JOURNAL OF VIROLOGY, Nov. 1995, p. 6917–6924 Vol. 69, No. 11
0022-538X/95/$04.00+0


# Recombinant Junctions Formed by Site-Specific Integration of Adeno-Associated Virus into an Episome

CATHERINE GIRAUD,[1] ERNEST WINOCOUR,[2] AND KENNETH I. BERNS[1]*

*Department of Microbiology, Hearst Microbiology Research Center, Cornell University Medical College, New York, New York 10021,[1] and Department of Molecular Genetics and Virology, Weizmann Institute of Science, Rehovot, Israel[2]*

Received 5 June 1995/Accepted 14 August 1995

**A model system using an episomal Epstein-Barr virus shuttle vector was recently developed to study the adeno-associated virus (AAV) site-specific integration event in chromosome 19q13.3-qter (C. Giraud, E. Winocour, and K. I. Berns, Proc. Natl. Acad. Sci. USA 91:10039–10043, 1994). In this study, we analyze the recombinant junctions generated after integration of the AAV genome into an Epstein-Barr virus shuttle vector carrying 8.2, 1.6, or 0.51 kb of the chromosome 19 preintegration sequence (AAVS1 locus). In most of the recombinants, one end of the viral genome was joined to a portion of the AAVS1 DNA previously shown to be a minimum target for AAV integration. Within this AAVS1 segment, the AAV insertion points were strikingly clustered around a binding site for the AAV regulatory protein. In all cases, the second junction with AAV occurred with vector DNA outside of the AAVS1 segment. With respect to the viral genome, one junction with the shuttle vector DNA occurred either within the AAV inverted terminal repeat (itr), or near the P5 promoter, approximately 100 nucleotides distal to a modified itr. The modified itr in 5 of 11 recombinants involved a head-to-tail organization. In one such instance, the AAV insert contained slightly more than one genome equivalent arranged in a head-to-tail manner with a junction close to the P5 promoter; the AAV insert in this recombinant episome could be rescued by adenovirus infection and replicated to virus particles. The significance of the head-to-tail organization is discussed in terms of the possible circularization of AAV DNA before or during integration.**

Adeno-associated virus (AAV) has been recognized recently as a likely vector for gene delivery in mammalian cells (4, 21, 31, 43). One particularly attractive feature of this virus is its ability to integrate with a high degree of specificity in the AAVS1 locus (preintegration sequence) on chromosome 19q13.3-qter (23–25, 33, 36). By using an Epstein-Barr virus (EBV) shuttle vector (27), we have previously shown that an 8.2-kb AAVS1 preintegration DNA, propagated as an extrachromosomal episome, is a target for site-specific AAV DNA integration (12). Sequential deletion of the 8.2-kb sequence identified a minimum sequence of 510 nucleotides (nt) able to direct integration. Several signals potentially involved in the integration process were identified on this short piece of DNA: a Rep 78/68 binding site (6, 42), a potential terminal resolution site (38), and an M26 motif (32, 37). This 510-bp fragment was also associated with rearrangements of the shuttle vector genome (12).

To characterize the AAV site-specific integration event, we have analyzed the products of integration in EBV vectors carrying the AAVS1 preintegration locus. Sequencing of the junctions between AAV and the vector has shown that a highly prevalent (in 80% of the recombinants) insertion point is located close to the previously described Rep binding motif and terminal resolution sequence at the 5′ end of the AAVS1 preintegration locus. This finding provides further evidence underlining the importance of these AAV recognition signals in site-specific integration. In addition, the sequencing data have revealed that 5 of the 11 recombinant structures analyzed contain a head-to-tail junction.

* Corresponding author. Mailing address: Department of Microbiology, Hearst Microbiology Research Center, Cornell University Medical College, 1300 York Ave., New York, NY 10021. Phone: (212) 746-6505. Fax: (212) 746-8587. Electronic mail address: KBerns@mail.med.cornell.edu.

## MATERIALS AND METHODS

**Production and selection of EBV recombinant vectors.** The C17 cell lines (3) propagating the p220.2 EBV shuttle vectors (9, 44) carrying segments of AAVS1 DNA, designated C17-p220.2(AAVS1 kb 0-8.2), C17-p220.2(AAVS1 kb 0-1.6), and C17-p220.2(AAVS1 kb 0-0.51), have been described previously (12). These cell lines were infected at several passage levels with AAV at an input multiplicity of 20 infectious units per cell. The extrachromosomal DNA was extracted 48 h postinfection with a slightly modified Hirt extraction procedure (18) and transfected into the SURE strain of *Escherichia coli* (14; Stratagene). The recombinants were isolated after colony hybridization of a replica filter with a nonradioactive single-stranded AAV genomic DNA probe as previously described (12).

**Oligonucleotide probes and hybridization.** AAV oligonucleotides (see Fig. 1) (AAV1, 5′-CTCACgTgACCTCTAATACAgg-3′; AAV2, 5′-gggACCTTAATC ACAATCTCg-3′; AAV3, 5′-CTTCTCggCCACggTCAgg-3′; AAV4, 5′-gAAAT gTCCTCCACgggCTg-3′; AAV5, 5′-CTTgTCgAgTCCgTTgAAgg-3′; AAV6, 5′-CATgAATCCTCTCATCgACC-3′; AAV7, 5′-gTggAgaTCgAgTgggAgC-3′; AA V8, 5′-ggCACCAgATACCTgACTCg-3′; AAV9, 5′-CTAgTTTCCATggCTAC g-3′; AAVD, 5′-ggAACCCCTAgTgATggAg-3′) and AAVS1 oligonucleotides (oligo 1, 5′-ACTTgCTAgTATgCCgTggg-3′; oligo2, 5′-CTACCTgCCCAgCAC ACC-3′; oligo 3, 5′-CATCCTCTCCggACATCg-3′), synthesized by the Oligo Etc Company, were labeled with the Dig-Oligonucleotide 3′ end Labeling Kit (Boehringer Mannheim). Vector DNA (100 ng) was dotted onto a nylon membrane (Hybond-N; Amersham), denatured with NaOH at 0.5 M plus NaCl at 1.5 M, neutralized with 0.5 M Tris-HCl (pH 8.0)–1.5 M NaCl, and UV cross-linked. The membranes were subsequently prehybridized for 2 h at 68°C and hybridized for 3 h at a temperature determined as optimum for each nucleotide (5 to 20°C under the melting temperature). Specific hybridization was detected with the Genius system kit (Boehringer Mannheim).

**Sequencing reactions.** Plasmid DNA was prepared by the Qiagen kit plasmid purification procedure (Qiagen Inc.), and sequencing reactions were performed with the Sequenase Quick-Denature plasmid sequencing kit (U.S. Biochemical). Oligonucleotides spread throughout the AAV genome (see Fig. 1) were used as primers for the sequencing reactions. To determine the junction sequences, an additional 15 primers (data available upon request) were utilized.

**PCR for determination of head-to-tail organization of inverted terminal repeat (itr) and adjacent DNA.** Plasmid DNA (10 ng) was mixed with 200 pmol each of oligonucleotides AAV1 and AAV8. The reaction was performed with the Boehringer PCR buffer (final concentrations, 10 mM Tris-HCl [pH 8.4], 50 mM

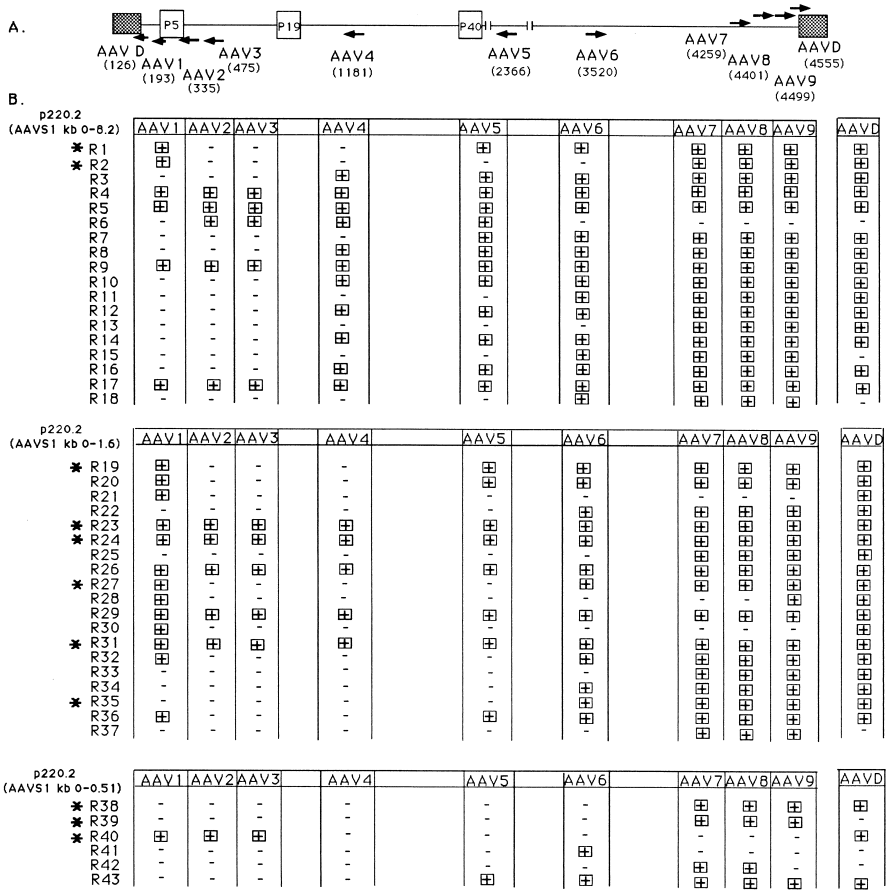


FIG. 1. Pattern of AAV sequences integrated in EBV recombinant vectors. (A) Schematic representation of the AAV genome showing the itr (▨), the promoters (P), and the oligonucleotides (→) spanning the AAV internal sequences and the D region of the itr used as probes. The gap indicates a disruption in the sequence with respect to the scale of the drawing. (B) Summary of dot blot hybridizations between AAV oligonucleotides and EBV recombinant vectors (indicated by the letter R followed by a number) isolated from the C17-p220.2(AAVS1 kb 0-8.2), C17-p220.2(AAVS1 kb 0-1.6), and C17-p220.2(AAVS1 kb 0-0.51) cell lines 48 h postinfection with AAV. The asterisks show the recombinants which were subsequently sequenced at their junctions with the AAV genome (see Fig. 2 to 5).

KCl, 1.5 mM MgCl2, and 0.01% gelatin) supplemented by the four deoxynucleoside triphosphates at a final concentration of 100 each μM and 2.5 U of *Taq* DNA polymerase (Boehringer) in a total volume of 100 μl. Thirty cycles (denaturation at 94°C for 45 s, annealing at 58°C for 45 s, and elongation at 72°C for 1 min) were carried out with a Perkin Elmer Cetus (Norwalk, Conn.) PCR apparatus, and 10 μl of the reaction mixture was loaded onto a 1.5% agarose gel. Reactions were also performed under the same conditions but with only one of the primers.

**Rescue of AAV from recombinant EBV shuttle vectors.** C17 cells were infected with adenovirus type 2 (Ad2) at an input multiplicity of 10 infectious units per cell, for 1 h at 37°C and subsequently transfected with 20 μg of purified EBV recombinant plasmids by using the cationic lipid *N*-[1(2,3-dioleoyloxy)propyl)]-*N*,*N*,*N*-trimethylammonium methyl sulfate (DOTAP) (10, 39). Forty hours later, the low-molecular-weight DNA was isolated (18) and analyzed for AAV DNA replication by the Southern blotting procedure.

To determine AAV capsid antigen synthesis, the cells were harvested 40 h after transfection, resuspended in radioimmunoprecipitation assay buffer (50 mM Tris-HCl [pH 7.5], 150 mM NaCl, 1% Triton X-100, 1% deoxycholate, 0.1% sodium dodecyl sulfate [SDS]) supplemented with 2 mM phenylmethylsulfonyl fluoride and 1 μg each of antipain, leupeptin, and pepstatin A, and centrifuged briefly at 4°C, and the supernatant fraction was kept at −20°C before analysis on an SDS–8% polyacrylamide gel electrophoresis gel. After electrophoresis, the polypeptides were transferred to a nitrocellulose membrane and immunodetection was performed with polyclonal capsid antibodies (kindly provided by Mertyn Malkinson) and a peroxidase label conjugate (Sigma). The chromogenic substrate 3-amino-9-ethylcarbazole (Sigma) was used to detect peroxidase activity.

To assay for infectious virus production, C17 cells were harvested 40 h post-transfection and lysed by several cycles of freeze-thawing. The extract was heated for 2 h at 56°C to inactivate adenovirus, clarified by centrifugation, mixed with a fresh sample of Ad2, and adsorbed to HeLa cells. At 40 h postinfection, the cells were lysed and Ad2 was inactivated as described above. One-tenth of the lysate was treated with DNase I (10 μg/ml) for 30 min at 37°C, denatured with 0.8 M NaOH, neutralized with 0.8 M $NH_4C_2H_3O_2$, dot blotted to nitrocellulose, and hybridized with a $^{32}$P-labeled AAV DNA probe.

## RESULTS

A useful feature of the EBV shuttle vector system is that the recombinant products can be retrieved in *E. coli* for structural analysis. Accordingly, we isolated EBV-AAV recombinants from cell lines that carried p220.2 shuttle vectors with 8.2-, 1.6-, and 0.51-kb segments of AAVS1 DNA and had been infected at various passage levels. The recombinants were analyzed with respect to AAV DNA content, the sequences at the junctions, and the capacity of the AAV genome to be rescued from the episomal shuttle vector by adenovirus infection.

**AAV DNA in recombinants.** To define the AAV sequences present on the EBV vectors after infection of the C17-p220.2 (AAVS1 kb 0-8.2), C17-p220.2 (AAVS1 kb 0-1.6), and C17-p220.2(AAVS1 kb 0-0.51) cell lines, we used a set of AAV oligonucleotides spanning the AAV genome to hybridize dot blots of the recombinant DNAs (Fig. 1). Some of the recombinants (9 of 43) reacted positively to all of the probes used for hybridization, suggesting that the entire AAV genome might be present. No evidence of the presence of several separate inserts of AAV DNA or several copies of the same DNA was

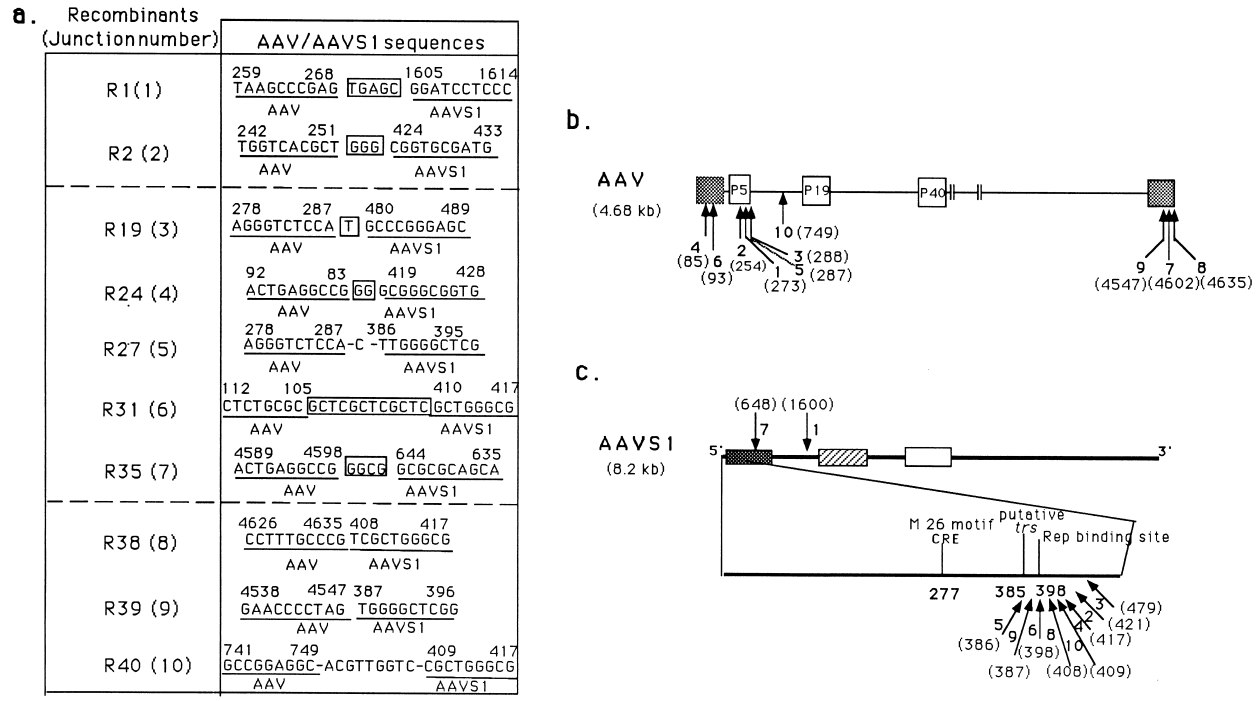


FIG. 2. Junctions between AAV and AAVS1 sequences. (a) Sequences at the nucleotide level. Boxed sequences at the junctions are common to both viral and vector genomes. A junction number (in parentheses) was given to each recombinant, and the same number is used in panels b and c. The horizontal dotted lines separate the three groups of recombinants isolated from three different cell lines (Fig. 1). (b) Schematic representation of the junctions with respect to the AAV genome. (c) Schematic representation of the junctions with respect to the AAVS1 segment of chromosome 19. The AAVS1 sequence (8.2 kb) shows a CpG island (▩), a region corresponding to a partial cDNA clone (▨), and a minisatellite repetitive DNA sequence (▭). The enlargement of the first 510 nt at the 5′ end of AAVS1 shows the M26 motif from *Saccharomyces cerevisiae*, a putative terminal resolution site (*trs*), and a Rep binding site. CRE, cyclic Amp response element.

found by restriction digestion (data not shown). Most of the recombinants (32 of 43) had deletions in the *rep* genes; the deletions extended to the capsid genes in at least 20 of these cases. Thirty-six of forty-three recombinants reacted positively to the oligonucleotide AAVD located in the D region of the itr. The D region of the AAV itr is outside of the 125-bp symmetrical sequences able to fold into a T-shaped structure. Hence, the itr, or at least part of it, is present in most of the recombinants. The pattern of the AAV sequences found in each recombinant (Fig. 1) was used to select AAV oligonucleotides for sequencing of the junctions between the viral and vector DNA segments.

**AAV and vector sequences at recombinant junctions.** The junctions of 11 recombinants (noted by asterisks in Fig. 1) were sequenced by using multiple oligonucleotides as primers (see Material and Methods and Fig. 2 to 5).

**(i) Junctions between AAV and AAVS1 sequences.** A junction between AAV DNA and the AAVS1 DNA segment of the p220.2 vector was found in all of the recombinants analyzed but one (R23; Fig. 2). Surprisingly, a second junction between these two DNAs was never found in the same recombinant. In AAV, the junctions were localized in the itr or near the P5 promoter (Fig. 2a and b). In one case (R40), the junction was found in the *rep* gene sequence.

In 80% of the cases, the junctions in the AAVS1 sequence were found in the 5′ 510-bp segment previously defined as the minimum chromosome 19 DNA sequence able to direct AAV integration (12). Strikingly, the junctions within the 510-bp segment were tightly clustered around the AAVS1 Rep binding site (Fig. 2a and c). Only two junctions were found 3′ to the 510-bp segment (in R1 and R35). Because the AAVS1 DNA at the junction was found to be oriented towards the 3′ end of AAVS1 (R35 is the sole exception; Fig. 2) and because a second junction with AAVS1 DNA was never found in the same recombinant DNA molecule, we wished to determine if AAVS1 sequences 5′ to the insertion might have been preferentially lost during integration. Therefore, we used three AAVS1 oligonucleotides (nt 60 to 79, 248 to 265, and 453 to 435) to screen by dot blot hybridization for the presence of AAVS1 DNA 5′ to the junction with AAV DNA. The 5′ AAVS1 DNA sequences were found in 7 of 11 cases (Table 1). Thus, although AAVS1 DNA 5′ to the AAV junction was present in 64% of the recombinant vectors, this DNA segment

TABLE 1. AAVS1 sequences from the 5′ 510-bp fragment present in EBV recombinant vectors[a]

| Recombinant (junction point in AAVS1 sequence [nt]) | Hybridization with AAVS1 oligonucleotide: | | |
|---|---|---|---|
| | 1 (nt 60–79) | 2 (nt 248–265) | 3 (nt 453–435) |
| R1 (1600) | − | − | − |
| R2 (421) | − | − | + |
| R19 (479) | + | + | + |
| R23 (?) | + | + | + |
| R24 (417) | + | + | + |
| R27 (386) | − | − | + |
| R31 (398) | − | − | + |
| R35 (648) | + | + | + |
| R38 (408) | + | + | + |
| R39 (387) | + | + | + |
| R40 (409) | + | + | + |

[a] The presence of sequences from the 5′ end of the preintegration locus was assessed by dot blot hybridization with oligonucleotide probes.

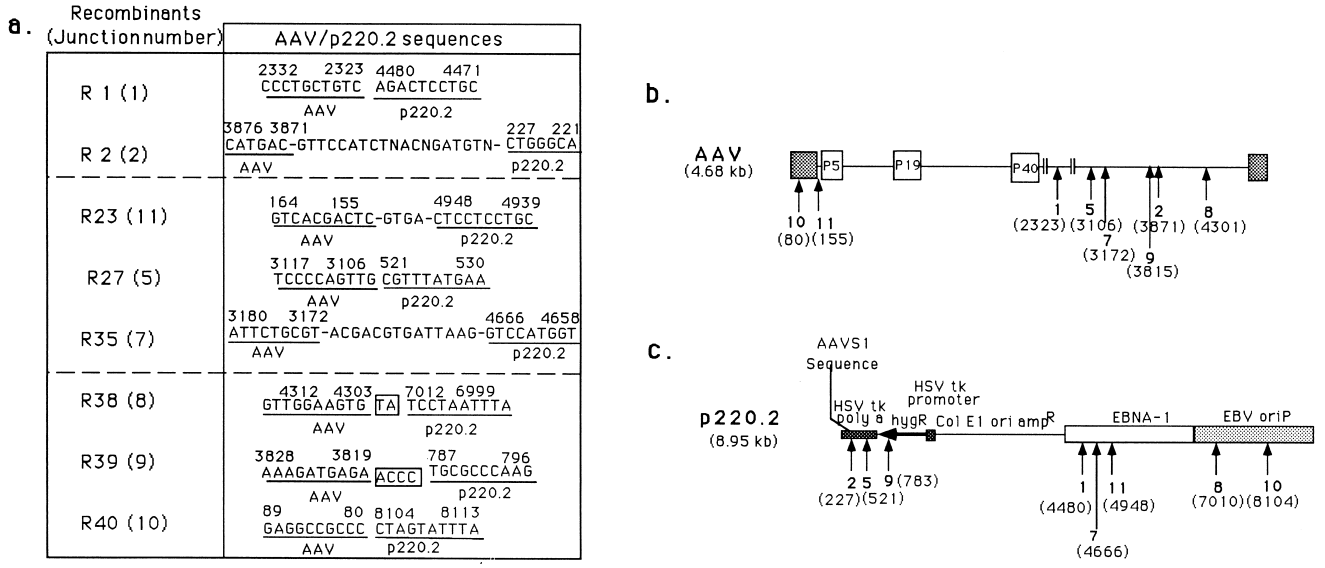


FIG. 3. Junctions between AAV and p220.2 sequences outside of the AAVS1 segment. (a) Sequences at the nucleotide level. Boxed sequences at the junctions are common to both viral and vector genomes. A junction number (in parentheses) was given to each recombinant, and the same number is used in panels b and c. The horizontal dotted lines separate the three groups of recombinants isolated from three different cell lines (Fig. 1). (b) Schematic representation of the junctions with respect to the AAV genome. (c) Schematic representation of the junctions with respect to the p220.2 plasmid genome. The 8.9-kb p220.2 plasmid contains the herpes simplex virus (HSV) thymidine kinase (tk) promoter and polyadenylation signal (poly a) flanking a hygromycin resistance-encoding gene (hygR), the EBV latent origin of replication (oriP), the EBV-encoded nuclear antigen EBNA-1, and *E. coli* sequences (Col E1 and ori amp$^R$).

was apparently translocated to other positions in the vector genome. Whether the rearrangements of AAVS1 DNA occurred during AAV integration or as a consequence of vector rearrangements which have been observed to occur prior to AAV infection (12) is not known, although the fact that the rearrangement always occurs immediately 5′ to the junction site suggests the former possibility.

Inspection of the AAV-AAVS1 junction sequences (Fig. 2a) revealed only limited patchy nucleotide sequence homology, indicating that integration into the vector occurs by nonhomologous recombination.

**(ii) Junctions between AAV DNA and p220.2 vector DNA outside of the AAVS1 segment.** As noted above, the second junction between AAV DNA and the vector genome occurred outside of the AAVS1 segment in each recombinant examined. With respect to the AAV genome, these junctions were located in the itr or in the *cap* gene (in the latter case, substantial portions of the *rep* and capsid genes were deleted in the recombinant genome) (Fig. 3a and b). With respect to the p220.2 vector, the junctions were found in the EBV *oriP* sequence, in the EBV EBNA-1 gene, and in the hygromycin resistance-encoding gene (Fig. 3a and c). Despite the crucial role of these vector elements for stable episome propagation in animal cells (27), the period of 48 h between AAV infection and recovery of the vectors in bacteria was probably too short for the cells to lose the plasmid by dilution. As expected, no junctions were found in the *E. coli* elements necessary for selection in bacteria by virtue of the screening procedure.

**(iii) Head-to-tail junctions between AAV sequences.** When primers located near the left and right itrs (AAV1 and AAV9; Fig. 1A) were used to sequence the junctions, it became evident that the itr and adjacent DNA were sometimes organized in a head-to-tail fashion, relative to the wild-type AAV genome (recombinants R1, R2, R19, R27, and R31 in Fig. 4). In the wild-type AAV genome, the first 125 nt of the 145-nt itr can form a T-shaped hairpin structure because of the presence of two small internal palindromes (B-B′ and C-C′) flanked by a larger palindrome (A-A′) (1). In addition, the complete itr element includes a single-stranded D region located outside of the palindromic hairpin. The head-to-tail AAV-AAV recombinant junctions are composed of one complete itr (D-A′-B′-B-C′-C-A) expanded by a 20-nt D′ region (D-A′-B′-B-C′-C-A-D′). Thus, in the recombinants with the head-to-tail orientation, the 3′ end of the AAV genome is joined to nt 126 at the 5′ end (Fig. 4). The R1, R2, R19, and R27 recombinants which display the unusual head-to-tail organization of the itr and adjacent DNA are also characterized by two additional common features; the junction with AAVS1 occurs near the P5 promoter, and a substantial segment of the AAV internal sequences (the *cap* and *rep* genes) is deleted (Fig. 2; see Fig. 6A). The fifth recombinant structure, which displays the head-to-tail organization of the itr and adjacent DNA at one end (R31), contains the internal AAV sequences and is linked to AAVS1 DNA via a disrupted itr at the other end (Fig. 6C); the AAV genome integrated into the R31 recombinant vector was rescued by adenovirus infection (see below).

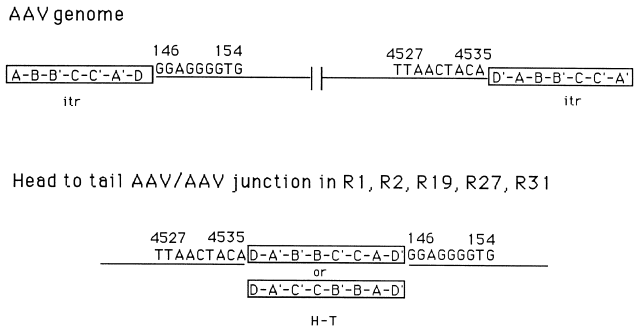


FIG. 4. Junctions between AAV sequences. The sequences D-A′-B′-B-C′-C-A-D′ and D-A′-C′-C-B′-B-A-D′ refer to the head-to-tail (H-T) organization of the recombinant AAV itr relative to the wild-type itr. The gap indicates a disruption in the sequence with respect to the scale of the drawing.

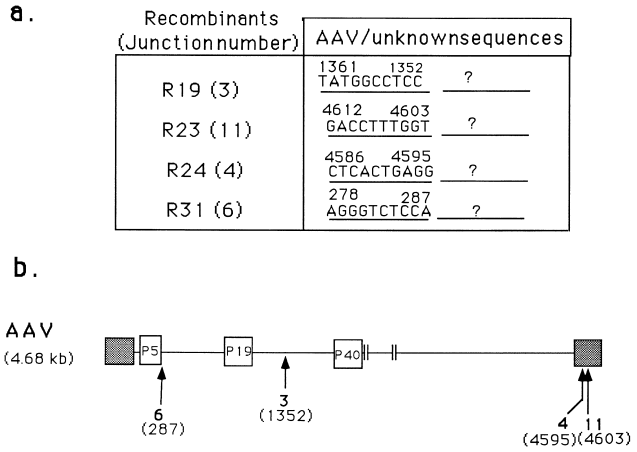


FIG. 5. Junctions between AAV and nonviral-nonvector sequences. (a) Sequences at the nucleotide level. A junction number (in parentheses) was given to each recombinant, and the same number is used in panel b. (b) Schematic representation of the junctions with respect to the AAV sequence.

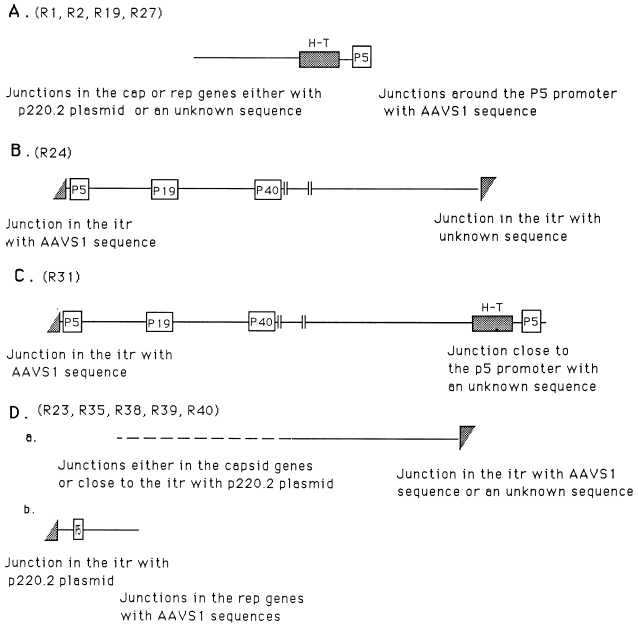


FIG. 6. Patterns of AAV inserts and junctions in EBV recombinant vectors. ▨, head-to-tail (H-T) organization of the itr (see the text for details) and adjacent DNA; ◤, disrupted itrs.

The head-to-tail organization of AAV DNA in recombinant R1, R2, R19, R27, and R31 DNAs was confirmed by PCRs with oligonucleotides AAV1 and AAV8 (Fig. 1A) as primers. The product was 230 nt long as judged by gel electrophoresis and hybridized with the AAV9 oligonucleotide (Fig. 1), thus confirming the presence of itr DNA.

When oligonucleotide primers AAV1 and AAV8 were used independently in PCRs, no products were detected, indicating that the organization of these AAV-AAV recombinant junctions is not head to head or tail to tail. The recombinants showing head-to-tail organization of the AAV itr and adjacent DNA were isolated in independent experiments from two different infected cell lines [R1 and R2 from C17-p220.2(AAVS1 kb 0-8.2) cells and R19 and R27 from C17-p220.2(AAVS1 kb 0-1.6) cells] (Fig. 1). The similarity of the head-to-tail organization of AAV DNA in each case is striking and suggests a common recombination-integration mechanism. The possibility that this involves circularization of AAV DNA prior to or during integration is discussed below.

**(iv) Junctions between AAV DNA and DNA unrelated to vector or viral DNA.** DNA unrelated to the vector or viral genomes was found at the junctions with AAV DNA in four recombinants (R19, R23, R24, and R31; Fig. 5). The unrelated sequence in R24 shows some homology with the DNA of the human Line-1 element (long interspersed repetitive sequence derived from a retrotransposon) (19). An oligonucleotide prepared from the R24 unrelated sequence hybridized with the DNAs of the R19 and R23 recombinants (but not with R31 DNA). All of the recombinants with nonviral-nonvector DNA, related to human Line-1 DNA, were derived from a single infected cell line propagating the p220.2(AAVS1 0-1.6 kb) episome. One vector isolate of this cell line which did not contain AAV DNA also hybridized with the oligonucleotide derived from the R24 nonviral-nonvector segment. These results suggest that some p220.2(AAVS1 0-1.6 kb) episomes acquired the human Line-1 DNA prior to recombination with AAV DNA. The human Line-1 element is known to be capable of existing as part of an extrachromosomal circular element (20).

**Summary of recombinant junctions.** On the basis of the sequencing data from 11 recombinants, the junctions can be grouped as shown in Fig. 6. In group A, one junction with the episome joins AAV DNA from around the P5 promoter to the episomal AAVS1 segment; the second junction occurs either between a p220.2 sequence (that is outside the AAVS1 segment) and the AAV capsid genes (R1, R2, and R27) or between an AAV *rep* gene sequence and nonviral-nonvector DNA (probably Line-1 DNA) (R19). All of the recombinants in group A show head-to-tail organization of the AAV itr and adjacent sequences, and much of the internal AAV sequences are deleted. In group B (R24), all of the internal AAV sequences appear to be present; one junction with AAVS1 is via a disrupted itr element; the second junction with human Line-1 DNA in the episome also occurs via a disrupted itr element. In group C (R31), the organization of the AAV insert is similar to that of group B except that one junction shows the head-to-tail organization of the viral itr and adjacent DNA and connects with episomal nonviral-nonvector DNA at a point near the P5 promoter. The AAV insert in R31 was rescuable by adenovirus infection (see below). Group D (R23, R35, R38, R39, and R40) is characterized by the presence of a segment of AAV DNA and a disrupted itr joined to either AAVS1, p220.2, or the nonviral-nonvector DNA.

**Rescue of AAV from EBV recombinant vectors.** It was of interest to determine if the AAV DNA in the recombinants which reacted with all of the virus-specific oligonucleotide probes (R4, R5, R9, R17, R23, R24, R26, R29, and R31 in Fig. 1) could be rescued by infection with adenovirus. Accordingly, C17 cells were first infected with Ad2 and then transfected with the plasmid-purified shuttle vector-AAV recombinants noted above. At 40 h posttransfection-infection, expression of the AAV capsid genes, rescue and replication of the AAV DNA, and production of infectious particles from these EBV recombinant vectors were assessed.

AAV capsid gene expression was determined by Western blot (immunoblot) analysis with a polyclonal AAV capsid antibody (Fig. 7). R31 and R26 (data not shown) expressed the capsid proteins at a level similar to that seen with the pSM620 plasmid (34), suggesting that the AAV DNA is rescued from the EBV recombinant vectors and replicated. A low expression

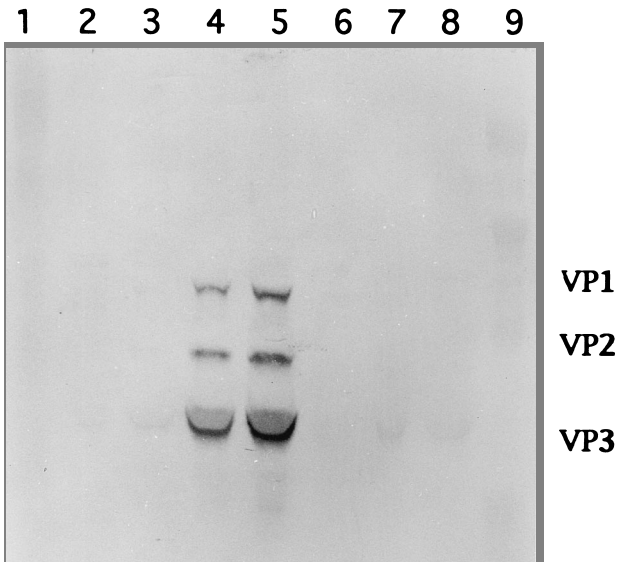


FIG. 7. Capsid gene expression of AAV rescued from EBV recombinant vectors. Immunoblot analysis of C17 cells infected with Ad2 and transfected with plasmid pAAV/Ad (35) (lane 3), plasmid pSM620 (34) (lane 4), recombinant R31 (lane 5), recombinant R9 (lane 6), recombinant R24 (lane 7), or recombinant R23 (lane 8). Lanes 1 and 9 contained, respectively, a Rainbow molecular weight marker (Amersham) and a prestained SDS-polyacrylamide gel electrophoresis standard high-range marker (Bio-Rad). The molecular masses of proteins VP1, VP2, and VP3 are 87, 73, and 62 kDa. Lane 2 contained a mock-transfected control. The antiserum used in the immunoblots was a polyclonal anti-AAV capsid serum.

level of the VP polypeptides was seen with some of the other recombinants (Fig. 7).

To assess rescue and DNA replication, the low-molecular-weight DNA was extracted from the transfected-infected C17 cells and analyzed on a 0.7% agarose gel (Fig. 8A). Only the extracts from cells transfected with recombinants R31 and R26 (data not shown) contained detectable amounts of the AAV monomer and dimer forms also seen when AAV DNA is rescued from the pSM620 plasmid. These products were resistant to *Dpn*I digestion (data not shown). Hybridization with an AAV probe (Fig. 8B) confirmed the nature of the DNA.

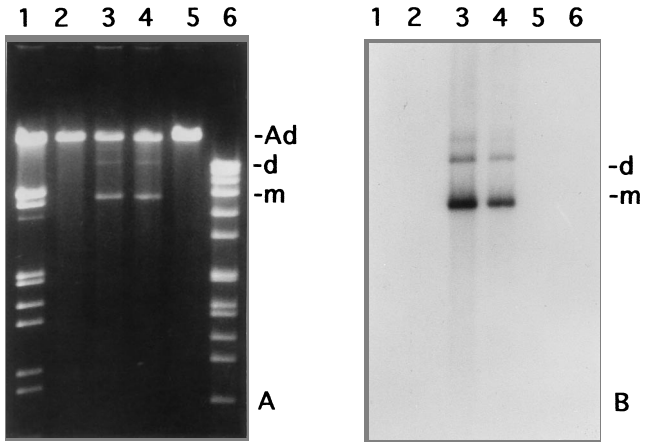


FIG. 8. Rescue and replication of AAV DNA from EBV recombinant vectors. Extrachromosomal DNA was extracted from C17 cells infected with Ad2 and transfected with recombinant R9 (lane 2), recombinant R31 (lane 3), or plasmid pSM620 (lane 4) or mock transfected (lane 5). Lanes 1 and 6 contained DNA molecular weight markers III and VII (Boehringer Mannheim) (A) Agarose (0.7%) gel stained with ethidium bromide. (B) Southern blot hybridization with an AAV probe. Abbreviations: m, monomer, d, dimer; Ad, adenovirus.

Production of infectious particles was assessed by adsorbing the lysates of transfected-infected C17 cells to adenovirus-infected HeLa cells. At 40 h postinfection, the presence of progeny AAV in the HeLa cells was determined by dot blot hybridization (AAV probe) of HeLa cell lysates treated first with DNase I (to screen for DNase-resistant particles) and then with NaOH to extract and denature the DNA. Two recombinants (R26 and R31) gave rise to progeny AAV in the above-described procedure. Both R26 and R31 contain the head-to-tail organization of the itr and adjacent sequences. Thus, despite this divergence from the wild-type organization of the AAV genome, the AAV inserts in R26 and R31 could be rescued and replicated to infectious particles.

## DISCUSSION

AAV is unique among animal viruses in its capacity to undergo site-specific chromosomal integration. To aid the investigation of the targetting mechanism, we previously developed an EBV-based shuttle vector system in which the chromosome 19 AAVS1 preintegration DNA propagates as an episome into which AAV can integrate (12). The objective of the present experiments was to define the structure of the recombinant junctions in episomal integration and to compare such junctions with AAV insertions into AAVS1 when it is part of the intact chromosome.

Although head-to-tail and tail-to-tail arrangements in chromosomal inserts containing multiple copies of the AAV genome have been reported previously (5, 22, 26, 30, 45), the finding of head-to-tail organization of the AAV itr and adjacent sequences in 5 of 11 recombinants sequenced (Fig. 6A and C) was not anticipated because only 1 of these contained an insert of greater than unit length. In these recombinants, the itr was modified by addition of a 20-nt D region and linked the 3′ end of the AAV genome to AAV nt 126 at the 5′ end. In all five of these recombinant junctions, the crossover point with the vector occurred at or slightly downstream of the P5 promoter at AAV nt 250 (Fig. 2b). Although the retention of viral DNA in this group of independently isolated recombinants varied from a rescuable, complete genome equivalent in R31 to extensive deletions in the *rep* and *cap* genes in R1, R2, R19, and R27, the structure of the head-to-tail modification was remarkably similar in each case. Multiple copies of the AAV genome in tandem array could occur in an insert as the consequence of either recombination or replication of the inserted sequence before or during integration. Recombination via the itr could result in either head-to-tail, tail-to-tail, or head-to-head orientation of adjacent copies of the genome. The former two possibilities have been reported (5, 16, 22, 26, 30, 45). Although the predominant frequency of head-to-tail orientation in chromosomal integration suggests the likelihood that a rolling-circle form of replication (11) is involved; even in this event, the first step would be formation of a circular intermediate by recombination via the itr. Integration of less than a full genome equivalent of any of the recombinants characterized in this report does not require replication but does suggest that a circular molecule is an intermediate in the process. As discussed elsewhere, (31), if such integrated forms do arise by replication, the mode of replication must differ from that of AAV during lytic infection, which gives rise only to head-to-head or tail-to-tail intermediates (1, 17, 40). The proposal that viral genomes may undergo an alternative mode of replication, quite different from that characteristic of replication during lytic infection, is not without precedent. Aberrant rolling-circle

replication intermediates—with high recombinogenic potential—have been described for papovavirus DNA (2, 7, 15), which normally replicates bidirectionally, giving rise to theta-type intermediates during lytic infection (41). Indeed, rolling-circle replication intermediates have been proposed as intermediates in papovavirus integration (8).

Previously we have shown that a 510-bp segment at the 5′ end of AAVS1 that has been cloned (23) is capable of directing site-specific integration of AAV into the EBV-based episomal vector (12). Likely recognition signals that might target the AAV genome to AAVS1 are the Rep binding site at AAVS1 nt 398 to 413 (6, 42), a potential terminal resolution site (38) at nt 384 to 389, and the yeast recombinogenic M26 motif (32, 37) at nt 277 to 283. Our analysis of the recombinant junctions in the episome has shown that in 80% of the cases examined, the AAV insertion points are tightly clustered around the Rep binding site (AAVS1 nt 386 to 479; Fig. 2c). However, two junctions with AAVS1 DNA, both 3′ to the 510-bp segment (AAVS1 nt 648 and 1600), were also detected. The tight clustering of a major proportion of crossover points around the potential terminal resolution site and Rep binding motif certainly underlines the importance of these signals in the targetting mechanism.

In latently infected cell lines, the few AAV insertion points that have been determined were located at AAVS1 nt 1026 to 1030 and 1144 to 1146 (23), in a region from nt 713 to 1303 (36, 45), and at position 727 (13). These insertion points in AAVS1, when it is part of intact chromosome 19, are somewhat 3′ to the tight cluster around positions 386 to 479 when AAVS1 is propagated in an episomal vector. The reason for this difference is not known. It should be noted, however, that mapping of AAVS1 recombinant junctions in latently infected cells can be performed only after many cell generations of growth subsequent to the initial integration event. In contrast, the recombinants generated in the EBV episomal system are not subject to this constraint.

Apart from the difference in the distance of the insertion points from the Rep binding motif, noted above, integration of AAV into AAVS1 in the manipulatable EBV-based episomal system shows many points of similarity to AAV integration into AAVS1 as part of intact chromosome 19. A genome equivalent of the viral DNA can be rescued by adenovirus superinfection in both cases. The AAVS1 target is substantially disrupted and rearranged both in the episome and in the chromosome. In the episome, AAVS1 sequences 5′ to the viral insertion are no longer present at the junction, having apparently been translocated to other vector positions. In the chromosome, disruption of AAVS1 is severe enough that it becomes diagnostic for AAV site-specific integration (25). The head-to-tail organization of the viral itr and adjacent DNA can also be encountered in both chromosomal integration (45) and episomal integration. Similarly, viral recombinant junctions localized to the P5 promoter region (which also contains a Rep binding sequence; 28, 29) have been noted in both systems (45). Thus, the highly manipulatable EBV-based episomal system reflects many aspects of AAV integration at the intact-chromosomal level.

## ACKNOWLEDGMENTS

We thank Erik Falck-Pedersen, Peter Ward, Saburo Kashii, and Michael Linden for helpful discussions and critical reading of the manuscript. We thank N. Cortez for excellent technical assistance.

This work was supported by grant AI122251 from the U.S. Public Health Service.

## REFERENCES

1. **Berns, K. I.** 1990. Parvovirus replication. Microbiol. Rev. **54:**316–329.
2. **Bjursell, G.** 1978. Effects of 2′-deoxy-2′-azidocytidine on polyoma virus DNA replication: evidence for rolling circle-type mechanism. J. Virol. **26:** 136–142.
3. **Canfield, V., J. R. Emanuel, N. Spickofsky, R. Levenson, and R. F. Margolskee.** 1990. Ouabain-resistant mutants of the rat Na,K-ATPase α2 isoform identified by using an episomal expression vector. Mol. Cell. Biol. **10:**1367–1372.
4. **Carter, B. J.** 1992. Adeno-associated virus vectors. Curr. Opin. Biotechnol. **3:**533–539.
5. **Cheung, A. K., M. D. Hoggan, W. W. Hauswirth, and K. I. Berns.** 1980. Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. J. Virol. **33:**739–748.
6. **Chiorini, J. A., M. D. Weitzman, R. A. Owens, E. Urcelay, B. Safer, and R. M. Kotin.** 1994. Biologically active Rep proteins of adeno-associated virus type 2 produced as fusion proteins in *Escherichia coli*. J. Virol. **68:**797–804.
7. **Deichaite, 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.
8. **Dorsett, D., I. Deichaite, and E. Winocour.** 1985. Cellular and linear simian virus 40 DNAs differ in recombination. Mol. Cell. Biol. **5:**869–880.
9. **DuBridge, R. B., P. Tang, H. C. Hsia, P.-M. Leong, J. H. Miller, and M. P. Calos.** 1987. Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. Mol. Cell. Biol. **7:**379–387.
10. **Felgner, P. L., T. R. Gadek, M. Holm, R. Roman, H. W. Chan, M. Wenz, J. P. Northrop, G. M. Ringold, and M. Danielsen.** 1987. Lipofection: a highly efficient, lipid-mediated DNA-transfection procedure. Proc. Natl. Acad. Sci. USA **84:**7413–7417.
11. **Gilbert, W., and D. Dressler.** 1968. DNA replication: the rolling circle. Cold Spring Harbor Symp. Quant. Biol. **33:**473–484.
12. **Giraud, C., E. Winocour, and K. I. Berns.** 1994. Site-specific integration by adeno-associated virus is directed by a cellular DNA sequence. Proc. Natl. Acad. Sci. USA **91:**10039–10043.
13. **Goodman, S., X. Xiao, R. E. Donahue, A. Moulton, J. Miller, C. Walsh, N. S. Young, R. J. Samulski, and A. W. Nienhuis.** 1994. Recombinant adeno-associated virus-mediated gene transfer into hematopoietic progenitor cells. Blood **84:**1492–1500.
14. **Greener, A.** 1990. E. coli SURE TM strain: clone "unclonable" DNA. Strategies **3:**5–6.
15. **Grossman, Z., K. I. Berns, and I. Winocour.** 1985. Structure of simian virus 40–adeno-associated virus recombinant genomes. J. Virol. **56:**457–465.
16. **Handa, H., K. Shiroki, and H. Shimojo.** 1977. Establishment and characterization of KB cells lines latently infected with adeno-associated virus type 1. Virology **82:**84–92.
17. **Hauswirth, W. W., and K. I. Berns.** 1979. Adeno-associated virus DNA replication: non unit-length molecules. Virology **93:**57–68.
18. **Hirt, B.** 1967. Selective extraction of polyoma DNA from infected mouse cell cultures. J. Mol. Biol. **26:**365–369.
19. **Hohjoh, H., R. Minakami, and Y. Sakaki.** 1990. Selective cloning and sequence analysis of the human L1 (Line-1) sequences which transposed in the relatively recent past. Nucleic Acids Res. **18:**4099–4104.
20. **Jones, R. S., and S. Potter.** 1985. L1 sequences in Hela extrachromosomal circular DNA: evidence for circularization by homologous recombination. Proc. Natl. Acad. Sci. USA **82:**1989–1993.
21. **Kotin, R. M.** 1994. Prospects for the use of adeno-associated virus as a vector for human gene therapy. Hum. Gene Ther. **5:**793–801.
22. **Kotin, R. M., and K. I. Berns.** 1989. Organization of adeno-associated virus DNA in latently infected Detroit 6 cells. Virology **170:**460–467.
23. **Kotin, R. M., R. M. Linden, and K. I. Berns.** 1992. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. EMBO J. **11:**5071–5078.
24. **Kotin, R. M., J. C. Menninger, D. C. Ward, and K. I. Berns.** 1991. Mapping and direct visualization of a region-specific viral DNA integration site on chromosome 19q13-qter. Genomics **10:**831–834.
25. **Kotin, R. M., M. Siniscalco, R. J. Samulski, X. D. Zhu, L. Hunter, C. A. Laughlin, S. McLaughlin, N. Muzyczka, M. Rocchi, and K. I. Berns.** 1990. Site-specific integration by adeno-associated virus. Proc. Natl. Acad. Sci. USA **87:**2211–2215.
26. **Laughlin, C. A., C. B. Cardellichio, and H. C. Coon.** 1986. Latent infection of KB cells with adeno-associated virus type 2. J. Virol. **60:**515–524.
27. **Margolskee, R. F.** 1992. Epstein-Barr virus based expression vectors. Curr. Top. Microbiol. Immunol. **158:**67–95.
28. **McCarty, D. M., D. J. Pereira, I. Zolotukhin, X. Zhou, J. H. Ryan, and N. Muzyczka.** 1994. Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein. J. Virol. **68:**4988–4997.
29. **McCarty, D. M., J. H. Ryan, S. Zolotukhin, X. Zhou, and N. Muzyczka.** 1994. Interaction of the adeno-associated virus Rep protein with a sequence within the A palindrome of the viral terminal repeat. J. Virol. **68:**4998–5006.
30. **McLaughlin, S. K., P. Collis, P. L. Hermonat, and N. Muzyczka.** 1988. Adeno-associated virus general transduction vectors: analysis of proviral structures. J. Virol. **62:**1963–1973.
31. **Muzyczka, N.** 1992. Use of adeno-associated virus as a general transduction vector for mammalian cells. Curr. Top. Microbiol. Immunol. **158:**97–129.
32. **Ponticelli, S. A., and G. R. Smith.** 1992. Chromosomal context dependence

of a eukariotic recombinational hot spot. Proc. Natl. Acad. Sci. USA **89:**227–231.
33. **Samulski, R. J.** 1993. Adeno-associated virus: integration at a specific chromosomal locus. Curr. Opin. Genet. Dev. **3:**74–80.
34. **Samulski, R. J., K. I. Berns, M. Tan, and N. Muzyczka.** 1982. Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. Proc. Natl. Acad. Sci. USA **79:**2077–2081.
35. **Samulski, R. J., L. S. Chang, and T. Shenk.** 1989. Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. J. Virol. **63:**3822–3828.
36. **Samulski, R. J., X. Zhu, X. Xiao, J. D. Brook, D. E. Housman, N. Epstein, and L. A. Hunter.** 1991. Targeted integration of adeno-associated virus (AAV) into human chromosome 19. EMBO J. **10:**3941–3950.
37. **Schuchert, P., M. Langsford, E. Kaslin, and J. Kohli.** 1991. A specific DNA sequence is required for high frequency of recombination in the ade 6 gene of fission yeast. EMBO J. **10:**2157–2163.
38. **Snyder, R. O., D.-S. Im, T. Ni, X. Xiao, R. J. Samulski, and N. Muzyczka.** 1993. Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. J. Virol. **67:**6096–6104.
39. **Stamatatos, L., R. Leventis, M. J. Zuckermann, and J. R. Silvius.** 1988. Interactions of cationic lipid vesicles with negatively charged phospholipid vesicles and biological membranes. Biochemistry **27:**3917–3925.
40. **Straus, S. E., E. D. Sebring, and J. A. Rose.** 1976. Concatemers of alternating plus and minus strands are intermediates in adenovirus-associated virus DNA synthesis. Proc. Natl. Acad. Sci. USA **73:**742–746.
41. **Tooze, J. (ed.).** 1980. Molecular biology of tumor viruses, 2nd ed. Part 2, DNA tumor viruses. Cold Spring Harbor Laboratory, Cold Spring Harbor, N.Y.
42. **Weitzman, M. D., S. R. Kyostio, R. M. Kotin, and R. A. Owens.** 1994. Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. Proc. Natl. Acad. Sci. USA **91:**5808–5812.
43. **Xiao, X., W. deVlaminck, and J. Monahan.** 1993. Adeno-associated virus (AAV) vectors for gene transfer. Adv. Drug Delivery Rev. **12:**201–215.
44. **Yates, J. L., N. Warren, and B. Sugden.** 1985. Stable replication of plasmids derived from Epstein-Barr virus in various mammalian cells. Nature (London) **313:**812–815.
45. **Zhu, X.** 1993. Characterization of adeno-associated virus proviral structure in latently infected human cells. Ph.D. thesis. University of Pittsburgh, Pittsburgh, Pa.