Infectious Molecular Clones of Adeno-Associated Virus Isolated Directly from Human Tissues

Bruce C. Schnepp, Ryan L. Jensen, K. Reed Clark, and Philip R. Johnson

The Research Institute, Nationwide Children’s Hospital, Columbus, Ohio, and Department of Pediatrics and Department of Molecular Immunology, Virology, and Medical Genetics, College of Medicine, Ohio State University, Columbus, Ohio

Received 7 August 2008/Accepted 14 November 2008

Adeno-associated virus (AAV) replication and biology have been extensively studied using cell culture systems, but there is precious little known about AAV biology in natural hosts. As part of our ongoing interest in the in vivo biology of AAV, we previously described the existence of extrachromosomal proviral AAV genomes in human tissues. In the current work, we describe the molecular structure of infectious DNA clones derived directly from these tissues. Sequence-specific linear rolling-circle amplification was utilized to isolate clones of native circular AAV DNA. Several molecular clones containing unit-length viral genomes directed the production of infectious wild-type AAV upon DNA transfection in the presence of adenovirus helper. DNA sequence analysis of the molecular clones revealed the ubiquitous presence of a double-D inverted terminal repeat (ITR) structure, which implied a mechanism by which the virus is able to maintain ITR sequence continuity and persist in the absence of host chromosome integration. These data suggest that the natural life cycle of AAV, unlike that of retroviruses, might not have genome integration as an obligatory component.

Since its initial description over 40 years ago (1, 2, 15), adeno-associated virus (AAV) has emerged as a unique member of the parvovirus family. AAV is replication incompetent and requires a helper virus (e.g., adenovirus) to produce new infectious virions (1, 6, 17, 25, 33, 34). In the absence of helper virus, the hallmark of AAV infection in cultured cells is site-specific integration of the viral genome into a specific locus (AAVS1) on human chromosome 19 (20, 21, 36). Virally encoded gene products, through the recognition and binding of similar viral and cellular sequences, mediate this unique site specificity. The integrated head-to-tail AAV DNA arrays contain rearranged viral inverted terminal repeats (ITRs) and flanking cellular sequences (22, 25, 43). Helper virus infection of cells harboring integrated AAV DNA results in rescue of the AAV genome, leading to the production of new infectious particles. This unique property of site-specific integration has been well documented in transformed cultured cells, but to date, there is only a single report of AAVS1 site-specific integration in two human testis tissue samples (24). We have not observed AAVS1 integration using a sensitive nested PCR assay in 54 AAV-positive human (38) and nonhuman (unpublished observations) primate nontesticular tissue samples.

While ubiquitous in nature, AAV has never been associated with any disease or pathological condition in humans (3–5, 13), and consequently, there have been relatively few studies of in vivo AAV biology. Most observations concerning AAV biology have utilized cultured cells, and these findings have been widely accepted as the model for natural infection in the primate host. The continued development of recombinant AAV vectors for human gene transfer has brought renewed interest in wild-type AAV infection in the natural host. A better understanding of the biology of natural infection will undoubtedly affect the design and use of recombinant AAV vectors in humans.

We recently characterized AAV genomes derived directly from human tissue samples (7). Most of the genomes were closely related to AAV serotype 2; a single isolate shared homology with both serotypes 2 and 3. Moreover, the majority of the predicted amino acid substitutions (relative to AAV2) were conserved among the individual samples, suggesting that a specific virus isolate was circulating in the local population during the time period of tissue sample procurement (winter 2002 to 2003). Interestingly, none of the AAV2 capsid sequences retained arginine residues at positions 585 and 588 (R585S and R588T substitutions), which have been shown to be critical for heparin sulfate proteoglycan receptor binding (18, 26). Consistent with this observation, we showed that two infectious AAV preparations generated using methods detailed herein failed to bind to heparin sulfate chromatography columns (7). These data suggest that this receptor is not required for natural infection in humans, at least for the viruses that we isolated for these studies (7). In contrast to the prevailing wisdom that AAV genomes integrate in humans, analyses of the molecular forms of the AAV DNA in these tissues revealed that the genome sequences in these tissues were extrachromosomal episomes (38). In fact, we did not identify any integration events at the AAVS1 locus, as judged by the failure to amplify AAV-AAVS1 junction fragments by using a sensitive nested-PCR-based assay.

In the present study, we set out to characterize extrachromosomal AAV genomes found in human tissues. In doing so, we developed a novel sequence-specific linear rolling-circle amplification method (SSLRCA) that amplified the extrachromosomal AAV genomes directly from tissue DNA. The resulting circular monomeric AAV genomes were shown to be biologically active when directly transfected into HeLa cells.

* Corresponding author. Present address: Room 1216B, Abramson Research Center, The Children’s Hospital of Philadelphia, 3615 Civic Center Blvd, Philadelphia, PA 19104-4318. Phone: (267) 426-0351. Fax: (267) 426-0363. E-mail: johnsonphi@chop.edu.

† Published ahead of print on 19 November 2008.
DNA was digested with a single-cut enzyme (XbaI or HindIII) and the unit products (40). Briefly, a complete ITR-ITR end-joined template was generated by digesting a psSub201-based rAAV vector plasmid (rAAV/CMV/eGFP) with PvuII to release the intact rAAV expression cassette (including ITRs). This fragment was self-ligated and then treated with PS-DNase to remove all linear forms. A portion of this circular template (diluted 1:100) was used as the template in an SSLRCA reaction as described above. Both self-ligated and SSLRCA products were linearized by Sau3A enzyme digestion, treated with alkaline phosphatase, and end-labeled using polynucleotide kinase and γ-[32P]ATP. After labeling, the products were digested with NdeI to create linear, end-labeled fragments. Selected molecular clones isolated from tonsil sample T88 also underwent SSLRCA and end labeling in a similar manner as described above, but with the following minor modifications. T88 clones were released from the plasmid vector backbone by Xbal digestion prior to self-ligation and treatment with PS-DNase, and then they were linearized with SnaBI before treatment with alkaline phosphatase and polynucleotide kinase. DraIII was used to generate linear end-labeled fragments. To determine ITR-ITR junction heterogeneity occurring in vivo, SSLRCA products were treated identically as described above for T88-derived molecular clones. DNA fragments were fractionated on 6% denaturing polyacrylamide gel and exposed directly to film. Specific band counts were determined using a Typhoon 9400 variable mode imager (GE Healthcare) with the ImageQuant 5.2 software.

**Functional analysis of AAV molecular clones.** To determine if SSLRCA-generated AAV molecular clones could produce infectious virions, the pBlueScript vector backbone and cap ORF were recreated by digestion with XbaI and self-ligation to form head-to-tail circular monomeric genomes. The ligated products (1 μg) were directly transfected into 4 × 10^6 HeLa cells by using the FuGene 6 transfection reagent (Roche Applied Science), followed by infection with Ad5. After maximum cytopathic effect development (48 h), the transfected cells (P0) were harvested and low-molecular-weight DNA prepared from half of the cells. The recovered DNA was subsequently treated with DpnI to digest input plasmid DNA. A clarified cell lysate was generated from the other half of P0 cells by three freeze-thaw cycles and a 60-min incubation at 50°C to heat-inactivate Ad5. The clarified lysate (1:10 dilution) was used to infect 4 × 10^6 HeLa cells (along with Ad5). After 48 h, the transduced cells (P1) were harvested and low-molecular-weight DNA isolated.

**Western blot analysis.** To detect the presence of Rep and Cap proteins, a cell lysate was generated from transfected cells (P0) after 48 h and electrophoresed on 10% sodium dodecyl sulfate-polyacrylamide gel electrophoresis and subsequently transferred to a polyvinylidene difluoride membrane (GE Healthcare). The membrane was blocked for 10 min at 25°C using TBST buffer (10 mM Tris [pH 8.0], 150 mM NaCl, 0.1% Tween 20) plus 5% nonfat dry milk. Primary antibodies (from American Research Products, Inc.) were used at a 1:100 dilution (Rep proteins, 1:1 mixture of anti-AAV Rep clone 303.9 and clone 226.7; Cap proteins, anti-AAV Cap clone B1) and hybridized in blocking solution for 1 h. The blot was washed three times for 10 min in TBST buffer and incubated for 30 min in TBST buffer plus 2.5% dry milk with a horseradish peroxidase-conjugated anti-mouse secondary antibody (Vector Laboratories) at a 1:10,000 dilution. The blot was washed as described above and developed using the ECL Plus detection system (GE Healthcare).

**Real-time PCR.** Wild-type AAV genomes and AAV plasmids were quantified by real-time PCR using the ABI Prism 7000 sequence detection system in a standard 25-μl TaqMan PCR with the following primers and probe: ForCapSS, 5'-AACG ACAACACTACITGGGTCG-3'; RevCapSS, 5'-AAGG GGCGATGAAATCCTGTTG-3', 900 nM final concentration; and CAPPSProbe, 5'-[FAM]CTACAGACACCCCCCTGGGGAATT'TTGAT[TAMRA-6-FAM]-3', 200 nM final concentration.

**RESULTS**

**SSLRCA.** In our previous work on characterization of AAV in human tissues, we found that the majority of AAV genomes were extrachromosomal and that none were integrated in the AAVS1 locus (38). As part of those efforts, we developed a novel method (SSLRCA) that allowed us to amplify extrachromosomal AAV genomes directly from tissue DNA. It occurred to us that these amplified products might represent intact, biologically active AAV genomes. To investigate this possibility, templates for SSLRCA were derived from total tissue DNA that was treated with a unique exonuclease (PS-DNase) which preferentially degraded linear DNA fragments,
leaving circular, double-stranded DNA intact (Fig. 1). This treatment enriched the substrate DNA for circular templates and facilitated amplification using AAV-specific primers and phi29 DNA polymerase. Following overnight isothermic incubation, the resulting high-molecular-weight amplicons were digested into unit-length fragments and cloned into a standard plasmid cloning vector for sequence and functional analyses. Unit-length AAV genomes were generated by releasing inserts from the vector backbone and performing an intramolecular ligation. The resulting circular monomeric AAV genome was then directly transfected into HeLa cells to evaluate infectivity.

To determine the limit of detection for circular AAV, decreasing amounts of a 15-kb plasmid containing AAV2 ITR and rep and cap genes were spiked into 100 ng of naive tonsil DNA and subjected to SSLRCA. The amplification over a range of input copies was determined for four independent DNA samples were subjected to SSLRCA, and six yielded AAV2-specific amplification. Digestion of the amplified products with a restriction enzyme that was predicted to cut once inside the AAV2 genome (HindIII) yielded the expected 4.7-kb unit-length fragment (Fig. 2). In addition, both high- and low-molecular-weight forms were present in samples T70 and T88. The high-molecular-weight forms might represent unresolved large concatemeric DNA that was undigested due to other genome species that do not contain the restriction site(s). The low-molecular-weight forms might correspond to genome species that contain additional restriction site(s) or could be derived from in vivo deleted viral genomes. The amount of amplified AAV2 product correlated with the input AAV copy number previously determined from the various samples (Table 1). The three AAV samples that failed to amplify (LG15, T17, and T41) harbored copy numbers that were at or below the limit of detection for the assay (see Table 2). The low level of amplification for T32 and T41 was also consistent with the low AAV2 copy numbers in these tissues.

As an important control, SSLRCA substrate specificity was examined using Detroit 6 cells (19), which are a well-characterized transformed cell line that contains integrated AAV2 genomes (head-to-tail tandem arrays) in the AAVS1 locus on chromosome 19 (q19.1). As noted above, AAV genomes integrated within the Detroit 6 DNA should not amplify using the SSLRCA method. Accordingly, 1,000 and 100,000 integrated AAV2 copies from Detroit 6 total DNA were spiked into 100 ng of naive tonsil DNA, and SSLRCA was performed. The AAV2 genome copy number was determined by cap gene quantitative PCR before and after SSLRCA, and the average amplification was 1.3 ± 0.0-fold over 10 independent replicate assays (data not shown). These data demonstrate the strict requirement for circular templates to achieve robust SSLRCA-dependent amplification.

**Direct amplification of AAV genomes from human tissue.** As noted above, we previously identified nine human DNA samples that contained variable amounts of AAV DNA ranging from 80 to 33,000 copies/μg of total cellular DNA (7). These nine DNA samples were subjected to SSLRCA, and six yielded AAV2-specific amplification. Digestion of the amplified products with a restriction enzyme that was predicted to cut once inside the AAV2 genome (HindIII) yielded the expected 4.7-kb unit-length fragment (Fig. 2). In addition, both high- and low-molecular-weight forms were present in samples T70 and T88. The high-molecular-weight forms might represent unresolved large concatemeric DNA that was undigested due to other genome species that do not contain the restriction site(s) or could be derived from in vivo deleted viral genomes. The amount of amplified AAV2 product correlated with the input AAV copy number previously determined from the various samples (Table 1). The three AAV samples that failed to amplify (LG15, T17, and T41) harbored copy numbers that were at or below the limit of detection for the assay (see Table 2). The low level of amplification for T32 and T41 was also consistent with the low AAV2 copy numbers in these tissues.

**Isolation of infectious AAV molecular clones.** Having amplified AAV genomes from human tissue samples, we next asked whether SSLRCA could be used to isolate intact full-length

---

**TABLE 1. SSLRCA assay sensitivity**

<table>
<thead>
<tr>
<th>Plasmid input (no. of copies)</th>
<th>Amplification (fold) for replicate:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1b</td>
</tr>
<tr>
<td>100</td>
<td>3.1 × 10^4</td>
</tr>
<tr>
<td>75</td>
<td>4.6 × 10^5</td>
</tr>
<tr>
<td>50</td>
<td>9.5 × 10^5</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
</tr>
<tr>
<td>10</td>
<td>0</td>
</tr>
<tr>
<td>7.5</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

* Copies of a circular 15-kb plasmid, containing AAV2 rep/cap genes, were spiked into 100 ng of total cellular DNA isolated from a naïve human tonsil sample and used as a template for SSLRCA.

b Amplification was determined by dividing the number of AAV cap copies present after SSLRCA by the number of cap copies present before the SSLRCA assay was performed.

c Number of total replicates scored as positive for the detection of circular template.
After adenovirus-induced cytopathic effect was observed, a cell lysate was prepared, clarified, and passed onto naïve HeLa cells (again, in the presence of adenovirus). Hirt DNA was isolated from both the transfected cells (P0) and the infected cells (P1) and analyzed by Southern blot hybridization. Analyses of nine clones isolated from three individual tonsil-adenoid tissue samples are shown in Fig. 3. DNA replication in P0 was observed for all clones, as judged by the presence of characteristic AAV monomeric (Rfm) and dimeric (Rfd) replicating DNA forms. Also present in the P0 Hirt DNA was a distinct 2.5-kb hybridizing fragment that was derived from the transfected DNA following DpnI enzyme digestion. Importantly, all the molecular clones gave rise to infectious virus, as judged by replicating AAV DNA in P1 cells. In summary, we were able to isolate infectious molecular clones from five of the six SSLRCA-positive tissue samples (Table 2). For sample S17, it remains possible that we simply did not isolate enough clones to identify an infectious clone.

Characterization of AAV produced from infectious molecular clones. To further characterize AAV produced by our clones, we serially passaged infectious virions derived from clones T70-43 and T88-41 and showed that particles were DNase I resistant and purified at the predicted density in iodixanol gradients (data not shown). The cap gene sequence for both of these isolates predicted replacement of charged arginine residues at positions 585 and 588 with serine and threonine, respectively. As predicted, purified virions representing both clones did not bind to a heparin sulfate matrix (7). We also analyzed AAV Rep and Cap protein expression profiles of these and selected other molecular clones (Fig. 4). As expected, all of the infectious molecular clones produced the four Rep and three Cap proteins. Interestingly, the Cap proteins (VP1, VP2, and VP3) derived from prototype AAV2 migrated faster than did the Cap proteins derived from the...
TABLE 2. Summary of molecular clones isolated by SSLRCA

<table>
<thead>
<tr>
<th>Tissue sample</th>
<th>No. of AAV copies&lt;sup&gt;a&lt;/sup&gt;</th>
<th>SSLRCA result&lt;sup&gt;b&lt;/sup&gt;</th>
<th>No. of clones analyzed&lt;sup&gt;c&lt;/sup&gt;</th>
<th>% Infectious clones&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>T17</td>
<td>56</td>
<td>−</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td>T32</td>
<td>21</td>
<td>+</td>
<td>1/1</td>
<td>100 (1/1)</td>
</tr>
<tr>
<td>T40</td>
<td>655</td>
<td>+</td>
<td>6/60</td>
<td>17 (1/6)</td>
</tr>
<tr>
<td>T41</td>
<td>79</td>
<td>+</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td>T70</td>
<td>780</td>
<td>+</td>
<td>15/80</td>
<td>93 (14/15)</td>
</tr>
<tr>
<td>T71</td>
<td>760</td>
<td>+</td>
<td>6/26</td>
<td>67 (4/6)</td>
</tr>
<tr>
<td>T88</td>
<td>3,300</td>
<td>+</td>
<td>14/80</td>
<td>71 (10/14)</td>
</tr>
<tr>
<td>S17</td>
<td>20</td>
<td>+</td>
<td>3/3</td>
<td>0 (0/3)</td>
</tr>
<tr>
<td>LG15</td>
<td>8</td>
<td>+</td>
<td>0</td>
<td>NA</td>
</tr>
<tr>
<td>Detroit 6</td>
<td>83,000</td>
<td>−</td>
<td>0</td>
<td>NA</td>
</tr>
</tbody>
</table>

<sup>a</sup> AAV copy number in 100 ng of total cellular DNA based on cap quantitative PCR.

<sup>b</sup> SSLRCA-positive result based on Southern blot detection of AAV genomes.

<sup>c</sup> The fraction represents the number of molecular clones that were analyzed for infectivity (passage) over the number of total clones isolated by bacterial colony hybridization.

<sup>d</sup> Since only a subset of clones was screened for passage, the percentage reflects our ability to obtain replicating monomeric and dimeric AAV genomic species following passage (P1). Fraction in parentheses represents the number of infectious clones over the number of clones analyzed for infectivity. NA, not applicable.

The phenomenon of double-D ITR structures present in tissues. Based on ITR sequence analysis, we concluded that infectious molecular clones contained a complete double-D ITR element. Conversely, limited ITR sequence analysis of noninfectious clones suggested that the deletion of internal ITR regions was responsible for the lack of infectivity, since these clones expressed the expected Rep and Cap proteins (data not shown). Consequently, we were interested to know if double-D ITR junction formation in vivo yielded predominately complete nonrearranged structures or if the ITR deletions were imprecise and only a subset contained the complete double-D ITR.

To answer this question, SSLRCA was performed directly on tissue genomic DNA because ITR rearrangements have been observed to occur during passage in bacteria. Preceding in vivo analysis of ITR forms, we characterized the fidelity of the SSLRCA assay by using several complex ITR structures, the first of which possessed two complete ITRs joined at their ends (Fig. 6). Following amplification, the ITR-ITR construct (input) and SSLRCA-amplified material were end-labeled and fractionated on an acrylamide gel. The end-joined ITR-ITR template was amplified faithfully, with no deletions evident based on identical mobility through the gel (Fig. 6, lane “TR-TR”). Additionally, when T88 molecular clones possessing four distinct double-D structures (Fig. 6, lanes “DD+,” “DD-,” “ADD,” and “D”) were similarly analyzed, all SSLRCA-amplified material was replicated faithfully.

After the preliminary analyses of these known substrates, SSLRCA was then performed on genomic DNA isolated directly from tonsil-adenoid sample T88 (Fig. 6, lane “T88”). Multiple discrete species were easily identified, and when these amplified fragments were compared on the same gel to the known tonsil T88 molecular clones, the predominant in vivo form corresponded in size to the complete double-D structure (Fig. 6, white arrowhead). Quantitative densitometry revealed that 48% of the total signal corresponded to intact double-D ITR structures. Significantly, there was no specific band present at the predicted size of an end-joined ITR-ITR structure (Fig. 6, white arrow), suggesting that in this particular sample, this was not an in vivo form. Moreover, the presence of numerous ITR species (some larger and many smaller) indicated that double-D formation is imprecise.

**DISCUSSION**

Molecular details regarding wild-type AAV replication, gene expression, and genome integration have been largely elucidated through the use of in vitro mammalian cell culture infection systems. These studies spawned a life cycle paradigm that holds site-specific virus integration as a central tenet.

**FIG. 4.** Rep and Cap expression of T88 molecular clones. Individual molecular clones were analyzed for AAV Rep (Rep78, 68, 52, 40) and Cap (VP1, VP2, VP3) protein expression. Shown are Western blot hybridization blots of transfected cell lysates (P0). “U” designates an untransfected cell lysate, while “wt” denotes a wild-type AAV2-infected lysate.
Integration in this model is mediated by the virally encoded large Rep proteins (Rep78 or Rep68) and involves recognition of cognate binding sites within the virus and host genome (GAGC tetra-nucleotide repeats). While targeted integration into chromosome 19 can be readily detected in cultured cells, it is driven by high MOIs and is actually a rare event that is estimated to occur once in every 1,000 productive transduction events in vitro (16). Only a single study to date has looked at experimental AAV infection in the primate host (14). The authors encountered difficulty in detecting the presence of the virus postinoculation in the oropharynx or respiratory tract, even in the presence of coinfecting adenovirus. This complexity illustrates the challenges associated with virus recovery and sampling effects when using large animal models. Studies using

\[
\begin{array}{cccccccc}
T88-41 & \text{REP} & D & A & C & C' & B & B' & A' & D & \text{CAP} \\
T88-79 & \text{REP} & D & A & B & B' & C & C' & A' & D & \text{CAP} \\
T88-62 & \text{REP} & D & A & B & B' & C & C' & A' & D & \text{CAP} \\
\end{array}
\]

FIG. 5. Structure of the AAV ITR junctions from selected molecular clones. Shown schematically are ITR structures observed in several infectious molecular clones, with the palindromic ITR sequences designated as A, B, B', C, C', A', and D', in accord with standard nomenclature. The complete ITR sequence was determined for clones T88-41, T88-62, and T88-79. Clones 41 and 62 contain a complete double-D ITR (in both flip and flop orientations). T88-62 contains an imperfect double-D ITR junction with two A' regions, both of which contain small deletions. In addition, three T70 clones (no. 18, 42, and 64) and three T71 clones (no. 1, 5, and 10) also possessed a complete double-D structure that was identical to that of T88-79. All ITR sequences were identical to the published AAV2 ITR sequence.

\[
\begin{array}{cccccccc}
\text{ITR Junction Structure} & \text{TR-TR} & \text{T88} & \text{DD+} & \text{DD} & \text{DD} & \text{D} & \text{SSLRCA} \\
\text{DD} & R & D & A & B & B & C & C & A & D & C \\
\text{ADD} & R & D & A & C & A & D & C \\
\text{D} & R & D & C \\
\end{array}
\]

FIG. 6. Episomal AAV contains double-D ITR structures in vivo. Various circular AAV templates were subjected to SSLRCA. The input templates (−) and the amplified products (+) were end-labeled and analyzed by acrylamide gel electrophoresis. Schematics of the ITR structure of various templates are shown to the left of the autoradiograph, with the various ITR regions denoted as A, B, C, D. The ITR junction is oriented with respect to the AAV rep (R) and cap (C) genes. Black arrowheads designate the specific input (−) and SSLRCA (+) bands. The ITR-ITR template contains two individual end-joined ITRs. The DD+ template was derived from clone T88-62, DD from clone T88-41, ΔDD from clone T88-16, and D from clone T88-78B. In all cases, the input template was amplified correctly by SSLRCA, as evidenced by identical band migration before and after SSLRCA. When SSLRCA products from tonsil sample T88 were analyzed in this fashion, several distinct bands were identified. The most prevalent band (48% of the overall signal; white arrowhead) corresponded in size to the complete DD template. The predicted size of an end-joined ITR-ITR junction from tonsil sample T88 is represented by a white arrow.
AAVS1 transgenic mice have documented targeted integration (32, 35), illustrating that there are probably inherent differences between cell culture and in vivo models of infection. We took an alternative approach to understanding molecular aspects of AAV natural infection by characterizing the molecular form of AAV genomes in human tissues. The SSLRCA method enabled the isolation of episomal molecular clones, from which structure function data were derived.

Our analyses led to four fundamental observations. First, based on our previous analysis (38), the predominant in vivo forms were extrachromosomal circles. Second, episomal circles possessed complete or partially deleted double-D ITR structures. Third, infectious DNA clones possessed complete double-D structures. Finally, the cellular mechanism for circularization appeared to be imprecise and resulted in multiple double-D ITR forms within a single tissue (but preferentially yielded the complete double-D element). The observation that almost 50% of the circular forms within a human tonsil sample were perfect double-D structures suggested sufficient recombinational specificity to maintain ITR continuity in a significant proportion of viral genomes to enable subsequent virus rescue. Additionally, the fact that infectious molecular clones were also adenovirus DNA negative (7) argues that the episomal forms amplified were not simply circular replicative forms.

These data have significant implications for our understanding of the AAV life cycle. The presence of the double-D ITR structure has been demonstrated to be the sole cis element required for AAV propagation (42). Upon double-D ITR resolution (which occurs in the presence of a coinfected helper virus), a no-end substrate is generated, and subsequent nicking at one of the two available terminal resolution sites by the large Rep proteins would allow AAV DNA replication to proceed via leading-strand synthesis and infectious virus formation. The direct implication is that AAV can proceed through its life cycle in the complete absence of an integration intermediate in vivo, and in fact, this has been shown in cell culture (30). Additionally, we and other investigators have reported identical double-D ITR elements for recombinant AAV vectors in muscle tissue, suggesting that AAV ITRs contain all the cis sequence information necessary to undergo rep-independent circularization (11, 37, 42). Thus, these data make the case for an additional cis ITR property, namely, circularization. It is logical that the viral ITR has evolved to promote accurate circularization via homologous recombination to preserve a critical structural element (double-D ITR) that appears essential for viral infectivity upon rescue following helper virus infection.

The mechanism of ITR circularization is unknown, but data using self-complementary vectors suggest that it occurs through a double-stranded intermediate, following second-strand synthesis (23). In cell culture, ITR circularization is a rapid process, and higher-order forms (concatemers) appear to form through the interaction of double-D ITR elements on separate circles (10, 11). Additional research using cells in culture has focused on characterizing those cellular enzymes involved in the circularization process. To date, multiple DNA double-strand break repair enzymes are implicated and include ATM, Mre11, DNA-PKcs, KU 70/86, and Rad 50, 51, and 52 proteins (8, 12, 39, 41). These recombination proteins could be envisioned to either promote circularization (e.g., Mre11 and ATM) or act to inhibit integration (e.g., DNA-PKcs). The multiple-sized double-D ITR structures observed in tonsil sample T88 are consistent with either homologous recombination or nonhomologous end-joining repair pathways facilitating ITR-ITR ligation. The failure to detect complete head-to-head ITR junctions and the presence of complete double-D isoforms argue that recombination via nonhomologous end joining may not be the favored pathway of ITR circularization but could still play a role in ITR-ITR ligation. In fact, it might be beneficial for AAV to use multiple pathways for circularization, thereby avoiding free hairpin ends that can be sensed as abnormal by the cell and, under specific circumstances, trigger apoptotic stimuli (31).

Under conditions of high virus inputs, the circularized viral form may act as an integration intermediate. Two observations are consistent with this supposition. First, most proviral integrants are arranged in head-to-tail arrays, which are thought to require a circular template and limited DNA replication to generate the arrays. Second, efficient AAVS1 integration in cell culture is heavily dependent on the MOI (28). Thus, at low MOI (<100 to 1,000), integration in cell culture is a rare event, and the viral genomes are lost during subsequent passage. It is unclear why a threshold MOI is needed for targeted integration; perhaps a relatively high intranuclear concentration of circular AAV molecules is needed for juxtaposition and formation of viral ITR/AAVS1 complexes. Whether similar input viral loads are necessary in vivo for integration is unknown, but we have observed no detectable integration into the AAVS1 locus by wild-type AAV genomes in excess of 400 copies per cell in nonhuman primate tissues (unpublished observation).

Although we were unable to isolate infectious molecular clones from episomal AAV genomes from all AAV-positive tissue samples (T17, T41, and LG15) by using SSLRCA, we were nonetheless unable to detect AAVS1 integration in these same tissue samples by using a very sensitive nested PCR assay (five integration events in 1.67 × 10^5 cell equivalents) (38). Thus, it seemed that integration within the AAVS1 locus was unlikely. Clearly, an integration event could have occurred at other sites within the genome, but we documented only a single event in sample T71 after use of a sensitive linear amplification-mediated PCR assay (one integration event in 1 × 10^3 cell equivalents) (38). Our inability to isolate infectious molecular clones seems more likely due to the fact that the AAV copy numbers in these tissues were at or below the limit of sensitivity for SSLRCA.

Taken together, our data compel us to take a fresh look at the role of integration in the AAV life cycle. On the one hand, we have shown that AAV DNA does not appear to integrate with any measurable frequency at the AAVS1 locus in human tissues (38). Importantly, we have now shown that integration is not required to generate infectious viral genomes. These observations have been confirmed in AAV-infected nonhuman primates (our unpublished observations). Furthermore, others have shown that in cell culture, only 0.1% of infectious AAV particles integrate into AAVS1 (28, 29). Thus, these considerations make it difficult to argue for a prominent role of virus integration in AAV natural infection. On the other hand, integration would clearly be of benefit in actively dividing cell populations, and perhaps this might occur at higher frequencies under “favorable” intracellular conditions (e.g., sufficient...
Rep expression and high virus inputs). Testing of this latter hypothesis will await controlled nonhuman primate infection studies.

In conclusion, our data convincingly suggest that the natural AAV life cycle, in contrast to that of retroviruses, does not have DNA integration as an obligatory component. Rather, AAV genomes appear to persist in tissues as monomeric and concatemeric circles. Importantly, the molecular form (double-D ITR) of these extrachromosomal circles probably permits virus rescue and replication upon helper virus infection, and it allows for a viral life cycle that does not require integration for propagation in the human host. Given that recombinant AAV vectors contain wild-type ITRs, it now seems likely that recombinant vector genomes behave in an analogous manner to wild-type genomes, and this has been borne out in the literature (11, 27, 37, 38).

ACKNOWLEDGMENTS

We thank Douglas McCarty for helpful discussions and critical reading of the manuscript. We also thank Robin Bolek for expert technical assistance.

This work was funded by the National Institutes of Health.

REFERENCES