Structure of Simian Virus 40-Adeno-Associated Virus Recombinant Genomes

ZEHAVA GROSSMAN,¹ KENNETH I. BERNIS,² AND ERNEST WINOCOUR¹

Department of Virology, Weizmann Institute of Science, Rehovot, Israel,¹ and Department of Microbiology, Cornell University Medical College, New York, New York 10021²

Received 14 May 1985/Accepted 16 July 1985

The structures of recombinant genomes formed by recombination between simian virus 40 (SV40) and adeno-associated virus 2 (AAV) DNAs after either DNA cotransfection or coinfection by virions were characterized. Two types of structures were found. Group A structures, found after cotransfection and in one of seven recombinants arising from coinfection, represented a simple deletion of SV40 sequences replaced by a slightly shorter AAV sequence. Group B structures were found in six of seven recombinants arising after virion coinfection. All contained either the left or right terminal sequences (approximately 250 to 450 bases) of the AAV genome adjacent to the SV40 origin of DNA replication. Only 350 to 650 bases (including the origin) remained of the SV40 sequence. The joined SV40-AAV sequences were present in the recombinant genome as a tandem repeat of a size that can be packaged into SV40 capsids.

To gain a better understanding of nonhomologous recombination in animal cells, we are exploiting the capacity of simian virus 40 (SV40) to recombine with a variety of other DNAs of both procaryotic and eucaryotic origins (11-13, 41, 42). In a previous study of recombination between defective adeno-associated virus (AAV) and SV40 (13), we compared recombination frequencies after both DNA cotransfection and virus coinfection as measured by the numbers of recombinant-producing cells identified by infectious center in situ plaque hybridization (40, 12). In most respects, the data obtained by both methods of DNA delivery were quite comparable. Independent of the method of delivery, the probability that a successfully infected cell would generate a recombinant was dependent on the initial SV40 DNA dosage as well as the AAV DNA concentration, and the maximum proportion of cells (relative to the total population) generating recombinants after virion coinfection was similar to the maximum proportion of recombinant-producing cells after DNA cotransfection, despite the 5- to 10-fold higher efficiency of infection in the former case.

In one respect, however, the recombination process that ensues after SV40 and AAV virus coinfection differed from that which occurs after SV40 and AAV DNA cotransfection. The terminal sequences of the AAV genome were found in inordinately high proportions (>90%) of recombinants generated after coinfection. In contrast, no such overrepresentation was observed in recombinants generated after DNA cotransfection. It was also found that, for DNA cotransfection systems, DNA sequences located at a free end possess no obvious advantage in recombination with circular SV40 DNA (11, 13).

In the present work, we analyzed in detail SV40-AAV recombinant structures. The recombinants can be divided into two groups according to their SV40 and AAV content and according to their general structure. The recombinants in group A (those produced after cotransfection and one of seven of those resulting from coinfection) retain more than 70% of the SV40 genome and carry single inserts derived from the inner parts of the AAV genome. The SV40-AAV recombinants in group B (six of seven of the recombinants

produced after coinfection) retain less than 20% of the SV40 genome (including the SV40 origin of replication) and contain an AAV segment that retains part of its terminal palindrome. The SV40 origin-containing segment is joined together with that of AAV in repeat units, usually in a head-to-tail tandem array.

Overrepresentation of AAV terminal sequences in recombinants generated after coinfection and the specific structure of the DNA of those recombinants containing the AAV terminal repeats may be related to certain features characteristic of the AAV genome when it is delivered to the cell in virions.

MATERIALS AND METHODS

Cells and viruses. SV40 (strain 777) was propagated in BSC-1 cells maintained in Dulbecco modified Eagle medium with 10% calf serum (21). AAV (2H) (17) was propagated in HeLa cells in suspension culture with adenovirus type 2 helper as described before (27).

Preparation of virion stocks containing recombinant genomes. Virion stocks containing SV40-AAV recombinant genomes (produced from a single transfected cell) were isolated from the agar overlays of the infectious center assay (42). Autoradiograms of filters hybridized to ³²P-labeled AAV DNA were aligned with the agar overlays, and plugs of agar corresponding to the autoradiographic signals were picked, freeze-thawed, and used to infect fresh BSC-1 cell monolayers in microwells; on the occurrence of full cytopathic effect, the microwell supernatants were assayed for the presence of AAV DNA sequences by dot-blot hybridization. The yields of positive wells were passaged by infection of fresh BSC-1 cells, and after the occurrence of full cytopathic effect, the supernatants were collected and used as inocula for progeny DNA preparation.

DNA preparations. SV40 DNA was prepared as described by Oren et al. (25). AAV virion DNA was purified by sedimentation through alkaline sucrose gradients and annealing of the complementary strands (2). Bacterial plasmids were prepared from *Escherichia coli* HB101 or DH1 cells grown to saturation in LB medium (11) containing the

FIG. 1. Restriction mapping, before cloning, of SV40-AAV recombinant DNA in isolates derived from single recombinant-producing cells. DNAs of isolates 4 and 5 (see the text), either undigested or digested with restriction enzymes which do not cut within the AAV DNA inserts, were fractionated on 1% agarose gels, transferred to nitrocellulose paper, and hybridized with ^{32}P -labeled AAV DNA. Lanes a to d, Isolate 4 DNA, undigested or digested with *BclI*, *NdeI*, and *PvuII*, respectively; lanes e to i, isolate 5 DNA, undigested or digested with *SmaI*, *NaeI*, *PvuII*, and *NdeI*, respectively. The symbols I, II, and III designate the positions of wild-type SV40 DNA forms I, II, and III deduced from the ethidium bromide stain of the gel before blot hybridization. The number 2,000 designates a marker fragment 2,000 bp long. Autoradiograms were exposed at -70°C for 40 h.

appropriate antibiotic (100 μg of ampicillin or 12.5 μg of tetracycline per ml) by the quick boiling method of Holmes and Quigley (18), by the alkali lysis method of Birnboim and Doly (4), or by the cleared lysate procedure described by Davis et al. (8). Plasmid DNA was purified by equilibrium sedimentation in ethidium bromide-containing cesium chloride density gradients. Single-stranded M13 DNA was prepared according to Hu and Messing (19).

Plasmid DNAs and restriction fragments were purified by separation on agarose gels and by adsorption to RPC-5 resin and elution with a solution with a high salt concentration (3). DNA from agarose gels was eluted either by the glass powder adsorption technique of Vogelstein and Gillespie (40) or by diffusion of the DNA (DNA fragments of less than 2,000 base pairs [bp]) from the agarose slices (which were disrupted by passing through a syringe) into the extraction solution (0.01 M EDTA, 0.2 M NaCl, 0.01 M Tris hydrochloride [pH 7.5 to 8.0], 0.5 g of agarose per 5 ml of extraction solution). Gel particles were removed by filtration through HA 0.45- μm (pore size) filters (Millipore Corp.), and the DNAs were concentrated by adsorption to RPC-5 resin (3).

Recombinant viral DNA was isolated from cells by a modification of the Hirt (16) procedure. After sedimentation of high-molecular-weight DNA from a total volume of 1.0 ml, 100 μl of 0.25 M EDTA–0.15 M Tris hydrochloride (pH 9.5)–3% sodium dodecyl sulfate was added to the superna-

tant which, after gentle mixing, was heated to 65°C for 15 min. The preparation was digested with RNase, proteins were removed by centrifugation for 10 min, and the supernatant was extracted by phenol-chloroform-isoamyl alcohol (50:48:2). DNA was precipitated with ethanol, and the dried pellet was dissolved in 200 μl of TE buffer (10 mM Tris hydrochloride [pH 7.4], 1 mM EDTA).

Restriction endonuclease digestion, electrophoresis, and blot hybridization. Restriction endonucleases were used according to the recommendations of the manufacturer (New England BioLabs, Inc.), and the products were separated by electrophoresis (40 V for 20 h) in Tris-acetate buffer on 1% agarose slab gels. The gel was stained with ethidium bromide (0.5 $\mu\text{g}/\text{ml}$) and photographed under UV light. The products were transferred to nitrocellulose sheets (Schleicher & Schuell, Inc.), and Southern (37) blot hybridization was carried out according to Smith and Summers (35) in Denhardt buffer (10).

DNA transfections. Cells were transfected in suspension in the presence of DEAE-dextran (23) by the method of Milman and Herzberg (24), as described previously (13).

Assay for recombination. The infectious center in situ plaque hybridization assay for recombination was performed as described previously (1, 13).

Molecular cloning. Cloning of AAV DNA and segments thereof has been previously described (13). The closed circular DNA yields of individual recombinant-producing cells were digested with the appropriate restriction enzyme and ligated to cleaved and alkaline phosphatase-treated (6) plasmid DNA (pBR322 or its derivatives Bd1H and pML2). *E. coli* HB101 or DH1 cells were transformed with the ligation products (11), and the SV40-AAV-containing colonies were isolated according to Hanahan and Meselson (15).

DNA sequencing. The nucleotide sequences of two SV40-AAV recombinant regions (see below) were determined by the dideoxynucleotide chain terminator method of Sanger et al. (30), as described by Smith et al. (36), with pBR322 *HindIII* and *EcoRI* primers obtained from New England BioLabs. Single-stranded DNA sequencing by the M13 system was done according to Hu and Messing (19). Acrylamide (6%) gels were used with Tris-borate buffer.

Source of bacterial plasmids. pML2 plasmid was obtained from M. Botchan, and Bd1H was from D. Dorsett.

RESULTS

Isolation and preliminary characterization of SV40-AAV recombinants. SV40-AAV recombinants formed after DNA cotransfection (isolate 4) and virion coinfection (isolates 5 through 40) were isolated from the agar overlay as described by Winocour et al. (42). The supercoiled DNA of SV40-AAV isolates consisted of a mixture of wild-type and recombinant DNAs. We determined the loss or retention of SV40 restriction sites in SV40-AAV recombinant genomes by digesting the DNA with restriction enzymes which cut wild-type SV40 DNA but not AAV DNA (*EcoRV*, *NaeI*, *NdeI*, *SphI*, and *PvuII*, which have no recognition sites in AAV DNA and one to three sites in SV40 DNA). The digestion products were separated by agarose gel electrophoresis and transferred to nitrocellulose paper (35), and those products containing AAV sequences were identified by hybridization with ^{32}P -labeled AAV DNA. Further analysis was then done with restriction enzymes which cut both SV40 and AAV DNAs (e.g., *BamHI*, *BglII*, *BclII*, and *EcoRI*). Figure 1 illustrates such an analysis.

Figure 2 schematically summarizes the retention of SV40 sequences in eight different isolates which were analyzed by

FIG. 2. Retention of SV40 sequences in SV40-AAV recombinants. We determined the loss or retention of SV40 restriction sites in SV40-AAV recombinant genomes by digesting the DNA with restriction enzymes, fractionating the digested DNA by agarose gel electrophoresis, and performing blot hybridization to ^{32}P -labeled AAV probe. The solid boxes denote SV40 sequences that were retained and are drawn relative to the map of the wild-type SV40 genome. Isolate 4 was derived from a DNA cotransfection experiment; all of the other isolates (5 to 40) were derived from virion coinfection. Isolates 4 and 22 retained more than 70% of the SV40 sequences; in the others, less than 20% of the SV40 genome (including the origin of replication) was conserved. The AAV segments contained about 800 nucleotides in isolate 4 and 500 or less in all of the others (see Fig. 6 and 7 and the text). Symbols: $\downarrow$, *Bgl*I; ∇ , *Hin*fI; N, *Nae*I; $\bullet$, *Nco*I; $\times$, *Nde*I; T, *Pvu*II; P, *Pst*I; $\uparrow$, *Sph*I; $\uparrow$, *Taq*I; Bam, *Bam*HI; RI, *Eco*RI; Rv, *Eco*Rv.

the above method. All of the recombinants were smaller than wild-type SV40 (genomes of 4,000 to 5,000 bp). The isolates could be clearly divided into two groups: (i) group A, containing isolates 4 (derived from a DNA cotransfection experiment) and 22 (derived from a virion coinfection experiment), which retained more than 70% of the SV40 genome; and (ii) group B, containing isolates 5, 9, 15, 16, 38, and 40 (all derived from virion coinfection experiments), which retained less than 20% of the SV40 genome. In all cases, the SV40 origin of replication was retained.

We identified the AAV inserts in the isolates by hybridizing blots to ^{32}P -labeled, defined AAV DNA probes. Group A isolates 4 and 22 contained inserts from the inner segments of the AAV genome; isolate 4 hybridized to the *Pst*I A fragment of AAV (nucleotides 1959 to 4258), and isolate 22 hybridized to the *Pst*I B fragment of AAV (nucleotides 496 to 1958). All of the group B isolates contained terminal sequences of the AAV genome. Since these latter isolates

hybridized to both the *Pst*I C (nucleotides 1 to 495; not shown) and *Pst*I D (nucleotides 4259 to 4675; Fig. 3 and 4) segments of AAV DNA, they probably contain at least part of the terminal inverted repeat of the AAV genome.

Group B recombinants contained less than 20% of the SV40 genome and no more than 15% of the AAV genome yet were 4,000 to 5,000 bp in length (Fig. 2). This apparent discrepancy can be explained if it is assumed that, in any given recombinant, the sequences that are retained appear more than once. To investigate this possibility, we digested the DNAs with restriction enzymes that cut once either in the SV40 sequence or in the AAV sequence that was retained in the recombinants, namely *Hind*III, which cuts SV40 in the origin region (nucleotide 5171), and *Bal*I, which cuts AAV at nucleotide 4554 (all other sites of these two enzymes were not preserved in these recombinants). Figure 4 summarizes this experiment. Since *Hind*III, which cuts only SV40 sequences, and *Bal*I, which cuts only AAV

FIG. 3. Restriction mapping (before cloning) of isolates 38 and 40. DNAs of isolates 38 and 40 were digested with restriction enzymes and fractionated on agarose gels, and the blot was hybridized with ^{32}P -labeled *Pst*I D fragment of AAV (nucleotides 4259 to 4675). Lanes: a to d, isolate 38; e to h, isolate 40, undigested or digested with *Eco*RI, *Pvu*II, and *Nde*I, respectively; i, wild-type SV40 DNA, digested with *Pvu*II. The symbols I, II, and III designate the positions of wild-type SV40 DNA forms I, II, and III. (A) Ethidium bromide staining of the gel. The digestion patterns of virus DNA (the only DNA visible) indicates that all reactions went to completion. (B) Autoradiogram of the blot after hybridization. Only recombinant DNA reacted; note that the recombinant DNAs were not digested by the enzymes.

sequences, generated small fragments of the same size in isolates 5, 16, and 40, it appears likely that a tandemly repeated structure containing SV40 and AAV sequences was present in these isolates. Isolate 38 also appears to have a repetitive structure, but of a more complex nature. It is noteworthy that the reiterated structure was found only when the recombinant contained inserts derived from the terminal repeat unit of the AAV genome.

Molecular cloning and structure of two SV40-AAV recombinants belonging to group A. We cloned the two isolates of group A (isolates 4 and 22) by inserting their DNA into the *Eco*RI site of the pBR-related plasmid Bd1H. Screening for the AAV-containing plasmids was performed according to Hanahan and Meselson (15). The proportions of colonies which reacted against ^{32}P -labeled AAV DNA were 30 and 0.01% in isolates 4 and 22, respectively, of those which reacted against ^{32}P -labeled SV40 DNA, indicating that the ratios of SV40-AAV recombinants to SV40 DNA molecules were 1:3 and 1:10,000. Cloned SV40-AAV recombinant DNAs were mapped by restriction endonuclease digestion analysis (Fig. 5).

AAV sequences in the recombinant genomes were identified by hybridization of ^{32}P -labeled recombinant DNAs to *Pst*I digests of plasmids containing different *Pst*I fragments of the AAV genome (13). Recombinant 4 DNA hybridized only to segment A of AAV DNA (nucleotides 1959 to 4258), and recombinant 22 DNA hybridized only to segment B of AAV DNA (nucleotides 496 to 1958). Figure 6 shows maps that were constructed on the basis of the data from these procedures.

The essential features of the group A recombinant structures are as follows. (i) The two recombinants carried inserts

derived from the inner parts of the AAV genome (rather than the terminal sequence). (ii) The AAV sequences (500 to 800 bp) were present as single inserts located within the early SV40 region. (iii) The deletion of SV40 sequences was greater than the size of the AAV insert, resulting in shortening of the recombinant genome relative to wild-type SV40 (4,800 bp for recombinant 4 and 4,200 bp for recombinant 22). (iv) In the two recombinants, the remainder of the SV40 DNA (other than the deletion which accommodated the AAV DNA insert and, in recombinant 4, the 600-bp SV40 duplication of the origin region) was retained in unaltered form, as judged by restriction mapping.

Molecular cloning and structure analysis of two recombinants containing AAV terminal repeat sequences (group B recombinants). The special structure of the recombinants in this group necessitated a different cloning strategy. We were unable to clone the entire recombinant genome in a plasmid, since all of the restriction enzymes that were tried either did not digest the recombinant DNA or digested the DNA into small fragments (Fig. 1 to 4). Therefore, we focused on the cloning of repeat units which contain the SV40-AAV junctions. *Hind*III-*Hind*III fragments of recombinants 16 (771 bp) and 40 (886 bp) (Fig. 7) were inserted into the *Hind*III site of pBR322. The two ^{32}P -labeled constructs containing the recombinant DNAs hybridized to both terminal segments C (nucleotides 1 to 425) and D (nucleotides 4255 to 4675) of *Pst*I-digested AAV DNA (13), indicating that the inserts carry the AAV terminal repeat; ^{32}P -labeled recombinant 16 also hybridized to segment A (nucleotides 1959 to

FIG. 4. Restriction analysis (before cloning) of SV40-AAV recombinant DNAs of group B. DNA of isolates 5 (a to c), 16 (d to f), 38 (g to i), and 40 (j to l), either undigested or digested with *Hind*III (cuts only in the SV40 segment of the recombinant) or *Ball*I (cuts only in AAV sequences), respectively. The DNA was fractionated on 1% agarose gel, transferred to nitrocellulose, and hybridized with ^{32}P -labeled *Pst*I D segment of AAV (nucleotides 4259 to 4675). *Hind*III and *Ball*I created the same size fragments in isolates 5, 16, and 40, indicating a tandem repeat structure. Isolate 38 DNA contains several sites for *Hind*III or *Ball*I, but apparently the units are not arranged in tandem.

4254). Schematic maps of the cloned repeat units are presented in Fig. 7.

Sequence analysis of the SV40-AAV junctions of recombinants 16 and 40. The repeat units of isolates 16 and 40 were inserted into the *Hind*III site of pBR322. Dideoxynucleotide sequencing (30) of the double-stranded template in pBR322 was done according to Smith et al. (36), with pBR322 *Hind*III and *Eco*RI primers obtained from New England BioLabs. For addition sequencing in M13, the inserts were recloned in the Mp9 vector (*Hind*III-*Pst*I fragment of recombinant 16 and the *Hind*III-*Bal*I fragment of recombinant 40; Fig. 7). The full nucleotide sequences of the repeated units were determined (771 and 886 bp for recombinants 16 and 40, respectively). The sequences of the junctions are shown in Fig. 8A, and several features of note can be summarized as follows. (i) In the two recombinants examined in each repetitive unit, the AAV sequences were inserted as one block, and in each case one of the junctions was within the inverted repeat (AAV nucleotide position 4591 in recombinant 16 and nucleotide 4604 [38] in recombinant 40 [Fig. 8A; junctions 16-1 and 40-1]). (ii) A portion of the AAV inverted repeat sequences was inserted into the SV40 genome near the SV40 origin of replication. (iii) The recombinant junctions in recombinant 16 were not identical with those in recombinant 40 with respect to either the SV40 or AAV sequence at the crossover points. (iv) In all four junctions, one or two bases were common to the parental DNAs. (v) A small palindromic structure could be drawn for the AAV sequences at the second junction in each recombinant (Fig. 8B; junctions 16-2 and 40-2). (vi) The crossover point at junction 40-1 was within a pentanucleotide (CTTTG) in the AAV sequence that overlapped the same pentanucleotide in the SV40 sequence just to the right of crossover. Junction 16-1 was flanked by three complementary bases to the left in the AAV sequence and five complementary bases to the right in the SV40 sequence (Fig. 8A). Other regions of direct or complementation homology between SV40 and AAV DNAs (5 to 10 nucleotides) were located within 35 nucleotides on either side of the crossover points in all four junctions (not shown in Fig. 8A). The significance of these homologies is discussed below.

DISCUSSION

The SV40-AAV recombinants classified as group A type, which arose from cells either cotransfected with SV40 and AAV DNAs (recombinant 4) or coinfecting with SV40 and AAV virions (recombinant 22), resemble the SV40- ϕ X174 simple insertion-deletion type described previously (41) in that they contained single AAV inserts (derived from any part of the authentic AAV genome), the deletion of SV40 DNA was larger than the insert, and at least 70% of the wild-type SV40 DNA was retained in the recombinant. SV40-AAV recombinant 4 contained one additional SV40 replication origin linked to the AAV insert. In contrast, a different picture emerged after analyzing the structure of the recombinants of group B. The six recombinants from this group were composed of repeating units which contained SV40 and AAV segments. Other than the origin of replication, very little of the SV40 genome was retained, and both the segment that was retained and the inserted AAV sequences were duplicated several times. All of the recombinants of group B contained terminal sequences of the AAV genome and were isolated from virion coinfection experiments. Previously, we have shown that, after SV40 and AAV coinfection, more than 90% of the recombinants con-

FIG. 5. Restriction mapping of SV40-AAV recombinant DNA derived from isolate 4 and cloned in *Eco*RI site of plasmid Bd1H. Lanes: a, undigested DNA; b to j, DNA digested with *Sma*I, *Eco*RI, *Kpn*I, *Nae*I, *Bam*HI plus *Sal*I, *Bgl*II, *Sph*I, *Hind*III, and *Taq*I, respectively. Note that *Bgl*II, *Hind*III, and *Sph*I, which cut in the SV40 origin region and do not cut AAV in segment A (see the text), create a fragment of 1,400 bp which contains the AAV sequences (suggesting a duplication of the SV40 replication origin region). The numbers to the right indicate size markers.

tained AAV terminal sequences (13). This was observed even when AAV infection preceded SV40 infection by as long as 10 days. In contrast, no overrepresentation of terminal sequences was observed in recombinants generated after DNA cotransfection. Similarly, after DNA cotransfection with circular SV40 DNA and circular ϕ X174 replicative form I DNA (11, 12), with circular SV40 DNA and ϕ X174 replicative form I DNA linearized at the single *Pst*I site, or with pBR322 plasmids containing AAV segments, it has been shown that all parts of the AAV or ϕ X174 genomes were equally represented in the population of recombinants (13). Hence, in general, with regard to cotransfection, DNA sequences located at a free end possess no obvious advantage in recombination with circular SV40 DNA.

Delivery of viral genomes into the cell as virions does not necessarily dictate recombinant structures of the group B type. Recombinants isolated from mouse cells after coinfection of polyomavirus and SV40 virus belong to group A, as do recombinants derived from DNA cotransfection experiments (Z. Grossman, unpublished data). Also, recombinant 22, cited above, which was derived from AAV-SV40 covirion infection, belongs to group A. Importantly, only recombinants that were derived from virion coinfection of SV40 and AAV and which contained the AAV terminal sequences exhibited the group B-type structure. Overrepresentation of the AAV terminal sequences and the repetitive structure may be related to certain features of the AAV genome when it is delivered to the cell in virions. The AAV genome in the virion is a linear, single-stranded DNA molecule that contains an inverted terminal repetition of 145 bases, the first 125 of which are palindromic and capable of forming a hairpin structure (1). One possibility is that insertion of the AAV palindromic inverted terminal repeat near the SV40 origin of replication triggers a special mode of

FIG. 6. Schematic structures of two SV40-AAV recombinant (Rec.) genomes belonging to group A. The maps of recombinant DNAs 4 and 22 are drawn relative to that of the wild-type (WT) SV40 genome and are based on data from restriction cleavage blot hybridization (see the text). The shaded boxes denote AAV inserts, the solid box denotes the duplicated area containing the SV40 replication origin, the open boxes denote SV40 DNA sequences, and dl refers to the deletion of SV40 DNA sequences. This deletion is greater than the AAV insert; as a consequence, the recombinant genomes are shorter than the SV40 WT genome by an amount corresponding to the dashed lines (the size order is WT > 4 > 22). The WT SV40 map also shows the sites of the restriction enzymes used to analyze the structures of recombinant genomes. Symbols: B, *Bam*HI; Bc, *Bcl*I; ↓, *Bgl*II; □, *Hind*III; ■, *Hinf*III; H, *Hpa*II; N, *Nae*I; ○, *Neo*; X, *Nde*I; ●, *Sph*I; P, *Pst*I; T, *Pvu*II; †, *Taq*I.

replication, leading to the structure observed for group B recombinants. Normally, SV40 replication occurs via Cairns-type theta structures (32). The tandem repeats of group B recombinants may arise by a rolling-circle mode of replication such as has been suggested to occur late during SV40 DNA replication (26) and which has been shown to occur when linear forms of origin-containing SV40 DNA are transfected into monkey cells (9). Moreover, the reiterated

structure of SV40-AAV group B recombinants is similar to that described recently for a special class of SV40-φX174 recombinants which arise when linear SV40 origin-containing DNA is cotransfected with circular SV40 and φX174 replicative form I DNAs (under these conditions, the frequency of SV40-φX174 recombination is enhanced) and which are postulated to result from the entry of the linear SV40 DNA into an aberrant rolling-circle type of replication

FIG. 7. Schematic structures of the SV40-AAV repeat units of recombinants (Rec.) 16 and 40. The maps of the repeat units from isolates 16 and 40 are derived from restriction cleavage blot hybridization and nucleotide sequencing. The solid boxes denote the SV40 duplicated area, including the origin of replication. The shaded boxes denote AAV inserts. The repeat unit (771 bp) of isolate 16 was recloned into the *Hind*III site of pBR322, as was the repeat unit of isolate 40 (886 bp). Upper numbers denote the SV40 nucleotides at SV40-AAV junctions; lower numbers denote AAV nucleotides. The 771-bp repeat unit from isolate 16 consists of 410 bp of AAV (nucleotides 4182 to 4591) and 362 bp of SV40 (nucleotides 4926 to 44). The repeat unit from isolate 40 (886 bp) is composed of 270 bp of AAV (nucleotides 4334 to 4604) and 617 bp of SV40 (nucleotides 4828 to 201). Symbols: □, *Hind*III; P, *Pst*I; ↓, *Bgl*II; Bal, *Bal*I.

FIG. 8. Sequence of the SV40-AAV junctions in the repeat units of recombinants 16 and 40. Nucleotide sequences were determined according to the dideoxynucleotide method of Sanger et al. (30). The nucleotide sequences of the complete repeat units were determined and are schematically summarized in Fig. 7. (A) The nucleotide sequences of the parental SV40 and AAV DNAs are compared with that of the recombinant DNA. The recombinant nucleotides are written in boldface. Note one or two shared bases in all junctions. Asterisks denote nucleotides identical to those in the recombinant. (B) Secondary structure of the AAV sequences near the SV40-AAV junctions. The arrows designate crossover points. In 40-1, the sequence at the joint was slightly ambiguous with respect to which of the two nucleotides indicated by the arrows represented the actual crossover point. Note that the AAV sequences represent the 5' to 3' complements of those reported by Srivastava et al. (38). The terminal palindromic sequence in the inverted repeat at the right end of the AAV 2 genome extends from nucleotide 4551 to nucleotide 4675. Only the internal portion of this sequence from nucleotide 4588 to nucleotide 4638 is illustrated in the 16-1 and 40-1 junctions.

(11). Possibly, the SV40 origin region or the AAV terminal repeat in the conformation in which it is released from the virion, or both, predispose to recombination in or adjacent to the origin of replication in a circular SV40 molecule and, by so doing, trigger a unidirectional mode of DNA synthesis, leading to a rolling-circle intermediate as opposed to the more usual theta-type structures. It is not known whether the relatively specific site of the recombination event or the presence of AAV terminal sequences at the joint in the recombinant is the critical factor leading to the tandem structure. In this respect, it will be of interest to analyze the structure of those SV40-AAV recombinants containing AAV terminal sequences which arise from DNA cotransfection to see if they too exhibit a reiterated structure.

The difference in structure between group A and B recombinants raises the question of whether both are the results of equivalent initial recombination events. The similarity in kinetics and dependence on initial DNA concentrations observed after both cotransfection and coinfection (13) suggest equivalence, even though the structures of these recombinants are strikingly different. Many replication cycles of the recombinant molecules take place before the recombinant is isolated, cloned, and analyzed. In addition, the infectious-center assay detects only those recombinant structures which retain a functional SV40 origin of replication and which are of the right size to be encapsidated. Although there is no reason to assume that the conditions required for amplification of the group B structures are radically different from those required for group A structures, it is not known to what extent the amplified and selected groups of recombinants represent the initial populations of recombinant structures. To examine this question, it will be necessary to clone out low-molecular-weight DNA molecules within a few hours of coinfection or cotransfection and to characterize the spectrum of recombinant structures that are found.

The four SV40-AAV junctions that were analyzed manifest some interesting features. The AAV terminal palindrome was inserted always near the SV40 origin of replication, as if it were a preferred site for recombination. This may reflect the more open configuration of this region reported in SV40 transcriptional complexes (31). SV40-AAV junctions were not identical with respect to the SV40 nucleotides or AAV nucleotides involved at the crossover points. The palindrome of the AAV terminal inverted repeat was involved in the two recombinants sequenced, and a small palindrome could be formed at the second end of the AAV insert (junctions 16-2 and 40-2). These results may imply a significance of the secondary structure of the DNA near the crossover points in nonhomologous recombination. Short regions (5 to 10 nucleotides) of homology between SV40 and AAV DNAs were observed at or near the crossover points. The clustering of these short, homologous stretches near the junctions appears to be statistically significant, since computer analyses with nucleotide shuffling tests indicate that they appear with twice the frequency expected on a purely random basis (D. Dorsett, Y. Shaul, and E. Winocour, unpublished data). The four SV40-AAV junctions sequenced thus share some features which fit the patchy homology model of Gutai and Nathans (14) and which are in agreement with the analyses of Bullock et al. (5) of integrated SV40 DNA-host cell DNA junctions.

Six of the seven group B recombinants contain the right terminus of the AAV genome. Whether this is significant is unclear. If the bias were real, it might suggest some special property of the right end of the genome. A similar over-

representation of right terminal sequences was noted in AAV-defective interfering particle genomes (33).

Of special interest is the resemblance of the structure of the group B recombinants to the structure of the integrated copies of the AAV genome found in latently infected cells (7). In the latter case, the junction between AAV DNA and cell DNA also involved the AAV terminal sequences, and the integrated viral DNA was present as a linear head-to-tail tandem repeat. The terminal repeat has been suggested to serve not only a role in integration but also to function as a primer for DNA replication (1, 38). The AAV genome has recently been cloned into pBR322 (28) as a biologically active molecule, so that it has been possible to characterize the requirements for the terminal sequences. Three points have emerged from these studies. (i) The sequence is required for rescue or replication from the integrated state or both (20, 29, 34). (ii) The sequence has an extensive capacity for self repair (29, 34). (iii) As long as the overall conformation of the terminal structure is maintained, the sequence can be discretely altered without loss of biological activity (22).

ACKNOWLEDGMENTS

We thank Tamar Koch and Bernard Danovitch for technical assistance and Rachel Ben-Levy for help with the DNA sequencing. We also thank Dale Dorsett and Yosef Shaul for their valuable advice and help with the computer analysis of DNA sequences.

Part of this work was supported by U.S. Public Health Service grant AI 22251 from the National Institutes of Health.

LITERATURE CITED

- Berns, K. I., and W. W. Hauswirth. 1982. Organization and replication of parvovirus DNA, p. 3-35. In A. S. Kaplan (ed.), *Organization and replication of viral DNA*. CRC Press, West Palm Beach, Fla.
- Berns, K. I., J. Kort, K. H. Fife, E. W. Grogan, and I. Spear. 1975. Study of the fine structure of adeno-associated virus DNA with bacterial restriction endonucleases. *J. Virol.* **16**:712-719.
- Best, A. N., D. P. Allison, and G. D. Novelli. 1981. Purification of supercoiled DNA of plasmid ColE1 by RPC-5 chromatography. *Anal. Biochem.* **114**:235-243.
- Birnboim, H. C., and J. Doly. 1979. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res.* **7**:1513-1523.
- Bullock, P., W. Forrester, and M. Botchan. 1984. DNA sequence studies of simian virus 40 chromosomal excision and integration in rat cells. *J. Mol. Biol.* **174**:55-84.
- Chaconas, G., and J. H. Van de Sande. 1980. 5'-[³²P] labelling of RNA and DNA restriction fragments. *Methods Enzymol.* **65**:75-85.
- Cheung, A. K. M., M. D. Hoggan, W. Hauswirth, and K. I. Berns. 1980. Integration of the adeno-associated virus genome into cellular DNA in latently infected Detroit 6 cells. *J. Virol.* **33**:739-748.
- Davis, R. W., D. Botstein, and J. R. Roth. 1980. Isolation of plasmids and bacterial DNA, p. 116-125. In R. Davis et al. (ed.), *Advanced bacterial genetics*. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
- Deichate, I., Z. Laver-Rudich, D. Dorsett, and E. Winocour. 1985. Linear simian virus 40 DNA fragments exhibit a propensity for rolling-circle replication. *Mol. Cell. Biol.* **5**:1787-1790.
- Denhardt, D. 1966. A membrane filter technique for the detection of complementary DNA. *Biochem. Biophys. Res. Commun.* **23**:641-646.
- Dorsett, D., I. Deichate, and E. Winocour. 1985. Circular and linear simian virus 40 DNAs differ in recombination. *Mol. Cell. Biol.* **5**:869-880.
- Dorsett, D. L., I. Keshet, and E. Winocour. 1983. Quantitation of a simian virus 40 nonhomologous recombination pathway. *J. Virol.* **48**:218-228.
- Grossman, Z., E. Winocour, and K. I. Berns. 1984. Recombination between simian virus 40 and adeno-associated virus: virion coinfection compared to DNA cotransfection. *Virology* **134**:125-137.
- Gutai, M. W., and D. Nathans. 1978. Cellular DNA sequences and sequences at the recombinant joints of substituted variants. *J. Mol. Biol.* **126**:275-288.
- Hanahan, D., and M. Meselson. 1980. Plasmid screening at high colony density. *Gene* **10**:63-67.
- Hirt, B. 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. *J. Mol. Biol.* **26**:365-369.
- Hoggan, M. D., N. R. Blacklow, and W. P. Rowe. 1966. Studies of small DNA viruses found in various adenovirus preparations: physical, biological and immunological characteristics. *Proc. Natl. Acad. Sci. USA* **55**:1457-1471.
- Holmes, S. D., and M. Quigley. 1972. A rapid boiling method for the preparation of bacterial plasmids. *Anal. Biochem.* **114**:193-197.
- Hu, N., and J. Messing. 1982. The making of strand specific M13 probes. *Gene* **17**:271-277.
- Laughlin, C. A., J. D. Tratschin, H. Coon, and B. J. Carter. 1983. Cloning of infectious adeno-associated virus genomes in bacterial plasmids. *Gene* **23**:65-73.
- Lavi, S., and E. Winocour. 1972. Acquisition of sequences homologous to host deoxyribonucleic acid by closed circular simian virus 40 deoxyribonucleic acid. *J. Virol.* **9**:309-316.
- LeFebvre, R. B., S. Riva, and K. I. Berns. 1984. Conformation takes precedence over sequence in adeno-associated virus DNA replication. *Mol. Cell. Biol.* **4**:1416-1419.
- McCutchan, J. A., and J. S. Pagano. 1968. Enhancement of the infectivity of simian virus 40 deoxyribonucleic acid with diethylaminoethyl dextran. *J. Natl. Cancer Inst.* **41**:351-357.
- Milman, G., and M. Herzberg. 1981. Efficient DNA transfection and rapid assay for thymidine kinase activity and viral antigenic determinants. *Somatic Cell Genet.* **7**:161-170.
- Oren, M., S. Lavi, and E. Winocour. 1978. The structure of a cloned substituted SV40 genome. *Virology* **85**:404-421.
- Rigby, P. W. J., and P. Berg. 1978. Does simian virus 40 DNA integrate into cellular DNA during productive infection? *J. Virol.* **28**:475-489.
- Rose, J. A., K. I. Berns, M. D. Hoggan, and F. J. Kocot. 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. USA* **64**:863-869.
- Samulski, R. J., K. I. Berns, M. Tan, and N. Muzyczka. 1982. Cloning of adeno-associated virus into pB322: rescue of intact virus from the recombinant plasmid in human cells. *Proc. Natl. Acad. Sci. USA* **79**:2077-2081.
- Samulski, R. J., A. Srivastava, K. I. Berns, and N. Muzyczka. 1983. Rescue of adeno-associated virus from recombinant plasmids. Gene correction within the terminal repeats of AAV. *Cell* **33**:135-143.
- Sanger, R., S. Nicklen, and A. R. Coulson. 1977. DNA sequencing with chain-terminating inhibitors. *Proc. Natl. Acad. Sci. USA* **74**:5463-5467.
- Scott, W. A., and D. J. Wigman. 1978. Sites in simian virus 40 chromatin which are preferentially cleaved by endonucleases. *Cell* **15**:1511-1518.
- Sebring, E. D., T. J. Kelly, Jr., M. M. Thoren, and N. P. Salzman. 1971. Structure of replicating simian virus 40 deoxyribonucleic acid molecules. *J. Virol.* **8**:478-490.
- Senapathy, P., and B. J. Carter. 1984. Molecular cloning of adeno-associated virus variant genomes and generation of infectious virus by recombination in mammalian cells. *J. Biol. Chem.* **259**:4661-4666.
- Senapathy, P., J. D. Tratschin, and B. J. Carter. 1984. Replication of adeno-associated virus DNA. Complementation of naturally occurring rep⁻ mutants by a wild type genome or an ori⁻ mutant and correction of terminal palindrome deletions. *J. Mol. Biol.* **179**:1-20.
- Smith, G. E., and M. D. Summers. 1980. The bidirectional transfer of DNA and RNA to nitrocellulose or diazobenzyloxymethyl paper. *Anal. Biochem.* **109**:123-129.
- Smith, M., D. W. Leung, S. Gillam, G. R. Asfell, D. L. Mont-

- gomery, and B. D. Hall. 1979. Sequencing of the gene for iso-1-cytochrome C in *Saccharomyces cerevisiae*. *Cell* **16**:753-761.
37. Southern, E. M. 1975. Detection of specific sequences among DNA fragments separated by gel electrophoresis. *J. Mol. Biol.* **98**:503-517.
38. Srivastava, A., E. W. Lusby, and K. I. Berns. 1983. Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J. Virol.* **45**:555-564.
39. Straus, S. E., E. Sebring, and J. A. Rose. 1976. Concatemers of alternating plus and minus strands are intermediates in adeno-virus-associated virus DNA synthesis. *Proc. Natl. Acad. Sci. USA* **73**:742-746.
40. Vogelstein, B., and D. Gillespie. 1979. Preparative and analytical purification of DNA from agarose. *Proc. Natl. Acad. Sci. USA* **76**:615-619.
41. Winocour, E., and I. Keshet. 1980. Indiscriminate recombination in simian virus 40 infected monkey cells. *Proc. Natl. Acad. Sci. USA* **77**:4861-4865.
42. Winocour, E., V. Lavie, and I. Keshet. 1983. Structure of SV40- ϕ X174 recombinant genomes isolated from single cells. *J. Virol.* **48**:229-238.