

Molecular Structure of Adeno-associated Virus Variant DNA*

(Received for publication, October 3, 1979)

Luis M. de la Maza and Barrie J. Carter

From the Laboratory of Experimental Pathology, National Institute of Arthritis, Metabolism, and Digestive Diseases,
National Institutes of Health, Bethesda, Maryland 20205

When lysates of human cells, infected jointly with the defective parvovirus, adeno-associated virus (AAV), and a helper adenovirus, are banded to equilibrium in CsCl buoyant density gradients, virus particles of various densities are obtained. Infectious AAV particles mainly band at a density of 1.41 g/cm³ with a minor component at 1.45 g/cm³. Noninfectious AAV particles band at densities between 1.41 and 1.32 g/cm³. We have analyzed, by mapping with site-specific endodeoxyribonucleases, the molecular structure of the variant AAV DNA molecules obtained from these light density particles. The size of variant DNA molecules ranged from 100 to 3% of genome length. In general, the variant DNAs are deleted for internal regions but retain the genome termini. Some of the variant DNAs appear to be cross-linked, spontaneously renaturing molecules having structures analogous to replicating forms of AAV DNA.

Adeno-associated virus (AAV) is a defective parvovirus which requires certain helper functions provided by adenovirus in order to multiply (1, 2). The genome of AAV¹ is a single, linear DNA strand about 4700 nucleotides long (3-5). Complementary DNA strands of "plus" or "minus" polarity are separately encapsidated into individual virions (3). Upon isolation of the DNA, these strands anneal to form linear duplex molecules. The individual DNA strands contain the same terminal-inverted repetition within the first 145 nucleotides at both 5' and 3' ends (4, 6, 7). The first 125 nucleotides of this sequence are palindromic and can fold into a T-shaped hairpin structure. The internal portion of this sequence (nucleotides 44 to 83) can occur in either of two orientations with opposite polarity. This sequence inversion has been termed a "flip-flop" (7) and we use the term solely for convenience. This sequence flip-flop may reflect events during replication of AAV DNA termini since the terminal palindrome apparently constitutes the origin of DNA replication (8). Also, parvovirus DNA appears to replicate by self-priming from the terminal hairpin and is followed by subsequent transfer (and sequence inversion) of the hairpin to progeny DNA (9-11). The consequence of this flip-flop is that some AAV DNA molecules have linear duplex termini, but others (when strands with termini of opposite polarity anneal) contain a nonlinear, base-paired cruciform structure (7, 12). This nonlinear structure gives rise to some terminal region restriction endonuclease

fragments which have aberrant electrophoretic mobility in acrylamide gels (12).

When lysates of cells infected with AAV and a helper adenovirus are banded to equilibrium in CsCl buoyant density gradients, a number of virus bands can be detected visually (13). The major and minor infectious AAV components band at densities of 1.41 and 1.45 g/cm³, respectively. In the electron microscope, these virions from both bands appear to be "full" particles. Infected cell lysates also contain a heterogeneous population of "light," noninfectious AAV particles which band at densities of less than 1.41 g/cm³ (13-17). The most prominent visible bands are those at 1.35 g/cm³ (at which density the helper adenovirus also bands) and 1.32 g/cm³. These light particles appear "empty" in the electron microscope and variously have been reported to contain either no DNA (15) or DNA of less than genome length (14).

Elsewhere we have described the biological properties of the various density classes of AAV particles and have shown that each class contains defective interfering particles (17, 18). In this report, we have analyzed the DNA structure of each density class of AAV particles. We show that all classes contain AAV DNA molecules. The size of these DNA molecules ranges from full length to about 3% of genome length. The density of the virus particle is generally correlated with the size of the variant DNA. These variant AAV DNAs are deleted for internal regions of the AAV genome but retain the genome termini. Some of the variant DNAs appear to be cross-linked, spontaneously renaturing ("snap-back" molecules) with structures analogous to AAV replicating intermediates (9). A preliminary account of some of the properties of AAV variant DNA has been presented before (16). The structure of these variant DNAs has important implications for AAV autointerference, replication, transcription, and virus particle assembly.

EXPERIMENTAL PROCEDURES

Virus Growth and Purification—Suspension cultures of KB cells were grown at 37°C in Eagle's medium supplemented with 5% horse serum. Cells at a density of 3×10^5 /ml were co-infected with AAV2 at a multiplicity of 40 infectious units/cell (17) and adenovirus type 2 at a multiplicity of 10 plaque-forming units/cell, then harvested 40 h after infection.

For production of ³²P-labeled DNA, cells were infected as above, then collected after 3 h and resuspended in phosphate-free Eagle's minimal essential medium containing 5% dialyzed horse serum; [³²P]orthophosphate (New England Nuclear) was added to a final concentration of 20 μCi/ml. The specific activity obtained ranged from 0.8 to 1.5×10^6 cpm/μg. To obtain ³H-labeled DNA, 5-fluorodeoxyuridine was added 13 h after infection (to a final concentration of 2×10^{-4} M) followed 30 min later by addition of [³H]thymidine (20 μCi/ml; 40 to 60 Ci/mmol, New England Nuclear). Cells from 1-liter cultures were resuspended in 10 mM Tris-HCl, pH 8.5, sonicated in 2 bursts of 15 s each at 50 watts (Branson Sonifier) and then treated with 1% sodium deoxycholate and 0.025% trypsin at 37°C for 30 min. The preparation was cooled on wet ice, sonicated again for 10 s, and

* The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "advertisement" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

¹ The abbreviations used are: AAV, adeno-associated virus; RF, replicating form; SS, single-stranded; DS, double-stranded; SB, snap-back; MVM, minute virus of mice.

CsCl was added to a final density of 1.40 g/cm³. Equilibrium centrifugation in CsCl was performed in the Beckman L5 ultracentrifuge in the SW 50 rotor. Individual virus bands were collected from the bottom of the tube and centrifuged twice more in CsCl using similar conditions.

DNA Purification—Visible bands of virus were dialyzed against 1 × SSC (0.15 M NaCl, 0.015 M sodium citrate) and adjusted to 0.1 M NaOH; the viral DNA was isolated by velocity sedimentation through a 5 to 20% alkaline sucrose gradient (0.3 M NaOH, 0.7 M NaCl, 10 mM EDTA, 0.15% Sarkosyl). Conditions for the sedimentation were 35,000 rpm for 15 h at 4°C using the Beckman SW 41 rotor. The gradients were collected from the bottom of the tube and the isolated DNA strands were dialyzed against 10 mM Tris-HCl, pH 8.0. To obtain double-stranded DNA, the complementary single strands were adjusted to 0.1 M NaOH and, after 10 min at room temperature, 1 M Tris-HCl, pH 8.0, and HCl were added to neutralize the solution. Purified formamide was added to a final concentration of 50% and the DNA reassociated at room temperature for 24 h (19). The DNA concentration during reassociation ranged from 5 to 20 µg/ml. After reannealing, the DNA was dialyzed, precipitated with alcohol, resuspended in 10 mM Tris-HCl, pH 8.0, and sedimented through a neutral sucrose gradient (1 M NaCl, 50 mM Tris-HCl, pH 8.0, 1 mM EDTA, 0.15% Sarkosyl). The DNA was sedimented in the Beckman L5 ultracentrifuge using a SW 41 rotor at 35,000 rpm at 4°C. The length of the run was adjusted to accommodate the different size molecules to be isolated. Specific DNA components were collected and precipitated with ethanol.

Digestion with Site-specific Endodeoxyribonucleases—All site-specific endodeoxyribonucleases (restriction enzymes) were obtained from New England Biolabs (Beverly, MA) and were used according to the supplier's specifications. The enzymes used were *Hha* I (EC 3.1.12.19), *Alu* I (EC 3.1.23.1), *Kpn* I (EC 3.1.23.26), *Eco* RI (EC 3.1.23.13), *Bam* HI (EC 3.1.23.6), *Mbo* II (EC 3.1.23.28), *Hae* III (EC 3.1.23.17), *Hae* II (EC 3.1.23.16), and *Hpa* II (EC 3.1.23.24). Enzyme digestions were carried in a final volume of 20 µl with DNA concentrations ranging from 0.5 to 5 µg/ml. Enzymes were added to obtain a complete digestion of the DNA in 1 to 2 h at 37°C. All the reaction mixtures contained 10 mM Tris, pH 7.4, 10 mM MgCl₂, 10 mM 2-mercaptoethanol, and NaCl at either 50 mM (*Hha* I, *Alu* I, *Kpn* I, *Eco* RI, *Bam* HI, *Mbo* II) or 6 mM (*Hae* III). For *Hpa* II digestions, 6 mM KCl was used instead of NaCl. For *Hae* II, NaCl was omitted.

Labeling of the 5' Terminus of DNA—This was performed essentially as described before (12, 20). DNA was first treated with bacterial alkaline phosphatase (EC 3.1.3.1, Worthington Biochemical Corp., Freehold, NJ) at 65°C. Following this, the DNA was purified by phenol extraction and, in some cases, by velocity sedimentation in neutral sucrose gradients. DNA was then incubated with phage T4 polynucleotide 5'-hydroxyl-kinase (EC 2.7.1.78, New England Biolabs) and [³²P]ATP (1000 to 3000 Ci/mmol; New England Nuclear) at 37°C as described (21). The 5' terminally labeled DNA was then again purified by phenol extraction and velocity sedimentation.

Digestion of DNA with Endonuclease S1—The endonuclease S1 (EC 3.1.30.1), purified from *Aspergillus oryzae* according to the procedure of Vogt (22), was obtained from Miles Laboratories (Elkhart, IN). DNA was digested with endonuclease S1 as described before (20, 23) at a temperature of 48°C.

Limited Digestion of DNA with Exonuclease III—Escherichia coli exonuclease III (EC 3.1.4.27, New England Biolabs) was used as described (24). Reaction mixtures (50 µl) contained 0.4 µg of AAV [³²P]DNA (specific activity, 2 × 10⁵ cpm/µg), 30 mM Tris-HCl, pH 8.0, 1 mM MgCl₂, 1 mM 2-mercaptoethanol, and 1 unit of exonuclease III. Reactions were incubated at 28°C and stopped at appropriate times by addition of EDTA to 10 mM. The DNA was then recovered by ethanol precipitation. The extent of digestion was monitored by release of acid-soluble, ³²P radioactivity.

Gel Electrophoresis and Autoradiography—Samples were analyzed by vertical polyacrylamide slab gel electrophoresis in 40 mM Tris-HCl, pH 7.8, 5 mM sodium acetate, and 1 mM EDTA buffer (25). The gels contained the designated percentage of acrylamide as indicated in the figure legends and *N,N'*-methylenebisacrylamide at 5% the concentration of acrylamide. Polymerization was initiated with 0.07% ammonium persulfate and 0.05% *N,N''*,*N'*-tetramethylethylenediamine. The gels were prerun for 1 h before electrophoresis. Restriction endonuclease digests were adjusted to 1% sodium dodecyl sulfate, 10% sucrose, and layered under the upper buffer. Bromophenol blue was added as a marker. The DNA fragments were electrophoresed at room temperature with a 100 V potential across the electrodes until the marker dye reached the desired position. After electrophoresis,

the gels were dried and the labeled DNA restriction fragments were located by autoradiography using SB-5 Kodak x-ray film (26).

Isolation of DNA from Gels—Samples were analyzed by disc electrophoresis in the previously described buffer through a 3% acrylamide, 0.7% agarose gel. After electrophoresis, the gels were sliced into 1-mm fragments and counted by Cerenkov radiation. Specific fragments were recovered from the gels by electrophoresis of the DNA into dialysis bags. The DNA was extracted with phenol and recovered by precipitation with ethanol.

Fractionation of DNA on Hydroxylapatite—DNA was fractionated in hydroxylapatite (Bio-Rad, DNA grade) using a batch-washing procedure. [³²P]DNA in 0.05 M sodium phosphate buffer was mixed at 65°C with excess hydroxylapatite. The hydroxylapatite was then pelleted by centrifugation and the supernatant was removed. The hydroxylapatite was then washed successively with phosphate buffer of increasing molarities, as described in the text, to elute single-stranded and double-stranded DNA. At least 95% of the DNA was bound to hydroxylapatite in 0.05 M phosphate buffer and could be subsequently eluted with buffer of increased molarity. Heat-denatured DNA was prepared by heating at 100°C for 10 min then rapidly chilled in a dry ice acetone bath. DNA from hydroxylapatite elutions was recovered by dialysis and ethanol precipitation.

Analysis of DNA Reassociation Kinetics—DNA was degraded to fragments of 250 to 300 nucleotides long by heating at 100°C for 5 min in 0.2 N NaOH then reneutralized and recovered by ethanol precipitation. The fragmented DNA was resuspended in 10 mM NaCl, 50 mM Tris-HCl, pH 8.0, and denatured by boiling for 10 min and quenching in ice. The reaction mixture was then adjusted to 0.21 M NaCl and immediately annealed at 68°C. Reactions were stopped at various times by chilling in ice and the fraction of the [³²P]DNA that was double-stranded (*f*_{DS}) was immediately determined using the single-stranded DNA-specific endonuclease S1 as described before (20, 23). The proportion of DNA remaining single-stranded (*f*_{SS}) was then calculated according to *f*_{SS} = 1 - *f*_{DS}. The resulting data were then plotted as 1/*f*_{SS} versus incubation time. This yields a straight line plot according to the relationship 1/*f*_{SS} = (1 + *kC*₀*t*)⁻¹ where *k* is a constant, *C*₀ is the initial concentration of single-stranded DNA, *n* = 0.45, and *t* is the time (27). In these conditions, when individual DNA preparations are annealed at the same initial DNA concentration, the slope of the straight line (i.e. the relative rate of reassociation) is inversely proportional to the relative genetic complexity of the DNAs.

RESULTS

Centrifugation of AAV Particles in CsCl Buoyant Density Gradients—Lysates of KB cells obtained 40 to 48 h after infection with both AAV and adenovirus yield four major visible bands of virus particles at densities of 1.45, 1.41, 1.35, and 1.32 g/cm³ when banded to equilibrium in a CsCl density gradient. The AAV particles from each of these visible bands are referred to by the buoyant density of the band as AAV(1.45), AAV(1.41), AAV(1.35), and AAV(1.32), respectively. The AAV(1.35) band also contains mature adenovirus particles. The four visible bands were separately taken and DNA was extracted as described below. It should be noted that while the visible bands represent the predominant species, there is a spectrum of AAV particles throughout the gradient in the region 1.41 to 1.32 which can be demonstrated by pooling the interband regions from bulk preparations and banding in CsCl again.

Other properties of the various AAV particle density classes have been described elsewhere (17, 18). In this report, we describe some physical properties of the DNA extracted from AAV(1.41), AAV(1.35), and AAV(1.32) particles. The DNA from AAV(1.45) particles is similar to that from 1.41 particles as will be described elsewhere.²

Isolation of DNA from AAV Particles—DNA was extracted from virus particles by sedimentation in 5 to 20% alkaline sucrose gradients. Full length, single-stranded DNA from the AAV(1.41) particles sedimented as a single, sharp 15.5 S peak and was designated AAV DNA L. The more

² L. de la Maza and B. Carter, manuscript in preparation.

slowly sedimenting DNA, containing approximately 5 to 10% of the total DNA from the 1.41 band, was pooled separately and designated AAV DNA II. The 1.35 g/cm³ particles yielded adenovirus DNA which pelleted to the bottom of the tube and a discrete 10 S species which was pooled and designated AAV DNA III. The 1.32 g/cm³ particles yielded very small DNA molecules designated DNA IV.

The four DNA classes obtained from alkaline gradients were neutralized and reassociated and the resulting double-stranded DNA molecules were centrifuged in neutral sucrose gradients. The DNA L yielded 14.5 S duplexes of AAV DNA, and some faster sedimenting material containing 5 to 10% of the DNA which corresponds to oligomeric AAV DNA duplexes previously characterized (5, 28). A small amount of material sedimenting at 12 to 13 S was pooled and is designated AAV DNA I. This DNA I was further purified in a second neutral sucrose gradient. Type II DNA was freed from contaminating L DNA in a neutral sucrose gradient. The reassociated DNA III molecules exhibited a heterogeneous sedimentation profile in neutral sucrose; about one-half sedimented at around 8 to 10 S (DNA III), while the remainder sedimented faster and appears to be oligomeric structures (16). The DNA IV molecules sedimented very slowly in neutral sucrose.

To facilitate the discussion, we will use the arbitrary designations L to refer to standard AAV DNA and I, II, III, and IV to refer to the variant DNA species.

Molecular Weight Estimation of AAV DNAs—The AAV DNA L, I, II, III, and IV preparations obtained from the neutral sucrose gradients were resedimented in both neutral and alkaline sucrose gradients (16). The molecular weight of each DNA species was calculated from both neutral and alkaline sucrose gradients and is listed in Table I. Also shown in Table I is the ratio of the molecular weight of the duplex DNA in neutral sucrose to that of the single strand in alkaline sucrose. For a normal linear DNA duplex, this ratio is 2.0. For a palindromic molecule, which could fold back by spontaneous self-renaturation, this ratio would approach 1.0. We refer to such molecules as snap-back molecules. DNA III, like DNA L, had a ratio of 1.87 (Table I) and thus appeared to comprise mainly linear duplex molecules but includes some snap-back molecules as well. DNA I, II, and IV had ratios close to 1.0 and thus appeared to be mainly self-renaturing, snap-back palindromic molecules.

TABLE I
Molecular weights of AAV DNA molecules

DNA ^a	Molecular weight determined in ^b				M_r (neutral)/ M_r (alkaline) ^d
	Alkaline sucrose		Neutral sucrose		
	$M_r \times 10^{-6}$	% genome length ^c	$M_r \times 10^{-6}$	% genome length ^c	
L	1.55	100	3.1	100	2.0
I	1.55	100	1.47	47	0.95
II	0.90	58	0.94	30	1.04
III	0.31	20	0.56	18	1.87
IV	0.05	3.2	0.05	1.6	1.0

^a The molecular weights were calculated from sucrose gradients according to the relationships derived by Studier (29) and Danna and Nathans (30). The molecular weight of AAV DNA L is recalculated from previous data (4, 5, 12) using revised molecular weights of Φ X 174 DNA (31) and SV40 DNA (32) calculated from the complete nucleotide sequence. All three sets of data yield the same estimate for AAV DNA, i.e. a length of approximately 4700 nucleotides.

^b The DNA preparations were those described in the text.

^c Molecular size is calculated on a percentage of the single strand (alkaline sucrose) or double strand (neutral sucrose) molecular weight of DNA L.

^d This is the ratio of the molecular weight in neutral sucrose to that determined in alkaline sucrose.

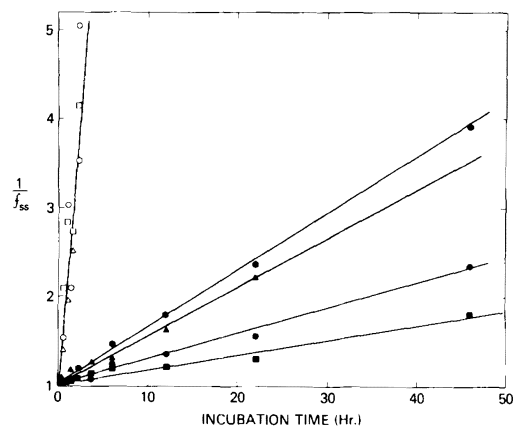


FIG. 1. Preparations of ³²P-labeled AAV DNA L, I, II, or III (specific activity, 360,000 cpm/μg) were separately denatured and then annealed at a concentration of 20 ng/ml at 68°C. At various times, the proportion of DNA that had reassociated to form duplex molecules was determined using the single strand-specific nuclease S1 (see "Experimental Procedures"). Each incubation mixture contained a 100-fold excess amount (2 μg/ml) of unlabeled heterologous DNA (closed symbols) or unlabeled AAV DNA L (open symbols). The ³²P-labeled AAV DNA in each reaction was L (■, □), I (●, ○), II (▲, △), or III (●, ○), respectively.

In the succeeding sections, we will describe first the genetic complexity of the variant AAV DNA species and then the structure of their molecular termini. Finally, we will analyze the overall physical structure, including the snap-back properties, of these DNA molecules.

Reassociation Kinetics of AAV DNA—The relative reassociation rates of individual preparations of ³²P-labeled AAV DNA L, I, II, or III, respectively, were measured (Fig. 1). In the conditions used, the rate of reassociation of the individual preparations of [³²P]DNA was inversely proportional to the genetic complexity of the DNA. Since some of the DNA species are snap-back molecules, and would otherwise anneal with unimolecular kinetics (see below), all the DNA preparations were fragmented by boiling in alkali (see "Experimental Procedures"). As shown in Fig. 1, the relative reassociation rates (closed symbols) of the AAV [³²P]DNAs L, I, II, or III were 1.0, 1.7, 3.3, and 3.9, respectively. These results indicate that DNAs I, II, and III contain not more than approximately 60, 30, and 25%, respectively, of the AAV genome sequence. DNA IV is too small to analyze readily its reassociation kinetics.

When an amount of unlabeled AAV DNA L equivalent to about a 100-fold molar excess of the amount of [³²P]DNA L was included in the reaction, the reassociation rate of the [³²P]DNA was increased by the expected amount (Fig. 1, open symbols). When the same amount of unlabeled AAV DNA L was added to the reactions of [³²P]DNAs I, II, or III, the reassociation rate of these [³²P]DNAs was also increased by the expected amounts and in each case the rate was similar to that for the [³²P]DNA L probe (Fig. 1). From these results, we conclude that virtually all of the [³²P]DNAs I, II, or III preparations represent AAV DNA sequences.

Cleavage of Variant DNA with Restriction Endonucleases—To determine which regions of the AAV genome were represented in the variant DNAs, each preparation was analyzed by gel electrophoresis after cleavage with various restriction endonucleases. Restriction endonuclease cleavage maps for duplex AAV DNA are shown in Fig. 2. The electrophoretic profile of the uncleaved variant DNAs is shown in Fig. 3. The

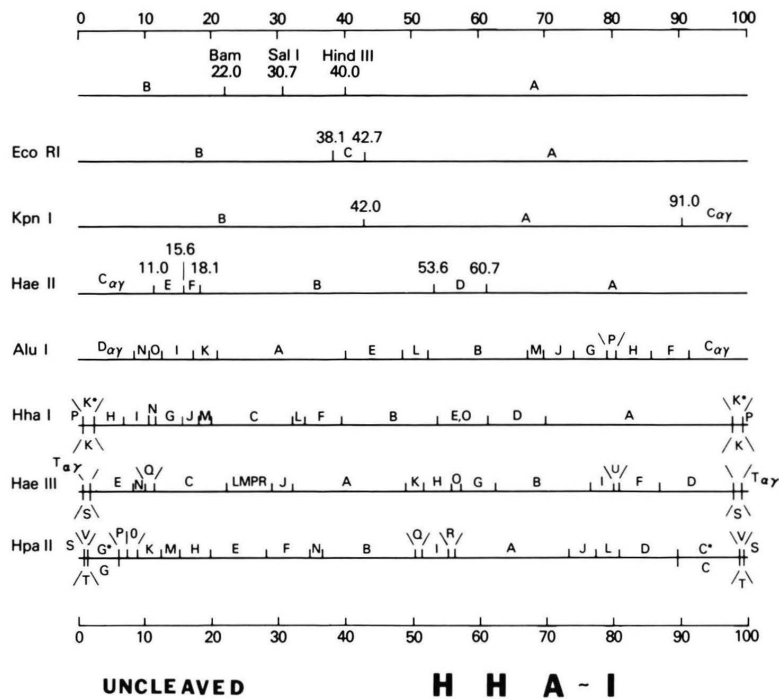


FIG. 2. A diagrammatic map of restriction endonuclease cleavage sites in AAV DNA duplexes. The map is drawn to a scale of 100 units (1 unit is equivalent to 1% of the genome length, i.e. about 47 nucleotides). The drawing is approximately to scale. The cleavage sites were determined previously (12, 16, 19, 20, 28, 33-37). Fragments are designated by the alphabetical letter code (38). In the text, we have also used the map coordinate nomenclature system adopted for parvoviruses (39). The subscripts α or γ for terminal fragments denote those containing a flip-flop structure or a normal linear duplex, respectively (see text for details). Similarly *Hha* K^* contains the flip-flop structure and *Hha* K is the linear duplex structure.

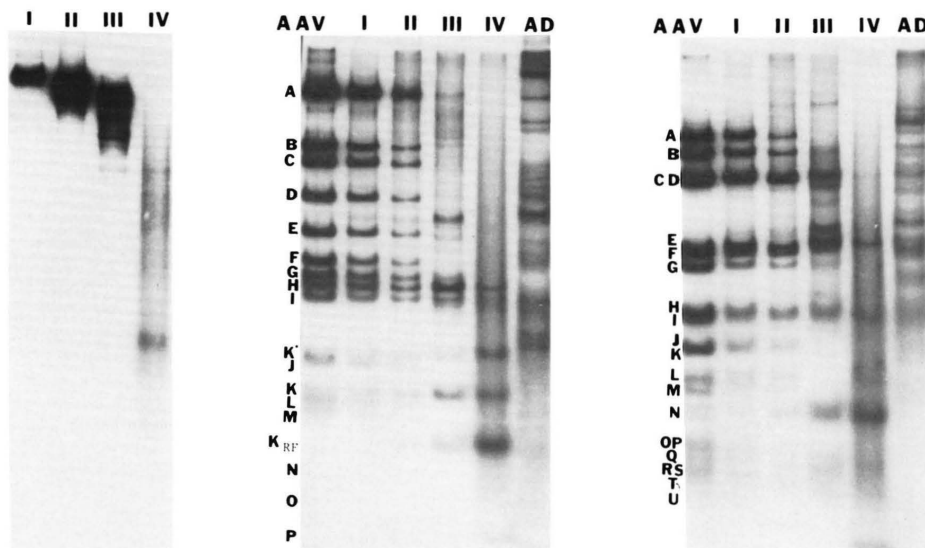


FIG. 3. Cleavage of variant DNA with restriction endonucleases. Preparation of AAV [^{32} P]DNA L, I, II, III, IV, and AD2 [^{32} P]DNA were cleaved with the restriction endonucleases *Hha* I or *Hae* III. The digestion products were electrophoresed in gels containing a gradient of 5 to 16% acrylamide and visualized by autoradiography. *Ad*, adenovirus.

molecular weights of the DNAs I to IV determined in acrylamide gels at a variety of concentrations (not shown) agreed well with those determined in neutral sucrose gradients (Table I). For the DNA IV preparation shown in Fig. 3, the fastest migrating discrete band is the major component with a molecular weight of 0.05×10^6 . The slower-migrating, more heterogeneous material in DNA IV (Fig. 3) was removed in

other preparations by a second neutral sucrose gradient purification step.

Cleavage of variant DNA with the restriction endonucleases *Hha* I, and *Hae* III, is shown in Fig. 3. Adenovirus 2 DNA, cleaved with the same enzymes, was run in parallel. In general, none of the fragments from the adenovirus cleavage appears to correspond to fragments from the AAV DNAs. Also, most

of the cleavage products of DNAs I to IV correspond to known AAV DNA fragments. These data, together with the reassociation kinetics described above, show that most of the ^{32}P -labeled variant DNA contain AAV-specific sequences. The restriction endonuclease cleavage patterns show that the variant DNAs are deleted for internal regions of the AAV genome (Fig. 3). In general, as the size of the variant DNA decreases, there is a progressive loss of the internal regions of the genome with a concomitant enrichment for the terminal regions. For example, the internal *Hha* fragments B through F are present in a low molar ratio in DNA I and II and are completely absent in DNA III and IV. On the other hand, *Hha* G, H, I, K, and K* that map in the terminal regions of the genome are represented in the DNA I and II at an increased molar ratio. In DNA III, *Hha* G and I are also mostly deleted and even H is mostly deleted from DNA type IV. Similarly, the internal *Hae* III fragments A, K, H, O, G, and B are progressively lost as the size of the variant DNA decreases, but *Hae* III D, E, and N from the ends of the AAV genome are present in a high molar ratio in DNA I, II, and III. This bias towards an increased representation of the terminal fragments in the variants could also be demonstrated by cleavage with *Alu* I, *Kpn* I, and *Hpa* II (not shown). It should be noted that both left (*Hha* H, *Hae* III E,) and right (*Hae* III D and F) termini are retained.

Analysis of Oligomeric DNA III—The gel electrophoresis is of DNA III (Fig. 3) revealed a population of molecules having several discrete bands with sizes between approximately 10 and 25% of genome length. In addition, the reassociated DNA III in the preparation obtained in the neutral sucrose gradient described above contained an amount of faster sedimenting oligomeric material that was not taken for analysis. The oligomeric DNA III consists of noncovalent

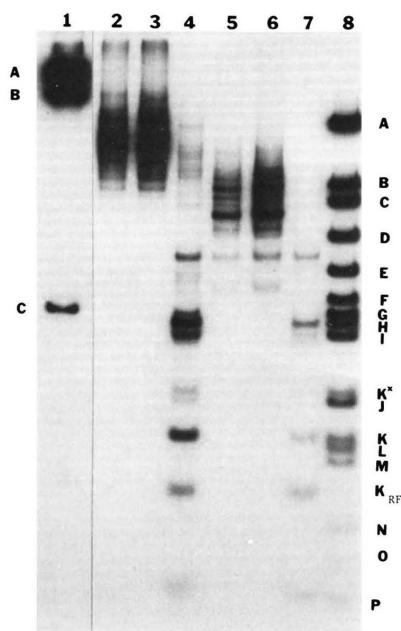


FIG. 4. Electrophoresis of DNA III. ^{32}P -labeled DNA was electrophoresed in a 3 to 10% acrylamide gradient gel and bands were detected by autoradiography. Gel tracks contained (1) AAV DNA L, *Eco* RI digest; (2) oligomeric DNA III; (3) oligomeric DNA III, *Eco* RI digest; (4) oligomeric DNA III, *Hha* digest; (5) DNA III; (6) DNA III, *Eco* RI digest; (7) DNA III, *Hha* digest; (8) DNA L, *Hha* digest.

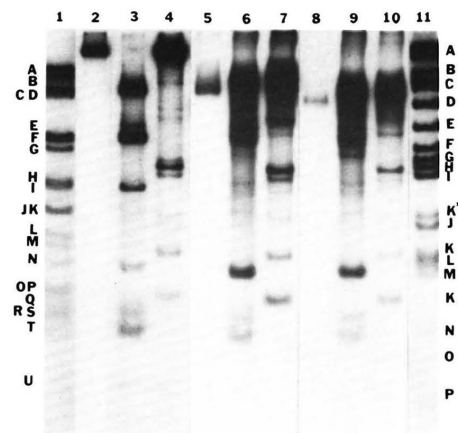


FIG. 5. Analysis of specific size classes of DNA III. ^{32}P -DNA III was electrophoresed in a 3% acrylamide gel and individual DNA III bands were eluted and recovered and analyzed in 3 to 10% acrylamide gradient gels as shown. Track 1, DNA L cleaved with *Hae* III. Track 11, DNA-L cleaved with *Hha*. Tracks 2, 3, and 4 show one DNA III band (28% of genome length) undigested and digested with *Hae* III or with *Hha*, respectively. Tracks 5, 6, and 7 show one DNA III band (12% of genome length) undigested and digested with *Hae* III or with *Hha*, respectively. Tracks 8, 9, and 10 show one DNA III band (9.5% of genome length) undigested and digested with *Hae* III or *Hha*, respectively.

oligomeric duplexes resulting, in part, from annealing of "cohesive ends" formed from the terminal palindromic sequences of AAV DNA (16). The cohesive ends of the AAV DNA terminal sequences have been previously described (4, 5, 28). It was also shown previously that in a neutral sucrose gradient, the trailing edge of reassociated DNA III was enriched for snap-back molecules. Therefore, a preparation of reassociated DNA III was sedimented in a neutral sucrose gradient and the gradient was divided at the peak fraction into leading and trailing regions designated oligomeric DNA III and DNA III, respectively. Both preparations were then analyzed by gel electrophoresis (Fig. 4). The oligomeric DNA III (Track 2) migrated as a series of discrete bands equivalent to 12 to 28% of genome size, whereas the DNA III (Track 5) showed a series of bands of 6 to 17% of genome size. Neither set of DNAs was cleaved by *Eco* RI (Tracks 3 and 6) but both yielded similar products when digested with *Hha* I (Tracks 4 and 7). Both preparations yielded *Hha* fragments H, K_{RF}, and P and a fragment migrating between D and E which probably represents the right-hand region of *Hha* A (*i.e.* from map positions 90 to 92, approximately). In addition, the oligomeric DNA contained a higher molar ratio of fragments I, G, J, and K*. The structure and map location of the *Hha* K_{RF} fragment is described below. Note, however, that the DNA III has a higher molar amount of K_{RF} relative to K or K* than does oligomeric DNA III.

Restriction Endonuclease Cleavage of Specific Size Classes of DNA III—To determine whether specific-sized bands from DNA III contained fragments from either the left or right, or both, termini of the AAV genome, ^{32}P -DNA III was fractionated by electrophoresis and specific size classes of DNA were recovered (see "Experimental Procedures"). Each DNA sample was then analyzed by restriction cleavage. Three discrete size classes of DNA III, with sizes equivalent to 28, 12, and 9.5% of genome length, respectively, were analyzed by cleavage with *Hha* or *Hae* III (Fig. 5). As previously shown for the cleavage of unfractionated DNA III molecules, most internal fragments of the AAV genome were not present in

either cleavage, and fragments closer to the ends were progressively lost as the size of the band decreased. More importantly, both left- and right-hand terminal fragments were present in the digests of each specific band of DNA III.

Cleavage of the largest (28%) band (Fig. 5, *Tracks 3 and 4*) gave digestion products consistent with the presence of about 28% of the genome at both left or right ends, *i.e.* with map positions 0/28 and 72/100. The cleavages of the 12 and 9.5% DNA bands (Fig. 5) gave analogous results to those above. The 12% band contained sequences from both the left- (0/12) and right- (88/100) hand 12% of the genome. Similarly, the 9.5% DNA III contained sequences from the left- (0/9) and right- (91/100) hand 9.5% of the genome.

Inspection of the cleavage patterns such as those in Fig. 5 showed that, even for each discrete size class of DNA III that was cleaved, there was a gradient of relative molar amounts of the left- or right-hand termini. Fragments nearer to the terminus were present in greater molar amounts than fragments mapping further inboard. This is best illustrated for the left-hand end by *Hha* H-I-G which have relative molar amounts $H > I > G$ (Fig. 5, *Tracks 4, 7, and 10*) which correspond to their map order. Analysis of *Hae* III E-N-C at the left, and *Hae* III D-F-I at the right-hand ends, gave similar results (Fig. 5). These gradients of fragment representation in discrete size classes of DNA are not readily accounted for by size heterogeneity within individual DNA III bands. This observation is best explained if most of the variant DNAs contain left and right terminal regions joined in molecules having internal deletions. If the precise location of the internal deletion is variable, then even for discrete size classes of variant DNA, a gradient of fragment molarities would be generated at both left- and right-hand ends (see "Discussion").

Restriction Endonuclease Cleavage Fragments of Variant DNA—The two possible orientations (*i.e.* flip-flop) of the terminal repetitious sequence result in two types of termini in AAV DNA duplex molecules (7, 12, 19, 20, 34, 37). This is shown schematically in Fig. 6. When the sequences are in the same orientation in both strands, a linear duplex terminus results. When the sequences are in opposite orientations in the two strands, then a nonlinear, cruciform (flip-flop) duplex terminus is obtained. Restriction endonucleases that do not cleave AAV DNA within the terminal repetition yield two types of terminal structure; the α component, which corresponds to the cruciform duplex, and the γ component, corresponding to the linear duplex structure (19, 20, 34). This is shown by the terminal doublets *Kpn* C_{α} , γ and *Hae* II C_{α} , γ (Figs. 2, 3, and 7). For *Hha*, the terminal fragment P is unique

since this region of the terminal repetition is identical in either orientation. However, the *Hha* K and K* fragments contain the flip-flop region; K corresponds to the linear duplex and K* the cruciform structure (Fig. 6). The enzymes *Hae* III and *Hpa* II cleave at sites within the flip-flop region (*i.e.* within *Hha* K and K*) and yield a complex set of partial and submolar products (12, 34, 37).

A third type of terminal structure, designated RF, is also possible in AAV DNA (Fig. 6). This structure can be formed by self-renaturation of the terminal sequence within a single DNA strand. Also, as shown in Fig. 6, the RF structure is the configuration that would be predicted during replication of AAV DNA if this occurs by a Cavalier-Smith, hairpin transfer mechanism as suggested by others (7, 9, 11, 41).

All three types of terminal structures shown in Fig. 6 are observed in the population of variant DNA molecules. The presence of linear and flip-flop termini is indicated in Figs. 3, 4, and 5 by the presence of *Hha* K and K*. Also, comparison with cleavages of DNA L shows that in the variant DNAs, there is a decreased amount of *Hha* K* relative to K or of the α components relative to the γ components. Concomitant with the decrease in α components, there is the appearance of an equivalent amount of the RF component. This is most clearly shown for DNAs I and II in Fig. 7. In both the left (*Hae* II C) and right (*Kpn* C) termini, the C_{α} component is decreased and the C_{RF} component appears. This C_{RF} component appears to

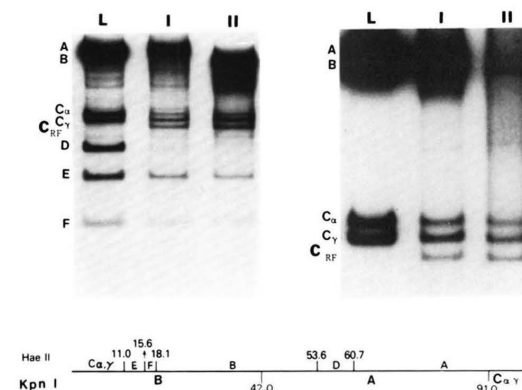


FIG. 7. [^{32} P]DNA L, I, or II cleaved with *Hae* II (left panel) or *Kpn* I (right panel) and analyzed on 3 to 10% acrylamide gradient gels.

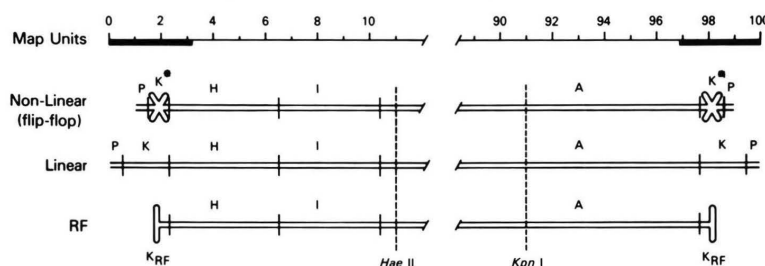


FIG. 6. Schematic representation of the structural configuration of AAV2 DNA duplex termini. The fragments derived by *Hha* cleavage are designated P, K, K*, K_{RF}, H, I, and A. The *Hae* II and *Kpn* I cleavage sites are indicated by broken lines. The sites have been mapped (see Fig. 2) and those bordering K and K* were confirmed by direct sequencing (7). The regions with map coordinates 0/3.1 and 96.9/100, indicated by the solid bar (upper line), contain the terminal repeated, palindromic sequences (145 nucleotides long) of which the first 125 nucleotides can fold into a T-shaped base-paired

hairpin. Because nucleotides 44 to 83 (map positions 1/2 and 98/99) can occur in two orientations (*i.e.* flip-flop) two forms of terminal duplexes occur: linear termini when the sequences in individual strands have the same orientation, and flip-flop termini when the orientations are opposite. The flip-flop configuration results in one of several nonlinear, base-paired cruciform structures, shown here schematically. The third configuration, RF, is believed to occur in replicating molecules (7, 9) and does occur at individual termini of single strands in a T-shaped hairpin (7, 40).

be about 60 nucleotides shorter than C_{γ} , which is the expected difference in size between the γ and RF configurations (Figs. 2 and 6).

The RF configuration present in the variant DNA is also reflected in the *Hha* K_{RF} fragment (Figs. 3, 4, and 5) which has an apparent length of about 1.2% (i.e. about 50 base pairs). The *Hha* K_{RF} fragment was designated *Hha* 1.2 in a preliminary report (16) and was demonstrated by densitometer scans of acrylamide gels to be present in some preparations of DNAs I, II, and III in a high molar ratio. As seen in Fig. 3, *Hha* K_{RF} is the most predominant fragment present in DNA IV.

Several lines of evidence support the interpretation of the RF terminal fragments as depicted in Fig. 6. First, isotopic labeling of the 5' termini of variant DNA using polynucleotide kinase resulted in the appearance of radioactive phosphorus only in the same termini as in DNA L (Fig. 8). None of the RF components were labeled in the kinase reaction. For instance, *Hha* cleavage of DNA L, I, II, or III yielded terminal label only in fragment P (Fig. 8). DNA IV did yield a large smear of terminal label which may reflect heterogeneity of 5' termini in this DNA. The *Mbo* cleavage of DNA L (Fig. 8) shows about equal amounts of 5'-terminal ^{32}P in the α and γ terminal fragments as expected. For DNA I, the terminal label in α was greatly reduced but there was no concomitant increase in labeling of any analogous RF fragment. This suggests that the RF fragment contained no single-stranded gaps since the terminal labeling was done in conditions that should allow labeling of internal single-stranded gaps. It is, of course, difficult to eliminate the possibility that there were gaps containing blocked 5' termini. However, the experiment described in the legend to Fig. 10 also argues against the presence of any gaps.

Second, digestion of duplex AAV DNA L with *E. coli* exonuclease III, to the extent that about 5% of the DNA was rendered acid-soluble, resulted in increased amounts of *Hha* K_{RF} and a concomitant decrease in K^* and K (Fig. 9). (The small amount of *Hha* K_{RF} observed in the absence of prior exonuclease III treatment was due to contamination of the

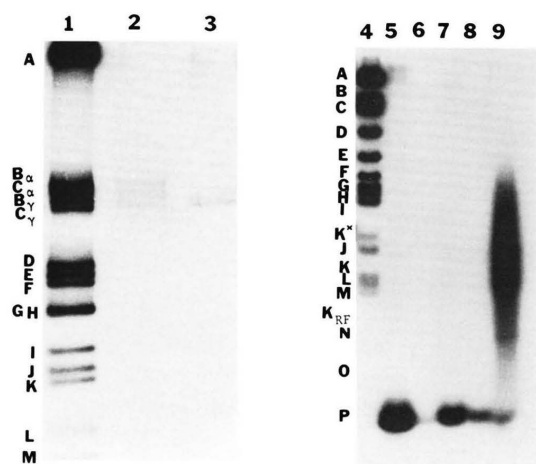


FIG. 8. Analysis of variant DNA labeled with a 5'-terminal ^{32}P . Left panel: *Mbo* I digest of uniformly ^{32}P -labeled DNA L (Track 1), 5'-terminal, ^{32}P -labeled DNA L (Track 2), or DNA I (Track 3). Electrophoresis was in a 3 to 20% gel. *Mbo* B contains the left-hand and *Mbo* C the right-hand termini, respectively. Right panel: *Hha* cleavage of uniformly labeled ^{32}P -DNA L (Track 4) or 5'-terminal, ^{32}P -labeled DNA L (Track 5), I (Track 6), II (Track 7), III (Track 8), or IV (Track 9). Electrophoresis was performed in a 5 to 16% acrylamide gel.

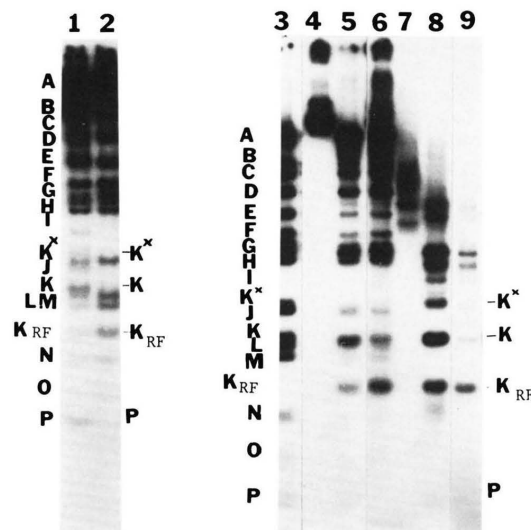


FIG. 9. Limited digestion of AAV ^{32}P -DNA L with exonuclease III prior to cleavage with *Hha* results in production of *Hha* K_{RF} . ^{32}P -DNA L was digested with *E. coli* exonuclease III (exo III) under conditions in which 5% of the DNA was converted to acid-soluble material. ^{32}P -DNA II or III was also digested under analogous conditions (i.e. with the same concentration of enzyme and 3' termini) as for the DNA L digestion. DNA samples were then cleaved with *Hha* and electrophoresed on a 5 to 16% acrylamide gel. Individual reactions show: DNA L, *Hha* digestion (Track 1); DNA L, exo III + *Hha* digest (Track 2); DNA L, *Hha* digest (Track 3); DNA II, undigested (Track 4); DNA II, *Hha* digest (Track 5); DNA II, exo III + *Hha* digest (Track 6); DNA III, undigested (Track 7); DNA III, *Hha* digest (Track 8); DNA III, exo III + *Hha* digest (Track 9).

DNA L with some variant DNA.) Similar results were obtained when preparations of DNA II or III were digested with exonuclease III under conditions analogous to those for DNA L (i.e. with the same relative concentrations of enzyme and DNA 3' termini). Again, exonuclease digestion caused increased amounts of K_{RF} relative to K and K^* (Fig. 9). This experiment confirms that *Hha* K_{RF} is located within 5% of the genome termini and can be formed when the 5' end of the DNA strand spontaneously renatures after exonuclease III digestion of the 3'-ended strand. It is of interest that prior to exonuclease digestion, the DNA III preparation of Fig. 9 yielded significant amounts of *Hha* K^* whereas DNA II did not. This may indicate that DNA II had a large proportion of snap-back molecules, whereas DNA III did not (see "Discussion").

Third, when a *Hha* digest of AAV DNA IV was treated with alkali immediately prior to electrophoresis in a non-denaturing gel at room temperature, the K and K^* fragments were abolished and the K_{RF} fragment remained (data not shown). Alkali treatment of undigested DNA IV did not alter its migration rate. We interpret these data as indicating denaturation of K and K^* followed by spontaneous renaturation to yield K_{RF} in the non-denaturing conditions of the gel. This is exactly the behavior which would be predicted on the basis of the nucleotide sequence of K , K^* (i.e. the single strands of K , K^* , and K_{RF} have identical sequences; Ref. 7).

Analysis of Self-renaturing (Snap-back Molecules) in Variant DNA Preparations—In the sedimentation analysis described above, it was suggested that some of the variant DNA species were spontaneously renaturing, snap-back molecules which reannealed with unimolecular kinetics. To demonstrate this more clearly, we first fractionated the DNA by elution

from hydroxylapatite in phosphate buffer. While single-stranded DNA is generally eluted from hydroxylapatite with 0.15 M phosphate buffer, for parvovirus DNA we have used 0.2 M phosphate to avoid retention of single strands because of self-renaturation of the palindromic terminal sequence (10, 23). Double-stranded DNA was eluted with 0.4 M phosphate. In these conditions, duplex AAV DNA L was eluted almost entirely in the 0.4 M phosphate, whereas the single-stranded, heat-denatured *Pst* A DNA fragment (map coordinates 42/91) was almost entirely eluted in the 0.2 M phosphate (Table II). DNA I and II also eluted in 0.4 M phosphate and appear to be mainly double-stranded (Table II). Sixteen per cent of the DNA III (III_{SS}) eluted in 0.2 M phosphate and the remainder (84%) eluted in 0.4 M phosphate. This latter portion of DNA III, when heat-denatured and refractionated on hydroxylapatite, yielded DNA III_{DS} (69% of total) in 0.2 M phosphate and DNA III_{SB} (15% of total) in 0.4 M phosphate (Table II). The DNA IV was eluted from hydroxylapatite in 0.2 M phosphate (Table II) but not with 0.15 M phosphate, which is consistent with the self-renaturation of the sequences contained in this DNA.

Each of the major DNA fractions from hydroxylapatite was then treated with endonuclease S1 either before or after heat denaturation (Table II). These results showed that most of the DNA L was denatured by heat, whereas DNA I, II, and IV contained a large proportion of molecules, which apparently were snap-back molecules that reannealed with unimolecular kinetics. DNA III_{DS} was mostly denaturable duplexes, whereas DNA III_{SB} was mostly snap-back molecules. DNA III_{SS} appeared to be DNA which was partly single-stranded. For the denaturable molecules in DNA L, III_{SS}, and III_{DS} the proportion of DNA that was still resistant to endonuclease S1 after heat-denaturation was about equivalent to the amount expected for the self-reannealing of the palindromic termini.

These results clearly show that snap-back molecules are present in variant DNA. That DNA I is mostly snap-back is

TABLE II
Variant DNA preparations contain spontaneously renaturing (snap-back) molecules

[³² P]DNA	Proportion (% of total) of DNA eluting from hydroxylapatite at		Proportion of [³² P]DNA (% of total) resistant to digestion with nuclease S1	
	0.2 M PB ^a	0.4 M PB	Before heat denaturation ^b	After heat denaturation ^c
L	1.0	99.0	99.0	9.0
I	1.0	99.0	98.0	75.0
II	1.0	99.0	84.0	71.0
III _{SS} ^d	16.0	(84.0) ^d	38.0	20.0
III _{DS} ^d	69.0		89.0	21.0
III _{SB} ^d		15.0	96.0	102.0
IV	100.0	0	92.0	105.0
(<i>Pst</i> A heat-denatured)	100.0	0	N.D. ^e	N.D.

^a PB, sodium phosphate buffer, pH 6.8.

^b DNA L, I, II, and IV that eluted from hydroxylapatite in 0.4 M PB and also DNA III_{SS}, III_{DS}, and III_{SB} were digested with nuclease S1 and the proportion remaining acid-insoluble was determined (see "Experimental Procedures"). Total acid-insoluble radioactivity (cpm) in the absence of S1 digestion was as follows: 2144 (DNA L), 750 (DNA I), 443 (DNA II), 195 (DNA III_{SS}), 186 (DNA III_{DS}), 150 (DNA III_{SB}), or 161 (DNA IV).

^c DNA was denatured by boiling for 5 min and quenching in ice immediately prior to S1 digestion.

^d DNA III was fractionated into III_{SS} (16% of total) that eluted in 0.2 M PB and the remainder (84% of total) that was eluted in 0.4 M PB. This latter fraction was recovered, heat-denatured, and quick-cooled, then refractionated on hydroxylapatite to yield III_{DS} (69% of total or original DNA III) and III_{SB} (15% of total).

^e N.D., not done.

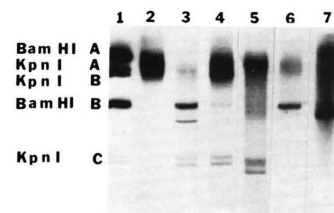


FIG. 10. Selection of snap-back endonuclease restriction fragments on hydroxylapatite. A preparation of [³²P]DNA II was cleaved with *Kpn* I or *Bam* HI. DNA from each digestion was extracted with phenol and recovered by ethanol precipitation. The DNA was resuspended, heat-denatured by boiling, quenched in ice, then fractionated on hydroxylapatite and eluted successively with 0.2 and 0.4 M phosphate buffer, pH 6.8. DNA from the 0.4 M phosphate eluate was recovered and analyzed by electrophoresis in a 3 to 10% acrylamide gel. Track 1, DNA L, cleaved separately with *Bam* or *Kpn* then mixed prior to electrophoresis; Track 2, DNA II; Track 3, DNA II digested with *Bam* and *Kpn* together; Track 4, DNA II, *Kpn* digest; Track 5, DNA II, *Kpn* digest, fragments heat-denatured and eluted from hydroxylapatite in 0.4 M phosphate; Track 6, DNA II, *Bam* digest; Track 7, DNA II, *Bam* digest, fragments denatured and eluted from hydroxylapatite in 0.4 M phosphate.

not surprising because it is precisely this property which allowed their separation from DNA L (see above). Similarly, DNA IV is a snap-back structure because the mapping experiments (Figs. 3 to 5) show that it contains primarily the terminal palindromic sequence. DNA III is a mixture of both normal duplexes (III_{DS}) and snap-back duplexes (III_{SB}). We believe that DNA III_{SS} probably represents heteroduplex molecules consisting of plus and minus strands containing internal deletions at different sites in each strand as suggested above. Since these duplexes are, therefore, annealed mainly at their termini they elute from hydroxylapatite at 0.2 M phosphate as does DNA IV. Elution of DNA from hydroxylapatite is partly dependent on the length of DNA duplex (42). In the DNA II preparation of Table II, a large proportion of the molecules appears to be snap-back molecules although this proportion, like that of DNA III, is highly variable in different preparations (see below).

Location of the Cross-link in Self-renaturing Variant DNA Molecules—To determine where the cross-link is located in the snap-back molecules, the experiment shown in Fig. 10 was performed. A preparation of [³²P]DNA II that had a mean molecular size estimated from gels of about 50% of genome length was cleaved jointly with *Kpn* and *Bam* HI. This yielded the left terminal *Bam* B_{a,γ,RF} fragments (0 to 22) and the right terminal fragments *Kpn* C_{a,γ,RF} (91 to 100) and one new fragment approximately 19 to 20% of genome length and a small proportion of uncleaved molecules. Cleavage with *Kpn* alone yielded *Kpn* C, *Kpn* B, and some uncleaved molecules. Similarly, cleavage with *Bam* alone yielded *Bam* B_{a,γ,RF} and uncleaved molecules. These results suggest that most of this preparation of DNA II was duplex molecules containing internal deletions of about 50% of the genome. The amounts of the RF fragments were very low indicating that in this preparation of DNA II there were few snap-back molecules.

To enrich for fragments containing the snap-back region, the preparations of DNA II cleaved individually by *Kpn* or *Bam* were heat-denatured and fractionated on hydroxylapatite. The DNA which still eluted at 0.4 M phosphate was highly enriched for the *Kpn* C_{RF} and *Bam* B_{RF} fragments and for some uncleaved molecules. These results suggest that the RF fragments described here, and above, can account for the properties of the variant DNA molecules. Thus, the cross-linked region in the variant DNA appears to be at the molec-

ular termini with respect to the complete genome (see "Discussion").

DISCUSSION

The variant DNA molecules extracted from AAV particles are a heterogeneous population of strands ranging in size from nearly full genome length (4700 nucleotides) to only about 3% of genome length (150 nucleotides). These molecules appear to be encapsidated in AAV virions with usually only one strand per particle; thus, the particle density has an approximately direct linear relationship to the strand length.³ The variant genomes contain only AAV DNA sequences as shown by reassociation kinetics and restriction endonuclease cleavage. The restriction endonuclease analysis also revealed a gradient in the distribution frequency of individual restriction fragments. This gradient also was seen in variant DNAs of a discrete size. There was a progressive enrichment for fragments near the genome termini and a concomitant deficiency or absence of fragments from interior portions of the genome. Heat denaturation followed by fractionation on hydroxylapatite and digestion with endonuclease S1 confirmed the suggestion, from the velocity sedimentation analysis, that a proportion of variant DNAs are cross-linked molecules which spontaneously reannealed (snap-back) with unimolecular kinetics to yield hairpin duplexes.

Terminal labeling of the variant molecules with a 5'-[³²P] followed by restriction endonuclease cleavage detected only the normal AAV genome terminal fragments in either the linear (γ) or flip-flop (α) configuration. However, all endonuclease cleavages of variant DNA yielded a new fragment (RF) which maps at the genome termini but which cannot be labeled with 5'-terminal phosphate in the polynucleotide kinase reaction. This RF terminus contained a cross-linked hairpin region as shown by heat-denaturation and selection on hydroxylapatite. This RF structure is consistent with the hairpin structure that can be formed by the AAV DNA terminal palindromic sequence and which is believed to exist in AAV replicating DNA (7, 9). This RF terminal structure appears to be sufficient to account for the proportion of snap-back molecules.

The simplest general model which can account for all our data is that shown in Fig. 11A. The variant DNAs are shown as molecules containing left and right molecular termini and having an internal deletion. The size of the deletion is variable. Also, the location of the deletion, even for a given size class of variant DNA molecules, is variable; i.e. in the model, for a given size of deletion, the distance a to b is constant but the location of positions a and b are variable. The snap-back molecules would consist of those in which one terminus (left or right) is in the covalent hairpin RF conformation. For the linear duplex variants, the proportion of linear (γ) or flip-flop (α) terminal fragments would be approximately equal, as for the normal AAV genome. For the snap-back variant molecules, the γ and RF conformations would be equivalent. Thus, for variant DNA populations, there would be an inverse relationship between the relative amounts of RF and terminal fragments, as is observed.

An alternative possible conformation for the snap-back molecules is that shown in Fig. 11B in which individual molecules contain sequences from left- or right-hand termini only. This model is less likely because, for a discrete size class of variant DNA, there should not be a gradient of molar amounts of fragments for left- or right-hand fragments. Also, this model would not readily explain the inverse relationship between the α and RF terminal fragments. Furthermore, our

³ L. de la Maza, M. Myers, F. Jay, and B. Carter, unpublished observations.

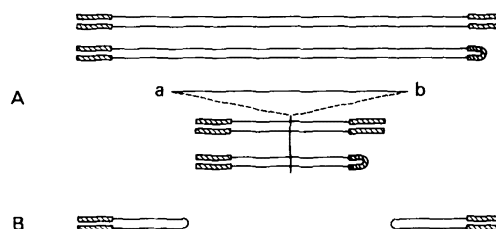


FIG. 11. General model for structure of AAV variant DNA. The AAV genome is represented in its duplex form. Boxed regions represent the terminal palindrome in its linear form or RF forms. The lower section of Panel A illustrates two possible configurations (linear duplex or snap-back), variant DNA having an internal deletion in the region a to b . The lower panel (B) shows alternative structures for variant DNA molecules containing only left or right ends and having the snap-back (or cross-linked) region internal to the normal genome termini. The model is discussed further in the text.

interpretation of the variant molecules as being mostly internally deleted molecules of the type shown in Fig. 11A is supported by electron microscope studies of DNA heteroduplexes formed from variant DNA.⁴ There may, however, be a small proportion of the type shown in Fig. 11B.

The variant DNA described here was obtained from virus particles. Similar molecules, some of which were snap-back structures, have been observed in Hirt supernatant extracts of intracellular AAV DNA (17, 43). Recently, in one study (44), it was suggested that those intracellular variant DNAs which were snap-back molecules had the conformation shown in Fig. 11B. However, in that study (44) only the right-hand terminus of the molecules could be easily analyzed and the data obtained are in fact consistent with either type of structure shown in Fig. 11. It is also possible that there may be selective encapsidation of certain types of variant molecules, or that different virus stocks will yield different variant molecules. Recently, it has been shown that minute virus of mice (MVM) produces variant DNA molecules which include internally deleted molecules analogous to those in Fig. 11A or snap-back structures as shown in Fig. 11B (45). In addition, for MVM the internally deleted variants preferentially accumulated with successive undiluted passages of the virus stock (45). Our stocks of AAV used in this study have been passaged several times at relatively high multiplicities (10 infectious units/cell). At these multiplicities, AAV rapidly accumulates variant DNA molecules (17, 18).⁵

The internally deleted molecules of the type described in Fig. 11A do not require any special or aberrant replication scheme for their production. Rather, they would appear to arise by recombination events (perhaps involving replicating molecules). We and others (16, 43-45) have previously proposed, for AAV or adenovirus, various abortive replication mechanisms involving a template strand switch by the DNA polymerase or template strand breakage. Such abortive replication would generate molecules of the type in Fig. 11B.

The structure of the variant AAV DNAs has a number of important biological implications. Our data support previous evidence (8) that one *cis*-acting function of AAV, namely the origin of DNA replication, is contained within the terminal repetitious sequence since this region appears to be retained in all variant molecules. Since these DNAs contain the AAV DNA replication origin, but are deleted in the regions coding viral mRNA (46) and proteins, they would be expected to be defective interfering genomes as we have recently demonstrated (17, 18). Also, preliminary evidence⁶ indicates that

⁴ L. de la Maza and B. Carter, unpublished observations.

⁵ M. Myers and B. Carter, unpublished observations.

⁶ C. Laughlin and B. Carter, unpublished observations.

some variant DNAs can be transcribed. This is consistent with evidence that transcription of at least some AAV RNAs may begin at a promoter very near the left-hand terminus (46, 47).

Since only one variant DNA molecule is packaged per particle, this indicates that a full length genome is not required for assembly. This is consistent with recent evidence that AAV is assembled by packaging DNA into a preformed capsid rather than by crystallization of capsid subunits around the DNA (48). Also, production of single AAV DNA strands from concatemeric intermediates does not appear to involve a length-measuring mechanism such as the head-full mechanism for phage T4 DNA (49, 50). AAV DNA replication is believed to occur by strand displacement from concatemeric RF molecules which requires a sequence-specific cleavage at both the 3' and 5' ends of the strand at sites defined by the terminal palindromic sequence (7, 9, 11). If the snap-back structures of Fig. 11A are packaged according to the general models proposed for parvovirus DNA replication, this suggests that as the DNA strand is displaced from its 5' end, cleavage does not always occur at the first terminal palindromic sequence reached. This would result in packaging of structures which are exactly analogous to AAV RF molecules. Thus, AAV appears to be unique among viruses in its ability to encapsidate an RF structure. We note that while the isolated snap-back variant DNA clearly has a structure analogous to a duplex RF molecule, it is not known whether this molecule has a single- or double-stranded conformation in the virion.

If a *cis*-acting sequence is required for encapsidation of AAV DNA then this must also be present in the terminal repetition. Our results, together with an analogous study of variant molecules of MVM DNA (45), suggest that such sequences may exist. This could explain the fact that AAV encapsidates either plus or minus DNA strands while nondefective parvoviruses such as MVM package only minus strands. According to this model, it is possible to have molecules of much longer than a normal genome. Such molecules are found in Hirt supernatants (8, 9, 17). We have not found any molecules longer than genome length that were encapsidated in particles. This presumably indicates a restriction on the size of the DNA that can be packaged into an AAV particle.

Acknowledgments—We thank Amos Panet and Catherine Laughlin for critical review and Joan Mok and Diane Shield for preparation of the manuscript.

REFERENCES

- Atchison, R. W., Casto, B. C., and Hammon, W. M. (1965) *Science* **149**, 754-756
- Hoggan, M. D., Blacklow, N. R., and Rowe, W. P. (1966) *Proc. Natl. Acad. Sci. U. S. A.* **55**, 1467-1471
- Rose, J. A., Berns, K. I., Hoggan, M. D., and Koczot, F. J. (1969) *Proc. Natl. Acad. Sci. U. S. A.* **64**, 863-870
- Koczot, F. J., Carter, B. J., Garon, C. F., and Rose, J. A. (1973) *Proc. Natl. Acad. Sci. U. S. A.* **70**, 215-219
- Gerry, H. W., Kelly, T. J., and Berns, K. I. (1973) *J. Mol. Biol.* **79**, 207-225
- Berns, K. I., and Kelly, T. J. (1974) *J. Mol. Biol.* **82**, 267-271
- Berns, K. I., Hauswirth, W. W., Fife, K. H., and Lusby, E. (1978) *Cold Spring Harbor Symp. Quant. Biol.* **43**, 781-787
- Hauswirth, W. W., and Berns, K. I. (1977) *Virology* **78**, 488-499
- Straus, S. E., Sebring, E. D., and Rose, J. A. (1976) *Proc. Natl. Acad. Sci. U. S. A.* **73**, 742-746
- Bourguignon, G. J., Tattersall, P. J., and Ward, D. C. (1976) *J. Virol.* **20**, 290-306
- Tattersall, P. J., and Ward, D. C. (1976) *Nature* **263**, 106-109
- Denhardt, D. T., Eisenberg, S., Bartok, K., and Carter, B. J. (1976) *J. Virol.* **18**, 672-684
- Hoggan, M. D. (1971) in *Comparative Virology* (Maramorosch, K., and Kurstak, E., eds) pp. 49-78, Academic Press, New York
- Torikai, K., Ito, M., Jordan, L. E., and Mayor, H. D. (1970) *J. Virol.* **6**, 363-369
- Johnson, F. B., Whitaker, C. W., and Hoggan, M. D. (1975) *Virology* **65**, 196-203
- de la Maza, L. M., and Carter, B. J. (1978) in *Replication of Mammalian Parvoviruses* (Ward, D., and Tattersall, P., eds) pp. 193-204, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
- Carter, B. J., Laughlin, C. A., de la Maza, L. M., and Myers, M. (1979) *Virology* **92**, 449-462
- Laughlin, C. A., Myers, M. W., Risin, D. L., and Carter, B. J. (1979) *Virology* **94**, 162-174
- Berns, K. I., Kort, J., Fife, K. H., Grogan, E. W., and Spear, I. (1975) *J. Virol.* **16**, 712-719
- Carter, B. J., Fife, K. H., de la Maza, L. M., and Berns, K. I. (1976) *J. Virol.* **19**, 1044-1053
- Weiss, B., Live, T. R., and Richardson, C. C. (1968) *J. Biol. Chem.* **243**, 4530-4542
- Vogt, V. M. (1973) *Eur. J. Biochem.* **33**, 192-198
- Carter, B. J., Khoury, G., and Rose, J. A. (1972) *J. Virol.* **10**, 1118-1125
- Richardson, C. C., Lehman, I. R., and Kornberg, A. (1964) *J. Biol. Chem.* **239**, 251-258
- Dingman, C. W., and Peacock, A. C. (1968) *Biochemistry* **7**, 659-668
- Laskey, R. A., and Mills, A. D. (1977) *FEBS Lett.* **82**, 314-316
- Smith, M. J., Britten, R. J., and Davidson, E. H. (1975) *Proc. Natl. Acad. Sci. U. S. A.* **72**, 4805-4809
- Carter, B. J., and Khoury, G. (1975) *Virology* **63**, 523-538
- Studier, F. W. (1965) *J. Mol. Biol.* **11**, 373-390
- Danna, K., and Nathans, D. (1971) *Proc. Natl. Acad. Sci. U. S. A.* **68**, 2913-2917
- Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., III, Slocumbe, P. M., and Smith, M. (1977) *Nature* **265**, 687-695
- Reddy, V. B., Thimmappaya, B., Dhar, R., Subramanian, K. N., Zain, B. S., Pan, J., Gosh, P. K., Celma, M. L., and Weissman, S. M. (1978) *Science* **200**, 494-502
- de la Maza, L. M., and Carter, B. J. (1976) *Nucleic Acids Res.* **3**, 2605-2616
- de la Maza, L. M., and Carter, B. J. (1977) *Virology* **82**, 409-430
- Carter, B. J., Khoury, G., and Denhardt, D. T. (1975) *J. Virol.* **16**, 559-568
- Carter, B. J., de la Maza, L. M., and Jay, F. T. (1977) in *DNA Insertion Elements, Plasmids and Episomes* (Burkhari, A., Shapiro, J., and Ahlady, S., eds) pp. 477-484, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
- Spear, I. S., Fife, K. H., Hauswirth, W. W., Jones, C. J., and Berns, K. I. (1977) *J. Virol.* **24**, 627-634
- Smith, H. O., and Nathans, D. (1973) *J. Mol. Biol.* **81**, 419-423
- Armentrout, R., Bates, R., Berns, K., Carter, B., Chow, M., Dressler, D., Fife, K., Hauswirth, W., Hayward, G., Lavelle, G., Rhode, S., III, Straus, S., Tattersall, P., and Ward, D. (1978) in *Replication of Mammalian Parvoviruses* (Ward, D., and Tattersall, P., eds) pp. 523-526, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
- Fife, K. H., Berns, K. I., and Murray, K. (1977) *Virology* **78**, 475-487
- Cavalier-Smith, T. (1974) *Nature* **250**, 467-470
- Wilson, D. A., and Thomas, C. A., Jr. (1973) *Biochim. Biophys. Acta* **331**, 333-340
- Hauswirth, W. W., and Berns, K. I. (1978) in *Replication of Mammalian Parvoviruses* (Ward, D., and Tattersall, P., eds) pp. 257-267, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
- Hauswirth, W. W., and Berns, K. I. (1979) *Virology* **93**, 57-68
- Faust, E. A., and Ward, D. C. (1979) *J. Virol.* **32**, 276-292
- Laughlin, C. A., Westphal, H., and Carter, B. J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 5567-5571
- Jay, F. J., de la Maza, L. M., and Carter, B. J. (1979) *Proc. Natl. Acad. Sci. U. S. A.* **76**, 625-629
- Myers, M., and Carter, B. (1980) *Virology*, **80**, in press
- Streisinger, G., Edgar, R. S., and Denhardt, G. H. (1964) *Proc. Natl. Acad. Sci. U. S. A.* **51**, 775-779
- Streisinger, G., Emrich, J., and Stahl, M. M. (1967) *Proc. Natl. Acad. Sci. U. S. A.* **57**, 292-295