoviruses, and features which might explain the unique pathogenic properties of ADV are discussed.

MATERIALS AND METHODS

Materials. All enzymes, isotopes, and other chemicals were obtained from sources previously listed (3, 4, 11).

Preparation and analysis of RNA. Total cellular RNA was prepared from control and ADV-G infected CRFK cells (14, 23) by using the guadinium thiocyanate method (33). Following ethanol precipitation, the RNA was dissolved in distilled water at 40 μg/ml, and contaminating DNA was digested by adding $MgCl_2$ to 10 mM, Tris hydrochloride (pH 7.9) to 40 mM, dithiothreitol to 5 mM, RNasin (Promega Biotec) to 400 U/ml, and RNase-free DNase (Promega Biotec) to 20 U/ml and incubating the mixture at 37°C for 30 min. This DNase step was necessary to avoid contamination of the final RNAs with virion DNA. Samples were then extracted with phenol and chloroform and ethanol precipitated. Poly(A)$^+$ RNA (mRNA) was purified from total RNA on an oligo(dT)-cellulose column as described (20).

For Northern (RNA) blot analysis, samples were electrophoresed into a 0.7, 1, or 1.5% agarose gel containing 6% formaldehyde and blotted (20) onto nylon membranes (Hybond-N; Amersham Corp.), and the RNA was fixed by baking. Selected filters were hybridized either with a 15- to 88-mu ADV probe (12) nick translated with [^{32}P]dCTP as previously described (13), with strand-specific RNA probes as previously described (2, 16), or with specific oligonucleotides as described below. An RNA size marker was included in each blot and consisted of 1 μg of a bacteriophage lambda RNA ladder (Bethesda Research Laboratories, Inc.). This part of the filter was probed with nick-translated bacteriophage lambda DNA (13).

Construction of cDNA libraries. cDNA libraries were constructed by using poly(A)$^+$ RNA from ADV-infected CRFK cells and a cDNA synthesis kit from Bethesda Research

* Corresponding author.

Laboratories. This system uses a cloned Moloney murine leukemia virus reverse transcriptase and involves the use of RNase H in the second-strand reaction. This strategy made it easy to obtain blunt-ended cDNAs, making subsequent cloning with linkers very efficient, although sequence information at the extreme 5′ end of the RNA is inevitably lost. Two libraries were constructed. Library 1 was prepared from 5 μg of poly(A)$^+$ RNA purified from infected cells at 24 h postinfection. cDNA was synthesized from this RNA with oligo(dT) as the primer, and the synthesized DNA was methylated with *Eco*RI methylase, filled in with the Klenow fragment of DNA polymerase, ligated to *Eco*RI 8-mer linkers, cloned into the *Eco*RI site of λgt11, packaged, and used to infect *Escherichia coli* Y1088 (20). Plaques were screened by blotting onto nitrocellulose filters and were probed with a 15- to 88-mu ADV probe nick translated with [^{32}P]dCTP (20). This library resulted in 2×10^4 independent plaques, of which 69 (0.35%) were ADV specific. Selected ADV-positive plaques were picked and purified by replating twice. Bacteriophage DNA was isolated from positive phages by the Lambdasorb solid-phase immunoadsorbent method (Promega Biotec). Initially, five of these clones were nucleotide sequenced, but we were unable to obtain long stretches of readable sequence, and ultimately ADV inserts were subcloned into the *Eco*RI site of plasmid pUC8.

The percentage of ADV-positive clones in library 1 was relatively high, and we therefore prepared a second library in a plasmid vector. We also wanted to use RNA isolated during late infection to pick up cDNAs representing late transcription products. Thus, library 2 was constructed from 10 μg of poly(A)$^+$ RNA from infected cells harvested 40 h postinfection. The cDNAs from this reaction were ligated to *Xba*I 8-mer linkers and cloned into the *Xba*I site of plasmid p-GEM3 (Promega Biotec). The ligation mixtures were then used to transform competent *E. coli* JM109, ED8654, and DB1256 (11, 16, 24). This library resulted in 3×10^4 colonies, of which 600 (2%) were ADV specific. Ten of these clones were nucleotide sequenced.

Hybridization with specific oligonucleotides. Oligonucleotides, 20 to 22 bases in length, based on either the published sequence of ADV-G (11) or the sequence information derived from the cDNAs sequenced in the present paper, were prepared as noted (11). For hybridization with oligonucleotides, filters were prehybridized at 30°C for 3 h in 6× SSPE (33) containing 0.1% sodium dodecyl sulfate, 5× Denhardt solution, and 100 μg of sheared and denatured salmon sperm DNA per ml. Oligonucleotides were 5′ end labeled with ^{32}P (20) and allowed to hybridize to the filters in the same solution at 2 ng/ml (approximately 10^6 cpm/ml) at 30°C for 16 h. The filters were washed twice in 2× SSPE–0.1% sodium dodecyl sulfate for 30 min at 37 to 48°C according to the calculated T_m of the oligonucleotides (33, 36, 50) and were exposed to film at −70°C with a Cronex Lightning-Plus (E. I. du Pont de Nemours & Co., Inc.) intensifying screen.

Sequencing. DNA nucleotide sequencing was done by using the dideoxynucleotide chain termination method of Sanger et al. (46) with commercially available primers or ADV-specific oligonucleotide primers as described previously (11).

Primer extension. Poly(A)$^+$ RNA isolated from control cells and from cells harvested 24 or 40 h after infection was hybridized to ^{32}P-labeled oligonucleotide primers and extended with reverse transcriptase essentially as described (36). Extended primers were then suspended in 90% formamide, denatured by heating, and electrophoresed on 8% polyacrylamide–7 M urea sequencing gels. φX174 DNA cut with *Hae*III and 5′ end labeled with ^{32}P (20) as well as a standard sequencing reaction were electrophoresed in parallel as size markers.

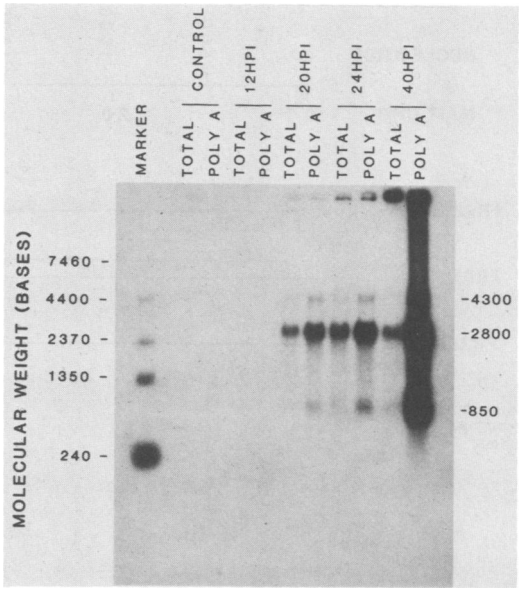

FIG. 1. Northern blot analysis of RNA isolated from CRFK cell cultures. Whole-cell RNA and poly(A) RNA (mRNA) from uninfected (control) and ADV-G-infected CRFK cells were extracted as described in Materials and Methods, and 1-μg samples were electrophoresed into a formaldehyde-agarose gel and Northern blotted onto a nylon filter. The filter was probed with an ADV-specific probe as described in Materials and Methods. The time after infection and type of RNA are indicated. An RNA size marker is included in the blot (left), and the sizes of the RNAs are indicated in bases.

RESULTS

Northern blot analysis of RNA. Our initial step in analyzing the transcription pattern of ADV infection in CRFK cells was to characterize the sizes of the mRNAs. When total and poly(A)$^+$ containing RNA were examined by Northern blotting, ADV-specific RNA was not detected in uninfected cultures or in cell cultures harvested 12 h after infection. At 20 h after infection and later, three size classes of ADV RNA could be detected (Fig. 1) in ADV-G-infected CRFK cells; 0.85, 2.8, and 4.3 kilobases (kb). All three bands were detected at each time point, and the relative abundances remained the same. The same three bands were noted in both total RNA and poly(A)-selected mRNA. Furthermore, the relative amount of ADV-specific mRNA increased dramatically with time after infection, suggesting preferential transcription of ADV mRNA. Northern blots were also hybridized to strand-specific RNA probes (2, 16); only RNAs hybridizing to the minus-sense probe could be detected (data not shown), indicating that all detectable ADV RNA is plus in sense.

Analysis of ADV RNA by cDNA cloning. Two cDNA libraries were prepared from poly(A)-selected mRNA isolated from ADV-infected CRFK cells. Of more than 600 ADV-positive clones, a total of 20 cDNA clones were nucleotide sequenced. The sequence analysis provided data on two aspects of ADV transcription. First, several consistent patterns of molecules were observed in cDNA libraries. Second, only four different splices were identified on the basis of juxtaposition of bases that were noncontiguous in

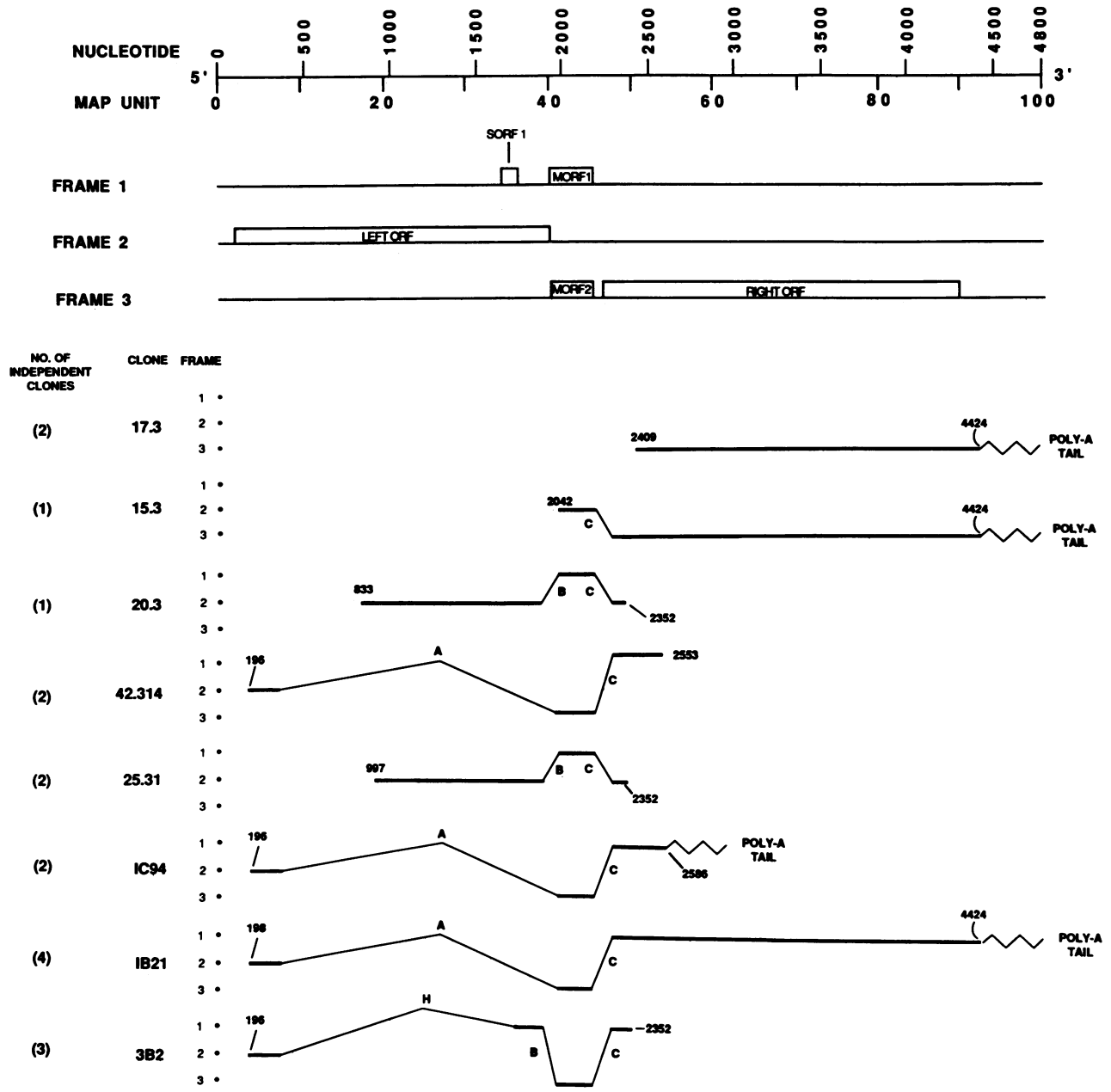

FIG. 2. Nucleotide sequence map of selected cDNAs. A diagram showing selected cDNA clones was prepared and aligned on a schematic indicating the major open reading frames of ADV-G. The number of independent cDNA clones having a particular structure is indicated. ▬, Actual sequences; ——, spliced introns; ∿∿, poly(A) tails. The nucleotide numbers at the 5′ end (left) and 3′ end (right) of the clones, as well as the designation of introns, are indicated. The span of the introns was as follows: intron A, nucleotides 385 to 2041; intron B, nucleotides 1962 to 2041; intron C, nucleotides 2214 to 2286; intron H, nucleotides 385 to 1736. For illustration purposes, the clones are arranged to fit the open reading frames shown at top. cDNA clones 17.3, 15.3, and 20.3 are in λgt11; 42.314 and 25.31 are lambda clones subcloned into pUC8; and IC94, IB21, and 3B2 are cDNA clones in pGEM-3.

the genomic sequence. Representative cDNA clones are depicted in Fig. 2. The nucleotide numbering is based on the ADV genomic sequence (11). To indicate shifts in potential translational frames caused by splicing, we have arranged the sequences to fit into probable open reading frame (see below).

A large number of clones (seven) had poly(A) tails (10 to 20 bases long) originating from ADV nucleotide 4424. Five of these clones (represented by 15.3 and IB21) extended uninterrupted for 2,138 bases in the left-hand direction to nucleotide 2287, at which point nucleotides 2214 to 2286 had been spliced out and a 171-base segment (nucleotides 2042 to 2213) was juxtaposed. This splice was designated splice C and was found in all cDNA clones spanning this region (15 clones). To the left of splice C, three different splices were found. Four clones (represented by 20.3 and 25.31) had nucleotide 2042 juxtaposed to nucleotide 1961, splicing out nucleotides 2041 to 1962. This splice was designated splice B. Three other clones (represented by 3B2) also had splice B, but had, in addition, nucleotide 1737 juxtaposed to

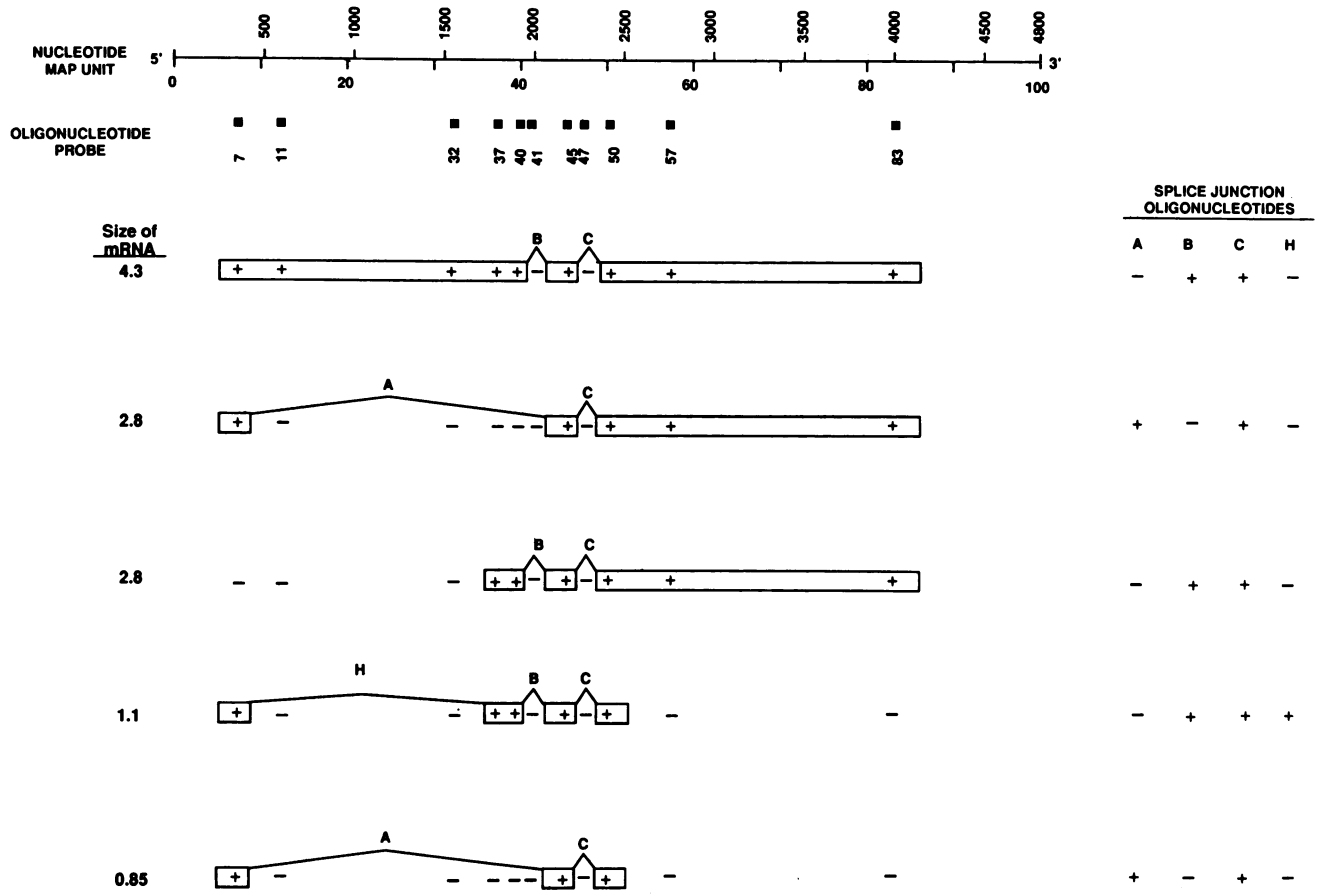

FIG. 3. Detailed oligonucleotide mapping of ADV-G mRNAs. Replicate Northern blots were probed with a series of 5'-end-labeled oligonucleotides specific for either the mRNA splices or positions on the ADV-G genome. The splices are designated A, B, C, and H, and the splice found in each mRNA is indicated. The positive (+) or negative (−) reaction of ADV mRNA species with oligonucleotides is indicated. The exact nucleotide coordinates of the oligonucleotides used as probes for genomic localization are as follows: 7 (323 to 342), 11 (523 to 542), 32 (1553 to 1572), 37 (1753 to 1772), 40 (1931 to 1950), 41 (1964 to 1983), 45 (2185 to 2204), 47 (2230 to 2249), 50 (2395 to 2414), 57 (2733 to 2752), and 83 (3953 to 3972). The oligonucleotides specific for the splice junctions were constructed to contain 10 or 11 bases on each side of the splice junctions as follows: oligonucleotide A (374 to 384 plus 2042 to 2052), oligonucleotide B (1952 to 1961 plus 2042 to 2051), oligonucleotide C (2204 to 2213 plus 2287 to 2296), and oligonucleotide H (375 to 384 plus 1737 to 1746).

nucleotide 384 (splice H). Eight clones (represented by 42.314, IC94, and IB21) had nucleotide 2042 juxtaposed to nucleotide 384 (splice A). All 11 clones having either splice A or splice H had their 5' ends at nucleotides 196 to 198.

One other position of polyadenylation other than the one at nucleotide 4424 was found. Two clones (represented by IC94) had a poly(A) tail of 10 to 20 bases originating from ADV nucleotide 2586 (Fig. 2).

Six clones (represented by 20.3, 25.31, and 3B2) had no poly(A) tail, and their 3' ends were at nucleotide 2352. These clones were probably derived from RNAs at which the oligo(dT) inadvertently had primed at the A-rich stretch around nucleotide 2352 (13 adenine residues of 17 total).

Detailed Northern mapping. The results obtained from cDNA analysis suggested that there were four types of splices in the various ADV mRNA molecules. However, because most of the cDNAs probably did not reflect complete mRNA molecules, we could not unambiguously determine which splices were in which sizes of mRNA. Consequently, we constructed a series of oligonucleotides specific for either the four splice junctions or different short segments of the ADV genome and used them to probe Northern blots. A schematic representation of the reaction of Northern blots probed with the different oligonucleotides is depicted in Fig. 3, and the actual reaction of Northern blots with selected oligonucleotides is shown in Fig. 4. Although three bands were detected on Northern blot analysis, these experiments led us to conclude that there are a total of five separate RNA species. We have designated these RNAs according to the convention for the prototype parvovirus, minute virus of mice (MVM) (6, 9, 19, 21, 22, 28, 32, 35, 36, 40, 47).

0.85-kb RNA. The 0.85-kb band consisted of a single RNA species, in this paper designated R2', having splice pattern A and C, a 5' end upstream of nucleotide 323, and a 3' end between nucleotides 2414 and 2733.

2.8-kb RNAs. The probe reaction on the 2.8-kb RNA band could be explained only by the existence of two different RNA species of the same size. One species, designated R2, presumably had the same 5' end and A,C splice pattern as R2', but extended throughout the right-hand end of the genome. Another species also banded at 2.8 kb but was different from R2. This species was designated R3 and apparently had the same 3' end as R2 but had the splice pattern B,C (reacted with oligonucleotide B and C). The size

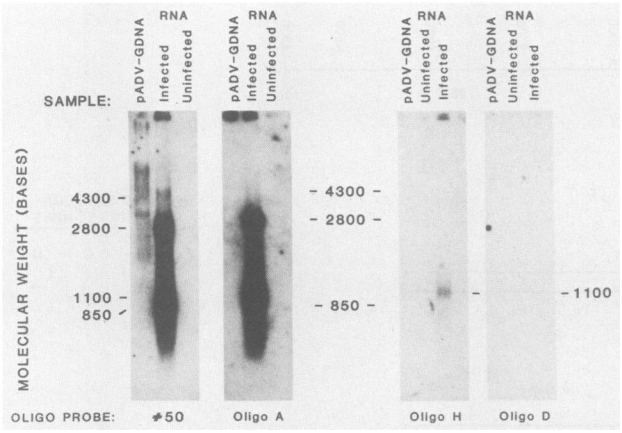

FIG. 4. Selected Northern blots performed with using specific oligonucleotides. Poly(A)-selected RNA from ADV-G-infected and uninfected CRFK cells was electrophoresed into formaldehyde-agarose gels and probed with specific oligonucleotides as described in Materials and Methods and in the legend to Fig. 3. A purified 15- to 88-mu ADV DNA fragment (pADV-G DNA) was included in each blot as a positive control for oligonucleotides spanning this segment of the genome and as a negative control for the splice-specific oligonucleotides. The sizes of the RNAs and the oligonucleotides used as probes are indicated. For details on probe designations, see Fig. 3. Oligonucleotide D was constructed to span a hypothetical splice junction (nucleotides 2162 to 2172 + 2287 to 2297).

(2.8 kb) and this B,C splice pattern indicated a 5′ end somewhere between nucleotides 1500 and 2000.

4.3-kb RNA. The 4.3-kb RNA (designated R1) apparently consisted of a single RNA species having a 5′ end upstream of nucleotide 323 and having the same B,C splice pattern as R3 and a 3′ end in the right-hand end of the genome.

1.1-kb RNA. When Northern blots were hybridized with oligonucleotide H (specific for splice H), a fifth RNA could be detected (Fig. 4), and this RNA was designated RX, since no RNAs with such a structure has been described for the other parvoviruses. The RX RNA was a 1.1-kb species with splice pattern B,C,H, a 5′ end upstream of nucleotide 323, and a 3′ end around nucleotide 2500 (Fig. 3).

Interestingly, when Northern blots were probed with oligonucleotide 50, which binds to all the RNA species, a different relative intensity of the bands was observed compared with those in blots hybridized with the 15- to 88-mu probe (Fig. 1 and 4). With oligonucleotide 50 the 4.3-kb RNA was faint (<5%), the 2.8-kb RNAs were more abundant (35 to 45%), and the 0.85-kb RNA was very abundant (50 to 60%). Close inspection of the blot showed that the 1.1-kb RX RNA was present as a faint band (<5%). This apparent discrepancy probably occurs because the 20-nucleotide oligonucleotide hybridizes stoichiometrically to specific RNA molecules independent of size, whereas the 3.5-kb partial genomic probe favors large RNAs.

Primer extension studies. As discussed above, the precise 5′ boundaries of the mRNAs cannot be inferred from cDNA analysis or Northern blotting. Therefore, we used primer extension studies with oligonucleotides whose domains were close to the presumed 5′ ends of the mRNAs. The extended products were separated in polyacrylamide sequencing gels to obtain superior resolution of relatively small products. The results of these experiments are depicted in Fig. 5 and 6. When the data obtained from cDNA sequencing and RNA mapping were correlated with the length of the extended products, the 5′ ends of the mRNAs could be estimated to an accuracy of ±1 nucleotide. When oligonucleotide 7 (spanning nucleotides 323 to 342 and annealing to the R1, R2, R2′, and RX RNAs) was used as the primer, a major doublet band of 162 and 163 bases was found (Fig. 6A). This extended-product size indicated that the 5′ end of the RNAs was at nucleotide 180 ± 1 (Fig. 5). Minor bands were also evident; however, several of these were also present in extended RNA from uninfected control cells, and smaller bands than those of the major products might represent premature termination of the reverse transcriptase. When oligonucleotide 45 (spanning nucleotides 2185 to 2204) was used, major bands of 371 and 365 ± 10 nucleotides were found, together with a minor band of 620 ± 30 nucleotides. These data also would suggest a 5′ end around 180 for an RNA with splice pattern A,B (365-nucleotide product) or splice pattern B,H (620-nucleotide product). A product of 371 nucleotides indicates a 5′ end around nucleotide 1750 for an RNA having splice pattern B (Fig. 5), consistent with the mapping data.

Primer extension was then done with oligo(A) as a primer. This experiment resulted in a major extended product of 216 ± 1 nucleotides (Fig. 6B). Oligo(A) reacts only with the 2.8-kb R2 and the 0.85-kb R2′ RNAs, indicating a 5′ end at nucleotide 180 for both of these two mRNAs. Finally, the RNA was extended with oligonucleotide 40 (spanning nucleotides 1931 to 1950) as a primer. This resulted in two major products of 207 ± 1 and 420 ± 10 nucleotides (Fig. 6C). This oligonucleotide hybridizes to the 4.3-kb R1, the 2.8-kb R3, and the 1.1-kb RX RNA, indicating a 5′ end around nucleotide 180 for the RX mRNA (420-nucleotide product) and at nucleotide 1744 ± 1 for the R3 mRNA (207-nucleotide product) (Fig. 5). The expected product for the R1 RNA would be around 1.7 kb, which is too big to be detected by this technique.

No difference in the pattern of extended products could be found between RNA isolated 24 or 40 h after infection, although extended products were much more abundant at 40 h postinfection (data not shown).

Search for alternative splice patterns. In MVM, which is the parvovirus whose transcription pattern has been the most intensively investigated, alternative splice patterns have been implied (6, 28, 36, 40). No evidence for such a mechanism was found in ADV transcription. All cDNAs sequenced had a definite set of splices (either splice A plus C or splice B plus C or splice H plus B plus C), and primer extension studies did not indicate alternative splices in the areas analyzed. Furthermore, Northern blotting with oligonucleotides specific for three alternative splice patterns were all negative (Fig. 4; data not shown), indicating that, if present, such RNA species are of very low abundance. The possibility of RNAs of very low abundance and with a different splice pattern cannot be totally excluded. However, a major role for such RNAs in the production of significant amounts of viral structural proteins as described for MVM is very unlikely to occur in ADV transcription and translation.

DISCUSSION

Previous studies, involving in situ hybridization (3, 4), have suggested that ADV transcription is modulated during in vivo persistent infection, and we have suggested that restricted transcription might be involved in the development of persistent infection and chronic immune dysfunctions (3). However, to verify this hypothesis, we must define ADV transcription under defined conditions of infection, and we have chosen to do this during ADV-G infection of cultured CRFK cells. Once this study is completed, we can

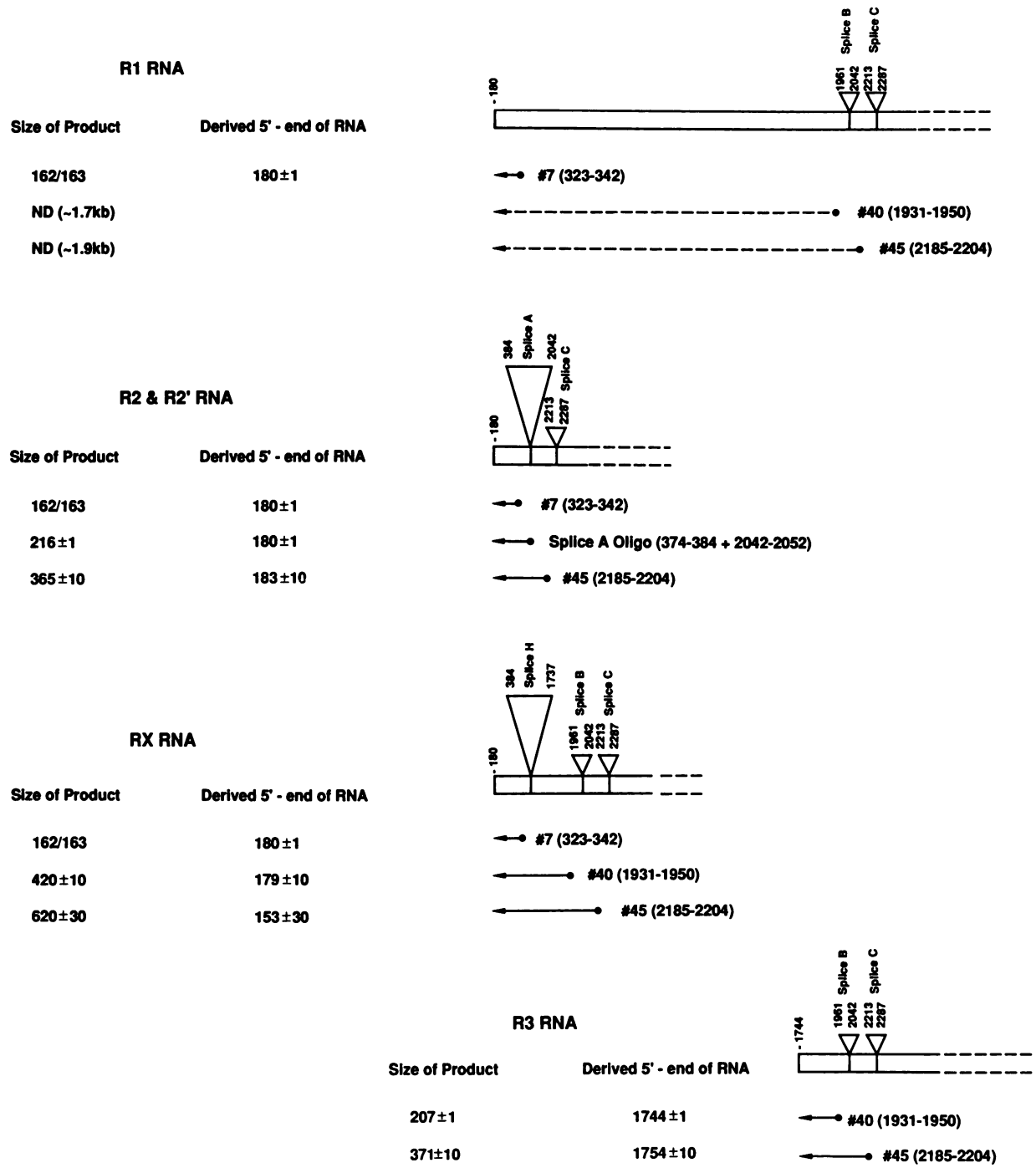

FIG. 5. Summary diagram of the primer extension experiments. Primer extension was performed on poly(A) mRNA from ADV-G-infected CRFK cells. The length of the major extended products was determined by gel analysis, and the presumed 5′ ends of the RNAs were calculated. The results of these experiments are schematically depicted. The 5′ ends of the different RNAs with their characteristic splices, the oligonucleotides used for primers, the length of extended products, and the derived 5′ ends of the RNAs are indicated. Abbreviation: ND, not detected.

begin to analyze in vivo ADV transcription in different experimental systems.

On the basis of the data presented in the present report, a model for ADV transcription and translation can be proposed (Fig. 7). The model accounts for the RNAs seen in Northern blot analysis, as well as the known structural and nonstructural proteins of ADV, and, furthermore, indicates a possibility for the translation of one hitherto unrecognized nonstructural protein, tentatively designated NS-3. The nucleotide sequences of selected areas around initiation sites, splices, polyadenylation sites, and putative promoters are shown in Fig. 8. In this section, we will discuss the various mRNAs and their probable gene products.

R1. The size of R1 RNA is 4.3 kb in Northern blot analysis

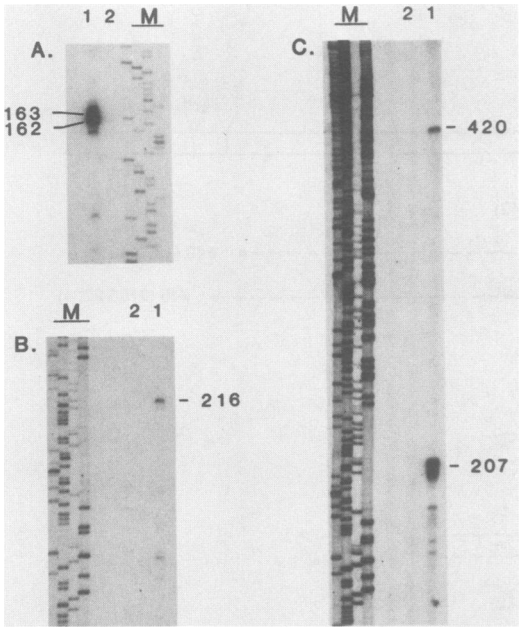

FIG. 6. Primer extension analysis on poly(A) RNA from infected (lane 1) and uninfected (lane 2) CRFK cells with oligonucleotide primers as described in Materials and Methods and in the legend to Fig. 5. A standard sequencing reaction (lanes M) was included in each electrophoresis and served as a precise size marker. The sizes of the extended products are indicated in nucleotides. (A) Oligonucleotide 7 (323 to 342) as primer; (B) oligonucleotide A (specific for splice A) as primer; (C) oligonucleotide 40 (1931 to 1950) as primer.

and 4.1 kb as calculated from cDNA and primer extension data. This indicates a poly(A) tail of about 200 bases, consistent with the size of the poly(A) tail reported for other parvoviruses (19). From primer extension data, the 5′ end of this RNA is at nucleotide 180 ± 1. This location is consistent with a promoter (GTATATAAGC) at nucleotides 151 to 160 (−29 from initiation of transcription [Fig. 8]) and a CAT enhancer motif at nucleotides 97 to 99 (−83 from initiation of transcription [Fig. 8]). This location of a promoter and upstream enhancer element conforms to the published consensus for initiation of transcription (8, 29–31). Moreover, this ADV promotor (nucleotides 151 to 160) is identical to the putative Kilham rat parvovirus P4 promoter and differs by only one nucleotide from the adeno-associated virus P5 promoter (19). Initiation of transcription at nucleotide 179 or 180 would be consistent with the most common initiation at a purine (19), whereas initiation at nucleotide 181 would indicate initiation at a pyrimidine, which has been observed for certain cellular and viral mRNAs (19) and has also been indicated for parvovirus RNA (19). RNA mapping and cDNA sequencing show that the R1 RNA has a 3′ end at nucleotide 4424. This is consistent with a poly(A) signal (AATAAA) at nucleotides 4394 to 4399 and the general rule of poly(A) addition at a CA, GA, or UA 11 to 30 bases downstream of the polyadenylation signal (10, 34). One cDNA clone had a poly(A) tail originating from the middle of the genome (data not shown), suggesting that in some instances this RNA polyadenylates after the poly(A) signal at nucleotide 2564, as described below for the R2′ mRNA. However, an mRNA of the predicted size of such a short R1 RNA (around 2.4 kb) could not be identified in Northern blots, suggesting that such an mRNA is of low abundance.

From oligonucleotide mapping and sequence analysis data, the R1 RNA can unambiguously be shown to have two splices, giving the final structure shown in Fig. 7. Both introns have the canonical GT at their 5′ ends (donor site) and AG at their 3′ ends (acceptor site), and the base composition around the donor and acceptor sites correlates well with the established consensus (37). On the basis of the structure of this RNA, the large nonstructural protein (NS-1 or p71) is encoded from the left-hand open reading frame (M. E. Bloom, S. Alexandersen, and J. B. Wolfinbarger, Abstr. II Parvovirus Workshop, abstr. no. MA9, p. 32, 1987), and by analogy to the other parvoviruses (18, 21, 22), the R1 RNA most probably codes for the 71-kilodalton (kDa) NS-1 nonstructural protein described for ADV (5, 15). Translation of this protein probably initiates at the ATG in reading frame 2 at nucleotides 206 to 208. This ATG is in a very favorable context (ANNATGG) for initiation of translation (30, 31) and is followed by an open reading frame all the way to the splice at nucleotide 1961, at which point R1 splices to reading frame 1 at nucleotide 2042. Translation of NS-1 then may proceed in reading frame 1, which is open until nucleotide 2208, where the translation terminates. This gives an estimated molecular mass for NS-1 of 72,500 Da, which is in agreement with the experimental value of 71,000 Da for the ADV NS-1 protein p71 (5, 15).

R3. The R3 RNA is a 2.8-kb RNA, including the poly(A) tail. This RNA initiates at nucleotide 1744 ± 1 (Fig. 7), which correlates with a promoterlike sequence (TATTAAA) at nucleotides 1729 to 1735 (−15 from initiation of transcription) and an enhancer CCATT motif at nucleotides 1674 to 1678 (−70 from initiation of transcription [Fig. 8]). Initiation at this position, mu 36, correlates well with a P38 promoter described for the other parvoviruses (9, 17, 19, 21, 22, 28, 32, 40, 45), although initiation of transcription only 15 bases downstream of the putative promoter is unusual for eucaryotic transcription (8, 19, 29–31). However, the predicted site of initiation has the characteristics of a cap box in which six of nine nucleotides are identical to the adeno-associated virus P40 cap structure (19), further suggesting this site as a bona fide site of ADV P36 initiation of transcription. Translation would initiate on R3 at the ATG in reading frame 2 at nucleotides 2204 to 2206. This would give a potential protein of 3 amino acids in reading frame 2 spliced to 686 amino acids in reading frame 3 and probably terminating at nucleotide 4346. The resulting protein of 689 amino acids would have an estimated molecular mass of 78,500 Da, in fairly good agreement with an experimental value of 85,000 Da for the ADV VP-1 protein p85 (5, 15). Since there was no evidence for other mRNAs coding for the other known ADV structural protein (VP-2 or p75) (5, 15), it is most likely that the R3 RNA also codes for VP-2. Initiation of translation at the ATG at nucleotides 2406 to 2408 in reading frame 3 would give a protein of 646 amino acids with an estimated molecular mass of 73,500 Da, in good agreement with the experimental value of 75,000 Da for the ADV VP-2 structural protein, p75 (5, 15). The ATG at nucleotide 2421 (VP-2) is in a much more favorable context (GNNATGG) (30, 31) than the ATG at nucleotide 2204 (VP-1), which could explain how two different proteins can be translated from the same mRNA, and, furthermore, this provides a reasonable mechanism for the observed constant higher ratio of p75 to p85 being produced (15). This mechanism would rely on the strength of the individual initiators of translation as the control of viral protein abundance, as opposed to a splice control mechanism proposed for MVM (36). In the present study no evidence of such a splice mechanism involving alternative splices around the initiation of translation of the large structural protein (VP-1) was found for ADV transcrip-

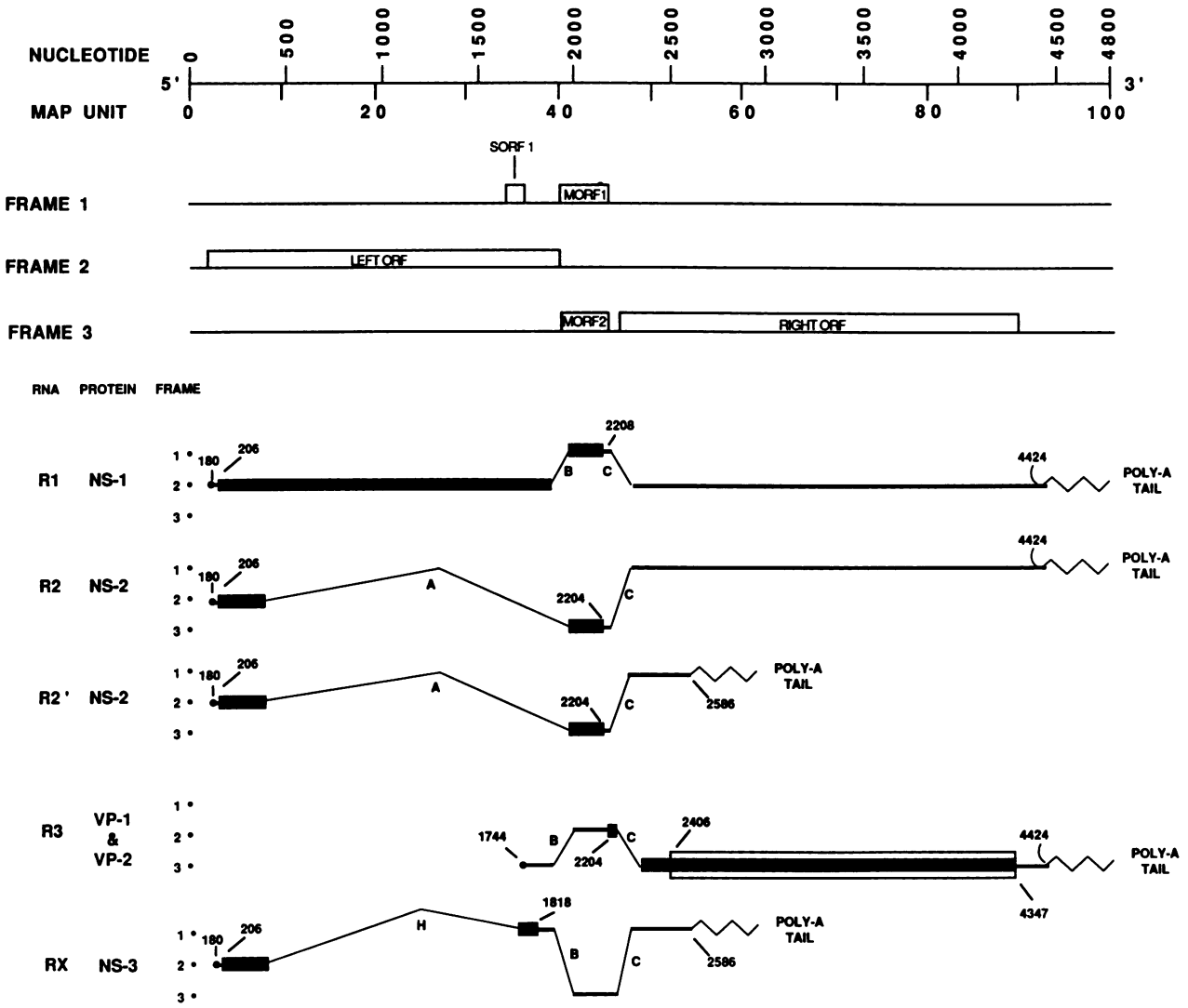

FIG. 7. Transcriptional and translational schemes of ADV-G. The predicted structures of the five mRNAs and four hypothetical protein products are diagrammed. The upper portion of the figure depicts the five open reading frames used in ADV-G. The predicted structure of the mRNAs is shown with the 5′ end at the left. Initiation of transcription and translation is indicated in nucleotides. Symbols: ——, RNA sequence; ∿∿, spliced introns; ▬, ▭, open reading frames presumably used in translation of the RNAs: the R3 mRNA probably codes for both VP-1 (▬) and VP-2 (▭). Nucleotide numbers at the start and end of RNA transcription and putative translation are indicated, as is the splice designation.

tion. A mechanism of initiation at two different ATGs at the same mRNA, as suggested here for ADV, has previously been suggested for AAV (7) and has also been observed for other viruses; it has been explained by a mechanism of leaky scanning by the ribosomes on the RNA (31).

About 20 amino acids downstream of the supposed initiation of the VP-2 protein is a glycine-rich stretch of 23 amino acids (containing 19 glycines). A similar motif has been described for the other parvovirus VP-2s and may reinforce the probability of VP-2 initiation at this point.

A third structural protein has been described for the other autonomous parvoviruses, but not for ADV. The VP-3 for the other parvoviruses is derived by proteolytic cleavage of VP-2 at an arginine residue immediately upstream of the glycine-rich stretch (19, 22), and perhaps ADV does not have such a VP-3 because in ADV the glycine-rich stretch is preceded by an alanine. If the protein is processed by a trypsinlike protease, such an enzyme would not cut in ADV, because trypsin cuts only at the carbonyl groups of arginine and lysine, explaining the absence of an ADV VP-3.

R2. R2 is a 2.8-kb RNA including a poly(A) tail of approximately 200 bases. The 5′ end of this RNA is similar to the R1 RNA at nucleotide 180, and the 3′ end is also similar to the R1 RNA, i.e., nucleotide 4424. The R2 RNA has splice pattern A,C, consisting of a big splice covering nucleotides 385 to 2041 (A) and a small splice similar to the second R1 splice covering nucleotides 2214 to 2286 (C). Both splice donors and acceptors are in agreement with established consensus (37). The protein encoded by this RNA is probably NS-2, the second parvoviral nonstructural protein (21, 22). This protein most probably initiates as NS-1 in reading frame 2 at nucleotide 206 and reads in frame to nucleotide 384, where it splices to reading frame 3 at nucleotide 2042. This reading frame is open until nucleotide

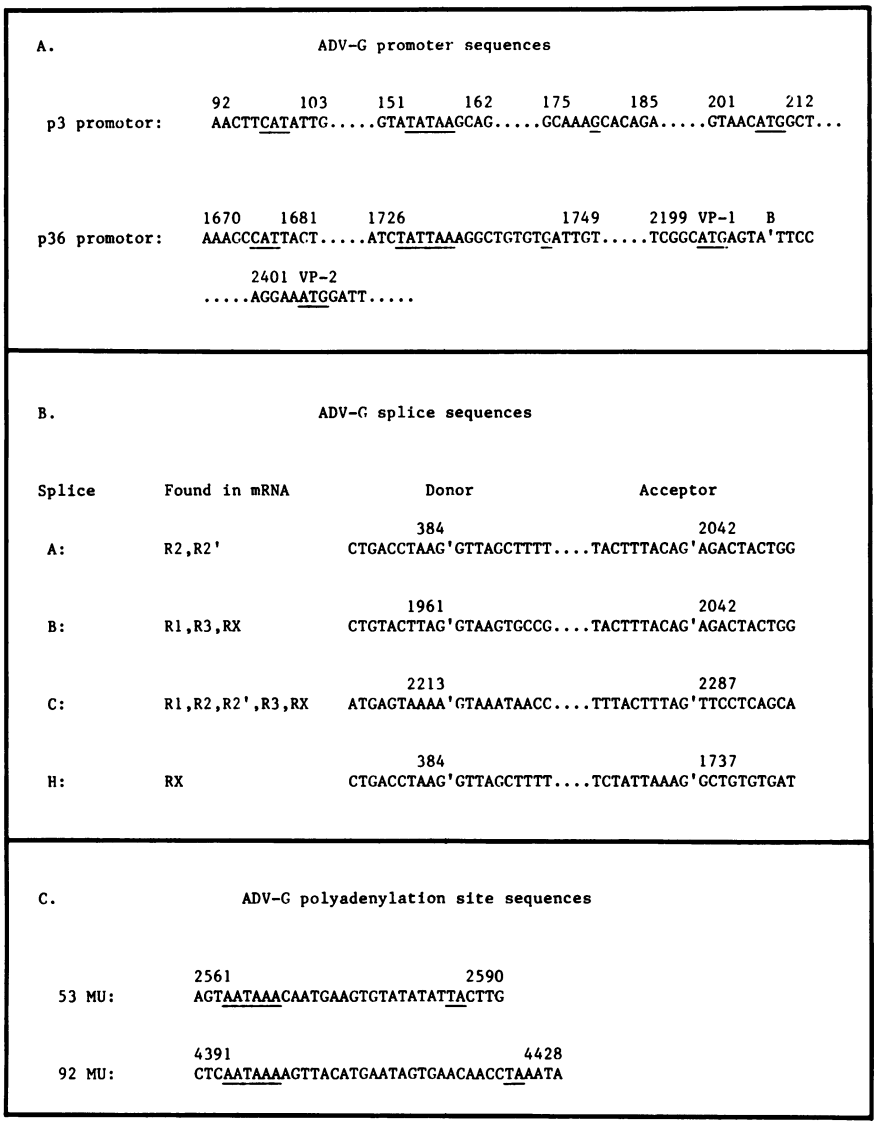

FIG. 8. Selected areas of ADV-G nucleotide sequence in the regions of promoters, initiation sites, splices, and polyadenylation sites. The nucleotide sequence of the ADV-G genome and the numbering are derived from Bloom et al. (11). (A) The position and nucleotide sequence of the P3(TATAA) and the p36 (TATTAA) promoters with upstream enhancer elements (CAT), as well as initiation sites of transcription (G) and translation (ATG), are shown. For the P36 promoter, the position of the B splice ('), as well as the presumed site of initiation of VP-1 and VP-2, is depicted. (B) The nucleotide sequence around the splices is listed. The actual splice sites are depicted ('). (C) The position and nucleotide sequence of the two polyadenylation signals (AATAAA) and the two actual sites of poly(A) tail addition (TA) are shown. Sequences not shown are indicated by dotted lines.

2204, where the protein translation probably is terminated. The predicted size of this NS-2 protein is 13,400 Da, which is somewhat smaller than the reported size for other parvoviral NS-2 proteins (21, 22). A protein of this size has not been reported for ADV; however, Western blot experiments do show the presence of one or more proteins in this size range (M. E. Bloom and S. Alexandersen, unpublished observations). Further experiments are in progress to identify an NS-2 for ADV.

R2'. The R2' RNA is a very abundant species with a size of 0.85 kb, including the poly(A) tail. This RNA is very similar to the R2 RNA; i.e., it has the same 5' end and the same splice pattern but differs from R2 at the 3' end, which, for R2', is at nucleotide 2586 (Fig. 7 and 8). The presence of a 3' end at this site is in excellent agreement with a poly(A) signal (AATAAA) at nucleotides 2564 to 2569 and the addition of the poly(A) tail at a UA 17 bases downstream from this signal. This RNA can code for the same NS-2 protein as described for R2.

RX. The RX RNA is a 1.1-kb species having three splices, giving it a unique structure (Fig. 7 and 8) without similarities to other parvoviral RNAs described so far. The possibility that this RNA is an artifact could be vigorously excluded by the facts that three independent cDNA clones had the same structure and that an RNA of the predicted size and structure could be found in Northern blots probed with a splice-specific oligonucleotide. This RNA is potentially able to code for a third, probably nonstructural, protein, tentatively designated NS-3. This protein could initiate at the same ATG as NS-1 and NS-2 at nucleotide 206 in reading frame 2. NS-3

would thus share the amino-terminal part with NS-1 and NS-2 until nucleotide 384, where it splices to reading frame 1 at nucleotide 1738 and picks up a small open reading frame of 27 amino acids until terminated at nucleotide 1818. This gives a potential nonstructural protein 3 (NS-3) of 86 amino acids with an estimated molecular mass of 10,000 Da.

Alternatively, this RNA might initiate translation at one of the multiple ATGs in reading frame 2 at nucleotide 1829, 1886, or 1943 and then splice to reading frame 1 at nucleotide 1961, pick up reading frame 1 at nucleotide 2042, and then terminate at nucleotide 2208. This would give a protein sharing the carboxy terminus with NS-1 and having a maximum of 99 amino acids with an estimated molecular mass of 11,000 Da. However, a protein of this structure could also be encoded by either of the R1 or R3 mRNAs, making it more likely that if expressing protein, this RNA (RX) most probably codes for an NS-3 nonstructural protein of the structure given in Fig. 7. No nonstructural proteins of this predicted size or mRNAs of the structure described for the RX RNA has been described for ADV or for the other parvoviruses. Fairly detailed transcription maps have been published for MVM (9, 19, 21, 22, 28, 32, 35, 36, 40) without description of such an RNA, although the other RNAs have a high degree of similarity to those described for ADV in the present paper. Perhaps a similar RNA does exist for MVM but has been overlooked, or perhaps the unique features of ADV pathogenesis (1–4) are controlled in part by the hypothetical protein translated from this RNA. Any specific role of the nonstructural proteins of the parvoviruses has yet to be identified, although these proteins have been implicated in viral replication and termed REP proteins (9, 19, 49). As our understanding of the complexity of parvoviral control mechanisms for gene transcription and replication emerges, it is not surprising that several independent REP proteins are necessary to efficiently control these events.

In the present report, detailed mapping of ADV transcription resulted in a scheme of ADV transcription including five mRNA species. The general picture of ADV transcription showed extensive homologies to the transcription maps of the other autonomous parvoviruses, especially MVM, which is the most characterized (9, 19, 21, 22, 28, 32, 35, 36, 40, 45). Like the other parvoviruses, ADV had two major promoters around mu 3 and 36 (Fig. 7 and 8). However, in contrast to the MVM and the H-1 parvovirus P3 promoter, the ADV P3 promoter appeared to be functionally stronger than the P36 promoter even in late infection. The P36 promoter has been described to be transactivated by the P3 promoter protein products during parvovirus infection (44, 49). Perhaps this transactivation does not occur in ADV, causing decreased expression of viral structural gene products. If so, this decreased expression could be an explanation for the nonlytic growth of ADV in cell culture and for the development of restricted persistent infection in mink. Another explanation might be that either the R2′ or the RX RNA or protein product in some way competes or interferes with the transactivation mechanism.

Analogs to the two transcription products R2′ and RX have not been found in the prototype parvovirus MVM or in the other rodent parvoviruses (9, 19, 21, 22, 28, 32, 35, 36, 40, 45, 48). However, the structures of these products have similarities to those of abundant products found in the human B19 parvovirus transcription (38), suggesting that ADV uses a complex RNA transcription program with strategies including those found in the rodent parvoviruses and in B19 to compress the coding information into a short, 4.8-kb genome.

ACKNOWLEDGMENTS

We thank James Wolfinbarger for excellent technical assistance, Irene Cook Rodriguez for typing the manuscript, and Robert Evans and Gary Hettrick for preparing the figures.

The study was supported in part by the Danish Fur Breeders Association Research Foundation and the Danish Council for Veterinary and Agricultural Research. Soren Alexandersen is a National Institutes of Health visiting associate on leave from the Institute of Veterinary Pathology, Royal Veterinary and Agricultural University of Copenhagen, Denmark.

LITERATURE CITED

1. **Alexandersen, S.** 1986. Acute interstitial pneumonia in mink kits: experimental reproduction of the disease. Vet. Pathol. **23:** 579–588.
2. **Alexandersen, S., and M. E. Bloom.** 1987. Studies on the sequential development of acute interstitial pneumonia caused by Aleutian disease virus in mink kits. J. Virol. **61:**81–86.
3. **Alexandersen, S., M. E. Bloom, and J. Wolfinbarger.** 1988. Evidence of restricted viral replication in adult mink infected with Aleutian disease of mink parvovirus. J. Virol. **62:**1495–1507.
4. **Alexandersen, S., M. E. Bloom, J. Wolfinbarger, and R. E. Race.** 1987. In situ molecular hybridization for detection of Aleutian mink disease parvovirus DNA by using strand-specific probes: identification of target cells for viral replication in cell cultures and in mink kits with virus-induced interstitial pneumonia. J. Virol. **61:**2407–2419.
5. **Alexandersen, S., A. Uttenthal-Jensen, and B. Aasted.** 1986. Demonstration of non-degraded Aleutian disease virus (ADV) proteins in lung tissue from experimentally infected mink kits. Arch. Virol. **87:**127–133.
6. **Astell, C. R., E. M. Gardiner, and P. Tattersall.** 1986. DNA sequence of the lymphotropic variant of minute virus of mice, MVM(i), and comparison with the DNA sequence of the fibrotropic prototype strain. J. Virol. **57:**656–669.
7. **Becerra, P., J. A. Rose, M. Hardy, B. M. Baroudy, and C. W. Anderson.** 1985. Direct mapping of adeno-associated virus capsid proteins B and C: a possible ACG initiation codon. Proc. Natl. Acad. Sci. USA **82:**7919–7923.
8. **Bensimhon, M., J. Gabarro-Arpa, R. Ehrlich, and C. Reiss.** 1983. Physical characteristics in eucaryotic promoters. Nucleic Acids Res. **11:**4521–4540.
9. **Berns, K. I., and M. A. Labow.** 1987. Parvovirus gene regulation. J. Gen. Virol. **68:**601–614.
10. **Birnstiel, M. L., M. Busslinger, and K. Strub.** 1985. Transcription termination and 3′ processing: the end is in site! Cell **41:** 349–359.
11. **Bloom, M. E., S. Alexandersen, S. Perryman, D. Lechner, and J. B. Wolfinbarger.** 1988. Nucleotide sequence and genomic organization of Aleutian mink disease parvovirus (ADV): sequence comparisons between a nonpathogenic and a pathogenic strain of ADV. J. Virol. **62:**2903–2915.
12. **Bloom, M. E., O.-R. Kaaden, E. Huggans, A. Cohn, and J. B. Wolfinbarger.** 1988. Molecular comparisons of in vivo- and in vitro-derived strains of Aleutian mink disease parvovirus. J. Virol. **62:**132–138.
13. **Bloom, M. E., R. E. Race, B. Aasted, and J. B. Wolfinbarger.** 1985. Analysis of Aleutian disease virus infection in vitro and in vivo: demonstration of Aleutian disease virus DNA in tissues of infected mink. J. Virol. **55:**696–703.
14. **Bloom, M. E., R. E. Race, and J. B. Wolfinbarger.** 1980. Characterization of Aleutian disease virus as a parvovirus. J. Virol. **35:**836–843.
15. **Bloom, M. E., R. E. Race and J. B. Wolfinbarger.** 1982 Identification of a nonvirion protein of Aleutian disease virus: mink with Aleutian disease have antibody to both virion and nonvirion proteins. J. Virol. **43:**606–616.
16. **Bloom, M. E., R. E. Race, and J. B. Wolfinbarger.** 1987. Analysis of Aleutian disease of mink parvovirus infection using strand-specific hybridization probes. Intervirology **27:**102–111.
17. **Blundell, M. C., C. Beard, and C. R. Astell.** 1987. In vitro identification of a B19 parvovirus promoter. Virology **157:**534–

538.
18. **Carlson, J. O., M.-K. Lynde-Maas, and S. Zheng-da.** 1987. A nonstructural protein of feline panleukopenia virus: expression in *Escherichia coli* and detection of multiple forms in infected cells. J. Virol. **61**:621–624.
19. **Carter, B. J., C. A. Laughlin, and C. J. Marcus-Sekura.** 1984. Parvovirus transcription, p. 209–258. *In* K. I. Berns (ed.), The parvoviruses. Plenum Publishing Corp., New York.
20. **Chesebro, B., R. Race, K. Wehrly, J. Nishio, M. Bloom, D. Lechner, S. Bergstrom, K. Robbins, L. Mayer, J. M. Keith, C. Garon, and A. Haase.** 1985 Identification of scrapie prion protein-specific mRNA in scrapie-infected and uninfected brain. Nature (London) **315**:331–333.
21. **Cotmore, S. F., and P. Tattersall.** 1986. Organization of nonstructural genes of the autonomous parvovirus minute virus of mice. J. Virol. **58**:724–732.
22. **Cotmore, S. F., and P. Tattersall.** 1987. The autonomously replicating parvoviruses. Adv. Virus Res. **33**:91–174.
23. **Crandell, R. A., C. G. Fabricant, and W. A. Nelson-Rees.** 1973. Development, characterization, and viral susceptibility of a feline (*Felis catus*) renal cell line (CRFK). In Vitro **9**:176–185.
24. **DeLange, A. M., M. Reddy, D. Scraba, C. Upton, and G. McFadden.** 1986. Replication and resolution of cloned poxvirus telomeres in vivo generates linear minichromosomes with intact viral hairpin termini. J. Virol. **59**:249–259.
25. **Gorham, J. R., J. B. Henson, T. B. Crawford, and G. A. Padgett.** 1976. The epizootiology of Aleutian disease, p. 135–158. *In* R. H. Kimberlin (ed.), Slow virus diseases of animals and man. North-Holland Publishing Co., Amsterdam.
26. **Haagsma, J.** 1969. Epizootiology of Aleutian disease (plasmacytosis) in mink. Neth. J. Vet. Sci. **2**:19.
27. **Hadlow, W. J., R. E. Race, and R. C. Kennedy.** 1983. Comparative pathogenicity of four strains of Aleutian disease virus for pastel and sapphire mink. Infect. Immun. **41**:1016–1023.
28. **Jongeneel, C. V., R. Sahli, G. K. McMaster, and B. Hirt.** 1986. A precise map of splice junctions in the mRNAs of minute virus of mice, an autonomous parvovirus. J. Virol. **59**:564–573.
29. **Khoury, G. and P. Gruss.** 1983. Enhancer elements. Cell **33**: 313–314.
30. **Kozak, M.** 1983. Comparison of initiation of protein synthesis in procaryotes, eucaryotes, and organelles. Microbiol. Rev. **47**:1–45.
31. **Kozak, M.** 1986. Regulation of protein synthesis in virus-infected animal cells. Adv. Virus Res. **31**:229–292.
32. **Labieniec-Pintel, L., and D. Pintel.** 1986. The minute virus of mice P_{39} promoter can encode both capsid proteins. J. Virol. **57**: 1163–1167.
33. **Maniatis, T., E. F. Frisch, and J. Sambrook.** 1982. Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
34. **McDevitt, M. A., M. J. Imperiale, H. Ali, and J. R. Nevins.** 1984. Requirement of downstream sequence for generation of a poly(A) addition site. Cell **37**:993–999.
35. **Merchlinksy, M. J., P. J. Tattersall, J. J. Leary, S. F. Cotmore, E. M. Gardiner, and D. C. Ward.** 1983. Construction of an infectious molecular clone of the autonomous parvovirus minute virus of mice. J. Virol. **47**:227–232.
36. **Morgan, W. R., and D. C. Ward.** 1986. Three splicing patterns are used to excise the small intron common to all minute virus of mice RNAs. J. Virol. **60**:1170–1174.
37. **Mount, S. M.** 1982. A catalogue of splice junction sequences. Nucleic Acids Res. **10**:459–472.
38. **Ozawa, K., J. Ayub, H. Yu-Shu, G. Kurtzman, T. Shimada, and N. Young.** 1987. Novel transcription map for the B19 (human) parvovirus. J. Virol. **61**:2395–2406.
39. **Padgett, G. A., J. R. Gorham, and J. B. Henson.** 1967. Epizootiologic studies of Aleutian disease. I. transplacental transmission of the virus. J. Infect. Dis. **117**:35–38.
40. **Pintel, D., D. Dadachanji, C. R. Astell, and D. C. Ward.** 1983. The genome of minute virus of mice, an autonomous parvovirus, encodes two overlapping transcription units. Nucleic Acids Res. **11**:1019–1038.
41. **Porter, D. D.** 1986. Aleutian disease: a persistent parvovirus infection of mink with a maximal but ineffective host immune response. Prog. Med. Virol. **33**:42–60.
42. **Porter, D. D., A. E. Larsen, and H. G. Porter.** 1969. The pathogenesis of Aleutian disease of mink. I. In vivo viral replication and the host antibody response to viral antigen. J. Exp. Med. **130**:575–589.
43. **Porter, D. D., A. E. Larsen, and H. G. Porter.** 1973. The pathogenesis of Aleutian disease of mink. III. Immune complex arteritis. Am. J. Pathol. **71**:331–344.
44. **Rhode, S. L., III.** 1985. *trans*-Activation of parvovirus P_{38} promoter by the 76K noncapsid protein. J. Virol. **55**:886–889.
45. **Rhode, S. L., III, and P. K. Paradiso.** 1983. Parvovirus genome: nucleotide sequence of H-1 and mapping of its genes by hybrid-arrested translation. J. Virol. **45**:173–184.
46. **Sanger, F. S., S. Nicklen, and A. R. Coulson.** 1977. DNA sequencing with chain-terminating inhibitors. Proc. Natl. Acad. Sci. USA **74**:5463–5467.
47. **Siegl, G., R. C. Bates, K. I. Berns, B. C. Carter, D. C. Kelly, E. Kurstak, and P. Tattersall.** 1985. Characteristics and taxonomy of parvoviridae. Intervirology **23**:61–73.
48. **Srivasta, A., E. W. Lusby, and K. I. Berns.** 1983. Nucleotide sequence and organization of the adeno-associated virus 2 genome. J. Virol. **45**:555–564.
49. **Tratschin, J. D., J. Tall, and B. J. Carter.** 1986. Negative and positive regulation in *trans* of gene expression from adeno-associated virus vectors in mammalian cells by a viral *rep* gene product. Mol. Cell. Biol. **6**:2884–2894.
50. **Wetmur, R. and R. Davidson.** 1968. Kinetics of renaturation of DNA. J. Mol. Biol. **31**:349–370.