JOURNAL OF VIROLOGY, Oct. 2007, p. 11304–11321 Vol. 81, No. 20
0022-538X/07/$08.00+0 doi:10.1128/JVI.01225-07

The Role of DNA-PKcs and Artemis in Opening Viral DNA Hairpin Termini in Various Tissues in Mice▽

Katsuya Inagaki,[1] Congrong Ma,[1] Theresa A. Storm,[2] Mark A. Kay,[2] and Hiroyuki Nakai[1]*

Department of Molecular Genetics & Biochemistry, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15261,[1] *and Departments of Pediatrics and Genetics, Stanford University School of Medicine, Stanford, California 94305*[2]

Received 4 June 2007/Accepted 30 July 2007

A subset of cellular DNA hairpins at double-strand breaks is processed by DNA-dependent protein kinase catalytic subunit (DNA-PKcs)- and Artemis-associated endonuclease. DNA hairpin termini of adeno-associated virus (AAV) are processed by DNA repair machinery; however, how and what cellular factors are involved in the process remain elusive. Here, we show that DNA-PKcs and Artemis open AAV inverted terminal repeat (ITR) hairpin loops in a tissue-dependent manner. We investigated recombinant AAV (rAAV) genome metabolism in various tissues of DNA-PKcs- or Artemis-proficient or -deficient mice. In the absence of either factor, ITR hairpin opening was impaired, resulting in accumulation of double-stranded linear rAAV genomes capped with covalently closed hairpins at termini. The 5′ end of 3-base hairpin loops of the ITR was the primary target for DNA-PKcs- and Artemis-mediated cleavage. In the muscle, heart, and kidney, DNA-PKcs- and Artemis-dependent hairpin opening constituted a significant pathway, while in the liver, undefined alternative pathways effectively processed hairpins. In addition, our study revealed a Holliday junction resolvase-like activity in the liver that cleaved T-shaped ITR hairpin shoulders by making nicks at diametrically opposed sites. Thus, our approach furthers our understanding of not only rAAV biology but also fundamental DNA repair systems in various tissues of living animals.

Adeno-associated virus (AAV) of various serotypes is a family of nonpathogenic replication-defective parvoviruses with a linear single-stranded DNA genome of approximately 5 kb. Recombinant AAV (rAAV) devoid of all the virally encoded genes is not only a promising viral vector for human gene therapy but also has recently emerged as a powerful tool for studying cellular biology, particularly DNA repair pathways (18). Upon infection in animals, rAAV genomes change their forms in infected cells by recombination through DNA hairpins at viral termini, named inverted terminal repeats (ITRs). DNA hairpins at AAV-ITR serve as a platform for all the rAAV genome recombinations, including self circularization (8), concatemerization (37), and integration (29). These processes are presumed to be mediated solely by cellular DNA repair machinery. Therefore, studying the interactions between viral DNA hairpins and host cellular DNA repair machinery will not only enhance our understanding of the biology of rAAV vectors, which is essential for successful human gene therapy, but also will provide important clues to elucidating fundamental cellular processes of DNA damage responses.

DNA hairpin structures in mammalian cells are often a target for cellular DNA repair machinery. The best example is hairpin-coding ends that emerge during the V(D)J recombination, a physiological and programmed DNA double-strand break (DSB) that occurs in developing B and T lymphocytes (23, 46). DNA hairpins at pathological DSBs caused by intrinsic and extrinsic sources such as reactive oxygen species and ionizing radiation also are processed by DNA repair machinery (22). It has been demonstrated that cellular endonuclease activity associated with DNA-dependent protein kinase catalytic subunit (DNA-PKcs) and Artemis opens and processes such DNA hairpin structures in the classical nonhomologous end-joining (NHEJ) repair pathway (26). The Rag proteins also have some roles in opening hairpin-coding ends in V(D)J recombination (13). Although our understanding of the metabolism of DNA hairpin structures at cellular DNA DSBs has progressed, little is known about how and what cellular factors process AAV-ITR DNA hairpin termini.

Artemis is a newly discovered cellular DNA repair endonuclease that possesses hairpin-loop-opening activity, identified as a protein deficient in human radiation-sensitive severe combined immunodeficiency (SCID) patients (30), in whom the V(D)J recombination is severely impaired due to a defect of the coding joint formation. In the past 5 years, our knowledge about DNA-PKcs and Artemis in the NHEJ pathway has been substantially enhanced by biochemical analyses of purified proteins, studies of various DNA repair pathway-defective cell lines, and investigation of the V(D)J recombination in various in vitro and in vivo systems (4, 14, 21, 25–28, 30, 41, 43, 44). It has been proposed that DNA-PKcs is recruited by Ku proteins bound at DNA DSBs and subsequently activates Artemis endonuclease activity, which otherwise remains inactive, by phosphorylation and/or direct interaction with autophosphorylated DNA-PKcs (14, 26). Activated Artemis/DNA-PKcs complex cleaves Rag-generated DNA hairpins and processes a diverse set of configurations at broken DNA ends (21, 26–28).

Previous studies by us and others have indicated the involvement of DNA-PKcs in rAAV genome metabolism. It has been shown that rAAV genome circularization via viral hairpin termini is impaired to some extent in the skeletal muscle (10, 48) but not in the liver (36, 49) of SCID mice deficient in DNA-

* Corresponding author. Mailing address: Department of Molecular Genetics & Biochemistry, University of Pittsburgh School of Medicine, W1244 BSTWR, 200 Lothrop St., Pittsburgh, PA 15261. Phone: (412) 648-8958. Fax: (412) 624-1401. E-mail: nakaih@pitt.edu.

▽ Published ahead of print on 8 August 2007.

PKcs. In addition, a study has indicated that integration of rAAV genomes into host chromosomal DNA is promoted in DNA-PKcs-deficient SCID mouse liver (49). These observations have implications in the interaction between AAV-ITR DNA hairpins and NHEJ repair machinery; however, their mechanistic interpretations have remained ambiguous.

To begin to elucidate the network of rAAV genome-DNA repair machinery interactions and understand DNA repair pathways in various tissues in animals, in the present study we investigated the roles of DNA-PKcs and Artemis in AAV-ITR hairpin metabolism in mice. Based on the previous observations that rAAV genome circularization is impaired in DNA-PKcs-deficient SCID mice (10, 36, 48) and the current knowledge about biological functions of DNA-PKcs and Artemis in opening DNA hairpins, we hypothesized that Artemis/DNA-PKcs endonuclease activity opens AAV-ITR hairpin termini and triggers rAAV genome recombinations in vivo. To test this hypothesis, we extensively characterized rAAV genomes processed by cellular DNA repair machinery in various tissues in DNA-PKcs- or Artemis-proficient or -deficient mice under various experimental conditions. The study demonstrated that Artemis/DNA-PKcs in fact cleaves the tips of viral DNA hairpins and triggers subsequent viral genome recombinations in vivo. In addition, the study indicated that there are Artemis- or DNA-PKcs-independent alternative pathways with a greater tissue-specific difference in processing viral DNA hairpins and implied that undefined resolvase activities, including Holliday junction resolvase-like activity, also participate in viral DNA hairpin metabolism.

MATERIALS AND METHODS

rAAV production. pAAV-ISce I.AO3 plasmid (accession number EU022316) was constructed to produce rAAV serotype 2 and 8, AAV2-ISce I.AO3 and AAV8-ISce I.AO3, as described elsewhere (16a). Plasmid pAAV-EF1α-nlslacZ, for production of AAV8- or AAV9-EF1α-nlslacZ, was described previously (16, 38). rAAV was produced using a triple-plasmid transfection method and was purified, and the titers were determined as previously described (5). AAV2- and AAV8-ISce I.AO3 did not express transgenes in mammalian cells. They carried a prokaryotic β-lactamase expression cassette and pUC ori (Fig. 1A), but this feature is not relevant to the present study. AAV8- or AAV9-EF1α-nlslacZ carried a human elongation factor 1α (EF1α) promoter-driven nuclear-localized bacterial β-galactosidase expression cassette.

In vitro synthesis of artificial no-end double-stranded linear monomer rAAV genome capped with covalently closed AAV-DDITR hairpins. No-end double-stranded linear monomer rAAV genome capped with covalently closed 165-base AAV-double-D ITR (DDITR) hairpins was synthesized according to the method of Snyder et al., with a modification (47). Briefly, approximately 50 μg of pVmeδDhF.IX plasmid was digested with PvuII (4 U per μg) at 37°C for 3 h. The plasmid pVmeδDhF.IX carried the AAV-EF1α-hF.IX (human coagulation factor IX-expressing AAV vector under the EF1α promoter) genome with two 130-bp AAV2-ITRs between the two PvuII sites. Therefore, PvuII digestion could excise the entire AAV-EF1α-hF.IX genome (31). Following phenol-chloroform extraction and ethanol precipitation, PvuII-digested plasmid DNA was digested with *Escherichia coli* exonuclease III (7.5 U per μg DNA; Exo III; New England BioLabs [NEB]) at 37°C for 50 s. Following heat inactivation of Exo III at 75°C for 15 min, DNA was treated with T4 DNA polymerase (0.5 U per μg DNA; T4 DNA Pol; NEB) and T4 DNA ligase (120 U per μg DNA; T4 DNA Lig; NEB) at 37°C for 3 h. Following phenol-chloroform extraction and ethanol precipitation, the DNA was again treated with Exo III (10 U per μg DNA) for 2 h. Dumbbell-shaped double-stranded linear rAAV genome, the 5′ and 3′ genome termini of which were not covalently closed but for which the whole molecular size was the same as that of the covalently closed no-end genome, was synthesized in the same manner as that used for the no-end genome, except that T4 DNA Lig was not added to the reaction mixture. All the reaction buffers used in the present study were exactly the same as those reported by Snyder et al. (47) unless otherwise noted.

Animal handling. All the animal experiments were performed according to the guidelines for animal care at Stanford University and the University of Pittsburgh. C57BL/6J mice and DNA-PKcs-deficient C57BL/6J SCID mice were purchased from Jackson Laboratory. Artemis-deficient mice developed by Frederick W. Alt's laboratory (44) were obtained from Susanna M. Lewis, Hospital for Sick Children Research Institute, Toronto, Canada. 129S6/SvEvTac mice were purchased from Taconic Farms Inc. The methods for tail and hepatic portal vein injections, intraperitoneal injection, and two-thirds partial hepatectomy have been described previously (16, 33, 40). Some of the rAAV-injected animals also were used for a separate study investigating rAAV integration (16a).

Southern blotting. Total genomic DNA was extracted from various mouse tissues by standard phenol-chloroform extraction. Genomic Southern blot analysis was performed according to our standard procedures as described previously (33, 37). Briefly, 10 to 20 μg of total DNA was digested with an appropriate restriction endonuclease(s) (4 U per μg) for 4 h. Following ethanol precipitation, DNA was dissolved in 20 μl of 1× Tris-EDTA (TE), electrophoresed on a 0.8% neutral agarose gel, blotted onto a nylon membrane, and probed with a ^{32}P-labeled probe. In several instances, DNA was treated with ATP-dependent exonuclease (4 U per μg; Plasmid Safe; Epicenter) at 37°C overnight after PstI digestion (4 U per μg, at 37°C for 4 h; PstI does not cut the AAV-ISce I.AO3 genome) or was treated with Exo III (10 U per μg) at 37°C for 2 h after EcoRI digestion (4 U per μg, at 37°C for 4 h; EcoRI digestion makes 5′ overhangs and does not cut the AAV-ISce I.AO3 genome). PS-negative and Exo III-negative control reactions were performed in parallel in the same reaction mixtures, except for the absence of each enzyme. For alkaline gel Southern blot analysis, 1.0% agarose gels were used.

The double-stranded AAV-ISce I.AO3 genome copy number in each sample was determined by Southern blot analysis. Total DNA was digested with BamHI and AlwNI, which excised a 1.3-kb AAV-ISce I.AO3 internal sequence, and was probed with a β-lactamase probe (Fig. 1A). Double-stranded rAAV genome copy number standards were prepared using pAAV-ISce I.AO3 plasmid.

Southern blot analysis of in-vitro-synthesized no-end double-stranded linear monomer rAAV genome and its intermediates produced during the enzymatic synthesis of the no-end genome was performed as follows. One microgram of DNA was collected from each step in the process of no-end genome synthesis, i.e., for step 1, PvuII was added; step 2, PvuII and Exo III were added; step 3, PvuII, Exo III, and T4 DNA Pol were added; and step 4, PvuII, Exo III, T4 DNA Pol, and T4 DNA Lig were added. A portion of each DNA collection either was incubated with or without Plasmid Safe (4 U per μg) at 37°C overnight or was incubated with or without Exo III (10 U per μg) at 37°C for 2 h in the appropriate reaction buffers, was separated on a 0.8% neutral agarose gel or a 1.0% alkaline agarose gel, and was subjected to Southern blot analysis as described above. The remaining portion of each DNA collection was digested with either DraIII or BamHI (left-side or right-side single cutter of the rAAV genome, respectively) and analyzed in the same way.

Southern blot analysis of Exo III-resistant no-end double-stranded linear monomer rAAV genomes recovered and purified from animal tissues was performed in essentially the same manner as that described above.

Probes used in the analyses were 0.7-kb β-lactamase probe, 2.1-kb *lacZ* gene probe, and 2.0-kb hF.IX cDNA probe. Densitometry was used to quantify Southern blot signals as previously described (36). All the copy number standard plasmids and other control DNA used in the Southern blot analyses were spiked into naïve liver genomic DNA and were treated in the same manner as that used for the experimental samples unless otherwise stated.

Gel purification of Exo III-resistant no-end double-stranded linear monomer rAAV genomes from mouse liver DNA. One hundred micrograms of total liver DNA recovered from rAAV-injected animals was digested with EcoRI or HindIII (4 U per μg) at 37°C for 4 h and was purified by phenol-chloroform extraction followed by ethanol precipitation. EcoRI and HindIII were used for AAV8-ISce I.AO3-transduced and AAV8-EF1α-nlslacZ- or AAV9-EF1α-nlslacZ-transduced samples, respectively. This enzyme-rAAV genome combination does not cleave the rAAV genomes. The DNA then was digested with Exo III (10 U per μg) at 37°C for 2 h. The treated DNA was separated on 1.0% low-melting-temperature agarose gel together with a no-end double-stranded linear monomer rAAV genome marker at a low voltage and at 4°C overnight. The regions corresponding to the position of each no-end double-stranded linear monomer rAAV genome were excised from the gel, and size-fractionated DNA was recovered from gel slices by β-agarase (NEB) treatment and purified by phenol-chloroform extraction followed by ethanol precipitation. The purified DNA was again treated with 100 U of Exo III at 37°C for 2 h, purified with phenol-chloroform extraction, and ethanol precipitated. The DNA pellet was dissolved in 32 μl of 1× TE and stored at −20°C. An agarose gel piece recovered

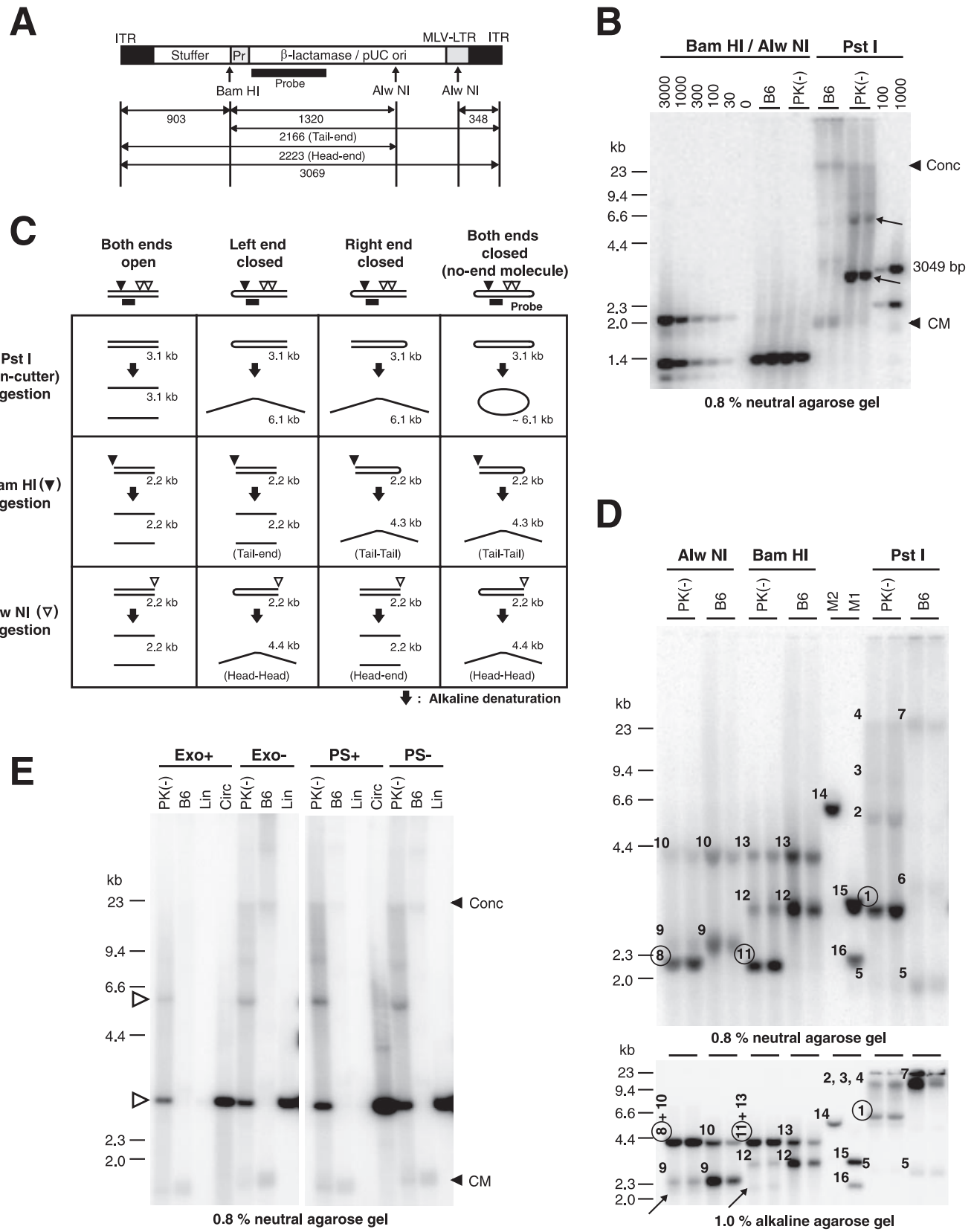

FIG. 1. Southern blot analyses of rAAV genomes in wild-type and DNA-PKcs-deficient mouse livers transduced with a high dose of AAV8-ISce I.AO3. (A) AAV-ISce I.AO3 genome map. The locations of restriction enzyme recognition sites and a β-lactamase probe used for Southern blot analyses and sizes of each digested DNA fragment (in bases or base pairs) are indicated. Pr, Tn*3* prokaryotic promoter; pUC ori, pUC plasmid origin of replication; MLV-LTR, a portion of the Moloney murine leukemia virus long terminal repeat. (B) Analysis of rAAV genome forms by

from a lane on which no DNA was loaded was processed in the same manner as a control.

Bisulfite DNA sequencing of rAAV viral hairpin termini. A bisulfite PCR and sequencing technique was used to determine rAAV viral terminal sequences. Bisulfite chemical modification was performed as previously described by Grunau et al. (15), with minor modifications. Briefly, 10 μg of total mouse DNA (sample DNA and control DNA) with or without a control plasmid, 5×10^{10} rAAV genomes with 10 μg of yeast tRNA (Invitrogen), or a 1/18 volume (2 μl) of gel-purified no-end double-stranded linear monomer rAAV genome preparation with 10 μg of yeast tRNA was alkaline denatured in 111 μl of freshly prepared 0.3 M NaOH solution at 42°C for 20 min. To the mixture was added 1,200 μl of 10 mM hydroquinone solution (Sigma-Aldrich H9003) saturated by sodium hydrogensulfite (Sigma-Aldrich 293937), and then it was covered with 200 μl of mineral oil and incubated at 55°C for 16 h. Following desalting with a QIAquick gel extraction kit (QIAGEN), the mixture was adjusted to 0.3 M NaOH, incubated at 37°C for 20 min, and neutralized with ammonium acetate. The chemically modified DNA was ethanol precipitated together with 1 μg of tRNA and 20 μg of oyster glycogen. The precipitated DNA was dissolved in 20 μl of 1× TE and stored at −20°C.

Bisulfite-modified AAV-ITR, AAV-DDITR, and rAAV viral terminal sequences in the samples were amplified by hot start PCR in 50-μl reaction mixtures containing 1× GeneAmp PCR Gold buffer, 2.5 mM $MgCl_2$, 20 pmol each of forward and reverse primers, and 1.25 U of AmpliTaq Gold (Applied Biosystems). PCR cycles were the following: 5 min at 95°C, 34 cycles of 15 s at 95°C, 30 s at 50°C, and 30 s at 72°C, and subsequently 5 min at 72°C. PCR products were checked by 2.5% agarose gel electrophoresis and were cloned into pCR2.1-TOPO vector using a TA TOPO cloning kit (Invitrogen). ElectroMax DH10B *E. coli* (Invitrogen) cells were transformed with the ligation mixture, and the resulting plasmid libraries were plated on Luria-Bertani (LB) agar plates containing ampicillin (50 μg/ml) and 5-bromo-4-chloro-3-indolyl-β-D-galactoside (X-Gal). Sequences of TOPO vector inserts were determined with a 3730x DNA analyzer (Applied Biosystems) using M13 reverse primer. PCR primer sequences were the following: DDITR F1, 5′-AGGAATTTTTAGTGATGGAGTT-3′; DDITR R1, 5′-AAAAACCCCTAATAATAAAATT-3′; ITR F2, 5′-AGGAATTTTTAGTGATGGAGTTGGTTAT-3′; and ITR R2, 5′-TTAACCACTCCCTCTCTACACACTCACTCA-3′.

PCR templates for AAV2-ITR were single-stranded rAAV viral genomes extracted from AAV8-ISce I.AO3 and AAV8-EF1α-nlslacZ viral particles used in the present study. The PCR templates used as controls for AAV-DDITR were a plasmid DNA carrying a 165-bp DDITR, pDDITRAAV-EF1α-hFAH.AO (35), and in-vitro-synthesized no-end double-stranded linear monomer AAV-EF1α-hF.IX genome.

RESULTS

Experimental approach. DNA repair pathways are redundant (1). When a cell is defective in one DNA repair pathway, up-regulation of an alternative pathway(s) may take place and compensate for the defect. This redundant nature of DNA repair machinery often precludes identification of cellular factors and characterization of their roles in a DNA repair pathway by a comparison between DNA repair protein-proficient and -deficient cells. To overcome this problem, we took an approach in which we saturated alternative DNA repair pathways that might compensate for a defective phenotype by introducing an excess of rAAV genomes that are, in fact, sensed as true DNA damage by cells (18). In the present study, we injected DNA repair protein-proficient or -deficient mice with rAAV serotype 8 (rAAV8) at several doses, including a very high dose. In one experiment, rAAV9 also was used. Satura-

neutral agarose gel Southern blot analysis with a β-lactamase probe. Wild-type C57BL/6J mice (B6) or DNA-PKcs-deficient mice [PK(−)] were injected with 7.2×10^{12} vg of AAV8-ISce I.AO3 via the tail vein ($n = 2$ mice each), and the liver tissues were harvested 6 weeks postinjection. rAAV genome copy numbers (BamHI/AlwNI digestion) and rAAV genome forms (PstI digestion) are shown with copy number standards (0 to 3,000 ds-vg/dge). Unusual bands that were not present in samples from wild-type mice were observed in PstI (noncutter) DNA digests in DNA-PKcs-deficient mouse livers (shown with arrows). The major unusual bands migrated slightly faster than a 3,049-bp marker, which is 20 bp shorter than full-length double-stranded linear AAV-ISce I.AO3 genomes (3,069 bp). Positions of large concatemeric rAAV genomes (Conc) and supercoiled circular monomer rAAV genomes (CM) are indicated with arrowheads. (C) Schematic explanation of how to discriminate between four different possible configurations of double-stranded linear monomer rAAV genomes by neutral and alkaline gel Southern blot analyses. Closed and open arrowheads indicate BamHI and AlwNI sites, respectively. Under neutral conditions, each linear monomer genome configuration digested with PstI, BamHI, or AlwNI would migrate as double-stranded DNA of the size indicated above each arrow, while under alkaline conditions it would migrate as single-stranded DNA of the size indicated below each arrow. The sizes of DNA fragments are calculated with the assumption that each rAAV genome terminus has a 145-base full-length ITR and each ITR hairpin cap at closed ends comprises two full-length ITRs (for more details, see Table 1). Only the DNA fragments detectable by the β-lactamase probe used in the study are shown. (D) Further characterization of the unusual monomer-length rAAV genomes that accumulated in DNA-PKcs-deficient mouse livers by neutral and alkaline gel Southern blot analyses. Numbers 1 to 16 are given to the left or upper left of each representative band on the blots, and the bands representing or derived from the unusual double-stranded linear monomer genomes, which were seen only in DNA-PKcs-deficient mice, are circled. Ten micrograms of total liver DNA ($n = 2$) was digested with the enzymes indicated above the lanes, separated on a 0.8% neutral agarose gel (upper panel) or a 1.0% alkaline agarose gel (lower panel), and subjected to Southern blot analysis. Samples were loaded in the same order for both blots. The result from PstI (non-rAAV genome cutter) digestion in the upper panel demonstrates that double-stranded linear monomer genomes (number 1 in the upper panel) are found only in DNA-PKcs-deficient mice and not in wild-type mice. The results from AlwNI digestion and BamHI digestion further confirmed that the bands (number 1 in the upper panel) are linear monomers, not circular genomes. The result from PstI digestion shown in the lower panel demonstrates that double-stranded linear monomer genomes in the DNA-PKcs-deficient mouse livers exhibited a dimer size (number 1 in the lower panel). AlwNI digestion and BamHI digestion excluded the possibility that the dimer-sized bands (number 1 in the lower panel) have a hairpin structure with an open end, because a majority of the bands migrated to head-to-head (number 8 in the lower panel) and tail-to-tail (number 11 in the lower panel) positions with AlwNI digestion and BamHI digestion, respectively. Based on these observations, we concluded that the unusual linear monomer genomes exhibit a no-end configuration (see panel C). For the expected size of each band, see Table 1. Very faint bands shown with arrows are the bands derived from linear genomes with open ends. Each band represents the following: 1, the unusual double-stranded (ds) linear monomer; 2, ds linear dimer; 3, ds linear trimer; 4 and 7, large concatemer; 5, supercoiled ds circular monomer; 6, relaxed ds circular monomer; 8, head-end; 9 and 12, head-tail; 10, head-head; 11, tail-end; 13, tail-tail; 14, 6,167-bp or 6,167-base dimer size marker; 15, 3,053-bp or 3,053-base monomer size marker; 16, 2,177-bp or 2,177-base size marker. (E) Southern blot analysis of Exo III-treated or Plasmid Safe-treated rAAV genomes from the liver. Total liver DNA (10 μg) was first digested with EcoRI or PstI and then with Exo III (Exo) or Plasmid Safe (PS), respectively. The digested DNA was analyzed by 0.8% neutral agarose gel Southern blot analysis along with naïve liver genomic DNA containing control linear (Lin) or supercoiled circular (Circ) double-stranded DNA. Lin, 2,943-bp linear DNA; Circ, a supercoiled circular plasmid containing the 2,943-bp fragment. After treatment with Exo III or Plasmid Safe, the supercoiled plasmid controls were treated with EcoRI to release the 2,943-bp fragment. The open arrowhead indicates the position of the unusual double-stranded linear monomer and dimer rAAV genomes. Positions of concatemers (Conc) and supercoiled circular monomers (CM) are indicated with closed arrowheads. For panels B and D, multiple lanes under the same experimental condition represent samples from different animals.

TABLE 1. Summary of the expected sizes of the bands in Southern blot analyses

Band	rAAV fragment size after digestion[a]			
	Neutral gel (bp)		Alkaline gel (bases)	
	BamHI	AlwNI	BamHI	AlwNI
Circular monomer				
Head-tail	3,069 (2,944)	2,571 (2,446)	3,069 (2,944)	2,571 (2,446)
Junction of linear and circular concatemer				
Head-tail	3,069 (2,944)	2,571 (2,446)	3,069 (2,944)	2,571 (2,446)
Head-head	No signal	4,446 (4,321)	No signal	4,446 (4,321)
Tail-tail	4,332 (4,207)	No signal	4,332 (4,207)	No signal
End of linear monomer and concatemer				
Head-end	No signal	2,223 (2,161)	No signal	4,446 (4,321)
Tail-end	2,166 (2,104)	No signal	4,332 (4,207)	No signal

[a] The sizes without parentheses represent those of rAAV genomes with no deletion at AAV-ITR junctions. The sizes in parentheses represent lengths of restricted fragments if AAV-ITR junctions or ends form a 165-bp DDITR. The probe used for the analysis is shown in Fig. 1A.

tion was possible with rAAV8 and rAAV9 because they could introduce a large amount of rAAV genomes into cells (16, 32). The advantage of our approach is that we could introduce rAAV genomes in most tissues in mice by a single noninvasive tail vein or intraperitoneal injection of these serotypes (16, 32), and therefore we could investigate DNA repair machinery in various tissues of interest in living animals.

Unusual double-stranded linear rAAV genomes accumulate in mouse liver in the absence of DNA-PKcs. We and others previously demonstrated that deficiency of DNA-PKcs activity did not significantly affect the processing of rAAV2 genomes in mouse liver, resulting in rAAV2 genome recombination via AAV-ITR hairpin termini (36, 49). This suggested that DNA-PKcs has no role in processing viral DNA hairpins in the liver. However, we were aware that uncoating of rAAV2 virions is a very slow process in mouse liver (51), taking several weeks, resulting in a relatively small viral genome load per cell in the studies with rAAV2 (36). In our previous study, we estimated that there was approximately ~0.1 rAAV vector genome (vg) per cell per day, which we assumed could be processed by DNA-PKcs-independent alternative pathways in the absence of DNA-PKcs (36).

To further investigate the role of DNA-PKcs in rAAV genome processing in the liver, we used rAAV8 in the present study because it can transduce the liver at very high levels, achieving 1,000 vg/cell or more in a relatively short period of time (less than a week) (32). This serotype primarily transduces hepatocytes in the liver (32). Wild-type C57BL/6J adult male mice and DNA-PKcs-deficient C57BL/6J SCID adult male mice were injected with AAV8-ISce I.AO3 at a dose of 7.2×10^{12} vg/mouse ($n = 2$) via the tail vein. Six weeks postinjection, the mice were sacrificed and various tissues, including liver, skeletal muscle, heart, and kidney, were collected for analyses. The processing of rAAV genomes in the liver was assessed by our standard genomic Southern blot analysis using total liver DNA samples. Double-stranded rAAV genome copy numbers in the livers were 2,291 ± 548 ds-vg/dge (double-stranded rAAV vector genome copy numbers per diploid genomic equivalent) for wild-type mice and 2,223 ± 101 ds-vg/dge for DNA-PKcs-deficient SCID mice (mean ± |mean-each value|; $n = 2$, where the vertical bars denote the absolute value of a number) (Fig. 1B), demonstrating that mouse hepatocytes were in fact exposed to a very high viral genome load. Southern blot analysis following digestion with PstI, which does not cut the rAAV genome, revealed the presence of double-stranded linear monomer-sized genomes only in DNA-PKcs-deficient SCID livers and not in wild-type mice (Fig. 1B). In DNA-PKcs-proficient mouse liver, rAAV genomes were processed into double-stranded circular monomers and concatemers, as we and others have previously reported (36, 49) (Fig. 1B). A substantial amount of double-stranded linear monomer-sized rAAV genomes in the liver is unusual and has not been demonstrated in any previous studies. The demonstration that there is a significant difference in rAAV genome processing in the presence and absence of DNA-PKcs activity in the liver at a high viral genome load verified that our approach was appropriate and provided evidence that DNA-PKcs had some roles in rAAV genome processing via AAV-ITR hairpin termini in mouse liver.

Excess viral DNA hairpin termini in cells remain closed in the absence of DNA-PKcs, forming no-end double-stranded linear genomes capped with covalently closed hairpins. (i) The unusual double-stranded linear rAAV genomes had covalently closed hairpin DNA termini exhibiting a no-end double-stranded linear DNA structure. To further characterize the unusual double-stranded linear monomer-sized rAAV genomes observed in DNA-PKcs-deficient SCID mouse liver, total liver DNA was digested with PstI, AlwNI, or BamHI and was separated on a 0.8% agarose gel under neutral conditions. The same digested DNA samples also were electrophoresed on a 1.0% agarose gel under alkaline conditions. AlwNI digests the AAV-ISce I.AO3 genome twice on the right side of the genome, while BamHI cuts the genome on the left side (Fig. 1A). The expected sizes of the bands from each possible form of rAAV genome in the neutral and alkaline gel Southern blot analyses are summarized in Fig. 1C and Table 1. The analyses revealed that (i) the unusual genomes were cut with restriction enzymes; therefore, they were double-stranded molecules (Fig.

1D); (ii) the unusual genomes digested with AlwNI or BamHI became shorter than 1 U of length and exhibited the sizes expected from double-stranded linear monomer genomes under neutral conditions; therefore, they were double-stranded linear monomer genomes (Fig. 1C and D); (iii) the unusual genome digested with PstI that does not cut the rAAV genome migrated at approximately a 2-U position under alkaline conditions; therefore, they are double-stranded linear monomer molecules with at least one end covalently closed (Fig. 1C and D); and (iv) a majority of the unusual genomes digested with AlwNI or BamHI migrated at a position corresponding to the head-to-head or tail-to-tail position in alkaline gel electrophoresis (Fig. 1C and D). These findings cannot be explained unless the unusual molecules were double-stranded linear monomers with both ends covalently closed, i.e., no-end double-stranded linear monomers (see Fig. 1C).

(ii) The unusual double-stranded linear rAAV genomes exhibit exonuclease insensitivity characteristic of no-end double-stranded linear DNA. To confirm the no-end structure of the unusual double-stranded linear monomer genomes observed in DNA-PKcs-deficient mouse livers, we exploited the catalytic activities of two different exonucleases, ATP-dependent exonucleases, Plasmid Safe and *E. coli* Exo III, to discriminate between no-end molecules and open-ended molecules. Plasmid Safe is a 5′-to-3′ and 3′-to-5′ exonuclease specific for linear double-stranded DNA with open ends and closed circular and linear single-stranded DNA, but it does not act at nicks in double-stranded DNA. Exo III is a 3′-to-5′ exonuclease that digests DNA from 3-hydroxil termini at blunt ends, recessed ends, and nicks in double-stranded DNA, but it is less active on single-stranded DNA. Therefore, no-end double-stranded linear molecules are resistant to both Plasmid Safe and Exo III digestions. Any nicks of DNA in double-stranded linear molecules with hairpin caps should result in substrate degradation with Exo III digestion, but this is not necessarily the case with Plasmid Safe. As shown in Fig. 1E, the double-stranded linear monomer rAAV genomes that accumulated in DNA-PKcs-deficient mouse livers were insensitive to both Plasmid Safe and Exo III, which was consistent with a no-end structure.

(iii) In-vitro-synthesized artificial no-end double-stranded linear rAAV genome exhibits exactly the same molecular characteristics as those for the unusual double-stranded linear monomer rAAV genomes. To further confirm the no-end structure of the unusual double-stranded linear monomer genomes, we synthesized dumbbell-shaped no-end double-stranded linear monomer rAAV genomes according to the method reported by Snyder et al. (47), and we characterized them in the same manner as that used for the unusual double-stranded linear monomer rAAV genomes in the actual mouse samples. For this, we synthesized no-end double-stranded linear monomer AAV-EF1α-hF.IX genome. As a control, we also synthesized double-stranded linear monomer AAV-EF1α-hF.IX genome without T4 DNA Lig. This control molecule is presumed to have exactly the same structure as that of the no-end genome, except for the presence of a nick at two sites. Figure 2A schematically shows the enzymatic synthesis of the two types of dumbbell-shaped molecules with and without nicks. The starting material was a PvuII-PvuII fragment of double-stranded linear AAV-EF1α-hF.IX with a 130-bp AAV2-ITR at each end (4,915 bp). Therefore, the size of synthesized dumbbell-shaped no-end double-stranded linear genome with a 165-base DDITR hairpin cap at each end is calculated as 4,820 bp [4,915 − (130 × 2) + 165 = 4,820], 95 bp shorter than the starting material.

The enzymatically synthesized no-end and nicked dumbbell-shaped genomes were treated with or without Plasmid Safe or Exo III, and then they were analyzed by neutral and alkaline gel Southern blot analyses (Fig. 2B, C, and D). The no-end genomes capped with 165-base DDITRs at both ends migrated slightly faster than the full-length double-stranded linear rAAV genomes on a neutral agarose gel under the conditions we used (Fig. 2B). Under alkaline conditions, the no-end molecule migrated at a nearly dimer position, while the nicked molecule stayed at a monomer position (Fig. 2C). As expected, the no-end genome was resistant to both Plasmid Safe and Exo III, while nicked dumbbell-shaped genome was resistant to Plasmid Safe but was sensitive to Exo III (Fig. 2B). The in-vitro-synthesized no-end genomes were further characterized following DraIII and BamHI digestion. DraIII and BamHI cut the double-stranded AAV-EF1α-hF.IX genome asymmetrically on the left and right sides, respectively (for the vector map, see Fig. 2E). Under the neutral condition, DraIII and BamHI digestion of the dumbbell-shaped no-end genomes resulted in tail-end and head-end molecules, respectively, smaller than the double-stranded linear monomer (Fig. 2D). Under the alkaline condition, the digested no-end molecules migrated at a tail-to-tail or head-to-head position (Fig. 2D).

All of these observations precisely mirror what we have observed with the unusual double-stranded linear monomer genome that emerged in DNA-PKcs-deficient mouse liver, which further confirmed the no-end structure of this unusual molecule. Because AAV-ITR hairpins became covalently closed and did not participate in rAAV genome recombination via AAV-ITR in the absence of DNA-PKcs, our results strongly indicate that DNA-PKcs-associated endonuclease activity opens AAV-ITR hairpin termini and triggers subsequent rAAV genome recombination and that if AAV-ITR hairpins are not cleaved, no-end double-stranded linear genomes capped with covalently closed DNA hairpins accumulate.

Diverse rAAV genomes should convert to no-end double-stranded linear DNA in the absence of DNA-PKcs. We next asked if the formation of no-end double-stranded linear rAAV genomes is a phenomenon unique to AAV-ISce I.AO3 or is a general phenomenon observed with other rAAV. We also wanted to know whether the conversion from single-stranded rAAV genomes to no-end double-stranded DNA molecules requires 6 weeks or whether they can be formed in a shorter period of time. To address these two questions, we intraperitoneally injected a DNA-PKcs-deficient C57BL/6J male mouse with AAV8-EF1α-nlslacZ at a dose of 7.2×10^{12} vg/mouse and sacrificed it 10 days after injection. Molecular forms of rAAV genomes in the liver were analyzed by Southern blot analysis. Accumulation of Exo III-resistant unusual double-stranded linear rAAV genomes were readily detectable, demonstrating the generality of the observations with AAV8-ISce I.AO3 and a short time requirement for the formation of no-end double-stranded linear rAAV genomes (data not shown).

DNA-PKcs-independent alternative pathways can effectively process AAV-ITR hairpins in the liver. Previous studies by us and others have shown no accumulation of double-stranded

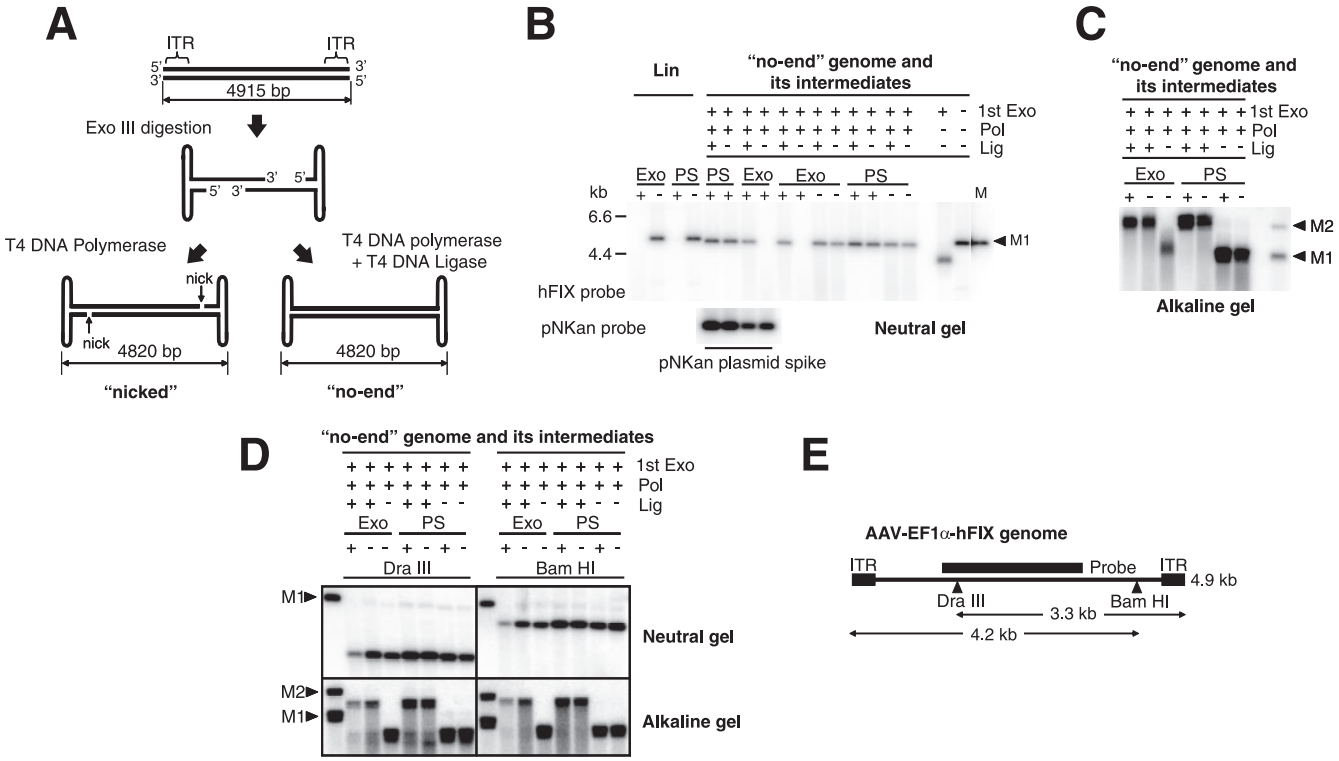

FIG. 2. Analyses of in-vitro-synthesized dumbbell-shaped no-end double-stranded linear monomer rAAV genomes. (A) Schematic representation of the method for synthesizing dumbbell-shaped nicked and no-end double-stranded linear rAAV genomes in vitro. The starting material is a 4,915-bp double-stranded PvuII-PvuII DNA fragment comprising AAV-EF1α-hF.IX genomes with a 130-bp AAV2-ITR at each genome terminus (for the vector map, see panel E). (B) Southern blot analysis of the in-vitro-synthesized dumbbell-shaped no-end rAAV genome and its intermediates. The top three rows of pluses and minuses indicate how each substrate (i.e., the no-end genome and its intermediates) was made, and the bottom row of pluses and minuses shows Exo III or Plasmid Safe (PS) treatment of the synthesized substrates. Ten nanograms of DNA of each product was separated on a 0.8% agarose gel and subjected to Southern blot analysis with an hF.IX cDNA probe. Some samples contained 400 ng of pNKan, a supercoiled circular plasmid containing the kanamycin resistance gene, and were treated in the same manner in order to verify that there was no irrelevant loss of sample DNA during the experimental procedures. Pol, T4 DNA Pol; M1, a 4,915-bp marker (M); Lin, a control double-stranded linear DNA, which is the same as M1 on this blot. 1st Exo indicates the Exo III treatment of the PvuII-PvuII AAV-EF1α-hF.IX fragment (see panel A). (C and D) Further characterization of the in-vitro-synthesized no-end genome and intermediates by neutral and alkaline agarose gel Southern blot analyses. Experimental conditions are indicated above each panel. (D) Pol-treated substrates and Pol- and Lig-treated substrates were treated with or without Exo III or Plasmid Safe and subsequently were treated with DraIII or BamHI. M2, 9,607-bp dimer marker. (E) Map of the AAV-EF1α-hF.IX vector genome.

linear rAAV genomes in DNA-PKcs-deficient SCID mouse livers following a portal vein injection of rAAV2. The viral genome load was ~4 ds-vg/dge in our previous study (36) and ~0.5 ds-vg/dge in a study by a different laboratory (49). With such low viral loads, rAAV genomes were efficiently processed via AAV-ITR by intra- and/or intermolecular recombination. This suggests that the DNA-PKcs-independent alternative pathways could process all AAV-ITR hairpin termini under such conditions. In the present study, we could achieve ~2,000 ds-vg/dge in the liver by rAAV8 injection. Thus, the present study provided an opportunity to determine the capacity of DNA-PKcs-independent alternative pathways that process AAV-ITR hairpins in mouse liver.

In this study, we injected wild-type C57BL/6J and DNA-PKcs-deficient C57BL/6J adult male mice with AAV8-ISce I.AO3 via the tail vein or with AAV2-ISce I.AO3 via the portal vein at various doses ranging from 8.3×10^9 to 7.2×10^{12} vg/mouse, as summarized in Table 2. Various tissues were harvested 6 weeks postinjection. Total liver DNA was extracted from the samples, and copy numbers of no-end double-stranded linear rAAV genomes, as well as total genome copy numbers in each sample, were determined by Southern blot analysis. For the quantification of no-end rAAV genomes, we digested DNA with PstI, which does not cut the rAAV genome, electrophoresed the DNA on an agarose gel, blotted it, and probed it with an rAAV genome-specific β-lactamase probe. The result clearly demonstrated viral-load-dependent emergence of no-end double-stranded linear rAAV genomes in the liver with a threshold of at least 62 ds-vg/dge (Table 2). The no-end genomes were readily detectable in the livers transduced with rAAV8 at doses of 1.8×10^{12} vg/mouse or higher but not in the livers transduced with rAAV8 at doses of 3.0×10^{11} vg/mouse or lower (Fig. 3 and Table 2). No no-end genomes were detected in any of the liver samples transduced with rAAV2, including those transduced with rAAV2 at 7.2×10^{12} vg/mouse, which had a viral load of only 30 ds-vg/dge in a wild-type mouse and 33 ± 2 ds-vg/dge in DNA-PKcs-deficient mice (mean ± |mean-each value|; $n = 2$), even though they were given a very high rAAV dose. These observations demonstrate that in the liver, AAV-ITR hairpin-opening pathways

TABLE 2. Summary of an rAAV8 dose-response study (rAAV serotypes, injected doses, number of animals in each group, and rAAV genome copy numbers in the liver)

Mouse type	rAAV serotype	Dose (10^{10})	No. of mice	Double-stranded genome copy numbers[a]			(%) No-end genome[d]
				Total	No-end linear genomes		
					Monomer	Dimer	
PK (−)[b]	rAAV8	0.83	6	1.3 ± 0.2	Not detected	Not detected	0
PK (−)	rAAV8	5	5	8.0 ± 3.5	Not detected	Not detected	0
PK (−)	rAAV8	30	5	62 ± 17	Not detected	Not detected	0
PK (−)	rAAV8	180	3	728 ± 76	107 ± 30	29 ± 25	15 ± 4
PK (−)	rAAV8	720	2	2,223 ± 101	863 ± 10	208 ± 33	39 ± 1
B6[c]	rAAV8	5	4	11 ± 4	Not detected	Not detected	0
B6	rAAV8	720	2	2,291 ± 548	Not detected	Not detected	0
PK (−)	rAAV2	30	4	7.4 ± 3.6	Not detected	Not detected	0
PK (−)	rAAV2	720	2	33 ± 2	Not detected	Not detected	0
B6	rAAV2	30	3	4.7 ± 2.1	Not detected	Not detected	0
B6	rAAV2	720	1	30	Not detected	Not detected	0

[a] Double-stranded genome copy numbers per diploid genomic equivalent (ds − vg/dge).
[b] DNA-PKcs-deficient C57BL/6J mice.
[c] Wild-type C57BL/6J mice.
[d] Percentage of no-end linear monomers and dimers among the total number of double-stranded rAAV genomes.

are redundant and DNA-PKcs-independent alternative pathways have a considerable capacity for DNA hairpin processing.

The DNA-PKcs-dependent pathway plays a significant and irreplaceable role in AAV-ITR hairpin opening in nonhepatic tissues. Our previous studies have demonstrated that systemic administration of rAAV8 at 7.2×10^{12} vg/mouse results in substantial transduction in various nonhepatic tissues at levels of 2 to 20 ds-vg/dge (32). Although the viral load in nonhepatic tissues was lower than that in the liver by 2 to 3 logs, it provided an opportunity to investigate DNA-PKcs-dependent and -independent pathways for DNA hairpin opening in various nonhepatic tissues. In the present study, we focused on three nonhepatic tissues: heart, skeletal muscle, and kidney. We obtained total DNA from these tissues, which were collected from wild-type and DNA-PKcs-deficient SCID mice injected with AAV8-ISce I.AO3 at 7.2×10^{12} vg/mouse. We determined molecular forms of rAAV genomes in each tissue by Southern blot analysis in the same way as that used for the liver samples. In all three tissues, we found that unusual double-stranded linear genomes accumulated only in DNA-PKcs-deficient SCID mouse tissues but not in DNA-PKcs-proficient tissues (Fig. 4A). The no-end nature of the unusual double-stranded linear genomes was confirmed by Exo III digestion followed by Southern blot analysis (Fig. 4B). It should be noted that in all three tissues, double-stranded circular monomers and concatemers were observed together with the no-end linear genomes, indicating the presence of DNA-PKcs-independent alternative hairpin-processing pathways in these tissues.

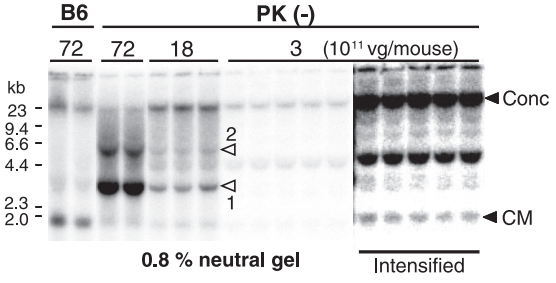

FIG. 3. rAAV dose-response study of the no-end double-stranded linear rAAV genomes in DNA-PKcs-deficient mouse livers transduced with AAV8-ISce I.AO3. Wild-type (B6) and DNA-PKcs-deficient [PK(−)] C57BL/6J mice were injected with AAV8-ISce I.AO3 via the tail vein at various doses as indicated above the lanes. rAAV genome forms in the livers were analyzed 6 weeks postinjection by Southern blot analysis with PstI (non-rAAV genome cutter) digestion and were probed with a β-lactamase probe. For the five right-most lanes, the same blot was intensified. Open arrowheads indicate no-end monomer (1) and dimer (2) rAAV genomes. Positions of concatemers (Conc) and supercoiled circular monomers (CM) are indicated with closed arrowheads. Multiple lanes under the same experimental conditions represent samples from different animals.

To assess the degree of dependency of DNA hairpin processing on the DNA-PKcs-independent alternative pathways in nonhepatic tissues, we analyzed tissues transduced with lower rAAV doses by Southern blotting. The no-end double-stranded linear genomes were observed when the viral genome load was as low as 2.2 to 6.5 ds-vg/dge (Fig. 4C). This demonstrates that in the heart, muscle, and kidney, although DNA-hairpin-opening pathways are redundant, the DNA-PKcs-dependent pathway is a significant pathway, and unlike in the liver, this pathway cannot be replaced with alternative pathways.

The role of Artemis faithfully mirrors that of DNA-PKcs in DNA hairpin metabolism in mice. DNA-PKcs by itself does not possess endonuclease activity. Artemis was recently proven to be a DNA repair endonuclease that is activated by DNA-PKcs; therefore, we hypothesized that it is Artemis endonuclease activity that cleaves AAV-ITR hairpin termini in concert with DNA-PKcs. To test this hypothesis, we injected Artemis-proficient ($n = 3$) and -deficient ($n = 3$) adult 129 male mice with AAV8-ISce I.AO3 at three different doses, 3.0×10^{11}, 1.8×10^{12}, and 7.2×10^{12} vg/mouse ($n = 1$ mouse per dose). Ten days postinjection, various tissues were harvested for molecular analyses.

To investigate molecular forms of rAAV genomes in the liver, total liver DNA was digested with PstI (noncutter), BamHI (left-side cutter), or AlwNI (right-side cutter) and

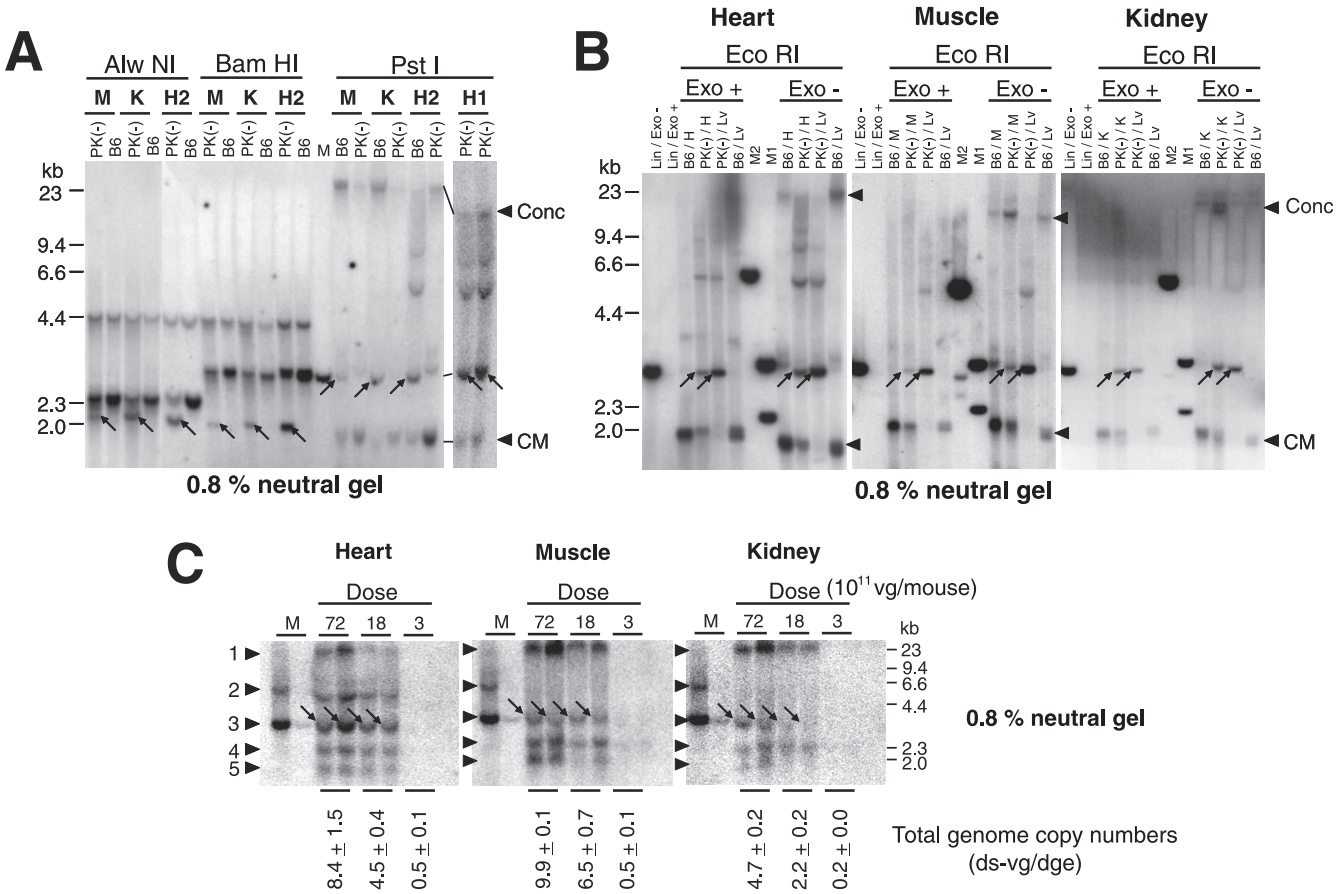

FIG. 4. Southern blot analyses of rAAV genomes in nonhepatic tissues of wild-type and DNA-PKcs-deficient mice injected with AAV8-ISce I.AO3. (A) rAAV genome forms in the muscle (M), kidney (K), and heart (H1 and H2) in wild-type (B6) and DNA-PKcs-deficient [PK(−)] mice injected with 7.2×10^{12} vg of AAV8-ISce I.AO3 via the tail vein. Total DNA was extracted from each tissue 6 weeks postinjection (for M, K, and H1) or 6 months postinjection (for H2). Sample H2 was prepared due to the insufficiency of sample H1. Ten micrograms of total DNA was digested with the restriction enzyme indicated above the lanes, was separated on a 0.8% neutral agarose gel, and was probed with a β-lactamase probe. The bands indicated with arrows represent those derived from no-end linear monomer rAAV genomes detected only in the absence of DNA-PKcs. Positions of concatemers (Conc) and supercoiled circular monomers (CM) are indicated with arrowheads. (B) Further characterization of the unusual monomer-length rAAV genomes found in the DNA-PKcs-deficient mouse tissues by Exo III digestion. Ten micrograms of tissue DNA first was digested with EcoRI followed by Exo III treatment, and then it was subjected to neutral agarose gel Southern blot analysis. As controls, an appropriate amount of liver DNA (Lv) from wild-type or DNA-PKcs-deficient mice injected with 7.2×10^{12} vg of AAV8-ISce I.AO3 also was processed in the same manner for each blot (0.05, 0.05, and 0.025 μg of liver DNA for the heart, muscle, and kidney blots, respectively). Arrows indicate no-end double-stranded linear monomer rAAV genome. Positions of concatemers (Conc) and supercoiled circular monomers (CM) are indicated with arrowheads. M1, 3,053- and 2,177-bp markers; M2, 6,167-bp marker; Lin, a control double-stranded linear DNA. The results demonstrate that the unusual linear monomer rAAV genomes were Exo III resistant. (C) An rAAV dose-response study of the no-end double-stranded linear rAAV genomes in DNA-PKcs-deficient mouse heart, muscle, and kidney transduced with AAV8-ISce I.AO3. Ten micrograms of tissue DNA was digested with PstI, was separated on a 0.8% neutral agarose gel, and was subjected to Southern blot analysis. Doses of rAAV are indicated above the lanes. No-end double-stranded linear monomer genome bands are indicated with arrows. Total double-stranded rAAV genome copy numbers determined by a separate Southern blot analysis are indicated below the lanes (means ± |mean-each value|; $n = 2$). Each arrowhead indicates the position of large concatemer (1), double-stranded linear dimer (2), double-stranded linear monomer (3), supercoiled double-stranded circular monomer (4), and single-stranded rAAV genome (5). M, total liver DNA (0.05 and 0.005 μg for DNA in the left and right lanes, respectively) extracted from the DNA-PKcs-deficient mouse liver transduced with 7.2×10^{12} vg of AAV8-ISce I.AO3. Multiple lanes under the same experimental condition represent samples from different animals.

was subjected to Southern blot analysis in the same way as that described above. In our preliminary analysis, we found that some rAAV genomes in this particular set of samples still remained as single-stranded DNA genomes, making it somewhat difficult to interpret the data. When samples were harvested 10 days postinjection, a considerable portion of rAAV particles remained coated in rAAV-transduced tissues (51). For this reason, we performed additional Southern blot analysis in which sample DNA was treated with Plasmid Safe to remove single-stranded rAAV genomes and artificial double-stranded genomes generated in test tubes by plus- and minus-strand annealing. The experimental results were exactly the same as what we had observed in the DNA-PKcs-deficient mouse experiments. Total rAAV genome copy numbers in the liver determined with Plasmid Safe-treated DNA were 37 ds-vg/dge (3.0×10^{11} vg/mouse), 568 ds-vg/dge (1.8×10^{12} vg/mouse), and 1,029 ds-vg/dge (7.2×10^{12} vg/mouse) in Artemis-deficient mice, and they

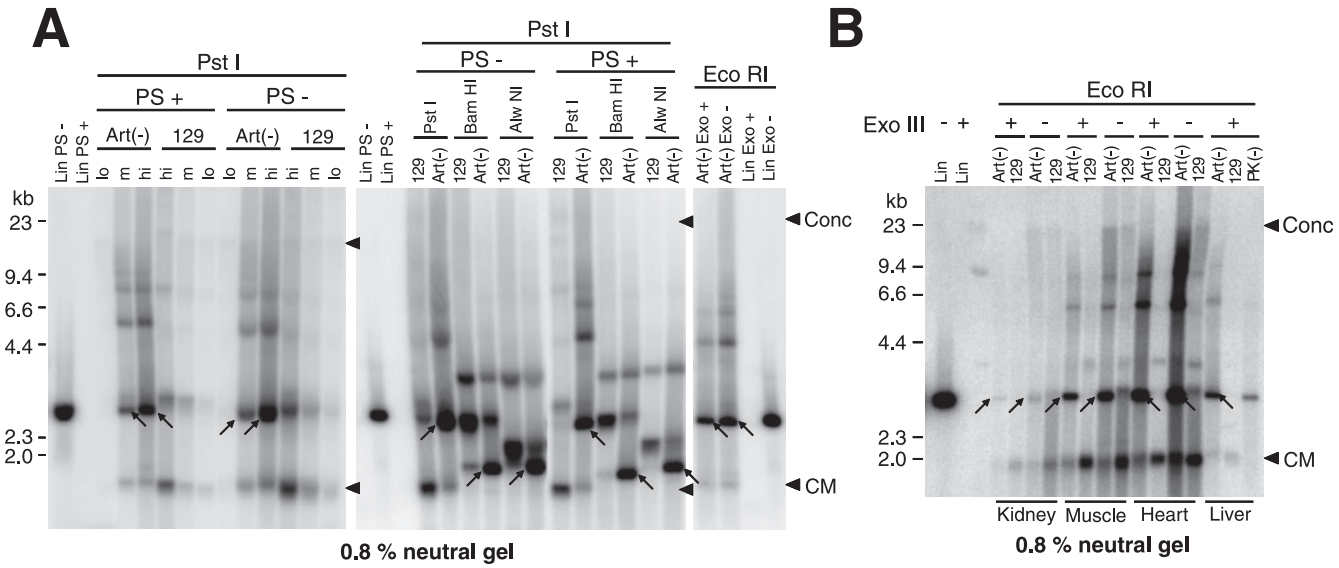

FIG. 5. Southern blot analyses of rAAV genomes in wild-type and Artemis-deficient 129 mouse tissues transduced with AAV8-ISce I.AO3. (A) rAAV genome forms in the livers of wild-type (129) and Artemis-deficient [Art(−)] mice injected with 3.0×10^{11} vg (lo), 1.8×10^{12} vg (m), or 7.2×10^{12} vg (hi) of AAV8-ISce I.AO3 via the tail vein. Total DNA was prepared from tissue samples 10 days postinjection. A Plasmid Safe digestion step was incorporated into the experimental procedure to remove single-stranded rAAV genomes and double-stranded genomes that may emerge in test tubes due to plus- and minus-strand annealing (see the text). Otherwise, the experimental procedure for Southern blot analysis was essentially the same as that for Fig. 1 and 4. For the left panel, sample DNA was digested with PstI first, and then it was treated with or without Plasmid Safe (PS). For the middle panel, sample DNA was further digested with PstI, BamHI, or AlwNI. For the right panel, sample DNA was first digested with EcoRI, and then it was treated with or without Exo III. For the middle and right panels, only the DNA samples prepared from the mice injected with 7.2×10^{12} vg of rAAV were used. The absence of double-stranded linear monomer genomes and the presence of a considerable amount of circular monomers and concatemers in the PstI- and PS-treated Art(−) lo sample have been confirmed in a separate Southern blot analysis (data not shown). The bands with arrows represent those derived from no-end linear monomer genomes. (B) The demonstration of no-end linear rAAV genomes (indicated with arrows) in nonhepatic tissues of Artemis-deficient mice, but not of wild-type mice, by Southern blot analysis. The experimental procedure was the same as that for Fig. 4B. Ten micrograms of total tissue DNA per lane was used, except for the liver, for which 0.05 μg per lane was used. Conc and CM in each panel indicate the positions of concatemers and supercoiled circular monomers, respectively. Lin, a 2,943-bp double-stranded linear DNA fragment.

were 69 ds-vg/dge (3.0×10^{11} vg/mouse), 293 ds-vg/dge (1.8×10^{12} vg/mouse), and 890 ds-vg/dge (7.2×10^{12} vg/mouse) in Artemis-proficient control mice. Unusual double-stranded linear rAAV genomes emerged in the Artemis-deficient mouse livers transduced with $\geq 1.8 \times 10^{12}$ vg/mouse (viral genome load, ≥568 ds-vg/dge) but did not emerge in any of the control mouse livers (Fig. 5A). The unusual double-stranded linear genomes were Exo III resistant, demonstrating the no-end structure (Fig. 5A). The no-end double-stranded linear genomes were observed in Artemis-deficient mouse heart, skeletal muscle, and kidney at a viral genome load of 46, 27, and 19 ds-vg/dge, respectively (Fig. 5B), which was much lower than that for the liver. We also injected an Artemis-deficient mouse with AAV9-EF1α-nlslacZ at a dose of 3.6×10^{12} vg/mouse via the tail vein and sacrificed it 2 weeks postinjection. Exo III-resistant no-end double-stranded linear rAAV genomes were readily detected in the liver (data not shown), demonstrating that this phenomenon was not exclusive for AAV8-ISce I.AO3, as has been demonstrated with DNA-PKcs-deficient mice.

Taken together, the observation that DNA-PKcs deficiency and Artemis deficiency result in the same phenotype in the context of hairpin DNA opening strongly indicates that DNA-PKcs and Artemis are on the same DNA repair pathway toward opening AAV-ITR hairpins. This lends strong support to our hypothesis that the Artemis/DNA-PKcs complex plays a crucial role in AAV-ITR hairpin opening.

No-end double-stranded linear genomes had 165-base or 143-base covalently closed AAV-DDITR hairpin caps: evidence of no hairpin opening and involvement of a resolvase activity in the formation of the no-end genomes. (i) Establishment of a methodology for sequencing AAV-ITR hairpin termini. To determine the terminal structure of no-end double-stranded linear monomer rAAV genomes that accumulated in the absence of DNA-PKcs or Artemis, we established a methodology by which AAV-ITR hairpin cap could be amplified by PCR, cloned into a plasmid vector, and sequenced. It has been very challenging to PCR amplify, clone, and sequence AAV-ITR due to a G+C-rich hairpin structure. To overcome this problem, we exploited a hairpin-bisulfite PCR technique that originally was designed for assessing CpG methylation patterns on both cDNA strands by one PCR (20). With this method, cytosine (C) residues convert to uracil (U) residues by chemical modification; therefore, the G+C richness of AAV-ITR can be significantly reduced and formation of DNA secondary structures can be minimized.

We first applied this technique to sequence a 165-bp double-stranded AAV-DDITR in a plasmid template, pDDITRAAV-EF1α-hFAH.AO (35). Bisulfite-PCR amplification of this tem-

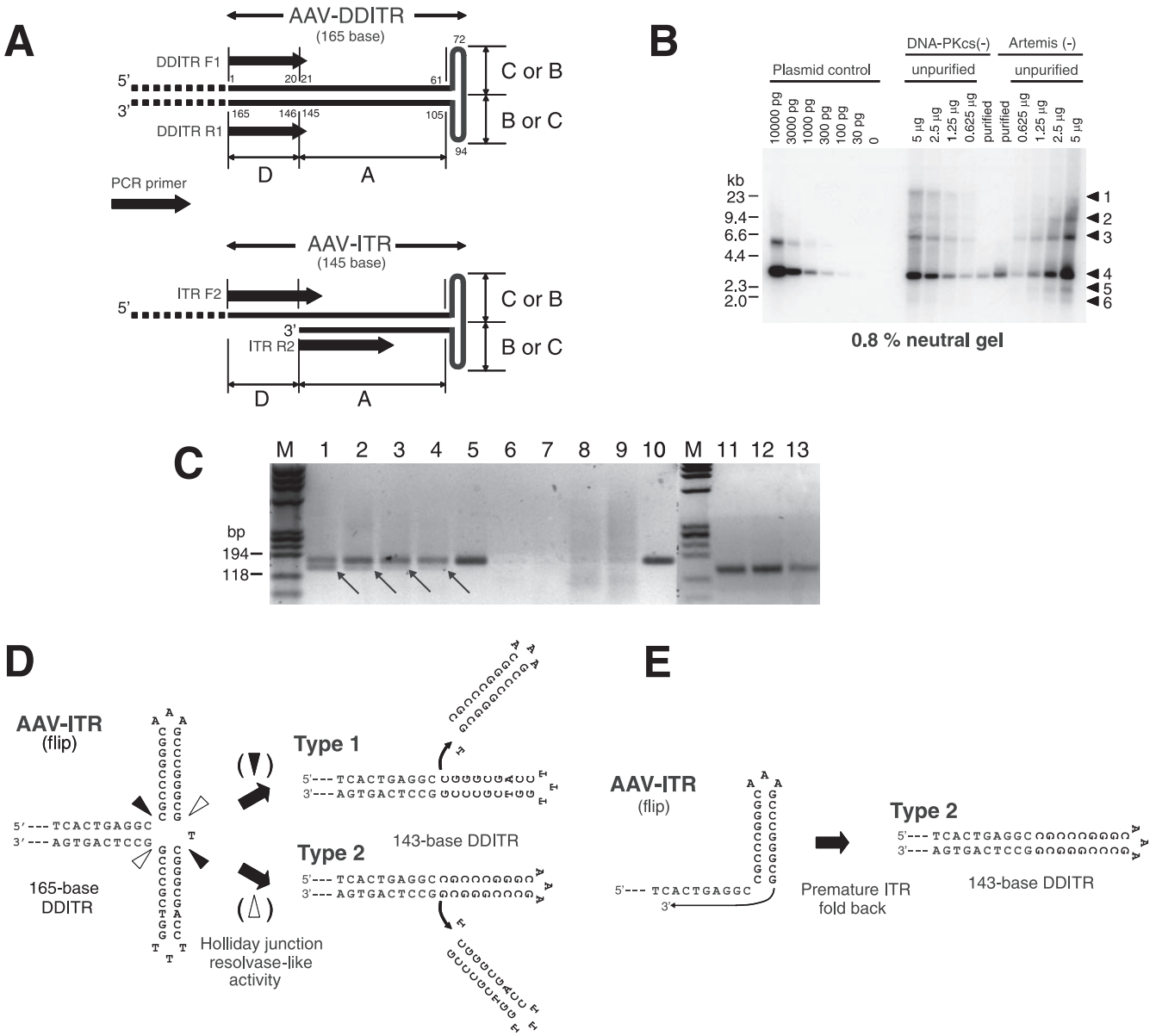

FIG. 6. Bisulfite PCR analyses of no-end rAAV genome terminal structures. (A) Location of bisulfate PCR primers. Hairpin ends of no-end rAAV genome (upper panel, 165-base DDITR) and single-stranded rAAV genome (lower panel, 145-base ITR) are shown. The letters A to D are ITR subregions. Numbers (1 to 165) indicate nucleotide positions in the ITRs. (B) Southern blot analysis of purified no-end double-stranded linear AAV8-ISce I.AO3 genomes recovered from DNA-PKcs-deficient [DNA-PKcs(−)] or Artemis-deficient [Artemis(−)] mouse livers. A portion of each purified DNA was separated along with 0.625 to 5 μg of unpurified liver DNA on a 0.8% neutral agarose gel and was subjected to Southern blot analysis with a β-lactamase probe. Purified no-end double-stranded linear monomer rAAV genomes (band position 4) are free from large concatemers (1), linear trimers (2), linear dimers (3), supercoiled circular monomers (5), and single-stranded genomes (6). (C) Bisulfite PCR products of AAV-ITR sequences. Chemically modified ITR sequences were PCR amplified from various templates with a primer set of DDITR F1 and DDITR R1 (lanes 1 to 10) or a set of ITR F2 and ITR R2 (lanes 11 to 13). The resulting PCR products were separated on a 2.5% NuSieve GTG agarose gel and stained with ethidium bromide. In lanes 1 to 4, a second band with various intensities (indicated with arrows) was observed in addition to the major band of the expected size. The bisulfite PCR templates are as follows: lanes 1 and 2, purified no-end linear AAV-ISce I.AO3 monomer genomes formed in DNA-PKcs-deficient (lane 1) and Artemis-deficient (lane 2) mice; lanes 3 and 4, purified no-end linear AAV-EF1α-nlslacZ monomer genomes formed in DNA-PKcs-deficient (lane 3) and Artemis-deficient (lane 4) mice; lanes 5 and 13, pDDITRAAV-EF1α-hFAH.AO plasmid carrying a 165-bp DDITR; lane 6, a control gel piece taken and processed as a no-end genome purification (see Materials and Methods); lane 7, no-DNA control; lanes 8 and 9, chemically modified total liver DNA from a C57BL/6 (lane 8) or 129 (lane 9) mouse transduced with 7.2×10^{12} vg of AAV8-ISce I.AO3; lane 10, in-vitro-synthesized no-end AAV-EF1α-hF.IX genome; lane 11, single-stranded AAV8-ISce I.AO3 viral genome DNA; lane 12, single-stranded AAV8-EF1α-nlslacZ viral genome DNA. M, HaeIII-digested ΦX174 DNA. (D and E) Two possible mechanisms for the formation of the 143-base AAV-DDITR: a model involving a Holliday junction resolvase-like activity (D) and a premature ITR fold-back model (E). In the first model (D), one arm of the T-shaped hairpin structure of AAV-ITR is cleaved at symmetrically related positions, as shown with open or closed arrowheads, by Holliday junction resolvase-like activity. This results in the removal of a 22-base DNA hairpin containing the B or C subregion of the ITR, forming two types of simple hairpin structures (types 1 and 2). In the second model (E), the leading strand folds back when AAV-ITR is half copied, skipping one arm of the T-shaped hairpin structure. The ITR shown in this figure is a flip-oriented ITR.

TABLE 3. Summary of the terminal structures of no-end rAAV genomes

Mouse type	rAAV	Dose (10^{12})[a]	Route[b]	No. (%) of no-end rAAV terminal structures determined					
				Total	No-end rAAV terminal structure				
					165-base DDITR	143-base DDITR			Others
						Total	Type 1	Type 2	
DNA-PKcs deficient	AAV8-ISce I.AO3	7.2	TV	22 (100)	8 (36)	12 (55)	11	1	2 (9)
Artemis deficient	AAV8-ISce I.AO3	7.2	TV	18 (100)	8 (44)	8 (44)	6	2	2 (11)
DNA-PKcs deficient	AAV8-EF1α-nlslacZ	7.2	IP	14 (100)	9 (64)	4 (29)	4	0	1 (7)
Artemis deficient	AAV9-EF1α-nlslacZ	3.6	TV	22 (100)	11 (50)	10 (45)	6	4	1 (5)

[a] rAAV vector genomes/mouse.
[b] TV, tail vein injection; IP, intraperitoneal injection.

plate with a set of primers that amplify the 165-bp DDITR (primers DDITR F1 and DDITR R1; see Fig. 6A) produced a single band of approximately 165 bp in length on an ethidium bromide-stained agarose gel (Fig. 6C, lane 5). PCR products then were cloned into pCR2.1-TOPO by TA TOPO cloning, and a PCR product plasmid library was generated on LB agar plates with X-Gal. Plasmid DNA isolated from 23 randomly picked-up white colonies was digested with EcoRI, and the insert was assessed by agarose gel electrophoresis. All the clones had an insert of the expected size. The sequence of the inserts from five clones was determined, showing an intact 165-bp DDITR. We also amplified AAV-DDITR hairpin termini of in-vitro-synthesized dumbbell-shaped no-end double-stranded linear monomer AAV-EF1α-hF.IX genome by this method and obtained a single band of the expected size (Fig. 6C, lane 10). Thus, our bisulfite PCR method could successfully amplify, clone, and sequence the 165-bp DDITR.

(ii) No-end double-stranded linear monomer rAAV genomes had 143- or 165-base AAV-DDITR hairpin caps. With the bisulfite PCR sequencing method described above, we investigated the structure of DNA hairpin caps of no-end double-stranded linear monomer rAAV genomes that emerged in DNA-PK-deficient or Artemis-deficient mouse livers transduced with AAV8-ISce I.AO3, AAV8-EF1α-nlslacZ, or AAV9-EF1α-nlslacZ. We purified the no-end monomer genomes free from double-stranded circular monomers and concatemers by Exo III digestion and gel fractionation as detailed in Materials and Methods (Fig. 6B). Using the purified no-end molecules as PCR templates, AAV-ITR hairpins at no-end genome termini were amplified by bisulfite PCR as described above. The PCR amplification yielded two products that differed by approximately 20 bp (Fig. 6C, lanes 1 to 4). The longer PCR products showed the same size as the 165-bp DDITR PCR product control (Fig. 6C, lane 5) and were, in general, more abundant than shorter products. The relative abundance of the longer bands compared to that of the shorter bands in the same sample was consistent in repeated PCR amplifications (data not shown) but varied among the samples analyzed (Fig. 6C, lanes 1 to 4). The PCR products were TA TOPO cloned, and PCR product libraries were constructed on LB agar plates. We determined the sequences of randomly selected clones from each library. The results showed that viral hairpin termini of the no-end genomes were exclusively either 165- or 143-base AAV-DDITR (Table 3). This was the case for both DNA-PK-deficient mice and Artemis-deficient mice and for both AAV-ISce I.AO3 and AAV8- or AAV9-EF1α-nlslacZ genomes.

(iii) Structure of the 143-base DDITR: an indication of the involvement of a Holliday junction resolvase-like activity. All 143-base DDITRs had a 22-base deletion of either arm of the T-shaped palindromic region corresponding to the AAV-ITR B or C subregion (Fig. 6D and E). Here, we propose two possible mechanisms for the formation of the 143-base DDITR: AAV-ITR hairpin cleavage by an undefined Holliday junction resolvase-like activity (Fig. 6D) and a premature ITR folding-back mechanism (Fig. 6E). In the first model, a resolvase activity cleaves AAV-ITR hairpins precisely at symmetrically related positions at the shoulder of T-shaped ITRs, resulting in a loss of one arm of T. The nature of perfect symmetry of nicking of two DNA strands allows sealing the nicks without the need for further processing or opening ITR hairpins, resulting in the formation of the covalently closed 143-base DDITR. In this model, two types of 143-base DDITR configurations, type 1 and type 2, would be possible depending on how the resolvase cleaves (Fig. 6D). In the second model, in which second-strand synthesis is involved, ITR folding back occurs with the 9-base ITR B or C subregion before the leading strand copies a full-length ITR. In this model, unlike the first model, only the type 2 configuration would be possible. To differentiate these two mechanisms, we took advantage of the bisulfite PCR sequencing method that can determine two complementary strands in a hairpin structure individually (20). In our strategy, 3-base loops (AAA or TTT) at the tips served as a signature to differentiate the two types. This signature functioned for both flip- and flop-oriented AAV-ITRs. As summarized in Table 3, both type 1 and type 2 configurations were detected. Interestingly, the type 1 configuration predominated (21 out of 27 [78%]) over the type 2 configuration. This lends strong support to the first model, in which a Holliday junction resolvase-like activity cleaves AAV-ITR, although the second model cannot be totally dismissed. In addition, this observation indicated that the undefined resolvase activity cleaves AAV-ITR hairpins more preferentially at the 5′ end of the T-arm hairpins than at the 3′ end. Although the exact mechanism for the 22-base deletion remains unclear, neither model should result in the formation of open ends. Therefore, the structures of DNA hairpin termini of no-end double-stranded linear rAAV genomes were quite consistent with our hypothesis that a subset of viral DNA hairpin termini remains closed in the

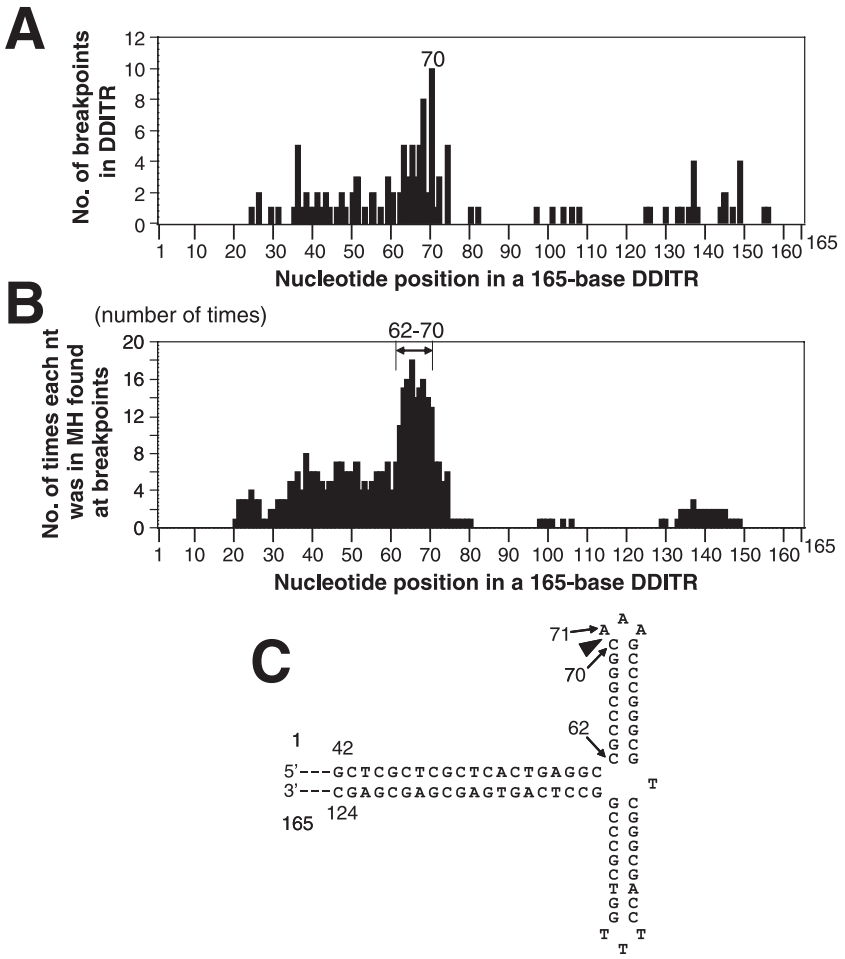

FIG. 7. Breakpoints of AAV-ITR in the wild-type mouse livers transduced with a very high dose of rAAV8. (A) The distribution of 120 breakpoints within the 165-base AAV-DDITR identified in 60 ITR-ITR recombination junctions in the wild-type mouse livers transduced with 7.2×10^{12} vg of AAV8-ISce I.AO3. Fifteen junctions that exhibited 165-bp DDITR are not included. The histogram collating the ITR breakpoints showed a clustering with a peak corresponding to nucleotide number 70. (B) Location of microhomology (MH) found at ITR-ITR junctions. Among 60 ITR-ITR junctions sequenced, 36 showed microhomology. The histogram gives the number of times each nucleotide was found within 36 microhomology regions at ITR-ITR junctions. Nucleotides 62 to 70 frequently are involved in ITR-ITR recombinations with microhomology. (C) Schematic representation of the hot spot of ITR breakage (between nucleotides 70 and 71, shown with an arrowhead) and the region frequently contributing to microhomology at ITR recombination (nucleotides 62 to 70).

absence of DNA-PKcs or Artemis, generating no-end linear rAAV genomes.

(iv) rAAV used for the present study did have intact 145-base AAV-ITR hairpin termini. In order to determine whether the 22-base deletion in AAV-ITR occurred in mice or whether the deletion originally was present in the rAAV preparations injected into mice, it was essential to determine the single-stranded rAAV viral genome terminal structure. To this end, single-stranded rAAV genomes were isolated from AAV8-ISce I.AO3 and AAV8-EF1α-nlslacZ virions used for the study. Full-length AAV-ITR was PCR amplified with a set of primers that amplifies 145-base ITRs (primers ITR F2 and ITR R2; see Fig. 6A). This PCR yielded a single PCR product of the expected size from both rAAVs and a control plasmid carrying a 165-bp DDITR (Fig. 6C, lanes 11 to 13). This provides evidence that the 22-base deletion occurred in mice. The PCR products were TA TOPO cloned, and PCR product plasmid libraries were constructed as described above. EcoRI digestion of plasmid DNA isolated from randomly selected white colonies in each library revealed that all 56 plasmid clones carrying an AAV-ITR derived from the AAV8-EF1α-nlslacZ genome, and 65 of 66 plasmids with an AAV8-ISce I.AO3-derived ITR, had an AAV-ITR PCR product of the expected size. We sequenced seven plasmid clones with a PCR product of the expected size, and all showed a 145-base intact AAV-ITR. One clone had a 22-base ITR deletion of type 2, suggesting the possibility that the undefined resolvase activity might cleave a very minor portion of ITR hairpins in rAAV vector production in human embryonic kidney 293 cells.

DNA-PKcs and Artemis-associated endonuclease activity cleaves at the tips of DNA hairpins. Finally, we addressed the question of how DNA-PKcs- and Artemis-associated DNA repair endonuclease activity might open AAV-ITR hairpin termini. At a very high rAAV8 dose (7.2×10^{12} vg/mouse), it was the DNA-PKcs- and Artemis-associated endonuclease activity that opened a majority of rAAV hairpin termini in wild-

type mouse livers. This is because at the high dose, a substantial amount of no-end linear rAAV genomes accumulated in the liver in the absence of either factor, but this was not the case for wild-type mice. Therefore, the structures of viral DNA hairpin termini in wild-type mouse livers transduced with 7.2×10^{12} vg/mouse should reflect the consequences of DNA hairpin cleavage by DNA-PKcs or Artemis. It should be noted that the same approach may not be applicable to other wild-type mouse tissues in which, if there were no DNA-PKcs or Artemis, no-end linear rAAV genomes would not be a major fraction of rAAV genomes or would not be present at all. This is because interpretation of the results may be confounding due to possible involvement of alternative pathways. In the study, total DNA isolated from the DNA-PKcs-proficient C57BL/6J and Artemis-proficient 129 mouse livers transduced with 7.2×10^{12} vg of AAV8-ISce I.AO3 were chemically modified by bisulfite. rAAV-rAAV junction sequences were amplified with the same primers as those used for DDITR amplification (primers DDITR F1 and DDITR R1; see Fig. 6A). The PCR products showed smearing on a gel (Fig. 6C, lanes 8 and 9). Plasmid libraries of the PCR products were constructed by TA TOPO cloning, and the sequences of randomly picked up clones were determined. Since the results obtained from C57BL/6J and 129 mice were essentially the same, we combined the results for the analysis. Fifteen of 75 junctions exhibited a 165-base DDITR. Because these samples did not contain no-end rAAV genomes detectable by Southern blot analysis, it is likely that the 165-base DDITR was, in fact, 165-bp double-stranded DDITR resulting from homologous recombination between two AAV-ITRs (9, 54, 55). The other 60 junctions showed various joining patterns of two AAV-ITRs with occasional microhomology at junctions. Importantly, when we mapped the breakpoints and microhomology regions, it became clear that it was the 5′ end of the 3-base AAA loop of AAV-ITR hairpin tips that DNA-PKcs and Artemis endonuclease activity primarily cleaved (Fig. 7). These observations strongly indicate that Artemis/DNA-PKcs endonuclease cleaves AAV-ITR hairpin tips, and various degrees of rAAV genome terminal deletions follow when ITRs are recombined.

No-end linear forms of rAAV genomes formed in quiescent cells completely disappear by passing through one or two cell divisions. Cellular pathways for repairing damaged DNA may change depending on the phase of the cell cycle that the cell is in. In the present study, parenchymal cells in the tissues studied are considered quiescent in general. To investigate the fate of no-end double-stranded linear rAAV genomes formed in DNA-PKcs-deficient SCID mouse tissues when rAAV-transduced cells reenter the cell cycle, we took advantage of two-thirds partial hepatectomy. This surgical procedure can stimulate cell cycling of hepatocytes that otherwise remain quiescent in adult animals (40). A majority of hepatocytes divide once or twice following the partial hepatectomy. In the study, DNA-PKcs-deficient SCID mice were injected with 1.8×10^{12} vg of AAV8-ISce I.AO3 via the tail vein. In the liver, no-end double-stranded linear rAAV genomes were stably maintained without further recombination over a period of 6 weeks. These mice underwent two-thirds partial hepatectomy 6 weeks postinjection ($n = 3$). rAAV genome forms and copy numbers in the livers at 8 weeks after hepatectomy were determined by our standard Southern blot analysis. rAAV genome copy numbers at hepatectomy and after hepatectomy were 396 ± 15 and 100 ± 4 ds-vg/dge, respectively (Fig. 8). No-end double-stranded linear monomer rAAV genomes were readily detectable at hepatectomy but disappeared completely after partial hepatectomy (Fig. 8). The magnitude of decrease in concatemers and circular monomer genomes by partial hepatectomy was not as drastic as the decrease in no-end double-stranded linear rAAV genomes. This observation indicates that no-end double-stranded rAAV genomes were more easily degraded and lost by hepatocyte division than circular or concatemeric rAAV genomes. Alternatively, it also may be possible that cell cycling recruited a DNA-PKcs-independent alternative hairpin DNA opening activity(ies) and triggered subsequent rAAV genome recombination, resulting in conversion from no-end genomes to circular monomers and/or concatemers. Further study is required to address the mechanism for the complete disappearance of no-end genomes following a minimum number of cell divisions.

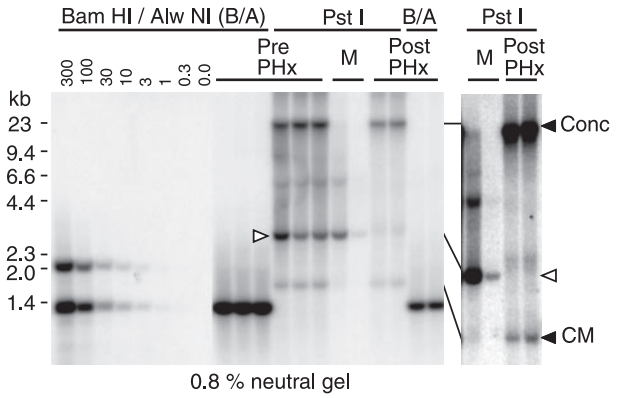

FIG. 8. Fate of the no-end double-stranded linear rAAV genomes in the DNA-PKcs-deficient mouse livers after two-thirds partial hepatectomy. DNA-PKcs-deficient mice injected with 1.8×10^{12} vg of AAV8-ISce I.AO3 via the tail vein underwent two-thirds partial hepatectomy (PHx) 6 weeks postinjection. The liver DNA samples at PHx (PrePHx) and 8 weeks after PHx (PostPHx) were analyzed for rAAV genome copy numbers and molecular forms by Southern blot analysis in the same manner as that described for Fig. 1. Ten micrograms of total DNA was digested with an appropriate enzyme, separated on a 0.8% neutral agarose gel, and probed with a β-lactamase probe. In the right panel, a separate analysis was performed to demonstrate that there was no no-end double-stranded linear monomer rAAV genomes in PostPHx samples. The bands at the positions indicated with open arrowheads represent no-end double-stranded linear genomes. Concatemers (Conc) and supercoiled circular monomer genomes (CM) are shown with closed arrowheads. M, 0.5 and 0.05 μg of liver DNA from a DNA-PKcs-deficient mouse injected with 7.2×10^{12} vg of AAV8-ISce I.AO3, serving as a no-end rAAV genome position marker. 0.0 to 300, rAAV vector genome copy number standards (double-stranded rAAV vector genome copy numbers per diploid genomic equivalents). Each lane represents individual mouse samples.

DISCUSSION

In the present study, we studied the interactions between rAAV genomes and host cellular DNA repair machinery in various tissues from mice by specifically focusing on the roles of DNA-PKcs and Artemis in opening AAV-ITR hairpin termini. We demonstrated that both DNA-PKcs and Artemis were required for efficient AAV-ITR hairpin opening in

rAAV-transduced mouse tissues. If the AAV-ITR hairpins were not opened, no-end double-stranded linear rAAV genomes with covalently closed hairpin caps accumulated. This is reminiscent of the observation that hairpin-shaped covalently sealed coding end molecules, V(D)J recombination molecules, accumulated in thymocytes in both DNA-PKcs-deficient mice (45) and Artemis-deficient mice (44). Bisulfite sequencing analysis of AAV-ITR hairpin termini strongly indicated that a DNA-PKcs- and Artemis-associated endonuclease activity opens an unpaired 3-base loop (AAA) at the AAV-ITR hairpin tips. These results provided a mechanistic interpretation of some of the effects caused by the absence of DNA-PKcs on rAAV genome processing in earlier studies (10, 36, 48, 49). The persistent presence of linear rAAV2 genomes in SCID mouse muscle (10, 48) can be explained by the lack of Artemis/DNA-PKcs endonuclease activity and the low capacity of DNA hairpin processing by the Artemis/DNA-PKcs-independent pathway in this tissue. This resulted in the formation of covalently closed AAV-ITR hairpin caps at rAAV termini that do not serve as a substrate for genome recombination. The absence of linear genomes and efficient self circularization of rAAV2 genomes in the liver (36, 49) can be explained by the large capacity of Artemis/DNA-PKcs-independent alternative DNA hairpin-processing pathways in the liver. Although our present study does not systematically address whether and how formation of no-end genomes can affect transgene expression from the rAAV vector, a previous study by Song et al. has suggested that no-end genomes are at least transcriptionally active (48). In their study, they observed no significant difference in rAAV-mediated transgene expression from muscle between DNA-PKcs-proficient and -deficient mice, even when a substantial amount of rAAV genomes exhibited linear forms (presumably no-end genomes) in DNA-PKcs-deficient mouse muscle (48). In the present study, livers of both DNA-PKcs-deficient mouse and Artemis-deficient mouse transduced with a high dose of AAV8- or AAV9-EF1α-nlslacZ, which had a considerable amount of no-end linear rAAV genomes, expressed transgene products at levels comparable to those found for wild-type mice (data not shown), which is concordant with the observation by Song et al. (48).

In addition to providing new insights into the biology of rAAV in animals, our study indicated that rAAV, with its ability to induce DNA damage responses upon infection (18), could be exploited to study the cellular biology of DNA damage in various tissues in living animals. Viruses recently have come under the spotlight as a powerful tool to study the cellular DNA damage responses and the repair processes of damaged DNA in cells (24). In our approach, rAAV served as a trigger of DNA damage responses as well as a straightforward readout to assess the effects of the evoked signals. The advantage of the use of this approach is that we could study cellular biology in the context of the entire body or in different organs of living mammals. rAAV8 and rAAV9 vectors transduce many organs globally by a simple tail vein injection (16, 32), and therefore they induce DNA damage responses in a variety of tissues in animals. rAAV does not express any virally encoded proteins; therefore, it can induce DNA damage responses without interference by viral proteins that would suppress host DNA repair pathways. It should be noted that many viruses are armed with machinery to inactivate cellular DNA repair machinery upon infection in a way that benefits their life cycles. Adenovirus blocks the cellular NHEJ pathway with viral proteins E1b55K and E4orf6, which degrade Mre11/Rad50/Nbs1 complex and prevent viral genome concatemerization so that viral genomes can be packaged into virions (50). Disruption of the ATR (ataxia telangiectasia-mutated and Rad3-related) pathway by sequestration of DNA repair components away from sites of viral replication has been reported in the herpes simplex virus type 1 life cycle (53). The distinct features of rAAV in cellular responses may provide a unique opportunity to study host cellular biology in response to DNA damage.

In the past 5 years, our knowledge about the roles of DNA-PKcs and Artemis in the NHEJ pathway has been substantially enhanced by biochemical analyses of purified proteins, studies of various DNA repair pathway-defective cell lines, and investigation of the V(D)J recombination in various in vitro and in vivo systems (4, 21, 25, 27). The major role of Artemis/DNA-PKcs in the NHEJ pathway has proven to be endonucleolytic cleavage of terminally blocked DSBs, including DNA hairpins (26). In vivo roles of Artemis/DNA-PKcs have been studied exclusively in the context of V(D)J recombination, in which RAG1 and RAG2 proteins and recombination signal sequences are involved. By contrast, in vivo roles and functions of Artemis/DNA-PKcs in DNA hairpin opening in a context other than V(D)J recombination is difficult to study, and therefore it has remained largely unknown. In this regard, the demonstration of the similarity of the observations from V(D)J recombination in lymphocytes and rAAV genome recombination in various mouse tissues strongly supports the notion that DNA hairpin opening with Artemis/DNA-PKcs-associated endonuclease activity is a universal mechanism for the metabolism of DNA hairpins at both intrachromosomal and extrachromosomal DSBs in mammals.

To what extent do cells depend on the Artemis/DNA-PKcs pathway for processing DNA hairpins and their equivalents in various tissues in animals? This question still remains unanswered. However, our study suggested that the capacity of Artemis/DNA-PKcs-mediated DNA hairpin processing should be substantially large. Based on our observations, we presume that a cell can process at least thousands of DNA hairpins within a relatively short period of time (e.g., within 10 days at most) by this pathway. In contrast, the capacity of Artemis/DNA-PKcs-independent alternative hairpin-processing pathways is very limited and readily saturated with a small number of DNA hairpin molecules, particularly in nonhepatic tissues. Although the nature of Artemis/DNA-PKcs-independent pathways remains to be elucidated, one of the pathways should be homologous recombination via AAV-ITR hairpin termini (54–56). In fact, we observed a 165-bp DDITR rAAV-rAAV junction, which is a hallmark of homologous recombination, in 16% of rAAV-rAAV junctions in wild-type mouse liver transduced with high doses of rAAV8. Nonetheless, the degree of dependency of the Artemis/DNA-PKcs pathway on DNA hairpin opening varies by tissue type, presumably hairpin types, and the context in which a hairpin resides. It should be noted that our recent study indicated that cleavage of DNA palindromes in the mouse genome, which potentially form DNA hairpins, occurs independently of DNA-PKcs activity (16a).

It is intriguing that single-stranded rAAV genomes convert to no-end double-stranded linear rAAV genomes in rAAV-

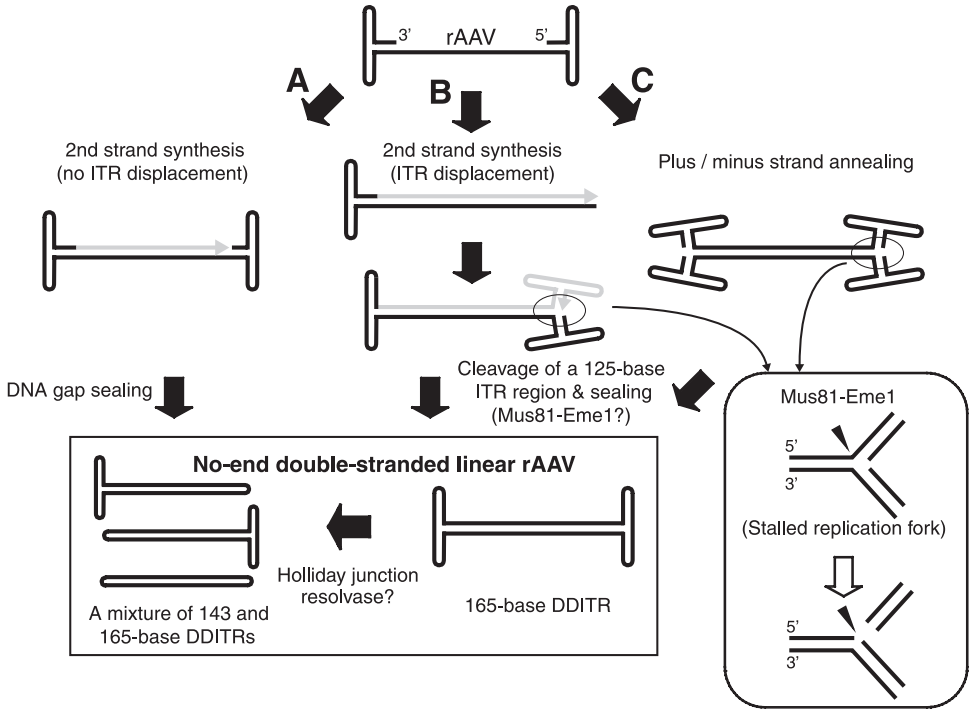

FIG. 9. Possible mechanisms for the formation of no-end double-stranded linear rAAV genomes with covalently closed DDITR hairpin caps in the absence of DNA-PKcs or Artemis. Three possible mechanisms (models A, B, and C) are presented. Model A is the simplest model, in which the leading strand hits the 5′ end of the AAV-ITR hairpin DNA and is ligated to it. Models B and C are different in how double-stranded rAAV genomes are formed (second-strand synthesis versus annealing), but they share the same hairpin terminal structure. An rAAV terminus with two 145-base ITRs is reminiscent of a stalled replication fork; therefore, it might be a target for the endonuclease activity of the Mus81-Eme1 complex, which cleaves the replication fork as illustrated. Since no active DNA replication is ongoing at the cleaved fork, the DNA gap presumably is sealed, forming DDITR hairpin caps. Holliday junction resolvase-like activity cleaves a 22-base hairpin region of the ITR (see Fig. 6D), making four types of no-end rAAV genomes in the absence of DNA-PKcs or Artemis. Whether this cleavage occurs before, at, or after forming no-end rAAV genomes is not known.

transduced mouse tissues if AAV-ITR hairpins are not opened. Based on the results of the bisulfite PCR analyses, it is most likely that no-end molecules are a mixture of four types of molecules: no-end molecules with 165-base DDITR caps at both ends, two types of molecules with a 165-base cap at one end and a 143-base cap at the other, and molecules with 143-base caps at both ends. This is because the relative amount of 143- and 165-base caps varied among samples, suggesting that cleavage of T-shaped ITR hairpin arms occurred randomly. At this point, although the mechanism of how the no-end molecules form remains unclear, we propose three possible models (Fig. 9). In the first model, the central single-stranded gap in the single-stranded rAAV genome is filled by second-strand synthesis (11, 12) and is sealed. In this model, DNA polymerase somehow does not displace ITRs, although DNA polymerase displaces ITRs when rAAV genomes replicate. In the second and third models, double-stranded linear rAAV genomes originally have two AAV-ITRs at one or two termini. Such genomes can be generated by either second-strand synthesis with ITR displacement (second model) or plus- and minus-strand annealing (37) (third model). This rAAV genome terminal configuration is recognized by cells as a stalled replication fork (18). An undefined cellular endonuclease activity removes a 125-base AAV-ITR region corresponding to A-B-C-A subregions, and then cellular DNA repair machinery seals the DNA break. A candidate for this activity is the Mus81-Eme1 complex (2, 6, 42). Mus81 is a xeroderma pigmentaosum, complementation group F-related cellular endonuclease that was initially identified in yeast (3, 17). Mus81, which is primarily inactive, acquires endonuclease activity when it is complexed with Eme1, and it cleaves replication forks, 3′ DNA flaps, and Holliday junctions (6, 19, 52). Although further studies are required, the demonstration by Jurvansuu et al. (18) that AAV-ITR triggers the same DNA damage responses as those evoked by stalled replication forks makes it likely that the Mus81-Eme1 pathway is involved in this process.

It also was intriguing that Holliday junction resolvase-like activity emerged in the absence of DNA-PKcs or Artemis, making a 143-base DDITR. This activity cuts off one of the arms of T-shaped ITR hairpins, presumably by making nicks exclusively at symmetrically related sites of the two strands. This nuclease activity is quite consistent with Holliday junction resolvase in the classic sense. Although AAV-ITR does not form a typical X DNA substrate, an unpaired A or T residue at the central position of the T-shaped hairpin forms a 1-base loop (see Fig. 6D), which may make AAV-ITR equivalent to a four-way junction. To date, two distinct resolvase activities have been identified in mammalian cells (7). One activity cleaves Holliday junctions at symmetrical positions, but the identities of the proteins are not known. The other activity is that of the aforementioned Mus81-Eme1 complex, which differs from all other Holliday junction resolvases in that this

complex preferentially cleaves replication forks and 3′ flaps with reduced cleavage site specificity. The first resolvase activity, therefore, appears to be involved in the process of making the 143-base DDITR. Further biochemical studies are required to determine whether Mus81-Eme1 endonuclease activity is involved in this process and whether AAV-ITR serves as a substrate for Holliday junction resolvase.

Our observations also have significant implications for the biology of rAAV vectors for human gene therapy. Our study demonstrates that the error-prone Artemis/DNA-PKcs-dependent NHEJ pathway is the major pathway to process rAAV genomes when the genome load is high. Activation of this pathway would be detrimental, because it may promote unwanted concatemerization, which reduces transcriptional ability or induces gene silencing (32, 38); various vector genome rearrangements, which often disrupt the transgene expression cassette in the vector; and vector genome integration, which may put cells at a risk of insertional mutagenesis (34, 39). In contrast, rAAV vector genome circularization via homologous recombination has many advantages. Double-stranded circular monomers are persistently transcriptionally active, and homologous recombination does not cause disruption of the transgene or result in dispersed chromosomal integration. Although it will be challenging, finding a way to control or regulate DNA repair pathways in rAAV vector transduction would be a substantial benefit for successful human gene therapy. Studies investigating the interactions between rAAV and host cellular biology have just begun, and we will need to develop biochemical approaches to outline the network of rAAV-evoked DNA damage signal transduction in animals in our future studies. Nonetheless, the present study has not only furthered our understanding of the in vivo rAAV biology but also has shown the power of virus-based approaches to elucidate the fundamental cellular biology of DNA damage responses in quiescent somatic cells in various tissues of living mammals.

ACKNOWLEDGMENTS

We thank Guangping Gao and James M. Wilson for providing the AAV8 and AAV9 packaging plasmids and Frederick W. Alt and Susanna M. Lewis for Artemis-deficient mice.

This work was supported by Public Health Service grants DK68636 and DK78388 (to H.N.) and HL64274 (to M.A.K.) from the National Institutes of Health and a Career Development Award from the National Hemophilia Foundation (to H.N.).

REFERENCES

1. **Bassing, C. H., and F. W. Alt.** 2004. The cellular response to general and programmed DNA double strand breaks. DNA Repair (Amsterdam) **3:**781–796.
2. **Boddy, M. N., P. H. Gaillard, W. H. McDonald, P. Shanahan, J. R. Yates III, and P. Russell.** 2001. Mus81-Eme1 are essential components of a Holliday junction resolvase. Cell **107:**537–548.
3. **Boddy, M. N., A. Lopez-Girona, P. Shanahan, H. Interthal, W. D. Heyer, and P. Russell.** 2000. Damage tolerance protein Mus81 associates with the FHA1 domain of checkpoint kinase Cds1. Mol. Cell. Biol. **20:**8758–8766.
4. **Burma, S., B. P. Chen, and D. J. Chen.** 2006. Role of non-homologous end joining (NHEJ) in maintaining genomic integrity. DNA Repair (Amsterdam) **5:**1042–1048.
5. **Burton, M., H. Nakai, P. Colosi, J. Cunningham, R. Mitchell, and L. Couto.** 1999. Coexpression of factor VIII heavy and light chain adeno-associated viral vectors produces biologically active protein. Proc. Natl. Acad. Sci. USA **96:**12725–12730.
6. **Ciccia, A., A. Constantinou, and S. C. West.** 2003. Identification and characterization of the human mus81-eme1 endonuclease. J. Biol. Chem. **278:** 25172–25178.
7. **Constantinou, A., X. B. Chen, C. H. McGowan, and S. C. West.** 2002. Holliday junction resolution in human cells: two junction endonucleases with distinct substrate specificities. EMBO J. **21:**5577–5585.
8. **Duan, D., P. Sharma, L. Dudus, Y. Zhang, S. Sanlioglu, Z. Yan, Y. Yue, Y. Ye, R. Lester, J. Yang, K. J. Fisher, and J. F. Engelhardt.** 1999. Formation of adeno-associated virus circular genomes is differentially regulated by adenovirus E4 ORF6 and E2a gene expression. J. Virol. **73:**161–169.
9. **Duan, D., Z. Yan, Y. Yue, and J. F. Engelhardt.** 1999. Structural analysis of adeno-associated virus transduction circular intermediates. Virology **261:**8–14.
10. **Duan, D., Y. Yue, and J. F. Engelhardt.** 2003. Consequences of DNA-dependent protein kinase catalytic subunit deficiency on recombinant adeno-associated virus genome circularization and heterodimerization in muscle tissue. J. Virol. **77:**4751–4759.
11. **Ferrari, F. K., T. Samulski, T. Shenk, and R. J. Samulski.** 1996. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. J. Virol. **70:**3227–3234.
12. **Fisher, K. J., G. P. Gao, M. D. Weitzman, R. DeMatteo, J. F. Burda, and J. M. Wilson.** 1996. Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. J. Virol. **70:**520–532.
13. **Fugmann, S. D., A. I. Lee, P. E. Shockett, I. J. Villey, and D. G. Schatz.** 2000. The RAG proteins and V(D)J recombination: complexes, ends, and transposition. Annu. Rev. Immunol. **18:**495–527.
14. **Goodarzi, A. A., Y. Yu, E. Riballo, P. Douglas, S. A. Walker, R. Ye, C. Harer, C. Marchetti, N. Morrice, P. A. Jeggo, and S. P. Lees-Miller.** 2006. DNA-PK autophosphorylation facilitates Artemis endonuclease activity. EMBO J. **25:** 3880–3889.
15. **Grunau, C., S. J. Clark, and A. Rosenthal.** 2001. Bisulfite genomic sequencing: systematic investigation of critical experimental parameters. Nucleic Acids Res. **29:**e65–5.
16. **Inagaki, K., S. Fuess, T. A. Storm, G. A. Gibson, C. F. McTiernan, M. A. Kay, and H. Nakai.** 2006. Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. Mol. Ther. **14:**45–53.

16a. **Inagaki, K., S. M. Lewis, X. Wu, C. Ma, D. J. Munroe, S. Fuess, T. A. Storm, M. A. Kay, and H. Nakai.** 2007. DNA palindromes with a modest arm length of ≥20 base pairs are a significant target for a recombinant adeno-associated virus vector integration in the liver, muscles, and heart in mice. J. Virol. **81:**11290–11303.

17. **Interthal, H., and W. D. Heyer.** 2000. MUS81 encodes a novel helix-hairpin-helix protein involved in the response to UV- and methylation-induced DNA damage in Saccharomyces cerevisiae. Mol. Gen. Genet. **263:**812–827.
18. **Jurvansuu, J., K. Raj, A. Stasiak, and P. Beard.** 2005. Viral transport of DNA damage that mimics a stalled replication fork. J. Virol. **79:**569–580.
19. **Kaliraman, V., J. R. Mullen, W. M. Fricke, S. A. Bastin-Shanower, and S. J. Brill.** 2001. Functional overlap between Sgs1-Top3 and the Mms4-Mus81 endonuclease. Genes Dev. **15:**2730–2740.
20. **Laird, C. D., N. D. Pleasant, A. D. Clark, J. L. Sneeden, K. M. Hassan, N. C. Manley, J. C. Vary, Jr., T. Morgan, R. S. Hansen, and R. Stoger.** 2004. Hairpin-bisulfite PCR: assessing epigenetic methylation patterns on complementary strands of individual DNA molecules. Proc. Natl. Acad. Sci. USA **101:**204–209.
21. **Le Deist, F., C. Poinsignon, D. Moshous, A. Fischer, and J. P. de Villartay.** 2004. Artemis sheds new light on V(D)J recombination. Immunol. Rev. **200:**142–155.
22. **Lieber, M. R., Y. Ma, U. Pannicke, and K. Schwarz.** 2004. The mechanism of vertebrate nonhomologous DNA end joining and its role in V(D)J recombination. DNA Repair (Amsterdam) **3:**817–826.
23. **Lieber, M. R., K. Yu, and S. C. Raghavan.** 2006. Roles of nonhomologous DNA end joining, V(D)J recombination, and class switch recombination in chromosomal translocations. DNA Repair (Amsterdam) **5:**1234–1245.
24. **Lilley, C. E., R. A. Schwartz, and M. D. Weitzman.** 2007. Using or abusing: viruses and the cellular DNA damage response. Trends Microbiol. **15:**119–126.
25. **Ma, Y., H. Lu, K. Schwarz, and M. R. Lieber.** 2005. Repair of double-strand DNA breaks by the human nonhomologous DNA end joining pathway: the iterative processing model. Cell Cycle **4:**1193–1200.
26. **Ma, Y., U. Pannicke, K. Schwarz, and M. R. Lieber.** 2002. Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. Cell **108:**781–794.
27. **Ma, Y., K. Schwarz, and M. R. Lieber.** 2005. The Artemis:DNA-PKcs endonuclease cleaves DNA loops, flaps, and gaps. DNA Repair (Amsterdam) **4:**845–851.
28. **Mansilla-Soto, J., and P. Cortes.** 2003. VDJ recombination: Artemis and its in vivo role in hairpin opening. J. Exp. Med. **197:**543–547.
29. **McCarty, D. M., S. M. Young, Jr., and R. J. Samulski.** 2004. Integration of adeno-associated virus (AAV) and recombinant AAV vectors. Annu. Rev. Genet. **38:**819–845.
30. **Moshous, D., I. Callebaut, R. de Chasseval, B. Corneo, M. Cavazzana-Calvo, F. Le Deist, I. Tezcan, O. Sanal, Y. Bertrand, N. Philippe, A. Fischer, and J. P. de Villartay.** 2001. Artemis, a novel DNA double-strand break repair/

V(D)J recombination protein, is mutated in human severe combined immune deficiency. Cell **105:**177–186.
31. **Nakai, H., S. Fuess, T. A. Storm, L. A. Meuse, and M. A. Kay.** 2003. Free DNA ends are essential for concatemerization of synthetic double-stranded adeno-associated virus vector genomes transfected into mouse hepatocytes in vivo. Mol. Ther. **7:**112–121.
32. **Nakai, H., S. Fuess, T. A. Storm, S. Muramatsu, Y. Nara, and M. A. Kay.** 2005. Unrestricted hepatocyte transduction with adeno-associated virus serotype 8 vectors in mice. J. Virol. **79:**214–224.
33. **Nakai, H., Y. Iwaki, M. A. Kay, and L. B. Couto.** 1999. Isolation of recombinant adeno-associated virus vector-cellular DNA junctions from mouse liver. J. Virol. **73:**5438–5447.
34. **Nakai, H., E. Montini, S. Fuess, T. A. Storm, M. Grompe, and M. A. Kay.** 2003. AAV serotype 2 vectors preferentially integrate into active genes in mice. Nat. Genet. **34:**297–302.
35. **Nakai, H., E. Montini, S. Fuess, T. A. Storm, L. Meuse, M. Finegold, M. Grompe, and M. A. Kay.** 2003. Helper-independent and AAV-ITR-independent chromosomal integration of double-stranded linear DNA vectors in mice. Mol. Ther. **7:**101–111.
36. **Nakai, H., T. A. Storm, S. Fuess, and M. A. Kay.** 2003. Pathways of removal of free DNA vector ends in normal and DNA-PKcs-deficient SCID mouse hepatocytes transduced with rAAV vectors. Hum. Gene Ther. **14:**871–881.
37. **Nakai, H., T. A. Storm, and M. A. Kay.** 2000. Recruitment of single-stranded recombinant adeno-associated virus vector genomes and intermolecular recombination are responsible for stable transduction of liver in vivo. J. Virol. **74:**9451–9463.
38. **Nakai, H., C. E. Thomas, T. A. Storm, S. Fuess, S. Powell, J. F. Wright, and M. A. Kay.** 2002. A limited number of transducible hepatocytes restricts a wide-range linear vector dose response in recombinant adeno-associated virus-mediated liver transduction. J. Virol. **76:**11343–11349.
39. **Nakai, H., X. Wu, S. Fuess, T. A. Storm, D. Munroe, E. Montini, S. M. Burgess, M. Grompe, and M. A. Kay.** 2005. Large-scale molecular characterization of adeno-associated virus vector integration in mouse liver. J. Virol. **79:**3606–3614.
40. **Nakai, H., S. R. Yant, T. A. Storm, S. Fuess, L. Meuse, and M. A. Kay.** 2001. Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction in vivo. J. Virol. **75:**6969–6976.
41. **Niewolik, D., U. Pannicke, H. Lu, Y. Ma, L. C. Wang, P. Kulesza, E. Zandi, M. R. Lieber, and K. Schwarz.** 2006. DNA-PKcs dependence of Artemis endonucleolytic activity, differences between hairpins and 5′ or 3′ overhangs. J. Biol. Chem. **281:**33900–33909.
42. **Oðrünç, M., and A. Sancar.** 2003. Identification and characterization of human MUS81-MMS4 structure-specific endonuclease. J. Biol. Chem. **278:**21715–21720.
43. **Pannicke, U., Y. Ma, K. P. Hopfner, D. Niewolik, M. R. Lieber, and K. Schwarz.** 2004. Functional and biochemical dissection of the structure-specific nuclease ARTEMIS. EMBO J. **23:**1987–1997.
44. **Rooney, S., J. Sekiguchi, C. Zhu, H. L. Cheng, J. Manis, S. Whitlow, J. DeVido, D. Foy, J. Chaudhuri, D. Lombard, and F. W. Alt.** 2002. Leaky Scid phenotype associated with defective V(D)J coding end processing in Artemis-deficient mice. Mol. Cell **10:**1379–1390.
45. **Roth, D. B., J. P. Menetski, P. B. Nakajima, M. J. Bosma, and M. Gellert.** 1992. V(D)J recombination: broken DNA molecules with covalently sealed (hairpin) coding ends in scid mouse thymocytes. Cell **70:**983–991.
46. **Schatz, D. G.** 2004. V(D)J recombination. Immunol. Rev. **200:**5–11.
47. **Snyder, R. O., R. J. Samulski, and N. Muzyczka.** 1990. In vitro resolution of covalently joined AAV chromosome ends. Cell **60:**105–113.
48. **Song, S., P. J. Laipis, K. I. Berns, and T. R. Flotte.** 2001. Effect of DNA-dependent protein kinase on the molecular fate of the rAAV2 genome in skeletal muscle. Proc. Natl. Acad. Sci. USA **98:**4084–4088.
49. **Song, S., Y. Lu, Y. K. Choi, Y. Han, Q. Tang, G. Zhao, K. I. Berns, and T. R. Flotte.** 2004. DNA-dependent PK inhibits adeno-associated virus DNA integration. Proc. Natl. Acad. Sci. USA **101:**2112–2116.
50. **Stracker, T. H., C. T. Carson, and M. D. Weitzman.** 2002. Adenovirus oncoproteins inactivate the Mre11-Rad50-NBS1 DNA repair complex. Nature **418:**348–352.
51. **Thomas, C. E., T. A. Storm, Z. Huang, and M. A. Kay.** 2004. Rapid uncoating of vector genomes is the key to efficient liver transduction with pseudotyped adeno-associated virus vectors. J. Virol. **78:**3110–3122.
52. **Whitby, M. C., F. Osman, and J. Dixon.** 2003. Cleavage of model replication forks by fission yeast Mus81-Eme1 and budding yeast Mus81-Mms4. J. Biol. Chem. **278:**6928–6935.
53. **Wilkinson, D. E., and S. K. Weller.** 2006. Herpes simplex virus type I disrupts the ATR-dependent DNA-damage response during lytic infection. J. Cell Sci. **119:**2695–2703.
54. **Yan, Z., D. C. Lei-Butters, Y. Zhang, R. Zak, and J. F. Engelhardt.** 2007. Hybrid adeno-associated virus bearing nonhomologous inverted terminal repeats enhances dual-vector reconstruction of minigenes in vivo. Hum. Gene Ther. **18:**81–87.
55. **Yan, Z., R. Zak, Y. Zhang, and J. F. Engelhardt.** 2005. Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes. J. Virol. **79:**364–379.
56. **Zentilin, L., A. Marcello, and M. Giacca.** 2001. Involvement of cellular double-stranded DNA break binding proteins in processing of the recombinant adeno-associated virus genome. J. Virol. **75:**12279–12287.