Adeno-associated virus (AAV) site-specific recombination does not require a Rep-dependent origin of replication within the AAV terminal repeat

Samuel M. Young, Jr.*†‡, and Richard Jude Samulski*†§¶

*Curriculum in Genetics and Molecular Biology, §Department of Pharmacology, and †Gene Therapy Center, University of North Carolina, Chapel Hill, NC 27599

Communicated by Clyde A. Hutchison III, University of North Carolina, Chapel Hill, NC, September 26, 2001 (received for review August 18, 2000)

Adeno-associated virus (AAV) is the only known eukaryotic virus capable of targeted integration in human cells. An AAV Rep binding element (RBE) and terminal resolution site (*trs*) identical to the viral terminal repeats required for AAV DNA replication are located on chromosome (ch) 19. Both ch-19 RBE and *trs* elements have been shown to be essential for viral targeting to this locus. To characterize the role of the AAV inverted terminal repeat (ITR) cis-acting sequences in targeted integration an AAV *trs* mutant incapable of supporting viral replication was tested. Wild-type and mutant substrates were assayed for targeted integration after cotransfection in the presence or absence of Rep. Our results demonstrated that, in the presence of Rep78, both ITR substrates targeted to ch-19 with similar frequency. Molecular characterization of the mutant ITR integrants confirmed the presence of the *trs* mutation in the majority of samples tested. Complementation analysis confirmed that the mutant targeted viral genomes were unable to rescue and replicate. In addition, Rep78 induced extensive rearrangement and amplification of ch-19 sequences independent of wild-type or mutant targeting substrate. These studies demonstrate that Rep-dependent nicking of the viral cis-acting *trs* sequence is not a prerequisite for site-specific recombination and suggests AAV targeting is mediated by Rep78/68-dependent replication from the ch-19 origin of replication (*ori*). These studies have significant impact toward the understanding of AAV site-specific recombination and the development of targeting vectors.

Adeno-associated virus (AAV) has a biphasic life cycle. In the presence of helper virus, Adenovirus (Ad), or Herpesvirus (HSV), AAV undergoes a productive infection. Lacking a helper virus, AAV latently infects by integration into the host genome. AAV is the only known eukaryotic virus capable of undergoing site-specific integration into the human genome (1). The ability to integrate site specifically is one of the attractive features for considering this virus as a vector for human gene therapy (2).

Wild-type (wt) viral DNA recombines at chromosome (ch)-19.13.3 qter at a frequency of >70% (3), a reaction that depends on the AAV replication (Rep) proteins (4, 5) Rep78 and Rep68 (6–8). Mutational analysis of Rep78 and Rep68 proteins indicates that DNA binding, site-specific endonuclease activity, and helicase activities are critical for site-specific recombination (1, 6, 9, 10).

The only viral *cis* elements necessary for targeted integration are the 145-bp inverted terminal repeats (ITRs). The ITRs also serve as the origins of replication (*ori*) for viral DNA synthesis and are the essential cis components required for generating AAV integrating vectors. In the absence of Rep, recombinant AAV (rAAV) integrates randomly. Although the mechanism of viral integration is unknown, wtAAV and rAAV proviral structures are similar because they are typically arranged in head to tail concatamers, have micro sequence homology at junctions, and possess variable ITR and chromosomal deletions (3, 11–14). These observations led to the conclusion that both Rep-dependent site-specific integration and random integration depend on the cellular recombination machinery (14). The ch-19 integration site, unique to primates (15), contains a Rep-responsive *ori* both *in vitro* (16) and *in vivo* (17). In the presence of Rep the ch-19 region is amplified and rearranged independently of site-specific recombination (17, 18). The ch-19 cis sequences that direct Rep dependent replication share 86% sequence identity to the AAV ITR (19–21). The ch-19 region contains a Rep binding element (RBE) and a terminal resolution site (*trs*) also found within the viral ITR (9, 19). Thirty-three base pairs of ch-19 containing the RBE and *trs* is sufficient for site-specific recombination (9, 20). Deletion of the *trs* resulted in a loss of viral targeting (9, 20). Episomal integration depends on both the sequence and position of the spacer DNA separating the RBE and *trs* motifs (21). Furthermore, in the presence of Rep, these ch-19 cis elements are sufficient to direct site-specific recombination in rodent models independent of chromosomal context (17, 22). Although these data support the role of critical ch-19 cis elements in targeted integration, a similar analysis of the required viral elements has not been performed.

The goal of this study is to determine whether AAV site-specific recombination is coupled to Rep-dependent replication by means of the AAV ITR *ori*. To test this, a modified viral ITR that contains an 8-bp insertion between the AAV RBE and *trs* (Fig. 1*A*) was used. Rep-dependent resolution in an *in vitro* nicking assay was over 100-fold less efficient when using this substrate compared with wt sequences (23). Consistent with the *in vitro* data, these modified ITRs were also deficient for AAV viral production *in vivo* (X. Xiao and R.J.S., unpublished data). Using transfection assays, the mutant ITR was tested for targeted integration in the presence of AAV Rep78. Here we show that mutant viral ITRs containing a nonfunctional *trs* exhibited the same targeting frequency as those containing a wt ITR. Furthermore, the ch-19 region was amplified and rearranged regardless of the targeting substrate used. These data support the conclusion that Rep-dependent replication of ch-19 and not the viral ITR is required for site-specific recombination.

Materials and Methods

Maintenance of Cells. HeLa cells were grown in DMEM supplemented with 10% heat-inactivated FBS. Cells were grown at 37°C in a 5% CO_2 humidified incubator.

Abbreviations: AAV, adeno-associated virus; ITR, inverted terminal repeat; wt, wild type; RBE, Rep binding element; ch, chromosome; *trs*, terminal resolution site.

‡Present address: The Salk Institute, Molecular Neurobiology Laboratories, La Jolla, CA 92037.

¶To whom reprint requests should be addressed at: 7119 Thurston Bowles CB 7352, University of North Carolina, Chapel Hill, NC 27599. E-mail: rjs@med.unc.edu.

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

Plasmids and DNA. pwtITR-neo plasmid was constructed by digesting psub201 (24) with *Sna*BI and ligating it to the 1-kb *Acc*I-*Xho*I fragment of MC Neo poly(A) (Stratagene). The *Xho*I-*Acc*I fragment was blunted and ligated to psub201 generating the following cassette, the TK promoter, neo resistance gene, and a poly(A) signal sequence between the AAV ITRs (1.3 kb). pHpaITR-neo was constructed identically to pwtITR-neo except that the pHpa8 plasmid (23) was used as the vector backbone. The plasmids were maintained in SURE cells (Stratagene).

Single Cell Cloning. To obtain clonal cell lines, cells were transfected with neo-resistant plasmids plus and minus Rep constructs (28). After 48 h, cells were trypsinized and plated at dilutions of 1:10, 1:50, and 1:100 under G418 selection (600 μg/ml). After 8 days of selection, colonies were isolated and propagated as described (15). All stable clones were maintained in 300 μg/ml G418.

AAV Replication Assay. Replication assays were carried out by transfecting HeLa cells with pwt/ITR-neo or pHpa/ITR-neo, in addition to XX6–80 (25) and Ad8 as described (25). Forty-eight hours after transfection, Hirt extracts were performed (26) and viral replication was characterized by sensitivity to Dpn I by using Southern blotting analysis.

Southern Blot Analysis. Southern blotting analysis was carried out as described (15). Briefly, ch-19 and neo sequence-specific probes were generated by a Random Primer Labeling kit (Boehringer). DNA was fractionated on agarose gels, transferred to filters, hybridized by using Quick-Hyb (Stratagene) according to the manufacturer's protocol, and finally exposed to x-ray film for data analysis.

PCR Assay for Targeted Integration. Genomic DNA from clonal cell lines transfected with the Rep expressing plasmid (pHIV78) and pwtITRneo or pHpaITR-neo constructs was PCR amplified following the protocol described by Samulski *et al.* (15), using a modification described by Urabe *et al.* (10). Ten microliters of the reaction was fractionated on a 1.2% agarose gel. Gels were subjected to Southern analysis by using a ch-19 or neo-specific probe. PCR analysis was carried out with single ch-19 primers to rule out any potential false positives. Controls included single primer reactions of ch-19 and Neo vector, respectively. Only positive PCR products obtained from reactions containing both primers [ch-19 primer 2649 and AAV A stem primer (15)] were identified as ch-19 targeted integration.

Results

***Hpa*I ITR Constructs Are Defective for Rep-Dependent Replication.** Three components are currently known to be required for AAV targeted integration: Rep proteins in trans, cis-acting AAV ITR and ch-19-specific sequences. Of the ch-19 cis-acting sequences, the RBE, *trs*, and the spacer element between these motifs are essential for site-specific recombination (20, 21). An understanding of the specific AAV cis-acting ITR sequences required, specifically the role of the *trs*, has not been established. Moving the viral *trs* 8 bp downstream of its normal position (Fig. 1*A*) results in a 100-fold reduction in Rep-mediated nicking in an *in vitro* assay (23). By using Hirt analysis (27) to assay the ITR mutant for AAV replication *in vivo*, these *in vitro* observations were extended. After transfection into permissive HeLa cells, replicated AAV DNA was assessed by sensitivity to DpnI digestion. In the absence of Dpn I treatment, both wt and mutant input plasmids were detected (Fig. 1*B*, lanes 1 and 3). However, in the presence of Dpn I, only the wt ITR plasmid substrate yielded characteristic AAV monomer and dimer replication intermediates (Fig. 1*B*, lane 2). These results corroborate our previous *in vitro* studies and confirm the mutant ITR as a replication negative substrate *in vivo* despite the presence of Rep (Fig. 1*B*, lane 4).

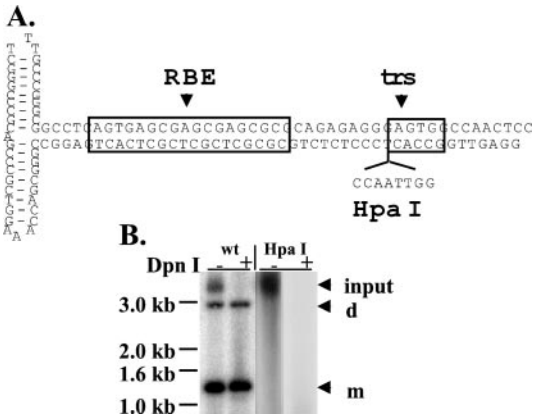

Fig. 1. AAV replication assay of wt and mutant ITR neo vectors. AAV ITR in hairpin configuration. RBE and *trs* are boxed and labeled. The 8-bp *Hpa*I linker insertion CCAATTGG and position of insertion is illustrated. The details for construction have been described (23). (*B*) Southern blot analysis of wt and pHpaITR vector substrate for replication. Hirt analysis of low molecular weight DNA after cotransfection with Ad helper plasmid XX-6–80, pAB11, and either the pwtITR-neo (wt) or pHpa-ITR neo (Hpa) construct was carried out as described in *Materials and Methods*. The positions of input AAV plasmids (lanes 1 and 3) and Dpn I-resistant (lanes 2 and 4) replicative forms are shown. The 1.3- and 2.6-kb Dpn I-resistant replicative forms are denoted by m (monomer) and d (dimer) and molecular weight markers are shown.

Role of Viral *trs* in ch-19 Site-Specific Recombination. Experiments have shown that AAV ITR-containing plasmids target ch-19 when either Rep78 or Rep68 is supplied *in trans* (6–8). To determine the requirement of a functional AAV ITR *ori* in site-specific recombination, neo-resistant plasmids containing the altered AAV replication-defective ITR (pHpaITR-neo) or wt ITR (pwtITR-neo) were transfected into HeLa cells with and without Rep-expressing constructs and single-cell clones were selected in G418 (6, 28). The frequency of neo-resistant colonies reflects the integration frequency. Both wt and mutant plasmids showed Rep-dependent increases in the integration frequencies, with 4.5- and 3-fold increases, respectively (Table 1). In the presence of Rep, only a 1.8-fold increase in the number of resistant colonies with the wt ITR compared with the mutant was seen (Table 1). In contrast to the dramatic inhibition of AAV replication by the mutant ITR, its effect on integration frequency appeared minimal.

To determine whether mutant ITR genomes inserted site-specifically into ch-19 sequences, random neo-resistant clones were selected and genomic DNA was analyzed by Southern blotting. DNA was digested with *Apa*I, an enzyme that does not cut the AAV ITR plasmids but releases a 2.8-kb ch-19 genomic fragment that contains the majority of AAV site-specific recombinants (7, 14, 15, 29). When using a ch-19 sequence specific

Table 1. Wild-type and mutant ITR neo-resistant clones

Neo clones	No.
pwt/ITR-neo+Rep78	32
pHPA/ITR-neo+Rep78	19
pwt/ITR-neo/No Rep78	7
PHpa/ITR-neo/No Rep78	6

(*A*) pwtITR-neo and pHpaITR-neo constructs were transfected into HeLa cells with (+Rep78) and with out (no Rep78) Rep78 expression plasmids. Experiments were done in triplicate and random colonies were selected and expanded in G418 for genomic analysis.

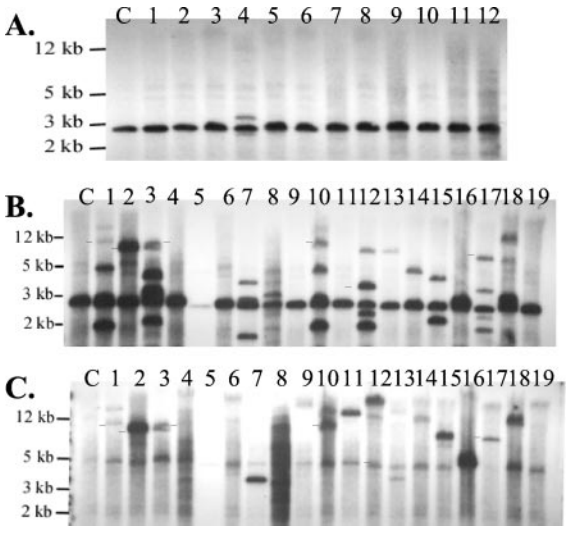

Fig. 2. Southern analysis of wt and pHpaITR neo-resistant genomic DNA. Creation of neo-resistant clonal cell lines and analysis was carried out as described in *Materials and Methods*. *Apa*I-digested genomic DNA releases a 2.8-kb fragment of the ch-19 preintegration region, but does not cut the vectors. *Apa*I digestion of HeLa cell genomic DNA before transfection assays was used as control (lane C) in all experiments (*A*–*C*). Molecular weight markers are indicated on the left of the gel. (*A*) *Apa*I-digested genomic DNA from wt (lanes 1–6) and mutant ITR (lanes 7–12) neo-resistant clones minus Rep78. DNAs were analyzed by using a ch-19-specific probe (see *Materials and Methods* for details). Presence of a 2.8-kb genomic signal is indicative of lack of site-specific integration (lanes C and 1–12). (*B*) *Apa*I-digested genomic DNA from randomly selected wt (lanes 1–9) and mutant ITR vector (lanes 10–19) neo-resistant clones cotransfected with Rep78-expressing plasmid pHIV78 (see *Materials and Methods* for details). Amplification and rearrangement of ch-19-specific region is indicated by multiple bands (lanes 1–19) compared with control (C). (*C*). Analysis of *B* by using a neo-specific probe after stripping of ch-19-specific sequences. Absence of the major 2.8-kb ch-19 signal and lack of neo-specific signal in the control lane (C) indicate specificity of the neo probe for vector sequences. Site-specific recombination was assigned to clones that contained bands that cohybridized both to the neo probe (arrowheads in *C*) and the ch-19 probe (arrowheads in *B*) as described (42).

probe, in the absence of Rep, only the expected 2.8-kb genomic band was observed (Fig. 2*A*). When either wt or mutant ITR plasmids were introduced in the presence of Rep, we detected amplification and rearrangement of the ch-19 target sequence in 84% pwtITR-neo clones and 69% pHpaITR-neo colonies (Table 2, Fig. 2*B*). Hybridization with a neo-sequence-specific probe (Fig. 2*C*) showed that 32% pwtITR-neo clones were targeted to ch-19 based on the comigration of ch-19 and neo-specific sequences (Table 2) and is in agreement with previous reports,

Table 2. Analysis of integration status of wtITR of HpaITR clones

Clones	#R	%R	#T	%T
pwt/ITR-neo+ Rep78	16/19	84	6/19	32
pHpa/ITR-neo +Rep78	13/19	69	9/19	47
Pwt/ITR-neo/−Rep78	0/6	0	0/6	0
PHpa/ITR-neo/−Rep78	0/6	0	0/6	0

Genomic DNA from 19 wild-type or 19 mutant ITR neo-resistant clones cotransfected with pHIV Rep78-expressing plasmid (+Rep78) were scored for ch-19 rearrangement after ApaI digestion (see Fig. 2 for illustration). The number rearranged (#R) and percent rearranged (%R) ch-19 sequences are indicated. Cohybridization to both ch-19 and vector-specific probe (as illustrated in Fig. 2 *B* and *C*) was used to determine the number targeted (#T) and percent targeted (%T). Similar analysis of wt and mutant ITR constructs in the absence of Rep expression (pwtITR-neo/No Rep78, pHpaITR-neo/No Rep78) served as controls.

Table 3. Analysis of pHpaITR genomic clones

Clone	Southern	PCR	No. of HpaI sites	Rescue
H#1	+	−	2	ND
H#2	+	+	1	+/−
H#3	+	+	2	−
H#4	−	−	ND	ND
H#5	−	−	ND	−
H#6	−	−	ND	−
H#7	+	+	0	−
H#8	+	+	1	−
H#9	−	−	ND	−
H#10	+	+	ND	ND
H#11	−	−	ND	ND
H#12	+	+	2	ND
H#13	−	−	ND	ND
H#14	−	−	ND	ND
H#15	−	+	2	ND
H#16	−	+	2	−
H#17	+	+	2	−
H#18	+	+	2	−
H#19	−	−	ND	ND

Detailed characterization of 19 pHpaITR genomic isolates for ch-19 targeting. Results from Southern (Southern blot analysis), PCR (PCR analysis), presence of mutated ITR after integration (# of HpaI Sites), and ability to re-enter the AAV lytic cycle (Rescue) after superinfection with helper viruses are described. Results are listed as clone number (Clone #, 1–19); +, positive; −, negative; ND, not done; and 0, 1, or 2, the number of HpaI sites present after Southern analysis using HpaI digestion.

using AAV plasmids for ch-19 targeting (6–8). In the case of pHpaITR-neo, 47% were targeted (Fig. 2*C*, Table 2). In the absence of Rep78, no ch-19 rearrangement or targeting with either ITR vector substrate was observed (Fig. 2*A*, Table 2).

Expression of Rep in the absence of AAV vector is sufficient to rearrange and amplify the ch-19 sequence (17). To rule out the possibility that unlinked neo-positive and rearranged ch-19 fragments comigrated in these experiments, genomic DNA was analyzed for the linkage of vector and ch-19 sequences by PCR. The primer pair consisted of a ch-19-specific and an AAV-vector-specific primer (10, 14, 15). PCR products were analyzed by Southern blotting, using both ch-19-specific and vector-specific probes (data not shown). The results confirmed those obtained by genomic Southern analysis; in the absence of Rep no clones (0/12) from either pwtITR-neo or pHpaITR-neo colonies scored positive for vector/ch-19 junctions (Tables 2 and 3). However, 53% pHpaITR-neo clones were positive for targeted integration in the presence of Rep78 (Table 3). Two clones (H#15 and H#16) were positive for site-specific integration by PCR, but negative by genomic Southern. One clone (H#1) was positive by Southern but negative when using this set of PCR primers (Table 3). In total, 11 pHpaITR-neo + Rep78 clones could be classified as targeted (58%), with 8 of 11 clones (42%) positive by both assays. For comparison, 32% of clones established by the wt ITR substrate were identified as targeted. Based on these observations, the mutant ITR showed no defect in AAV site-specific integration, indicating that a functional AAV ITR *ori* as defined by *in vitro* nicking and *in vivo* replication is not required for this process.

Mutant ITRs Remain Intact After Integration. The AAV ITR is highly recombinogenic and will gene convert very efficiently (27, 30). If efficient gene conversion of mutant to wt ITR sequences occurred before integration, it would appear as if mutant AAV ITR vectors could undergo efficient targeting. To ensure that ch-19 targeted pHpaITR-neo clones had not gene converted to a functional AAV ITR *ori* before integration, the status of integrated vector was determined by Southern analysis.

pHpaITR-neo clonal genomic DNA was digested with *Hpa*I and analyzed by using a neo-specific probe. The appearance of a 1.3-kb fragment is indicative of an integrated mutant vector containing a *Hpa*I site within each ITR. The majority of the clones (67%) retained both *Hpa*I sites. Two clones (22%) lost one *Hpa*I site, and 1 clone (11%) lost both *Hpa*I sites (Table 3). This analysis dismissed the possibility that pHpaITR-neo had gene converted to a functional AAV ITR *ori* before integration.

AAV Rescue Assay. A hallmark of the AAV life cycle is the ability of the proviral genome to rescue from the latent state and reenter the lytic life cycle when super infected with helper virus. Rescue occurs ≈50–60% of the time in latently infected cell lines (3). This aspect of AAV integration takes place with wtAAV or rAAV (14). Because the mutant ITR substrates were defective for AAV replication (Fig. 1*B*), but competent for efficient targeted integration (Fig. 2, Table 2), we assayed stable clones for rescue after infection with helper virus. Unlike wt ITR cell lines, all clones tested that retained two mutant ITRs were defective for rescue (Table 3). The inability to rescue is consistent with the mutant ITR as a nonfunctional template for replication. These data suggest that a functional ITR, as defined by AAV lytic replication, is not required for targeted integration. However, one clone that contained one defective ITR demonstrated a weak signal for rescue product. Differential hybridization with AAV and ch-19 probes demonstrated that the majority of the signal was for ch-19 sequences (data not shown). This rescue molecule appears to be an interesting variant of AAV linked to ch-19 sequences. Further studies are ongoing to determine the molecular configuration of this chimeric substrate and whether the ch-19 *ori* serves as the template for rescue.

Discussion

In this study, the role of an AAV mutant ITR for the ability to target ch-19 was evaluated. Although this mutant is defective for AAV Rep-dependent replication (Fig. 1), our results demonstrate that the modified terminal repeat sequence is completely competent for ch-19 site-specific integration (Fig. 2, Tables 1–3). Both the efficiency and frequency of targeting was equal to or better than wt substrates (Table 2). These studies are of significant importance because they begin to define the role of the viral cis sequences required for site-specific integration.

Currently, AAV targeted integration requires three known components: Rep proteins in trans, cis-acting ch-19-specific sequences, and the AAV ITR elements. Of the ch-19 cis-acting sequences, the RBE and *trs* element are essential for site-specific recombination (20). Recently, a third component of ch-19, the spacer element between these motifs, was shown to be critical for site-specific recombination and by increasing the distance between the ch-19 RBE and *trs* completely abolished targeted integration (21). Similar to these studies, an 8-bp increase in the spacer region between the AAV ITR RBE and *trs* resulted in a 100-fold decrease in Rep-dependent nicking *in vitro* (23) and complete absence of viral replication *in vivo* (Fig. 1*B*). Contrary to the loss of integration specificity observed with modified ch-19 spacer sequence substrates, the modified AAV ITR targeted ch-19 at a similar frequency to the wtITR (Fig. 2, Table 2). Also, the mutant ITR was retained after integration, ruling out the possibility of gene conversion to a wt ITR before targeting (Table 3). All clones assayed were negative for rescue when challenged with helper virus, with one exception (Table 3). These data suggests that Rep-dependent nicking of the AAV ITR as defined by *in vitro* assays and *in vivo* AAV replication assays is not required for site-specific integration.

The Role of Rep-Dependent Replication in Targeting. Loss of the *trs* in the ch-19 region results in a loss of targeting capability (9, 20) and suggests that the ability of Rep78 and Rep68 to initiate replication on ch-19 is essential to the integration reaction. This hypothesis is supported by Rep68's ability to initiate replication on ch-19 *in vitro* (16), and the fact that the majority of breakpoints identified *in vivo* are located downstream of the ch-19 *trs* (15). Furthermore, we demonstrated *in vivo* amplification and rearrangement of ch-19 in the presence of Rep78 when no targeting vector substrates were present (17).

What Is Unique About the ch-19 Site? In addition to carrying the cis-acting RBE and *trs*, the ch-19 target locus contains DNA elements that promote an open chromatin conformation (31). This DNA conformation may be essential for initiating Rep interaction with the ch-19 target locus. In support of this concept, the ch-19 cis elements have been able to direct site-specific recombination to non-ch-19 regions in animal models (17, 22). These models share 1.6 kb of ch-19 in common, additional experiments are required to determine whether the minimal cis elements (33 bp) identified in the EBV episomal integration assay system are sufficient for *in vivo* targeting. At present, these 33 bp do not include the DNase hypersensitive site.

Initial Steps in AAV Targeting. All of these data and numerous biochemical studies *in vitro* support a model for site-specific integration (Fig. 3) where Rep initiates replication on ch-19 (Fig. 3*A*) by interaction with the ch-19 RBE and *trs*. Most likely asymmetric replication on the ch-19 *ori* (16) begins by Rep unwinding the ch-19 RBE creating a replication bubble followed by site- and strand-specific nick at the ch-19 *trs*. Rep78/68 becomes covalently attached to the 5′ end of the cut site by means of tyrosine phosphate linkage (Fig. 3*B*; ref. 32) and generates a free 3′-OH substrate that can be used to initiate replication (33, 34). As replication proceeds, the 5′ strand gets displaced and generates a flap (Fig. 3*B*; ref. 1). The 5′ flaps that are generated during cellular replication are typically removed by the FEN-1 protein (35). However, FEN-1 protein only acts on single-stranded DNA substrates, and are blocked by flaps with double-stranded ends (36) or single-stranded flaps protected by 5′ protein (37, 38). In the absence of RAD27, yeast homologue of FEN-1, there is an increased rate of recombination, duplications, and double-stranded breaks in the host genome (39). In addition, trinucleotide repeat disease appears to be related to inability of FEN-1 enzyme to resolve flaps containing CAG repeats because of hairpin conformation (36), and results in large expansion of these sequences. It is possible that unscheduled DNA replication mediated at ch-19 by Rep would generate substrates (protein-linked 5′ flaps) that become recombination "hot spots" (Fig. 3*B*). Because Rep78/68 will be covalently attached to the displaced strand, it is likely that it will inhibit the ability of FEN-1 to process the displaced strand. The inability to process this displaced strand would lead to amplification and rearrangement of the chromosomal region (38). We recently documented amplification and rearrangement of ch-19 in the presence of Rep alone (17). These data suggest that Rep-mediated replication of ch-19 is sufficient to generate all chromosomal recombination products observed after site-specific integration and the extent of amplification and rearrangement of the ch-19 integration region may be subject to repeated Rep initiation on the ch-19 *ori* (Fig. 3*C*). Evidence supporting this idea comes from experiments using a hormone-regulated Rep68 construct that expressed Rep68 transiently at lower levels resulted in a decrease in the amount of amplification and rearrangement of the ch-19 region (18). We previously demonstrated that Rep78 and Rep68 are present at 1,000–4,000 copies per latently infected HeLa cell although this may vary from cell type to cell type. This amount of Rep may increase the rate of initiation and lead to an accumulation of Rep-covalently linked 5′ flaps (Fig. 3*C*) and ultimately lead to an increase of rear-

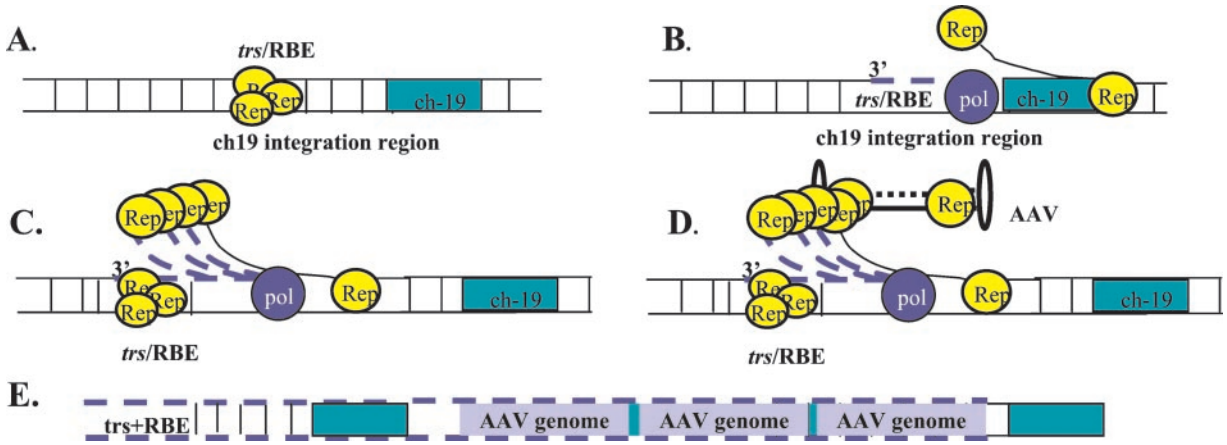

Fig. 3. AAV integration model. (*A*) Complex of Rep proteins (yellow circles labeled Rep) initiate unscheduled replication on ch-19 by interaction with the ch-19 RBE and *trs*. (*B*) Asymmetric replication on the ch-19 origin ensues by Rep unwinding the ch-19 RBE creating a replication bubble followed by site- and strand-specific nick at the ch-19 *trs*. Rep78/68 becomes covalently attached to the 5′ end of the cut site by means of tyrosine phosphate linkage and generates a free 3′-OH substrate that can be used to initiate replication (dashed blue line). As replication machinery proceeds (blue circle labeled pol), the 5′ strand gets displaced and generates a flap (Rep yellow circle attached to solid line). The DNA replication mediated at ch-19 by Rep would generate substrates (protein-linked 5′ flaps) that become recombination ''hot spots'' (see *Discussion*). (*C*) The extent of amplification and rearrangement of the ch-19 integration region may also be subject to repeated Rep reinitiation on the ch-19 origin. This would result in multiple 5′ flaps with covalently attached Rep proteins (yellow circles attached to dashed lines) that influence the extent of amplification and rearrangement seen with site-specific recombination. (*D*) Complex formation between actively replicating ch-19 sequences and Rep protein bound to AAV ITR sequences (hairpin structure). The dashed line representing the top strand of the AAV genome indicates that either single- or double-stranded templates may be substrates for Rep complex formation carrying duplex ITRs. (*E*) Resolution of the viral genome into the ch-19 locus is carried out by host enzymes that result in amplified (large green boxes) and rearranged (small green boxes) ch-19 sequences as well as the characteristic AAV head to tail (light blue boxes labeled AAV genome) concatamers.

rangement and amplification through the cell's inability to process these replication by-products. It is of interest to note that a minute virus of mice (MVM)-based targeted recombination system reported that the amplification and rearrangement of the target DNA region was similar to ch-19 region (40). Both reactions are dependent on their respective large nonstructural proteins and are similar in function and enzymatic activity (1). The similarities of the enzymatic activities of the large nonstructural proteins points to a common mechanism of amplification and rearrangement that may be a result of their respective enzymatic nature (i.e., covalent attachment after site-specific endonuclease activity; Fig. 3*B*). Whether reinitiation is required or not for more efficient targeting, our data supports site-specific integration at ch-19 in the absence of classical Rep-mediated nicking of the viral ITR substrate. Our data supports a concept first proposed by Owens and colleagues (19) where Rep facilitates site-specific integration by mediating complex formation between the viral ITR and the ch-19 locus (Fig. 3*D*). Because targeted integration depends on the *trs* being located in the correct proximity to the RBE on ch-19 but not on the viral ITR implies that the incoming substrate may only require Rep-binding activity. Further studies *in vivo* with additional ITR mutant substrates should help define the minimal requirements of the terminal repeat sequences for AAV site-specific integration.

ch-19 Amplification and Rearrangement. The enzymatic steps required after complex formation to establish stable ch-19 integrants have been previously alluded to (Fig. 3*E*; refs. 9, 17, 20, and 41). Proviral structures for AAV vectors devoid of Rep result in identical head-to-tail concatamers and region-specific deletions in the terminal repeats as well as microhomology at the viral/cellular junctions are also observed, albeit at random sites in the genome (12, 14, 42). The mutant ITR proviral genomes in this study were predominantly head-to-tail concatamers that also contained the plasmid backbone (data not shown; ref. 43). The formation of head-to-tail concatamers may be a general by-product of eukaryotic cells and not specifically related to AAV substrates. For example, episomal forms of AAV vectors readily form head-to-tail concatamers in the absence of Rep, a feature now being exploited to overcome AAV packaging limitations (44, 45). Head-to-tail concatamers is a distinguishing feature of all transgenes when generating transgenic animals and configuration of other DNA virus (SV40, Ad, etc.) proviral structures (46, 47), and was first described by Schimke and colleagues after transfection of naked DNA into eukaryotic cells (48–50). From this perspective, AAV carries many hallmarks of illegitimate recombination; deletions and duplications at the insertion sites, head-to-tail concatamers, and amplification and rearrangement of the integration locus. These proviral structural similarities between wt, mutant ITRs, and AAV vectors suggest a similar mechanism of concatamer formation and DNA amplification carried out by cellular factors (Fig. 3*E*). These observations help to define Rep's role in site-specific recombination, which appears to be to initiate unscheduled ch-19 replication and facilitate the interaction of viral and ch-19 sequences by Rep–Rep protein interaction.

Implications for Gene Therapy and Targeting Vectors. A key feature of the wt AAV biphasic life cycle and an aspect associated with current AAV vectors is the ability of the provirus to reenter the lytic pathway when all necessary helper functions are present. Our data showed that the modified ITR was capable of site-specific integration of plasmid substrates but deficient for rescue and replication when challenged with permissive conditions. The inability to rescue would eliminate the risk of vector dissemination after incidental superinfection with wt AAV and helper virus, a feature now associated with current AAV vectors. AAV plasmid vectors carrying the modified ITRs could exploit this aspect of targeted integration without the risk of rescue. Because a functional AAV ITR *ori* is not required for ch-19 targeting, we view the mutant ITR substrate as a second-generation AAV plasmid vector that has site-specific recombination-only activity. Characterization of additional mutations in the AAV ITR may result in more efficient targeting vectors (i.e., expanded RBE elements, modified hairpins, etc.). In addition, the mutant ITRs

may remove other unwanted features of the viral ITR. For example, the ITR has low-level basal transcription activity (51) that has impeded the development of highly regulated AAV expression vectors (52, 53). Based on the results of this study and the recent characterization of these transcriptional elements (51), it may be possible to generate variants vector substrates containing additional mutations that are highly efficient for integration (53). At minimum, the relevant cis elements (RBE, ITR hairpin, etc.) required for efficient targeting should be forthcoming from such studies. The mutant ITR integrating plasmid vector described in this study may be important in complementing efficient plasmid delivery systems currently compromised by transient transduction (54). Finally, *in vitro* integration systems aimed at defining the molecular steps required for AAV site-specific recombination are being developed (14, 55). The data established in this study point to important viral ITR cis-acting sequences required for targeting (i.e., RBE). Further mutational analysis of the viral ITR aimed at defining the role of the RBE element as described above should greatly facilitate development of AAV site-specific *in vitro* integration systems.

We thank Dr. Terry Van Dyke for technical reading of this manuscript, Xing-Hua Zeng for expert technical assistance, and Nathan Allen for assistance with the generation of Fig. 1. This work was supported by National Institutes of Health Grants DK51880 and HL 48347.

1. Berns, K. I. (1996) in *Fields Virology*, ed. Fields, B. N. (Raven, Philadelphia), Vol. 2, pp. 2173–2197.
2. Samulski, R. J., Sally, M. & Nuzyczka, N. (1999) in *The Development of Human Gene Therapy* (Cold Spring Harbor Lab. Press, Plainview, NY), pp. 131–172.
3. Muzyczka, N. (1992) *Curr. Top. Microbiol. Immunol.* **158,** 97–129.
4. Samulski, R. J. (1993) *Curr. Opin. Genet. Dev.* **3,** 74–80.
5. Walsh, C. E., Liu, J. M., Xiao, X., Young, N. S., Nienhuis, A. W. & Samulski, R. J. (1992) *Proc. Natl. Acad. Sci. USA* **89,** 7257–7261.
6. Xiao, W. (1996) in *Curriculum in Genetics and Molecular Biology* (Univ. of North Carolina, Chapel Hill), p. 179.
7. Surosky, R. T., Urabe, M., Godwin, S. G., McQuiston, S. A., Kurtzman, G. J., Ozawa, K. & Natsoulis, G. (1997) *J. Virol.* **71,** 7951–7959.
8. Lamartina, S., Roscilli, G., Rinaudo, D., Delmastro, P. & Toniatti, C. (1998) *J. Virol.* **72,** 7653–7658.
9. Linden, R. M., Ward, P., Giraud, C., Winocour, E. & Berns, K. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 11288–11294.
10. Urabe, M., Hasumi, Y., Kume, A., Surosky, R. T., Kurtzman, G. J., Tobita, K. & Ozawa, K. (1999) *J. Virol.* **73,** 2682–2693.
11. Cheung, A. K., Hoggan, M. D., Hauswirth, W. W. & Berns, K. I. (1980) *J. Virol.* **33,** 739–748.
12. McLaughlin, S. K., Collis, P., Hermonat, P. L. & Muzyczka, N. (1988) *J. Virol.* **62,** 1963–1973.
13. Rutledge, R. A. & Russell, D. W. (1997) *J. Virol.* **71,** 8429–8436.
14. Yang, C. C., Xiao, X., Zhu, X., Ansardi, D. C., Epstein, N. D., Frey, M. R., Matera, A. G. & Samulski, R. J. (1997) *J. Virol.* **71,** 9231–9247.
15. Samulski, R. J., Zhu, X., Xiao, X., Brook, J. D., Housman, D. E., Epstein, N. & Hunter, L. A. (1991) *EMBO J.* **10,** 3941–3950.
16. Urcelay, E., Ward, P., Wiener, S. M., Safer, B. & Kotin, R. M. (1995) *J. Virol.* **69,** 2038–2046.
17. Young, S. M. J., McCarty, D. M., Degtyareva, N. & Samulski, R. J. (2000) *J. Virol.* **74,** 3953–3966.
18. Rinaudo, R., Lamartina, S., Roscilli, G., Ciliberto, G. & Toniatti, C. (2000) *J. Virol.* **74,** 281–294.
19. Weitzman, M. D., Kyostio, S. R. M., Kotin, R. M. & Owens, R. A. (1994) *Proc. Natl. Acad. Sci. USA* **91,** 5808–5812.
20. Linden, R. M., Winocour, E. & Berns, K. I. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 7966–7972.
21. Meneses, P., Berns, K. I. & Winocour, E. (2000) *J. Virol.* **74,** 6213–6216.
22. Rizzuto, G., Gorgoni, B., Cappelletti, M., Lazzaro, D., Gloaguen, I., Poli, V., Sgura, A., Cimini, D., Ciliberto, G., Cortese, R., *et al.* (1999) *J. Virol.* **73,** 2517–2526.
23. Snyder, R. O., Im, D. S., Ni, T., Xiao, X., Samulski, R. J. & Muzyczka, N. (1993) *J. Virol.* **67,** 6096–6104.
24. Samulski, R. J., Chang, L.-S. & Shenk, T. (1987) *J. Virol.* **61,** 3096–3101.
25. Li, J., Samulski, R. J. & Xiao, X. (1997) *J. Virol.* **71,** 5236–5243.
26. Hirt, B. (1967) *J. Mol. Biol.* **26,** 365–369.
27. Samulski, R. J., Srivastava, A., Berns, K. I. & Muzyczka, N. (1983) *Cell* **33,** 135–143.
28. Gavin, D. K., Young, S. M., Xiao, W., Temple, B., Abernathy, C. R., Pereira, D. J., Muzyczka, N. & Samulski, R. J. (1999) *J. Virol.* **73,** 9433–9445.
29. Palombo, F., Monciotti, A., Recchia, A., Cortese, R., Ciliberto, G. & La Monica, N. (1998) *J. Virol.* **72,** 5025–5034.
30. Xiao, X. (1992) in *Microbiology* (Univ. of Pittsburgh), p. 203.
31. Lamartina, S., Sporeno, E., Fattori, E. & Toniatti, C. (2000) *J. Virol.* **74,** 7671–7677.
32. Snyder, R. O., Im, D.-S. & Muzyczka, N. (1990) *J. Virol.* **64,** 6204–6213.
33. Ni, T. H., Zhou, X., McCarty, D. M., Zolotukhin, I. & Muzyczka, N. (1994) *J. Virol.* **68,** 1128–1138.
34. Ni, T. H., McDonald, W. F., Zolotukhin, I., Melendy, T., Waga, S., Stillman, B. & Muzyczka, N. (1998) *J. Virol.* **72,** 2777–2787.
35. Gordenin, D. A., Kunkel, T. A. & Resnick, M. A. (1997) *Nat. Genet.* **16,** 116–118.
36. Leeflang, E. P., Tavare, S., Marjoram, P., Neal, C. O., Srinidhi, J., MacFarlane, H., MacDonald, M. E., Gusella, J. F., de Young, M., Wexler, N. S. & Arnheim, N. (1999) *Hum. Mol. Genet.* **8,** 173–183.
37. Waga, S. & Stillman, B. (1998) *Annu. Rev. Biochem.* **67,** 721–751.
38. Gary, R., Park, M. S., Nolan, J. P., Cornelius, H. L., Kozyreva, O. G., Tran, H. T., Lobachev, K. S., Resnick, M. A. & Gordenin, D. A. (1999) *Mol. Cell. Biol.* **19,** 5373–5382.
39. Tishkoff, D. X., Filosi, N., Gaida, G. M. & Kolodner, R. D. (1997) *Cell* **88,** 253–263.
40. Corsini, J., Tal, J. & Winocour, E. (1997) *J. Virol.* **71,** 9008–9015.
41. Chiorini, J. A., Wiener, S. M., L., Y., Smith, R. H., Safer, B., Kilcoin, N. P., Liu, Y., Urcelay, E. & Kotin, R. M. (1996) *Curr. Top. Microbiol. Immunol.* **218,** 25–33.
42. Kotin, R. M., Siniscalco, M., Samulski, R. J., Zhu, X. D., Hunter, L., Laughlin, C. A., McLaughlin, S., Muzyczka, N., Rocchi, M. & Berns, K. I. (1990) *Proc. Natl. Acad. Sci. USA* **87,** 2211–2215.
43. Young, S. (2000) in *Curriculum in Genetics and Molecular Biology* (Univ. of North Carolina, Chapel Hill), p. 251.
44. Samulski, R. J. (2000) *Nat. Biotechnol.* **18,** 497–498.
45. Sun, L., Li, J. & Xiao, X. (2000) *Nat. Med.* **6,** 599–602.
46. Syu, L. & Fluck, M. (1997) *J. Mol. Biol.* **271,** 76–99.
47. Botchan, M., Topp, W. & Sambrook, J. (1979) *Cold Spring Harbor Symp. Quant. Biol.* **43,** 709–719.
48. Schimke, R. T., Roos, D. S. & Brown, P. C. (1987) *Methods Enzymol.* **151,** 85–104.
49. Schimke, R. T. (1984) *Cell* **37,** 705–713.
50. Schimke, R. T., Sherwood, S. W., Hill, A. B. & Johnston, R. N. (1986) *Proc. Natl. Acad. Sci. USA* **83,** 2157–2161.
51. Flotte, T. R., Afione, S. A., Solow, R., Drumm, M. L., Markakis, D., Guggino, W. B., Zeitlin, P. L. & Carter, B. J. (1993) *J. Biol. Chem.* **268,** 3781–3790.
52. Haberman, R. P., McCown, T. J. & Samulski, R. J. (1998) *Gene Ther.* **5,** 1604–1611.
53. Haberman, R., McCown, T. & Samulski, R. (2000) *J. Virol.* **74,** 8732–8739.
54. Zhang, G., Song, Y. K. & Liu, D. (2000) *Gene Ther.* **7,** 1344–1349.
55. Dyall, J., Szabo, P. & Berns, K. I. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 12849–12854.