JOURNAL OF VIROLOGY, Oct. 1979, p. 276-292
0022-538X/79/10-0276/17$02.00/0

Vol. 32, No. 1

Incomplete Genomes of the Parvovirus Minute Virus of Mice: Selective Conservation of Genome Termini, Including the Origin for DNA Replication

EMANUEL A. FAUST AND DAVID C. WARD*

Department of Human Genetics and Department of Molecular Biophysics and Biochemistry, Yale University School of Medicine, New Haven, Connecticut 06510

Received for publication 16 April 1979

Deletion mutants of minute virus of mice arising during a single high-multiplicity passage and after serial undiluted passage have been isolated, and the incomplete viral genomes contained therein have been analyzed. The DNA isolated from incomplete virions derived from a single high-multiplicity passage was heterogeneous, ranging in size from 15 to 70% of the intact viral genome, with an average molecular length of approximately 2,000 nucleotides. Two distinct types of molecules, designated as type I D-DNA and type II D-DNA, could be distinguished on the basis of their degree of secondary structure, and these were present in roughly equal amounts. Type I D-DNAs were predominantly single-stranded, recombinant molecules in which the self-complementary sequences derived from both genomic termini were conserved. The 5′ terminus was modified relative to the analogous wild-type structure. Although virtually all of the wild-type genome sequence was seen in the total type I D-DNA population, sequences which map between coordinates 47.3 and 87.1 were clearly underrepresented. However, the extent and position of the deletions in individual molecules varied significantly. The shortest molecules in the population lacked between 90 and 95% of the internal wild-type genome sequence and consisted of sequences derived almost exclusively from within 5.0 map units (250 nucleotides) at both ends of the viral genome. Moreover, these miniature recombinant molecules were selectively amplified during serial undiluted passage and were therefore believed to contain all of the critical recognition sites necessary for the replication of minute virus of mice viral DNA. Type II D-DNAs were virus-specific, double-stranded hairpin molecules whose complementary strands were covalently continuous at variable sites distal to the 5′ end of the viral minus strand. In sharp contrast to the type I genomes, these hairpin molecules consisted of sequences which mapped entirely at the 5′ end of the viral genome between positions 85.0 and 100. Furthermore, type II molecules were gradually lost from the total D-DNA population during serial undiluted passage, suggesting that these molecules are not competent for DNA replication but arise as the result of fatal replication errors. Deletion mutants of the type described here for minute virus of mice should be valuable generally as aids to future studies on parvovirus DNA replication, transcription, and cell-virus interactions.

Parvoviruses have been proposed as simple model systems for the study of DNA replication in eucaryotes. The genome of minute virus of mice (MVM), an autonomous parvovirus, is a linear, single-stranded DNA molecule consisting of approximately 5,000 nucleotide residues (3). Short, self-complementary sequences form specialized duplex structures at each of the genome termini, and in vivo studies of replicating molecules of the autonomous parvoviruses H-1 and MVM, as well as the helper-dependent adeno-associated viruses (AAV), indicate that these regions are likely to be intimately involved in the initiation of viral DNA replication (14, 15, 26, 31, 33, 41). Direct sequence analysis at the 3′ terminus of virion DNA obtained from four antigenically distinct autonomous parvoviruses, including MVM (1, 1a), has shown that the first 115 nucleotide residues exist in a predominantly base-paired, hairpin configuration and that this region shares common features also found at the origin of DNA replication in several other animal and bacterial virus genomes (11, 20, 30, 32, 35).

In an attempt to define precisely the regions of the viral genome which are indispensable for DNA replication, spontaneously occurring deletion mutants of MVM have been isolated, and the structure and sequence organization of the deleted viral genomes have been determined. These studies provide evidence that the critical sites of MVM DNA replication lie entirely within 200 to 300 nucleotides encompassing the self-complementary sequences at both ends of the viral genome.

MATERIALS AND METHODS

Cells and virus. The virus strain MVM (T) was used throughout (36). The virus was propogated in Ehrlich ascites (EA) cells (39) and assayed for infectivity and hemagglutination as described previously (36, 37).

Incomplete virions were produced either during a single-cycle, high-multiplicity infection (multiplicity of infection = 10 PFU/cell) with wild-type (wt) virus or after multiple rounds of undiluted passage. Serial undiluted passage was carried out as follows. In the first passage, 10^7 cells were infected with wt virus at a multiplicity of 5 PFU/cell, and the infection was allowed to proceed for 24 h. The cells were harvested by centrifugation, the medium was decanted, and cell pellets were resuspended in 1 to 2 ml of sterile TE 8.7 buffer (0.05 M Tris-hydrochloride, 1.0 mM EDTA, pH 8.7) and dispersed by blending in a Vortex mixer. The cells were then lysed by three cycles of freeze-thawing. Approximately 0.1 ml of this cell lysate containing 3,000 hemagglutination units was used to infect 10^7 cells in the second passage. Adsorption was carried out at 37°C for 60 min at 10^7 cells per ml in 1.0 ml of phosphate-buffered saline containing 5 mM each of $CaCl_2$ and $MgCl_2$. After adsorption, the cells were diluted to 2×10^5 to 3×10^5 cells per ml in D medium (Dulbecco modification of Eagle minimal medium) supplemented with 10% fetal calf serum. Infected cells were again harvested at 24 h postinfection. Subsequent passages were carried out in a similar fashion. To examine the virus yield produced during a given passage, 5×10^7 to 1.0×10^8 cells were infected at 3,000 hemagglutination units per 10^7 cells with the cell lysate from the previous passage as inoculum. Care was taken to dilute the lysate in adsorption buffer sufficiently, i.e., at least 1:10, to maintain a pH of 7.0 to 7.4 during adsorption.

Radioactive labeling of virus. Virus labeled with [^{32}P]phosphate was prepared by infecting cells at a multiplicity of infection of 5 to 10 PFU/cell. After adsorption, the cells were resuspended in D medium supplemented with 10% fetal calf serum. At 8 h postinfection, the cells were pelleted and resuspended in fresh D medium containing 2% of the normal inorganic phosphate, 10% dialyzed fetal calf serum, and 10 μCi of [^{32}P]phosphoric acid per ml (carrier free, New England Nuclear). The cells were harvested at 48 h postinfection, and the virus was extracted as described below. Virus labeled with [^{3}H]thymidine was grown as above in the presence of [*methyl*-^{3}H]thymidine (1 μCi/ml, 2.5×10^{-6} M). The label was added at 1 h postinfection, and the cultures were harvested at 48 h.

Purification of virus. Generally, full and empty virions were purified as described previously (33), but the procedure was modified as described here to enable the resolution of defective virus from full virus in CsCl gradients.

Infected-cell pellets were resuspended in 5 to 10 volumes of TE 8.7 buffer, and the cells were lysed by three cycles of freeze-thawing. The lysate was centrifuged at 10,000 rpm in a Sorvall SS-34 rotor for 10 min at 5°C to remove cell debris. Virus in the supernatant fraction was then sedimented to equilibrium in a CsCl-sucrose step gradient consisting of 5 ml of CsCl solution (ρ = 1.40 g/ml) and 1 to 2 ml of a 1.0 M sucrose solution, made up in TE, pH 8.7. Centrifugation was carried out in a Beckman SW41 rotor at 38,000 rpm for 20 h at 4°C. After this initial purification step, incomplete virions are present in the CsCl layer but are poorly resolved from "full" virions. Full, incomplete, and empty virions were recovered from the CsCl layer in a single pool and dialyzed against TE, pH 8.7. Incomplete virions were then further purified by centrifugation in a second, shallower CsCl gradient. Shallow CsCl gradients were constructed as follows.

A 1-ml amount of CsCl solution, at a density of 1.40 g/ml, was placed in the bottom of an SW41 cellulose nitrate centrifuge tube. Over this was layered 3 ml of CsCl at a density of 1.35 g/ml, followed by an additional 3 ml of CsCl at a density of 1.30 g/ml. The three discrete layers of CsCl were then mixed slightly by brief swirling with a Pasteur pipette. Before centrifugation, contaminating cellular DNA present in the partially purified virus suspension was degraded by exhaustive digestion with micrococcal nuclease (5 to 10 μg/ml, 60 min, 37°C in the presence of 5 to 10 mM $CaCl_2$). The reaction was stopped by adjusting the virus suspension to a concentration of 0.01 M with respect to EDTA. After centrifugation to equilibrium, as above, three types of virus particles were well resolved in the lower half of the centrifuge tube: DNA-containing, full particles were found at a density of 1.40 to 1.41 g/ml; a broad band of incomplete virions was detected between densities of 1.38 and 1.33 g/ml; and a sharp band of empty particles was found at a density of 1.32 g/ml. The incomplete virions were pooled, dialyzed, treated with micrococcal nuclease for a second time, and sedimented to equilibrium in CsCl once again. For the third centrifugation, CsCl gradients were constructed of a lower layer of CsCl (3.5 ml) at a density of 1.35 g/ml and an upper layer of CsCl (3.5 ml) at a density of 1.30 g/ml. The layers were mixed briefly as described above. Incomplete virions were recovered after centrifugation to equilibrium for the third time and concentrated before layering on alkaline sucrose gradients (see below).

Isolation of D-DNA from purified virions. Incomplete virions were recovered from CsCl solution by adding 7 parts methanol to 1 part CsCl solution (adjusted to 1.40 g/ml) and letting the methanol-CsCl mixture stand at room temperature for 30 min (37). The virus was then pelleted by centrifugation at 5,000 rpm in a Sorvall type HS 4 swinging-bucket rotor for 15 min. The methanol-CsCl supernatant was decanted, and the walls of the glass Sorvall centrifuge tube were carefully wiped dry. Virus pellets were taken

up in 0.1 to 0.2 ml of 0.1 N NaOH, and the DNA was isolated by centrifuging the virus suspension in an alkaline sucrose gradient (5 to 20% [wt/vol] sucrose–0.3 N NaOH–0.7 M NaCl–0.15% Sarkosyl). Generally, centrifugation was carried out in an SW41 rotor at 38K for 24 h at 4°C. Under these conditions the DNA from incomplete virions sedimented to the bottom half of the centrifuge tube. The DNA was recovered from these gradients and dialyzed exhaustively against TE 7.4 buffer (0.01 M Tris-hydrochloride, 0.2 mM EDTA). DNA samples were concentrated by evaporation in a stream of nitrogen.

Enzymes. *Escherichia coli* DNA polymerase I (large fragment lacking 5′ to 3′ exonuclease activity) was obtained from Boehringer Mannheim Corp. Bacteriophage λ exonuclease (1 U gives 10 nm of acid-soluble nucleotides in 30 min at 37°C) was the generous gift of C. Radding. Nuclease S_1 from *Aspergillus oryzae*, purified by the method of Vogt (40), was obtained from Sigma Chemical Co. Mung bean nuclease, prepared by the method of Kowalski, et al. (18) (specific activity, 1.4×10^6 U/mg), was from P-L Biochemicals, Inc. Restriction endonucleases *Mbo*I from *Moraxella bovis*, *Hin*fI from *Haemophilus influenzae* R_f, *Hin*dII and *Hin*dIII from *H. influenzae* R_d, and *Hha*I from *Haemophilus haemolyticus* were obtained from New England Biolabs and Boehringer Mannheim.

Enzyme assay conditions. (i) In vitro synthesis of wt RF and D-DNA replicative forms (D-RF). Typically, 0.1 μg of input MVM DNA was incubated with 1 U of *E. coli* DNA polymerase I (large fragment, lacking 5′ to 3′ exonuclease activity) in a total volume of 0.1 ml at 37°C for 60 min. The reaction mixtures were buffered with 50 mM NaH_2PO_4 at pH 7.2 and contained in addition 10 mM $MgCl_2$, 5 mM dithiothreitol, 2 nm each of dGTP, dTTP, and dCTP, 0.2 nm of dATP, and 10 μCi of [α-^{32}P]dATP (specific activity, 400 Ci/mmol, Amersham). The reaction was monitored by measuring the increase in trichloroacetic acid-insoluble material as described previously (3).

(ii) λ Exonuclease. Immediately before use, the stock nuclease preparation (40,000 U/ml) was diluted 1:100 in buffer containing 0.05% bovine serum albumin, 0.001 M dithiothreitol, and 0.01 M Tris-hydrochloride, pH 7.5. Nuclease digestion mixtures contained 67 mM glycine-KOH buffer (pH 9.6), 4 mM $MgCl_2$, 3 mM dithiothreitol, 0.5 mM EDTA, 0.1 to 1.0 μg of DNA and 0.2 to 0.5 U of exonuclease. The reaction was stopped by adjusting the reaction mixture to 0.01 M with respect to EDTA and then extracting the mixture with phenol–chloroform (1:1).

(iii) S_1 nuclease. The reaction conditions used were the same as those described by Bourguignon et al. (3) except that dCTP was not included in the reaction mixture. The reactions were followed by monitoring the decrease in trichloroacetic acid-insoluble material with time as described previously (3).

(iv) Mung bean nuclease. Mung bean nuclease digestion buffer contained 0.05 M sodium acetate (pH 5.0), 0.05 M NaCl, and 0.004 M $ZnSO_4$. Reaction mixtures (0.2 ml) usually contained 7 U of enzyme solution, varying amounts of radiolabeled viral DNAs, 1.5 μg of native salmon sperm DNA, and 1.5 μg of denatured salmon sperm DNA. Digestion was carried out at 37°C for 30 min, and the reaction was stopped by adding EDTA to a final concentration of 0.01 M. Reaction mixtures were then extracted with phenol–chloroform (1:1).

(v) Restriction endonucleases. Reaction mixtures contained in a total volume of 0.05 to 0.1 ml: 0.05 M NaCl, 0.005 M $MgCl_2$, 0.01 M Tris-hydrochloride (pH 7.4), 0.005 M dithiothreitol, and 1 to 2 U of enzyme. Incubation was at 37°C for 30 min, and the reactions were stopped by the addition of EDTA to a final concentration of 0.01 M.

Hydroxylapatite chromatography. DNA samples were adsorbed to a 0.5-ml column of hydroxylapatite (Bio-Rad Laboratories, HTP grade) and then eluted with a 30-ml linear gradient of sodium phosphate buffer, pH 7.2 (0.05 to 0.30 M). Phosphate concentrations were measured by refractive index as determined against a standard curve. Step-wise elutions of single- and double-stranded DNA fractions were performed with 1.0-ml steps of 175 and 300 mM phosphate buffer, respectively. Under these step-wise elution conditions, approximately 10% of the single-stranded DNA elutes in the double-stranded fraction.

BD-cellulose chromatography. DNA samples were adsorbed to a 0.5-ml column of benzoylated DEAE (BD)-cellulose (cellex-BD 100 to 200 mesh, Bio-Rad Laboratories) and eluted step-wise with solutions of TE buffer, pH 7.4, containing either 0.3 M NaCl, 1.0 M NaCl, or 1.0 M NaCl plus 1.8% caffeine (Sigma Chemical Co.).

Purification of (+) strands. One microgram of in vitro-synthesized MVM double-stranded DNA, labeled with ^{32}P in the plus [(+)] strand, was purified by BD-cellulose chromatography. The material eluting in the 1.0 M NaCl fraction was dialyzed against TE buffer (pH 7.4) and then treated with λ exonuclease to selectively degrade the template minus [(−)] strand in the 5′ to 3′ direction (23). The reaction mixture was extracted with phenol–chloroform, and intact ^{32}P-labeled (+) strands were recovered by hydroxylapatite chromatography. Generally, between 40 and 50% of the DNA remains resistant to digestion, even after prolonged incubation with multiple additions of enzyme. Nevertheless, the (+)-strand DNA which is produced comigrates with virion DNA in agarose gels. In addition, hybrids formed between these (+)-strand preparations and virion DNA can be digested with a variety of restriction endonucleases, and by this criterion can be shown to contain all of the wt genome sequence (see Results).

Isolation of restriction fragments for hybridization-selection of type I D-DNA. A 5-μg amount of wt MVM virion DNA (unlabeled) was converted to the duplex form in vitro. The reaction mixtures contained 0.04 M sodium phosphate buffer (pH 7.2); 0.005 M dithiothreitol; 0.008 M $MgCl_2$; 10.0 nm each of dGTP, dCTP, dTTP, and dATP; 3 μCi of [α-^{32}P]dATP (23 Ci/mmol); and 10 U of DNA polymerase in a total volume of 0.31 ml. The reaction mixture was incubated at 37°C for 30 min, adjusted to a concentration of 0.3 and 0.01 M with respect to NaCl and EDTA, respectively, and then extracted with phenol–chloroform (1:1). The duplex MVM DNA was then purified by BD-cellulose chromatography. The 1.0 M NaCl fraction was recovered and dialyzed against TE buffer, pH 7.4, and the DNA was cleaved with the restriction endonuclease *Hin*dIII. This enzyme cleaves wt RF-DNA

twice, at map positions 52 and 78, yielding three restriction fragments (5). After digestion, the DNA was again extracted with phenol-chloroform (1:1), diluted in 2.5 ml of distilled water, and then heated at 96°C for 5 min and quick-cooled at 4°C in an ice-water bath. The DNA was then immediately adsorbed to a BD-cellulose column and eluted step-wise as described above. The snap-back *Hin*dIII fragment, 0/52, was recovered in the 1.0 M NaCl eluate, whereas the denatured fragments 52/78 and 78/100 eluted in the caffeine fraction. The purity of fragments isolated in this way was confirmed by gel electrophoresis. The 3′ terminal *Hin*dII fragment, 0/36 (Faust and Ward, unpublished data), was prepared by digesting the *Hin*dIII 0/52 fragment with *Hin*dII and again isolating the snap-back 3′-terminal fragment by BD-cellulose chromatography. Before use for hybridization with type I D-DNA, the purified *Hin*dII fragment 0/36 was treated with *Hha*I, which cleaves at a number of sites within this restriction fragment. The DNA fragments prepared in this way had a final specific activity of 3 $\times 10^4$ to 5 $\times 10^4$ cpm/μg with respect to the (+) strand.

Electron microscopy. The formamide modification of the Kleinschmidt technique described by Davis et al. (8) was used throughout. In general, the procedure used was similar to that described by Singer and Rhode (31). Before use, formamide (Fisher) was vacuum distilled at 112°C and 25 mm of Hg. The hypophase solution contained 10% formamide (vol/vol), 0.01 M Tris-hydrochloride (pH 8.5), and 0.001 M EDTA. The spreading solution was made up by mixing 20 μl of formamide and 25 μl of a solution containing 0.1 M Tris-hydrochloride (pH 8.5), 0.01 M EDTA, 5 μg of cytochrome *c*, and 25 to 50 ng of DNA. The total volume of the spreading solution was 0.05 ml. The DNA was rotary shadowed with platinum-palladium (80:20) and examined in a Philips 300 electron microscope with a lens current of 40 kV.

Gel electrophoresis. Polyacrylamide slab gels (4%, 150 ml) were formed by mixing 30 ml of acrylamide (20%), 10 ml of bis-acrylamide (2%), 15 ml of a 10× solution of Jeppesen buffer (17) and 90 ml of distilled water. The mixture was degassed, and the acrylamide was polymerized by adding 0.75 ml of an ammonium persulfate solution (10%) and 0.75 ml of *N,N,N′,N′*-tetramethylethylenediamine (10%). Stacking gels (3.5%) were formed as described by Jeppesen (17). Just before electrophoresis, DNA samples received ⅛ volume of a solution containing 0.05 M EDTA, 1.0 M urea, 1.0 M sucrose, and the tracking dyes bromophenol blue and xylene cyanol FF. Electrophoresis was carried out at 0.5 to 1.0 mA/cm in a 40-cm slab gel for 18 to 24 h. Gradient gels were formed as described by Jeppesen (17). For autoradiography, gels were dried under vacuum and exposed to Kodak XR-5 double-sided X-ray film. Gels were prepared for fluorography by the method of Bonner and Laskey (2a) and exposed to presensitized film (21) at −70°C.

RESULTS

Band purification of defective virions in gradients of CsCl and isolation of incomplete genomes. In previous studies, full infectious virions of MVM (ρ = 1.41 g/ml) were resolved from empty virions (ρ = 1.32 g/ml) by centrifugation to equilibrium in CsCl density gradients (37). In addition, it was noted that these gradients contained a population of virions banding at intermediate densities which, although labeled with [^{3}H]thymidine, lacked infectivity. In the present study, these noninfectious particles have been obtained free of both full and empty virions by repeated banding in CsCl gradients. Typically, defective virions band as a heterogeneous population with densities in CsCl ranging from approximately 1.33 to 1.38 g/ml. This distribution is seen by labeling in vivo with [^{3}H]thymidine or inorganic $^{32}PO_4$, by measuring the UV absorption at 260 nm, and also by measuring the hemagglutination activity (Fig. 1). In either case, discrete peaks are not observed, and the various measurements yield the same density distribution of viral particles. Also, on the basis of hemagglutination activity, the relative amounts of full and defective virions in these gradients are about the same.

An estimate of the average size of the D-DNA produced during a single high-multiplicity passage was obtained initially by sedimenting purified defective virions in alkaline sucrose gradients (Fig. 2). Under these conditions, the DNA appears relatively heterogeneous in size and, by comparison with wt virion DNA, has an average sedimentation coefficient of 11 to 12S (34). Thus, D-DNA has an average single-stranded molecular length of approximately 2,000 nucleotides, i.e., roughly 40% of the size of the intact viral genome. Additional measurements of the average size of this DNA population, which includes electron microscopic contour length measurements and sedimentation analysis in the presence of DNA markers of known molecular weight, confirm this initial size estimate (see Fig. 12 and 14). It should be noted in addition that the size of the DNA varies with particle density; particles which band at a greater density in CsCl contain larger DNA molecules than those which band at a lesser density (Table 1). Defective virions of the related parvoviruses Lu III and AAV have been reported to have similar characteristics (10, 22, 24). All further analyses of D-DNA were performed on the material which was pooled after the sedimentation of purified defective virions in preparative alkaline sucrose gradients.

Resolution of two types of D-DNA. The existence of two subpopulations of D-DNA was demonstrated initially by chromatography on hydroxylapatite and by digestion of the D-DNA with the single-strand-specific nuclease S_1. The latter experiments were performed by mixing ^{32}P-labeled D-DNA and [^{3}H]thymidine-labeled wt virion DNA and subjecting the mixture to

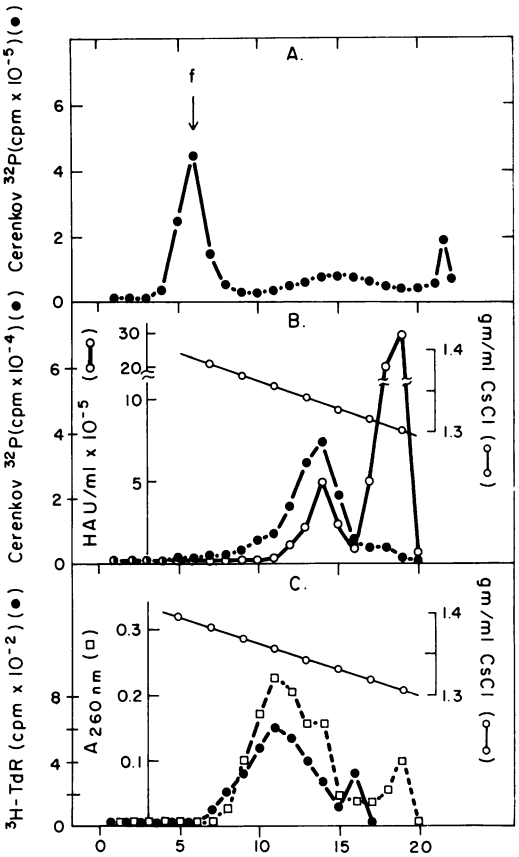


FIG. 1. *Analysis of incomplete MVM particles by buoyant density in CsCl. (A) Partially purified virions labeled in vivo with [^{32}P]phosphate were recovered from a CsCl-sucrose step gradient and dialyzed against TE buffer, pH 8.7. After micrococcal nuclease treatment, virions were sedimented to equilibrium for a second time, as described in the text. The position of full virus in the gradient is designated by (f). (B) Incomplete and empty virions were recovered after centrifugation as in (A) and banded to equilibrium for a third time as described in the text. Fractions were assayed for ^{32}P radioactivity and hemagglutination. (C) Incomplete virions labeled in vivo with [^{3}H]thymidine ([^{3}H]TdR) were recovered after the third cycle of CsCl equilibrium sedimentation and dialyzed against TE buffer, pH 8.7. A sample was then centrifuged to equilibrium in CsCl under the third-cycle conditions (see text). Gradient fractions were monitored for ^{3}H radioactivity by liquid scintillation counting and for UV absorption at 260 nm (A_{260}). HAU, Hemagglutinating units.*

digestion with nuclease S_1. As expected, roughly 7 to 10% of the wt DNA in this analysis remains S_1 resistant (Fig. 3). This is in agreement with the results obtained by Bourguignon et al. (3) and Chow and Ward (5), who demonstrated that this low level of S_1 resistance is due to the presence of two small duplex regions which comprise the hairpin termini of the viral genome. In contrast, D-DNA exhibits a considerably greater degree of S_1 resistance. The value obtained in the experiment illustrated in Fig. 3 is around 55%. Moreover, a significant portion of the D-DNA remains S_1 resistant even after heat denaturation and quick-cooling, suggesting in addition that a significant portion of the D-DNA renatures spontaneously. In a separate analysis and in the way of a control experiment, virtually all of the ^{32}P-labeled material in the D-DNA preparation was degraded by pancreatic DNase.

For chromatography on hydroxylapatite, ^{32}P-labeled D-DNA was adsorbed to the column together with [^{3}H]thymidine-labeled single-stranded and double-stranded positional markers, and the column was developed with a sodium phosphate gradient. As can be seen in Fig. 4, the D-DNA is resolved into two fractions. We conclude on the basis of these results and the results of Fig. 3 that the D-DNA population consists of single-stranded as well as double-stranded molecules and that these two subpop-

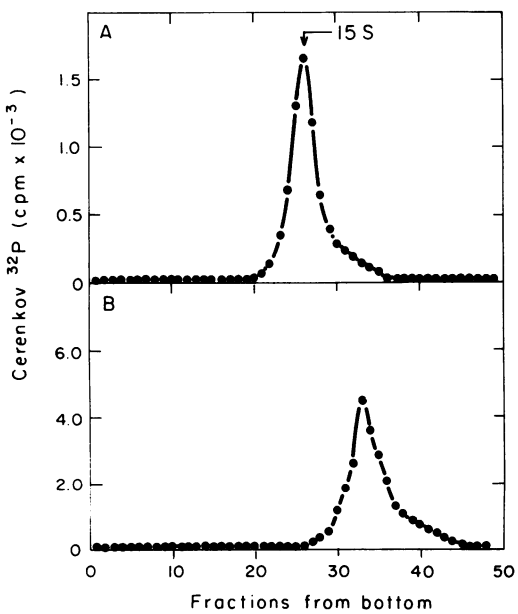


FIG. 2. *Alkaline sucrose gradient centrifugation of ^{32}P-labeled MVM DNA from purified wt full (A) and incomplete (B) virions. The gradients (5 to 20% sucrose) containing 0.3 N NaOH, 0.7 N NaCl, and 0.15% Sarkosyl were run for 16 h at 38,000 rpm (5°C) in a Beckman SW41 rotor. Fractions were collected by bottom puncture, and the distribution of radioactivity was determined by Cerenkov counting.*

TABLE 1. *Relationship between the buoyant density of defective particles in CsCl and the molecular weight of defective DNA*

ρ in CsCl (g/cm³)[a]	S value of DNA	Mol wt	% Genome
1.370–1.395	8–15	0.28×10^6–1.40×10^6	20–100
1.355–1.370	7–13	0.21×10^6–0.92×10^6	16–73
1.341–1.355	6–12	0.15×10^6–0.73×10^6	11–52
1.331–1.341	5–11	0.09×10^6–0.60×10^6	7–43
1.300–1.331	4–8	0.06×10^6–0.28×10^6	5–20

[a] The relative abundance of each subclass of defective virions varies from preparation to preparation; however, no fraction represents less than 10 to 15% of the total population.

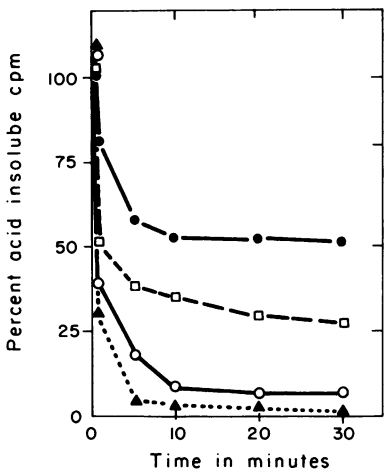


FIG. 3. *Kinetics of S_1 endonuclease digestion of wt MVM DNA* (○) *and incomplete virion DNA, before* (●) *and after* (□) *heat denaturation and quick-cooling. Pancreatic DNase digestion of incomplete virion DNA* (▲).

ulations are present in roughly equal amounts. This distribution has been observed for several independent D-DNA preparations derived from virus stocks passaged a single time at a high multiplicity of infection. Evidently, the double-stranded molecules exist in a hairpin configuration and consist of covalently linked (+) and (−) strands; further evidence establishing this point is presented in the experiments to be described below. The remainder of the paper is devoted, for the most part, to the analysis of the single-stranded (type I D-DNA) and double-stranded (type II D-DNA) fractions which were obtained after preparative fractionation of the total D-DNA population by hydroxylapatite column chromatography.

Mung bean nuclease digestion of type I D-DNA. Previous studies have established that the wt MVM genome contains hairpin duplexes at both genome termini which can be detected as nuclease-resistant fragments after digestion with single-strand-specific nucleases (5, 7). The smaller fragment is a hairpin 115 nucleotides in length and constitutes the 3′ terminus of the viral genome. The larger 130-base-pair fragment constitutes the hairpin structure derived from the 5′ terminus. To determine whether these small duplex regions are present in type I D-DNA, a sample of the DNA was treated with the single-strand-specific mung bean nuclease, and the digestion products were analyzed by electrophoresis in an 8% polyacrylamide gel. For purposes of comparison, wt virion DNA was analyzed in parallel (Fig. 5). In the case of the wt genome, nuclease digestion yields two nuclease-resistant fragments. Mung bean nuclease digestion of type I D-DNA also yields nuclease-resistant fragments. One of these comigrates precisely with the 115 nucleotide fragment seen in digests of wt DNA. The other nuclease-resistant band is seen as a doublet and migrates distinctly slower than the 130-base-pair fragment found in digests of wt DNA. Using *Hha*I restriction fragments of φX174 RF as size markers, we have estimated these larger fragments to be between 140 and 145 base pairs in length. Although the basis for the apparent difference in size between these fragments and those derived from the 5′ terminus of wt DNA is not known, several possibilities can be suggested (see Discussion). Comparing the band intensities of nuclease-resistant fragments which are obtained in the case of both the wt and defective genomes and taking into account the smaller size of the D-DNA molecules, we have estimated that these small duplex regions are contained in a majority and possibly all of the molecules in the D-DNA population. Moreover, since they remain associated with the D-DNA after exposure to alkaline conditions, it is also likely that these small duplex regions constitute hairpin termini in the D-DNA. This conclusion is supported by the results of the restriction enzyme analysis described below.

Self-primed in vitro synthesis of type I D-RF and cleavage of D-RF with restriction endonucleases. To examine the sequence organization within type I genomes, the DNA was

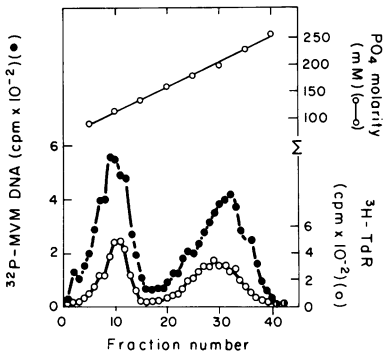


FIG. 4. *Hydroxylapatite chromatography of ^{32}P-labeled MVM D-DNA (●) in the presence of a mixture of wt MVM virion DNA and simian virus 40 RFII DNA (○), both labeled with [^{3}H]thymidine ([^{3}H]-TdR).*

first converted in vitro to a double-stranded form suitable for restriction enzyme analysis. Previous studies have shown that wt virion DNA is an efficient template for several DNA polymerases (3). The reaction proceeds by a self-priming mechanism and results in the conversion of wt virion DNA to a complete monomer length duplex consisting of covalently linked (+) and (−) strands. Type I D-DNA molecules also possess a 3′-terminal hairpin structure suitable for self-primed DNA synthesis, and this structure appears to be similar, if not identical, to that found in the wt genome. In this series of experiments, *E. coli* DNA polymerase I (lacking 5′ to 3′ exonuclease activity) was used to convert type I D-DNA into a duplex form (designated D-RF DNA) in vitro. The overall kinetics of this reaction are similar to those obtained when wt DNA is used as a template (data not shown). A restriction enzyme analysis of the DNA polymerase products is presented in Fig. 6, and physical maps for restriction endonucleases *Hin*fI and *Mbo*I are given in Fig. 7. When compared with the cleavage pattern of wt RF DNA, certain restriction fragments found in D-RF DNA digests are judged on a relative basis to be present in less than molar amounts. These include *Mbo*I fragments 47.3/53.6 (G), 53.6/67.0 (C), 67.0/72.2 (H), and 72.2/87.1 (B). These fragments map contiguously on the viral genome between positions 47.3 and 87.1. Examination of a parallel *Hin*fI digest confirms that the deletions occur predominantly in this region. Accordingly, in a *Hin*fI digest of D-RF DNA, *Hin*fI fragments 17.8/49.4 (A), 49.4/54.9 (G), and 54.9/86.5 (B) are severely underrepresented. [Although *Hin*fI fragments 17.8/49.4 (A) and 54.9/86.5 (B) are not resolved on the gels illustrated in Fig. 6, secondary digestion of D-RF DNA with *Hin*dIII confirms that both *Hin*fI fragments 17.8/49.4 (A) and 54.9/86.5 (B) are present.] In comparison with these internal deletions, restriction fragments which represent terminal portions of the viral genome appear to be relatively well conserved. For example, the 3′-terminal *Hin*fI fragment 0/4 (H) and the 3′-terminal *Mbo*I fragment 0/28.5 (A) are both well represented in these digests. Also seen in relatively good yield are the 5′-terminal *Mbo*I and *Hin*fI fragments 91.5/100 (D) and J and 86.5/94.2 (D) and 94.2/100 (F), respectively. Thus, restriction enzyme digests of the total D-RF DNA population exhibit a gradient of band intensities which, in terms of the physical map, decreases toward the internal regions of the genome and increases toward the molecular termini.

Essentially identical results were obtained in a separate restriction enzyme analysis involving hybrids formed between type I D-DNA and

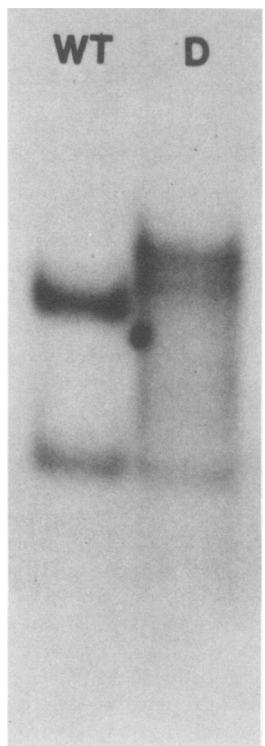


FIG. 5. *Polyacrylamide slab gel electrophoresis of MVM wt DNA and type I D-DNA, both labeled with [^{3}H]thymidine after digestion with mung bean nuclease. Nuclease-resistant digestion products were analyzed in an 8% polyacrylamide gel. Electrophoresis was for 6 h at 40 mA. The gel was dried and processed for fluorography.*

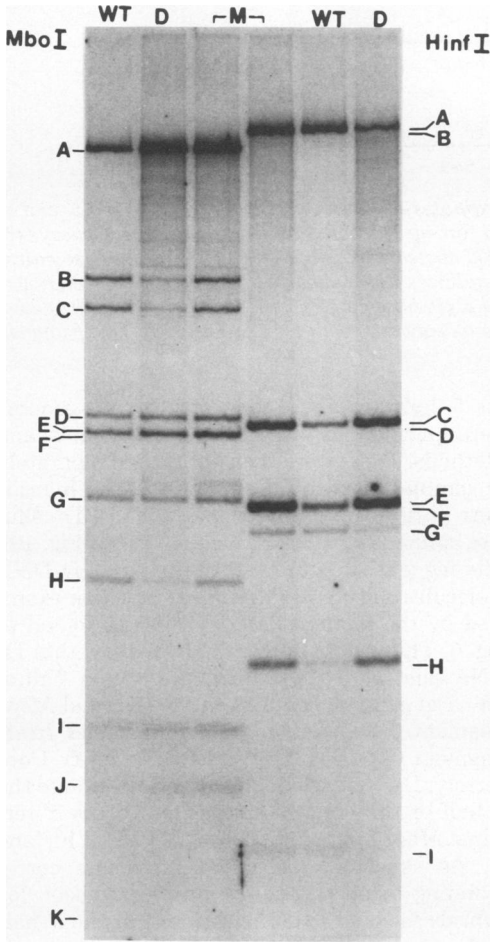


FIG. 6. *Polyacrylamide slab gel electrophoresis of duplex MVM DNA fragments generated by cleavage of wt RF DNA and D-RF DNA with MboI and HinfI restriction endonucleases. Both types of RF-DNA were synthesized in vitro with E coli DNA polymerase (see text) and contained ^{32}P label in the (+) strand. The channels denoted M contained a mixture of wt DNA and D-DNA as template for the polymerase reaction. The products were then digested with MboI and HinfI as before.*

purified, ^{32}P-labeled (+) strands. The *Mbo*I and *Hin*fI digests of these hybrids are illustrated in Fig. 8. Again, by comparing the relative intensities of different bands in these gels, it is apparent that virtually all of the wt sequence is present, but that certain internal regions are underrepresented. Note, for example, the relatively low intensities of the *Hin*fI fragments 17.8/49.4 (A) and 54.9/86.5 (B) in digests of the D-DNA-containing hybrids (Fig. 8, lane f). An additional noteworthy feature is that the 3′-terminal restriction fragments seen for the hybrid molecules formed either with wt or D-DNA have altered mobilities when compared with the corresponding fragments present in a digest of wt RF DNA (see arrows in Fig. 8). These changes in mobility correspond to an apparent increase in size of 50 to 100 base pairs. One would expect such an increase if the (+) strands used for hybridization contained short duplex regions at their 5′ termini. This is consistent with the action of λ exonuclease, which is known to leave a small region of residual double-helical structure in substrate molecules after maximal digestion (23).

Although an overall deletion of wt nucleotide sequence can be discerned in the total D-RF population by these techniques, further analyses are necessary to determine the sequence organization in individual molecules. Two experiments which address this problem are presented below.

Restriction enzyme analysis of type I D-DNAs selected by hybridization. Type I D-DNA was fractionated by size as indicated in Fig. 9 (inset), and the sequence organization of these individual D-DNA populations was then examined by restriction enzyme analysis after conversion of the D-DNA to the D-RF DNA form. It is evident that the sequence complexity of the D-DNA decreases as the DNA becomes smaller (see Fig. 9). Moreover, the decrease in complexity derives from a progressive loss of the DNA sequences which map at various internal positions in the viral genome. Accordingly, as the DNA becomes smaller, there is a rapid disappearance of the internal *Hin*fI fragments 17.8/49.4 (A), 49.4/54.9 (G), and 54.9/86.5 (B), whereas the terminal fragments are conserved. Thus, the smallest DNA population appears to consist almost entirely of the 3′-terminal *Hin*fI fragment 0/4 (H) and the 5′-terminal *Hin*fI fragment 94.2/100 (F). Since the *Hin*fI fragments 4.0/9.8 (E) and 94.2/100 (F) are not resolved on the gels shown in Fig. 9, it is possible that both of these fragments are present. However, we conclude that the majority, if not all, of this band represents the 5′-terminal *Hin*fI fragment F on the basis of a parallel analysis of the D-RF DNA after cleavage with *Mbo*I (data not shown). In this study, a similar reduction in the sequence complexity was observed as the size of the D-DNA decreased. The increase in the relative molar abundance of the 5′-terminal *Mbo*I fragment 91.5/100 (D) observed in the D-DNA from pools B and C can be rationalized only if the majority of the *Hin*fI-E/F band represents the 5′-terminal *Hin*fI fragment F.

Proof that type I D-DNAs are true recombinants, with sequences from both genomic ter-

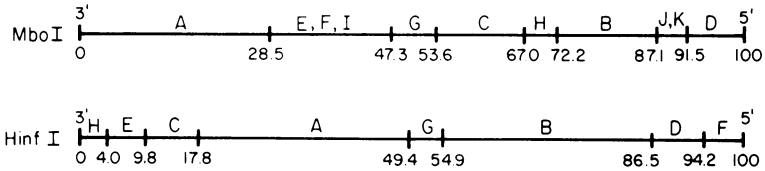


FIG. 7. *Cleavage maps of wt MVM RF DNA. The coordinates are numbered from the 3' to the 5' end in accordance with the convention adopted at the 1978 Cold Spring Harbor Symposium on Parvoviruses (42). The percentage values shown refer to a genome size of 5,000 nucleotides. MboI fragments E (390 base pairs), F (390 base pairs) and I (190 base pairs) map within the boundaries shown, but their precise order is unknown. The same is true for MboI fragments J (175 base pairs) and K (80 base pairs). A 100-base-pair HinfI fragment (I) has also not been mapped definitively, but it is known to lie between positions 17.8 and 94.2. This fragment has been omitted from the map shown.*

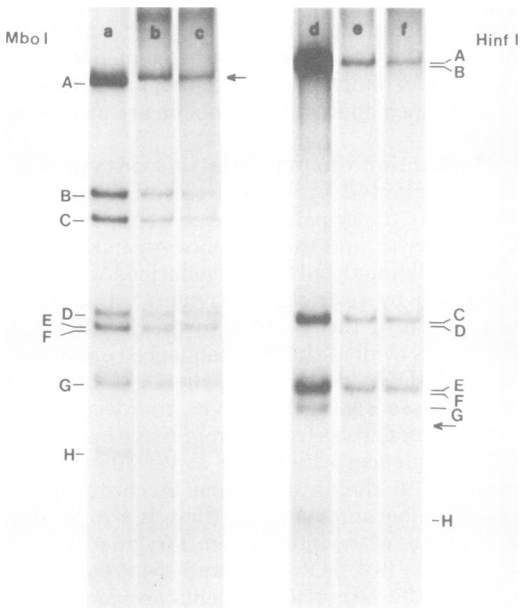


FIG. 8. *Restriction enzyme analysis of hybrids formed between type I D-DNA and purified ^{32}P-labeled (+) strands derived from wt RF. A 3-ng amount of purified (+) strand DNA was hybridized with either 50 ng of wt virion DNA or 70 ng of type I D-DNA. Hybrids were formed in TE buffer, pH 7.5, containing 0.1 M NaCl by incubating at 68°C for 3 h. Restriction enzyme digestion buffer was then added and the hybrids were incubated with either MboI or HinfI for 30 min at 37°C. Digests were analyzed by electrophoresis in a 4% polyacrylamide slab gel. (a) and (d), MboI and HinfI digests of wt RF; (b) and (e), MboI and HinfI digests of hybrids formed between purified (+) strands and wt virion DNA; (c) and (f), MboI and HinfI digests of hybrids formed between purified (+) strands and type I D-DNA. Arrows indicate position of altered 3'-terminal fragments (see text).*

mini present together on individual molecules, was obtained in the following experiment. Restriction fragments representing regions within the 5' half and the 3' half of the viral genome were purified as described in Materials and Methods. These were then hybridized separately to samples of type I D-DNA, and the hybrids were purified on hydroxylapatite and BD-cellulose columns (for details, see legend to Fig. 10). The sequence organization of the type I D-DNA molecules selected by hybridization was examined by the method described in the legend to Fig. 6. The results of this analysis show that D-DNA selected by hybridization to the 5' half of the viral genome contains the 3'-terminal *Mbo*I fragment 0/28.5 (A) and the 3'-terminal *Hin*fI fragment 0/4 (H) (Fig. 10, lanes c and f). Conversely, D-DNA selected by hybridization to the 3' half of the viral genome contains the 5'-terminal *Mbo*I fragment 91.5/100 (D) (Fig. 10, lane b). As expected, restriction fragments corresponding to both genome termini are seen for hybrids selected either with the 3' or the 5' half of the viral genome, and the deletions appear to be the same as those seen for the total D-DNA population. The extra bands seen in these gels are considered to represent incorporation of label into the DNA used for hybridization (see legend to Fig. 10).

These results, taken together with the data presented in Fig. 9, suggest that individual molecules of type I D-DNA are recombinant genomes which contain sequences derived from both genomic termini, although the extent and position of the deletions in individual molecules vary.

Structure of type II D-DNA. The results of S_1 nuclease digestion presented in Fig. 3 indicated that type II D-DNAs are double-stranded molecules which exist in a hairpin configuration. The following additional analyses were performed to further establish this point.

Type II D-DNA was recovered as the double-stranded fraction from a preparative hydroxylapatite column and subjected to a variety of denaturation and renaturation conditions. The hydroxylapatite chromatograms of these treated

samples of type II-DNA are shown in Fig. 11: (A) native DNA; (B) denatured DNA; (C) mung bean nuclease treated and denatured; (D) treatment as in C followed by complete renaturation; (E) treatment as in D with renaturation in the presence of a 20-fold molar excess of wt virion DNA. The following conclusions can be drawn from these experimental results. First, approximately 85% of D-DNA renatures spontaneously; second, this property of the DNA can be overcome for the most part by digestion with a single-strand-specific nuclease before denaturation; third, the spontaneously renaturable properties of type II D-DNA can be attributed to a covalent association of the complimentary strands via relatively large single-stranded loops; and fourth, at least 75% of the type II D-DNA is virus specified. These conclusions are supported by the electron microscopic observations and restriction enzyme analyses outlined below.

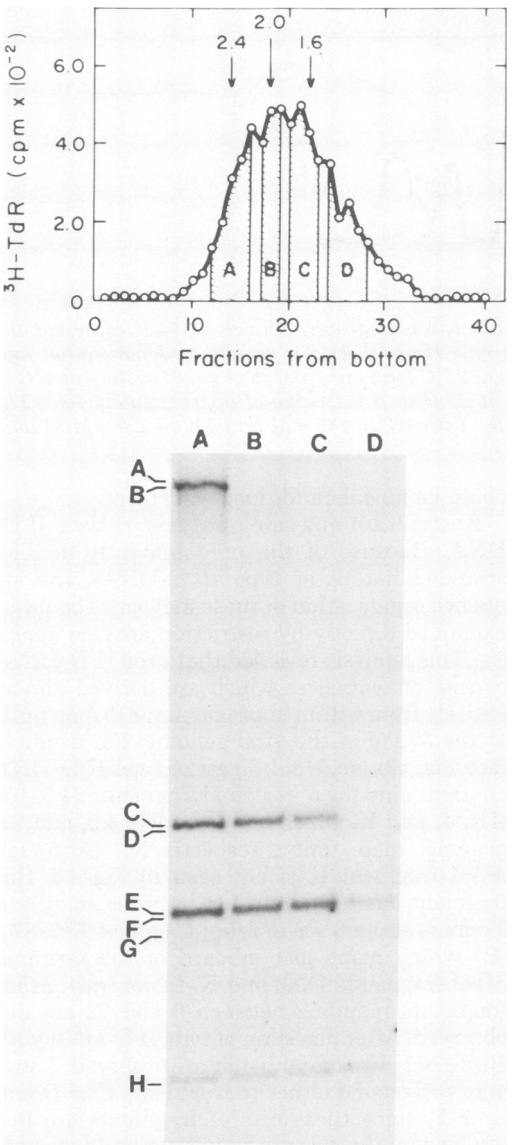


FIG. 9. *Restriction enzyme analysis of type I D-DNA fractionated on the basis of size. Various size classes of D-DNA were pooled from an alkaline sucrose gradient as indicated in the figure. After dialysis, the DNA was converted to D-RF in vitro and cleaved with the restriction endonuclease HinfI. The digests were then analyzed by electrophoresis in 4% polyacrylamide slab gels.*

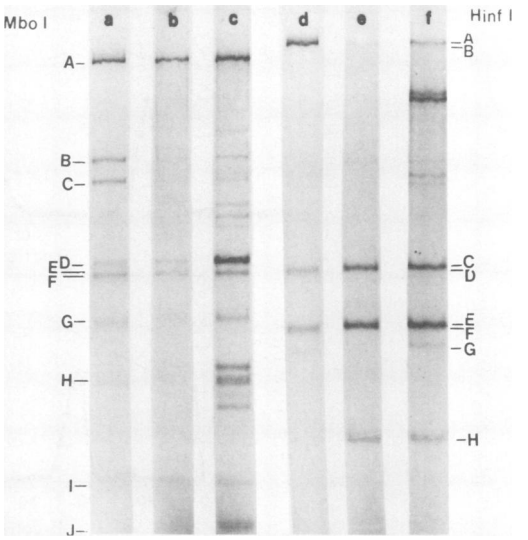


FIG. 10. *Restriction enzyme analysis of type I D-DNA selected by hybridization. Approximately 0.10 μg of ^{3}H-labeled type I D-DNA in TE, pH 7.4, was mixed with 0.10 to 0.20 μg of restriction fragments derived from the left-end (3') or right-end (5') portions of the wt genome. The mixtures were then heated at 96°C for 5 min and cooled to 4°C in an ice-water bath. The solution was adjusted to 0.1 M with respect to NaCl and incubated at 68°C for 2 h in a total volume of 0.5 ml. Hybrids were purified by step-wise elution from hydroxylapatite. The double-stranded hydroxylapatite fraction was then adsorbed to a BD-cellulose column, and the caffeine fraction containing the hybridized D-DNA was recovered and dialyzed against TE, pH 7.4. The type ID-DNA in this fraction was used as template and converted to D-RF DNA in the reaction with DNA polymerase I (lacking 5' to 3' exonuclease activity). The polymerase products, labeled with [α-^{32}P]dATP, were cleaved with restriction endonucleases, and the digests were analyzed by electrophoresis in a 4% polyacrylamide slab gel. (a) and (d), wt RF DNA; (b) and (e), D-DNA selected by hybridization to restriction fragments derived from the left-end (3') portion of the wt genome; (c) and (f), D-DNA selected by hybridization to restriction fragments derived from the right-end (5') portion of wt MVM DNA.*

Unfractionated D-DNA preparations contain-

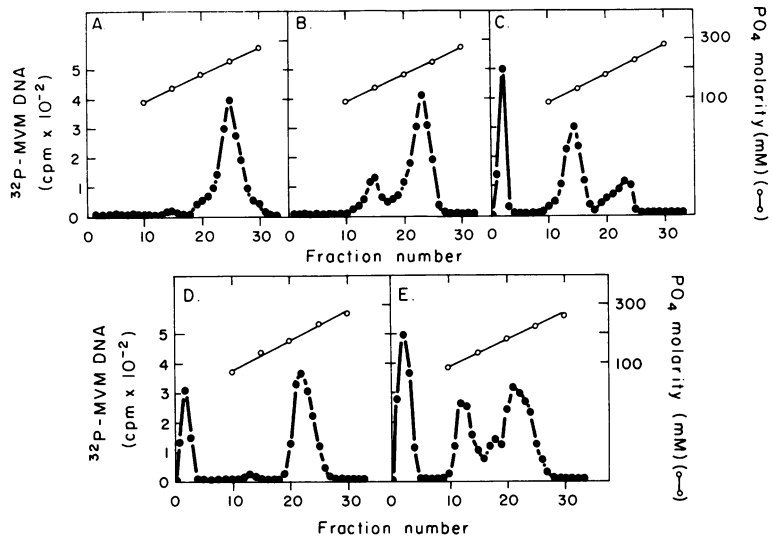


FIG. 11. *Analysis of type II D-DNA by hydroxylapatite chromatography. Panels A to E represent the hydroxylapatite elution profiles obtained by subjecting type II D-DNA to a variety of denaturation and renaturation conditions. DNA was denatured by heating at 96°C for 5 min, and then quick-cooling at 4°C in an ice-water bath. Renatured samples were incubated at 68°C to a C_0t value of approximately 10^{-2}. The single-stranded and double-stranded fractions were eluted with 100 to 140 mM and 190 to 220 mM sodium phosphate, respectively.*

ing both type I and type II molecules were examined in the electron microscope. The majority of the molecules observed are linear and have an average contour length of 0.2 to 0.3 μm (n = 65). For single-stranded DNA, this corresponds to a molecular length of ~1,800 nucleotides (31). The broad range of contour lengths observed (from 0.1 to 0.5 μm) relative to measurements made on molecules of φX174 DNA (0.9 ± 0.1 μm) confirms the size heterogeneity of the D-DNA population. A significant proportion of the molecules observed, about 30%, appear to consist of double-stranded stems and single-stranded loops (Fig. 12). These "panhandle" structures are also heterogeneous in size. In addition, there is considerable variation in the size of both the loops and the stems in these molecules. For example, based on contour length measurements, the loops range in size from 300 to 1,000 nucleotides. This interpretation of the structures depicted in Fig. 12 is supported by the following independent observations. DNA samples treated with S_1 nuclease contain a much lower proportion of molecules judged to have "panhandle" structures than do untreated samples. Circular structures lacking stems are not observed. The size distribution of the S_1-resistant molecules is similar to that seen for the stems in panhandle structures. Finally, only about 2% of the molecules examined appear as cross-over structures in which both ends of the molecules are free, making it unlikely that these could be mistaken for loop structures.

Restriction enzyme analysis of type II D-DNA. Because of the predominantly double-stranded nature of type II D-DNAs, the sequence organization of these molecules could be examined directly by restriction enzyme analysis. This analysis revealed that type II D-DNAs consist of sequences which are derived almost entirely from within approximately 15 map units at the 5′ end of the viral genome. For example (see Fig. 13), an *Mbo*I digest of type II D-DNA contains only the 5′-terminal fragments 91.5/100 (D), J, and K, which comprise 9.0, 4.0, and 2.0 genome map units, respectively. (Although *Mbo*I fragment K is not seen in Fig. 13, this fragment has been detected in other analyses). There is also a trace of *Mbo*I fragment 72.2/87.1 (B) which maps just inboard of the terminal *Mbo*I fragments D, J, and K. In contrast, *Mbo*I fragments mapping between 0 and 72 are not observed. After digestion of type II D-DNA with *Hin*fI, only two major bands are observed. These may correspond either to fragments C or D and E or F, since these sets of fragments are not resolved on the gels presented in Fig. 13. Nevertheless, since *Hin*fI fragments 86.5/94.2 (D) and 94.2/100 (F) map together at the 5′ end (see map, Fig. 7) and since we do not detect either the 3′-terminal *Hin*fI fragment 0/4 (H) or the internal fragments 17.8/49.4 (A), 49.4/54.9 (G), and 54.9/86.5 (B), this result is considered to be in support of the conclusion derived from the

*Mbo*I digestion data. In addition, there is one band seen in each of these digests (arrows in Fig. 13) which does not correspond to any of the restriction fragments found in digests of wt RF DNA. These bands may represent low yields of the 5'-terminal *Mbo*I and *Hin*fI fragments which exist in the "foldback" rather than the extended configuration (41).

Selective enrichment for the genome termini in type I D-DNAs during serial high-multiplicity passage. The incomplete genomes arising after several undiluted passages of a mixture of wt and incomplete virions were found to differ in several respects from the single-passage D-DNA already described. Whereas the single-passage yield of D-DNA consists of the type I and type II subpopulations in roughly equal amounts, we have observed a progressive loss of the type II molecules during successive cycles of infection. Accordingly, at passage 10, the D-DNA population consists almost entirely (95% or more) of type I structures as determined by hydroxylapatite chromatography (data not shown). A second major change concerns the average size of the D-DNA. Although still extensively heterogeneous, the D-DNA obtained after the seventh undiluted passage is on the average only about 1,200 nucleotides in length, as compared with an original size of 2,000 nucleotides (Fig. 14). To derive some information about the possible significance of these changes, we have also examined the sequence complexity of the

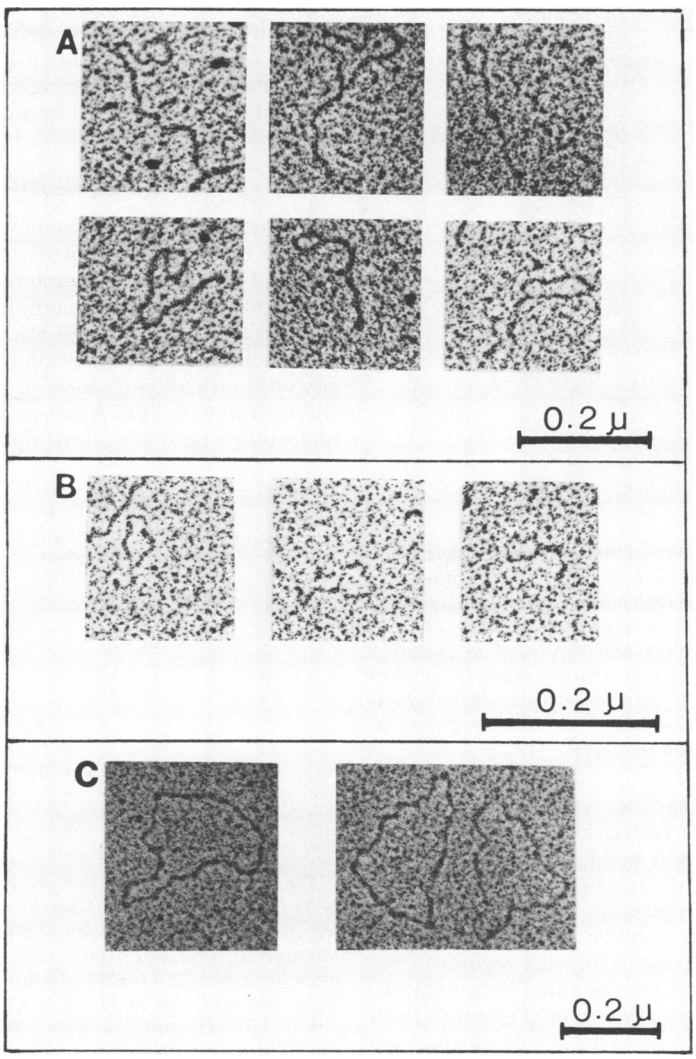


FIG. 12. *Electron micrograph of type II D-DNA molecules exhibiting a panhandle configuration. Molecules shown in (B) and (C) are single-stranded circular φX174 DNA.*

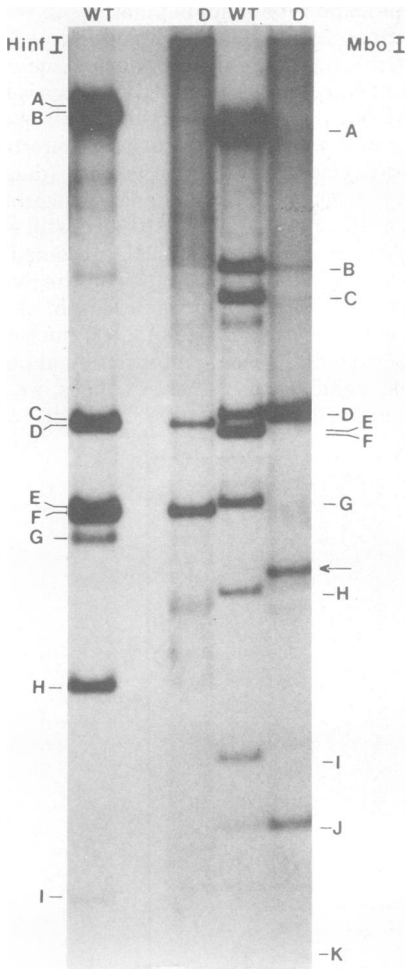


FIG. 13. *Restriction enzyme analysis of type II D-DNAs. Molecules labeled in vivo with ^{32}P-phosphate were cleaved with the restriction endonucleases HinfI and MboI. In vitro synthesized wt RF was cleaved with the same enzymes and analyzed in parallel. Restriction enzyme fragments were resolved by electrophoresis in a 3.5 to 7.5% polyacrylamide gradient slab gel. Arrow denotes the 5'-terminal fragment in foldback configuration (see text).*

type I D-DNA molecules as a function of passage. As shown by restriction enzyme analysis, there is a significant reduction in the overall sequence complexity in the type I population which is related to passage number. Moreover, this reduction in sequence complexity derives from a progressive loss of internal regions of the viral genome and is accompanied by an enrichment of the sequences which map closer to both genomic termini. Thus, in digests of late-passage D-RF DNA, we have observed a relative increase in the band intensities of the terminal *Mbo*I fragments 91.5/100 (D), J, and 0/28.5 (A) as compared with the internal fragments E + F, 47.3/53.6 (G), 67.0/72.2 (H), 53.6/67.0 (C), and 72.2/87.1 (B) (Fig. 15). This trend is confirmed by the *Hinf*I digestion pattern of late-passage D-DNA, since we see a decrease in intensity of the internal fragments 17.8/49.4 (A), 49.4/54.9 (G), and 54.9/86.5 (B), whereas the 3'-terminal *Hinf*I fragment 0/4 (H) is relatively well conserved. It is also apparent from the data presented in Fig. 15 that there are a number of bands in these gel profiles which do not correspond to restriction fragments found in digests of wt RF DNA. Moreover, these also are not present in digests of single-passage D-RF DNA. (For example, compare lane a with lanes b and c in Fig. 15.) Thus, after a number of high-multiplicity passages, the defective genomes of MVM may acquire cellular DNA or undergo extensive sequence rearrangements. The possible implications of these findings are presented below.

DISCUSSION

In the present study, we have obtained information relating to the structure and sequence organization of the incomplete genomes of the parvovirus MVM. The incomplete genomes arising during a single high-multiplicity passage were isolated by sedimenting purified, low-density virions in alkaline sucrose gradients, and, by this criterion, the DNA was assigned an average molecular length of approximately 2,000 nucleo-

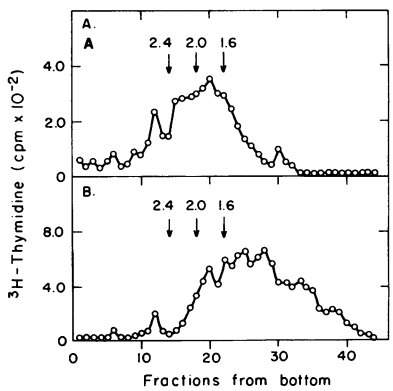


FIG. 14. *Alkaline sucrose gradient sedimentation of D-DNA obtained after serial undiluted passage. Purified incomplete virions obtained after the first passage (A) and after seven undiluted passages (B) were layered on a 5 to 20% alkaline sucrose gradient. Centrifugation was at 38,000 rpm and at 4°C with an SW41 rotor. The numbers above the arrows designate the DNA size in kilobases as determined by sedimentation of restriction fragments of wt RF DNA of known size.*

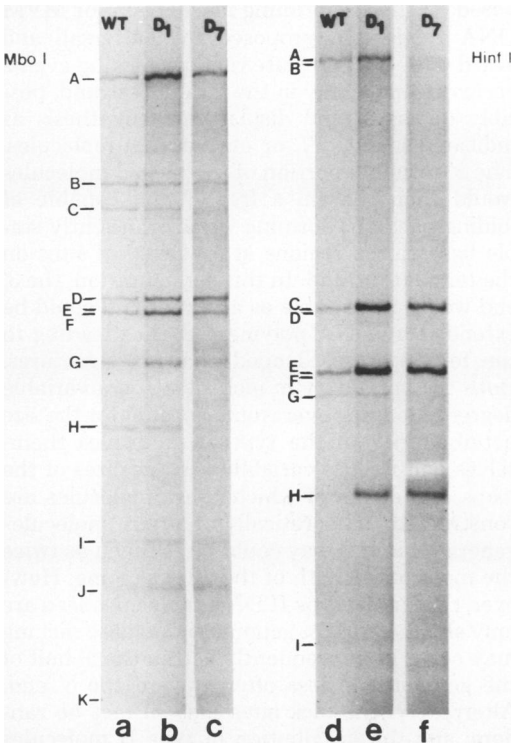


FIG. 15. *Restriction enzyme analysis of D-RF DNA prepared from type I D-DNA obtained after one (D_1) and seven (D_7) serial undiluted passages.*

tides. In addition, based on a variety of experimental observations, we have concluded that the incomplete genomes exist as two structurally distinct populations, which can be easily distinguished on the basis of their degree of secondary structure.

Type I D-DNAs are a heterogeneous collection of predominantly single-stranded, recombinant molecules differing in size and sequence content. As a whole, this D-DNA population contains virtually all of the sequences found in wt virion DNA, although certain regions are clearly underrepresented. These regions have been located based on the results of restriction enzyme analysis and were found to map between positions 47.3 and 87.1 in the 5′-terminal half of the viral genome. Significantly, the smallest molecules in the D-DNA population (pool D, Fig. 9) lack approximately 90% of the wt genome sequence and appear to consist exclusively of sequences which map within each of the terminal *Hin*fI restriction fragments, regions which include only about 200 and 300 nucleotides at the 3′ and 5′ termini of the viral genome, respectively. Moreover, based on studies in which incomplete genomes were selected by hybridization to specific portions of the viral genome, we have concluded that both genomic termini are present on individual molecules as part of a covalently continuous structure.

As in the case of wt DNA, type I genomes contain short, double-helical regions at their molecular termini. The 3′-terminal hairpin in these molecules appears to be identical to that found in wt DNA, with respect to size, ability to act as a primer for DNA polymerase I (lacking 5′ to 3′ exonuclease activity), and restriction enzyme cleavage. There is, however, some indication that the self-complementary region at the 5′ terminus is somehow modified in the type I genomes, since this region appears to be between 10 and 15 base pairs larger than the analogous structure derived from wt DNA. The basis for this difference is not known. However, it is interesting to note that the incomplete genomes of H-1 virus acquire tandemly repeated insertions of unknown origin which appear to map in the vicinity of the self-complementary region at the 5′ terminus of H-1 RF-DNA (27). Insertions at the 5′ end in H-1 D-DNA and possibly in MVM D-DNA may represent tandem duplications of certain critical regions involved in the replication of parvovirus genomes and thus deserve further study. Alternatively, the apparent difference in the size of the 5′-terminal hairpin duplex in D-DNA could reflect the presence of a 5′-terminal DNA-protein complex similar to those which have been observed for other viral DNAs (28).

A model for the sequence organization in the type I genomes which is consistent with the data presented in this paper is illustrated in Fig. 16. The essential features of this model are the conservation of terminal self-complementary nucleotide sequences in individual molecules and the existence of deletions which occur at various internal positions. Since the splice points are not unique, sets of molecules which are on the average only 50% genome-size can be expected to contain virtually the entire genome sequence. This aspect of the model is supported

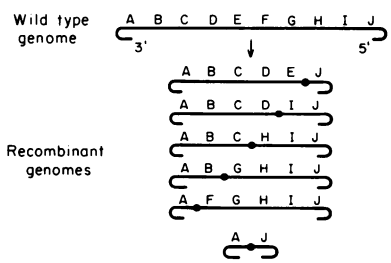


FIG. 16. *Model for the sequence organization in individual type I MVM recombinant genomes (see text).*

by the restriction enzyme analysis of the D-RF in pool A of Fig. 9. At the present time we do not know whether the type I genomes contain only single deletions as indicated in Fig. 16 or whether individual molecules may also exhibit multiple deletions, sequence inversions, reiterations, or even insertions of foreign, host-cell DNA. These kinds of changes have been described in some detail for the deleted genomes of simian virus 40 where they have been detected by altered restriction enzyme cleavage patterns of the DNA (9). It is of interest to note, therefore, that restriction enzyme digests of MVM D-RF DNA contain fragments in addition to those seen in digests of wt RF DNA. This is particularly evident in digests performed with D-RF DNA derived from late-passage D-DNA samples (for example, compare the gel profiles in Fig. 15). However, to decide whether this DNA is viral or cellular in origin and, if cellular, whether it is covalently linked to viral sequences will require an examination of cloned populations of type I D-DNA.

In contrast to the type I structures, type II molecules are hairpins, ranging in size up to 1,000 base pairs in length, and consist of sequences which map exclusively within approximately 0.2 fractional lengths at the 5′ terminus of the viral genome. In Fig. 17, we consider a simple mechanism by which type II molecules might arise. This model is similar in concept to the one proposed originally by Daniell to explain the aberrant adenovirus genomes (6) and is based on the self-priming mechanism for MVM DNA replication proposed by Tattersall and Ward (38). We postulate random nicking events occurring internally in the viral (−) strand, possibly during strand displacement synthesis as indicated in Fig. 17, or on parental molecules. The 5′-terminal portion of the nicked molecules would then contain a free 3′ end capable of folding back and forming weak, transiently stable base-paired regions at a variety of sites on the template strand. In this configuration, the 3′ end would be suitable as a primer and could be extended by DNA polymerase, thus leading to the formation of "looped" hairpin structures. Both the variability in nick sites and a variable degree of looping over would account for the size heterogeneity of the type II molecules themselves and for the variability in the sizes of the loops and stems of which these molecules are constructed. Theoretically, hairpin molecules generated in this way could be as much as twice the molecular length of the viral genome. However, the largest type II DNA molecules seen are only about 50 to 60% genome size, and so nicking may occur more frequently within the 5′ half of the genome and less often toward the 3′ end. Alternately, the nick sites may in fact be random, and the distribution of type II molecules that we have observed could be due to selective packaging of shorter molecules. Incomplete genomes of adenovirus type 3, which are encapsidated into virions, contain sequences derived only from the left half of the adenovirus DNA (6), whereas sequences from both ends of the DNA are seen in incomplete genomes extracted from infected cells (7). These results suggest that encapsidation of adenovirus DNA is selectively initiated at the left end of the genome. The total absence of sequences from the 3′ half of the MVM genome in type II D-DNA molecules is also consistent with the previous suggestion (38) that encapsidation of MVM DNA is initiated at the 5′ end of the virion DNA. An analysis of the incomplete MVM genomes isolated from infected cells will be required to substantiate this interpretation.

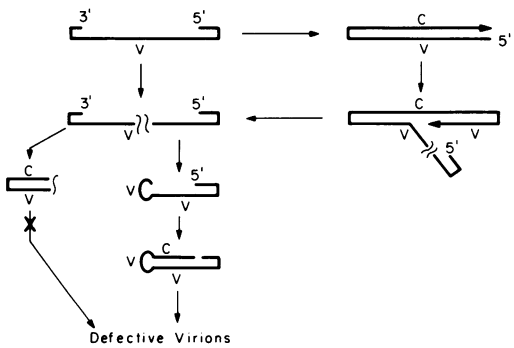
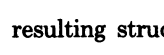


FIG. 17. *Model for the derivation of type II genomes from the 5′-terminal portions of nicked viral (−) strands. Type II genomes are considered to arise after nicking events which occur either in the parental strand or during strand displacement synthesis from the monomer RF as shown at the top right-hand part of the figure. 5′-Terminal fragments generated by either pathway would form a common pool for generating type II genomes by a loop-over mechanism. Remnant 3′-terminal fragments may be further processed to a duplex form, but these molecules are not packaged into virions. The letters v and c designate the viral (−) and complementary (+) strands, respectively.*

The incomplete genomes of the related parvovirus AAV have similar properties to those described here for MVM, although the fact that AAV packages both (+) and (−) strands (2, 29) has made the interpretation of the structural analysis in the AAV system somewhat more difficult (4, 10, 22). In the latter studies, truncated hairpin molecules analogous to the type II structures seen here for MVM were described. However, unlike MVM, the AAV hairpin molecules contain sequences derived from both genomic termini. Considering the model depicted in Fig. 17, this would be the expected result if each strand of the AAV genome could give rise

to hairpin molecules independently. Of the two resulting structures  , the latter would correspond to the type II structures seen with MVM, since this virus packages only the (−) strand. In contrast to the model presented in Fig. 17, Hauswirth and Berns (16) have proposed that the hairpin AAV genomes arise by a template strand-switch mechanism. Although some of the type II genomes of MVM may also be generated in this way, the majority of the molecules contain relatively large (300 to 1,000 nucleotides) single-stranded loops. This feature is not entirely consistent with a template strand-switching mechanism which would not be expected to yield hairpin duplexes with such large single-stranded regions. Finally, duplex molecules which are not in a hairpin configuration were also described for AAV. These structures may represent recombinant genomes of both (+) and (−) strands which have annealed; however, confirmation of this suggestion will have to await further analysis.

These studies on the deleted genomes of MVM were initiated to delineate required *cis*-acting functions, such as the origin for MVM DNA replication. A variety of independent observations have demonstrated the importance of the self-complementary sequences located at the termini of parvovirus genomes in the initiation of viral DNA replication. Important events such as site-specific nicking and "hairpin transfer" almost certainly occur at or near these specialized regions. The results obtained here provide additional support for this general view. For example, we have estimated that ~30% of the type I D-DNA molecules from a single high-multiplicity passage lack 85 to 90% of the genetic sequence of wt MVM DNA (see pool D, Fig. 9). Significantly, the regions which are conserved in these molecules map almost entirely within 200 nucleotides at both genomic termini. Furthermore, type I D-DNAs arising during serial undiluted passage are enriched for shorter molecules in which there is an increased loss of internal wt sequence but a conservation of the self-complementary sequences from both genomic termini. We have taken this to mean that type I genomes are part of an intracellular pool of actively replicating DNA at least in the presence of wt virus and therefore that the selective conservation of the terminal self-complementary sequences in these molecules is of direct functional significance for the process of DNA replication. For these reasons, we have concluded that the critical recognition sites involved in MVM DNA replication lie entirely within a region, about 200 nucleotides in length, at either end of the viral genome. It is interesting to note, in this regard, that the 3′-terminal 175 nucleotides of the DNA of the rodent parvoviruses MVM, H-1, H-3, and Kilham rat virus have almost the identical nucleotide sequence (1a).

In addition, we have noted a decrease in the relative amount of type II genomes present in late-passage virus stocks. This observation lends itself to several possible interpretations. The first of these is that type II genomes do not take part in DNA replication, but arise as dead-end replication errors. The absence from these molecules of sequences mapping at the 3′ terminus supports this view. In this event, the disappearance of type II genomes would be expected to be gradual, especially if these molecules arise from the type I genomes, as well as wt DNA, at each passage. This gradual disappearance of type II genomes is in fact observed. However, we would expect similar results if these molecules simply replicate at a much slower rate in comparison with the type I genomes. Clearly, this question can best be resolved by examining intracellular pools of replicating DNA.

The availability of a variety of deletion mutants, such as the ones described here for MVM, should be useful (i) in the study of MVM transcription and mRNA processing (19), (ii) as probes for examining the activity of proteins involved in MVM DNA replication, and (iii) as potential cloning vehicles for propagating foreign genetic elements in mammalian cells (12, 13, 25). For these reasons, elucidation of the structure of cloned populations of type I MVM genomes is of considerable interest.

ACKNOWLEDGMENTS

We thank Stewart Millward for generously providing temporary laboratory facilities to one of us (E.A.F.) during the initial stages of this study. We also thank L. Guluzian for his expert guidance in the spreading of DNA for electron microscopy and for preparing the electron micrographs. The assistance of D. Brooker and H. McGuire in the production of MVM is gratefully acknowledged. We thank also C. Cole for his gift of simian virus 40 DNA and C. Hours for providing φX174 DNA.

E.A.F. was a recipient of a Canadian Medical Research Council Postdoctoral Fellowship. This work was supported by United States Public Health Service grants GM-20124 from the National Institute of General Medical Sciences and CA-16038 from the National Cancer Institute.

LITERATURE CITED

1. **Astell, C. R., M. Smith, M. B. Chow, and D. C. Ward.** 1979. Structure of the 3′ hairpin termini of four rodent parvovirus genomes: nucleotide sequence homology at the origins of DNA replication. Cell **17**:691–703.

1a. **Astell, C. R., M. Smith, M. B. Chow, and D. C. Ward.** 1979. Sequence of the 3′ terminus of the genome from Kilham rat virus, a non-defective parvovirus. Virology **96**:669–674.

2. **Berns, K. I., and S. Adler.** 1972. Separation of two types of adeno-associated virus particles containing complementary polynucleotide chains. J. Virol. **9**:394–396.

2a. **Bonner, W. M., and R. A. Laskey.** 1974. Fluorographic detection of ^{3}H, ^{14}C, and ^{35}S in acrylamide gels. Eur. J.

Biochem. **46**:83–88.
3. **Bourguignon, G. J., P. J. Tattersall, and D. C. Ward.** 1976. DNA of minute virus of mice: self-priming, nonpermuted, single-stranded genome with a 5′-terminal hairpin duplex. J. Virol. **20**:290–306.
4. **Carter, B. J., C. A. Laughlin, L. M. de la Maza, and M. Myers.** 1979. Adeno-associated virus autointerference. Virology **92**:449–462.
5. **Chow, M. B., and D. C. Ward.** 1978. Comparison of the terminal nucleotide structures in the DNA of nondefective parvoviruses, p. 205–217. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses, Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
6. **Daniell, E.** 1976. Genome structures of incomplete particles of adenovirus. J. Virol. **19**:685–708.
7. **Daniell, E., and T. Mullenbach.** 1978. Synthesis of defective viral DNA in HeLa cells infected with adenovirus type 3. J. Virol. **26**:61–70.
8. **Davis, R. W., M. Simon, and N. Davidson.** 1971. Electron microscopic hetero-duplex methods for mapping regions of base sequence homology in nucleic acids. Methods Enzymol. **21**:413–428.
9. **Davoli, D., E. D. Ganem, A. L. Nussbaum, and G. C. Fareed.** 1977. Genome structures of reiteration mutants of simian virus 40. Virology **77**:110–124.
10. **de la Maza, L. M., and B. J. Carter.** 1978. DNA structure of incomplete adeno-associated-virus particles, p. 193–204. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
11. **Fife, K. H., K. I. Berns, and K. Murray.** 1977. Structure and nucleotide sequence of the terminal regions of adeno-associated virus DNA. Virology **78**:475–487.
12. **Ganem, D., A. L. Nussbaum, D. Davoli, and G. C. Fareed.** 1976. Propagation of a segment of bacteriophage λ DNA in monkey cells after covalent linkage to a defective simian virus 40 genome. Cell **7**:349–359.
13. **Hamer, D. H., D. Davoli, C. A. Thomas, Jr., and G. C. Fareed.** 1977. Simian virus 40 carrying an *Escherichia coli* suppressor gene. J. Mol. Biol. **112**:155–182.
14. **Hauswirth, W. W., and K. I. Berns.** 1977. Origin and termination of adeno-associated virus DNA replication. Virology **78**:488–499.
15. **Hauswirth, W. W., and K. I. Berns.** 1978. Initiation and termination of adeno-associated virus DNA replication, p. 257–267. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
16. **Hauswirth, W. W., and K. I. Berns.** 1979. Adeno-associated virus DNA replication: nonunit-length molecules. Virology **93**:57–68.
17. **Jeppesen, P. G. N.** 1974. A method for separating DNA fragments by electrophoresis in polyacrylamide concentration gradient slab gels. Anal. Biochem. **58**:195–207.
18. **Kowalski, D., W. D. Kroeker, and M. Laskowski, Sr.** 1976. Mung bean nuclease. I. Physical, chemical and catalytic properties. Biochemistry **15**:4457–4462.
19. **Lai, C. J., and G. Khoury.** 1979. Deletion mutants of simian virus 40 defective in biosynthesis of late viral mRNA. Proc. Natl. Acad. Sci. U.S.A. **76**:71–75.
20. **Langeveld, S. A., A. D. M. van Mansfeld, P. D. Baas, H. S. Jansz, G. A. van Arkel, and P. J. Weisbeek.** 1978. Nucleotide sequence of the origin of replication in bacteriophage φX174 DNA. Nature (London) **271**:417–420.
21. **Laskey, R. A., and A. D. Mills.** 1975. Quantitative film detection of ^{3}H and ^{14}C in polyacrylamide gels by fluorography. Eur. J. Biochem. **56**:335–342.
22. **Laughlin, C. A., M. Myers, D. L. Risin, and B. J. Carter.** 1979. Defective-interfering particles of the human parvovirus adeno-associated virus. Virology **94**: 162–174.
23. **Little, J. W.** 1967. An exonuclease induced by bacteriophage λ. J. Biol. Chem. **242**:679–686.
24. **Müller, H. P., M. Gautschi, and G. Siegl.** 1978. Defective particles of parvovirus Lu III, p. 231–240. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
25. **Mulligan, R. C., B. H. Howard, and P. Berg.** 1979. Synthesis of rabbit β-globin in cultured monkey kidney cells following infections with a SV40 β-globin recombinant genome. Nature (London) **277**:108–113.
26. **Rhode, S. L., III.** 1977. Replication process of the parvovirus H-1. VI. Characterization of a replication terminus of H-1 replicative-form DNA. J. Virol. **21**:694–712.
27. **Rhode, S. L., III.** 1978. Defective interfering particles of parvovirus H-1. J. Virol. **27**:347–356.
28. **Robinson, A. J., and H. J. D. Bellet.** 1974. A circular DNA-protein complex from adenoviruses and its possible role in DNA replication. Cold Spring Harbor Symp. Quant. Biol. **39**:523–531.
29. **Rose, J. A., K. I. Berns, M. D. Hoggan, and F. J. Koczot.** 1969. Evidence for a single-stranded adenovirus-associated virus genome: formation of a DNA density hybrid on release of viral DNA. Proc. Natl. Acad. Sci. U.S.A. **64**:863–869.
30. **Sherrat, D.** 1978. Replication origins. Nature (London) **271**:404–405.
31. **Singer, I. I., and S. L. Rhode III.** 1977. Replication process of the parvovirus H-1. VII. Electron microscopy of replicative-form DNA synthesis. J. Virol. **21**:713–723.
32. **Soeda, E., G. Kimura, and K. Miura.** 1978. Similarity of nucleotide sequences around the origin of DNA replication in mouse polyoma virus and simian virus 40. Proc. Natl. Acad. Sci. U.S.A. **75**:162–166.
33. **Straus, S. E., E. D. Sebring, and J. A. Rose.** 1978. Self-primed replication of adeno-associated virus DNA, p. 243–255. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
34. **Studier, F. W.** 1965. Sedimentation studies of the size and shape of DNA. J. Mol. Biol. **11**:373–390.
35. **Subramanian, K. N., R. Dhar, and S. M. Weissman.** 1977. Nucleotide sequence of a fragment of SV40 DNA that contains the origin of DNA replication and specifies the 5′ ends of "early" and "late" viral mRNA. J. Biol. Chem. **252**:355–367.
36. **Tattersall, P.** 1972. Replication of the parvovirus MVM. I. Dependence of virus multiplication on cell growth. J. Virol. **10**:586–590.
37. **Tattersall, P., P. J. Cawte, A. J. Shatkin, and D. C. Ward.** 1976. Three structural polypeptides coded for by minute virus of mice, a parvovirus. J. Virol. **20**:273–289.
38. **Tattersall, P., and D. C. Ward.** 1976. Rolling hairpin model for replication of parvovirus and linear chromosomal DNA. Nature (London) **263**:106–109.
39. **van Venrooij, W. J., E. C. Henshaw, and C. A. Hirsch.** 1970. Nutritional effects in the polyribosome distribution and rate of protein synthesis in Ehrlich ascites tumor cells in culture. J. Biol. Chem. **245**:5947–5953.
40. **Vogt, V. M.** 1972. Purification and further properties of single-strand-specific nuclease from Aspergillus oryzae. Eur. J. Biochem. **33**:192–200.
41. **Ward, D. C., and D. K. Dadachanji.** 1978. Replication of minute-virus-of-mice DNA, p. 297–313. *In* D. C. Ward and P. Tattersall (ed.), Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.
42. **Ward, D. C., and P. Tattersall (ed.).** 1978. A standardized nomenclature for restriction endonuclease fragments. An Appendix, p. 523–526. *In* Replication of mammalian parvoviruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N. Y.