

Identification of an Insulator in AAVS1, a Preferred Region for Integration of Adeno-Associated Virus DNA

Toshihiko Ogata,¹ Takuyo Kozuka,² and Tadahito Kanda^{1*}

Division of Molecular Genetics, National Institute of Infectious Diseases, Toyama, Shinjuku-ku, Tokyo 162-8640,¹ and Department of Radiotherapy, Cancer Institute Hospital, Japanese Foundation for Cancer Research, Kamiikebukuro, Toshima-ku, Tokyo 170-8455,² Japan

Received 17 March 2003/Accepted 13 May 2003

In latent adeno-associated virus (AAV) infection, the viral genome is integrated preferentially into the human chromosome 19 q arm at a specific region designated AAVS1, which has an open chromatin conformation as indicated by the presence of a DNase I-hypersensitive site (DHS-S1). We examined whether an insulator, which defines the domain of gene expression by directionally blocking the action of enhancers and by preventing the spread of heterochromatin, is present near the DHS-S1 in the middle of a 2.6-kbp AAVS1-related DNA fragment used in this study. The fragment, cloned into an Epstein-Barr virus (EBV)-based eukaryotic episomal plasmid, was introduced into HEK293 cells. The DHS-S1 on the plasmid replicating in the nuclei was hypersensitive to DNase I digestion, and thus, the EBV plasmid system was used in an enhancer-blocking assay with the 2.6-kbp DNA and two shortened DNAs, of 1.6 kbp and 336 bp, containing DHS-S1. The three DNA fragments, when inserted in the proper direction between the cytomegalovirus immediate-early enhancer and minimal promoter, repressed the expression of a reporter gene. Thus, the enhancer-blocking activity was located within the 336-bp DNA containing the entire region (300 bp) of DHS-S1. To investigate the prevention of repression caused by heterochromatin, a transgene-expressing cassette flanked by the two 336-bp DNAs placed in the enhancer-blocking direction was introduced into HEK293 and HeLa cells. All the cell clones examined with the cassette integrated into cell DNA continued to express the transgene, which indicates that the pair of 336-bp DNA apparently prevented the spread of heterochromatin. The results show that an insulator lies between nucleotides 17 and 354 near the DHS-S1 in AAVS1. In a gel shift test, the 336-bp DNA did not bind an *in vitro*-prepared CCCTC-binding factor that binds to the chicken β -globin insulator, suggesting that the AAVS1 insulator requires an as yet unidentified binding protein. The newly identified AAVS1 insulator is likely to contribute to the maintenance of an open chromatin conformation that affects the life cycle of AAV.

Adeno-associated virus type 2 (AAV), a human parvovirus, is a nonenveloped small virus with a 4.7-kb genome of single-stranded linear DNA (2, 3). Efficient propagation of AAV requires coinfection with a helper virus, such as adenovirus (6, 27, 32), herpes simplex virus (4, 23, 31), or human cytomegalovirus (CMV) (10, 22). Without a helper virus, the AAV genome is preferentially integrated into a specific target site (the preintegration site) on the q arm of chromosome 19 (11, 17, 18, 19) and is maintained as a latent provirus (3). AAV is readily rescued from the integrated genome when a latently infected cell is superinfected with a helper virus (8, 14) or stimulated by genotoxic (33) or apoptotic (24) stress.

The AAV preintegration site has been isolated, and its 4-kbp-long sequenced segment is called AAVS1, with nucleotide numbering that starts from the *EcoRI* site (17). Within AAVS1, the recombination between cell and viral DNAs occurs in the region spanning approximately nucleotides (nt) 700 to 2400 (12, 15), and a DNase I-hypersensitive site (DHS), designated DHS-S1, is located in the region from nt 50 to 350 (20). DHSs are generally considered to be a marker of open chromatin where *trans*-acting factors are allowed access to DNA. They are functionally associated with transcription, rep-

lication, recombination, and chromosome segregation (13, 28, 35) and are sometimes associated with insulators (21, 29).

Insulators are DNA sequences that define the domains of gene expression by directionally blocking enhancers from affecting promoters in different domains and by acting as boundaries to the surrounding heterochromatin that silences the genes located within (5). With isolated DNA fragments, these two insulator functions can be tested by using surrogate expression systems (to test the blocking of enhancer action and prevention of the spread of heterochromatin): when an insulator is inserted in the proper direction between an enhancer and a promoter in an episomal expression plasmid, the action of the enhancer is blocked in the cells harboring the plasmids; when cells are stably transformed by a transgene-expressing cassette flanked by two insulators, the cells continue to express the transgene irrespective of the site of integration of the cassette into the chromosomes (7, 16, 25, 26). Of all the known vertebrate insulators, the chicken β -globin insulator (cHS4) is the one that has been most extensively studied. It is located in the fourth DHS in the locus control region (LCR) of the chicken β -globin gene (9). The sequence-specific binding of CCCTC-binding factor (CTCF) to cHS4 is required for the insulation activity (1).

We have speculated that an insulator may be present near the DHS-S1 in AAVS1 and would prevent the spread of heterochromatin over the preintegration site and the latent provirus, providing a favorable environment for AAV to be readily

* Corresponding author. Mailing address: Division of Molecular Genetics, National Institute of Infectious Diseases, 1-23-1 Toyama, Shinjuku-ku, Tokyo 162-8640, Japan. Phone: (81) 3-5285-1111. Fax: (81) 3-5285-1166. E-mail: kanda@nih.go.jp.

integrated and rescued. In this study, we examined the presence of an insulator, using a 2.6-kbp-long AAVS1-related DNA fragment that contains the DHS-S1 and half of the region where the sites of recombination occur. The tests for the insulator functions, with the 2.6-kbp fragment and its two shortened fragments of 1.6 kbp and 336 bp, showed that an insulator lies between nt 17 and 354, in a region overlapping the DHS-S1. The AAVS1 insulator did not bind CTCF that bound specifically to the chicken insulator cHS4.

MATERIALS AND METHODS

Cells. HEK293 cells (human embryonic kidney cells immortalized by adenovirus E1), HeLa cells, T1G cells (primary human lung fibroblasts purchased from Health Science Research Resources Bank, Tokyo, Japan), and DT40 cells (chicken cell line purchased from Health Science Research Resources Bank) were grown in Dulbecco's modified Eagle's medium supplemented with 10% fetal calf serum.

DNA fragments used for plasmid construction. DNA fragments to be tested for insulator functions were cloned from genomic DNA extracted from T1G cells. The cell DNA was digested with *Bam*HI (Takara Co. Ltd., Kyoto, Japan) and then electrophoresed in an agarose gel. DNA fragments with approximate sizes of 2 to 3 kbp were extracted from the agarose gel and were inserted into pUC19 at the *Bam*HI site. Bacterial clones containing pUC19 with AAVS1 were screened by colony hybridization. A DNA fragment used to generate ³²P-labeled probe was obtained by PCR using the DNA extracted from T1G cells and primers consisting of 5'-GAATTCCTAACTGCC and 5'-TTGGTGGAGTC CAGCACGG. The nucleotide sequences of the primers were determined by referring to the AAVS1 nucleotide sequence previously reported (17). Thus, pUC19 containing a 2.6-kbp AAVS1-related *Bam*HI fragment was obtained. Then the plasmid was cleaved with *Bam*HI, *Eco*RI plus *Bam*HI, and *Sma*I to obtain AAVS1-*Bam*HI/*Bam*HI, AAVS1-*Eco*RI/*Bam*HI, and AAVS1-*Sma*I/*Sma*I, respectively. AAVS1-*Eco*RI/*Bam*HI and AAVS1-*Sma*I/*Sma*I were inserted into pBluescript (Stratagene, La Jolla, Calif.) between the *Eco*RI and *Bam*HI sites and at the *Sma*I site, respectively, and then the DNA was cleaved with *Bss*HI to obtain AAVS1-EB/*Bss*HI and AAVS1-S/*Bss*HI. pBluescript containing AAVS1-*Sma*I/*Sma*I was cleaved with *Bam*HI plus *Sal*I to obtain AAVS1-S/*Bam*HI/*Sal*I. pBluescript containing AAVS1-S/S was cleaved with *Bam*HI, ligated with a *Kpn*I linker, and then cleaved with *Kpn*I to obtain AAVS1-S/*Kpn*I.

The 1.2-kbp DNA fragment containing the LCR of the chicken β -globin gene (β G-LCR) in the DNA extracted from DT40 cells was amplified by PCR with primers consisting of 5'-CCTGGATCCGAGCTCACGGGGACAGCCCC and 5'-AGGATCCAATATTCTACTGACTCCGTCCTG. These primers had *Bam*HI recognition sites at their 5' regions. The amplified DNA was cloned into pGEM-T Easy vector (Promega Corp., Madison, Wis.) and sequenced. The DNA fragment, β G-LCR, was obtained by cleavage of the plasmid with *Bam*HI. The 254-bp DNA fragment containing the core region of the β -globin insulator was isolated from β G-LCR by cleavage with *Bam*HI plus *Hind*III and inserted between the *Bam*HI and *Hind*III sites of pBluescript. Then, DNA fragments were obtained by cleavage with *Bam*HI plus *Sal*I (cHS4-*Bam*HI/*Sal*I) and with *Bss*III (cHS4-*Bss*III). The plasmid was cleaved with *Bam*HI, ligated with *Kpn*I linker, and then cleaved with *Kpn*I to obtain cHS4-*Kpn*I.

The DNA fragments containing CMV minimal promoter (CMV-Pro) and CMV immediate-early enhancer (CMV-En) in plasmid pRL-CMV (Promega Corp.) were amplified by PCR with primers with the sequences 5'-TAGATCT TAGCGTGTACGGTGGGAGG and 5'-TAAGCTTGGGCGCGGAGGCT GGATCG and primers with the sequences 5'-CGGATCCCGCTTACATAA CTTACGGT and 5'-AAAGATCTCAAAACAACTCCCATTTGAC, respectively.

The DNA fragment containing Epstein-Barr virus (EBV) oriP and the expression cassette for EBNA-1 (oriP/EBNA-1) was isolated from pEBV-His-C (Invitrogen Corp., Carlsbad, Calif.) by digestion with *Eag*I and *Sal*I.

The 2.2-kbp expression unit for *Renilla* luciferase driven by the TK promoter was isolated by digestion with *Bam*HI plus *Bgl*II and inserted into the *Bam*HI site of pUC19. Then the resultant plasmid was digested with *Kpn*I and *Sal*I to obtain the DNA fragment for *Renilla* luciferase expression (TK-RLuc).

The 564-bp DNA fragment of λ -DNA digested with *Hind*III was inserted into pBluescript at the *Hind*III site. The plasmid was cleaved with *Bss*III and *Bam*HI plus *Sal*I to obtain λ -*Hind*III 564-*Bss*III and λ -*Hind*III 564-*Bam*HI/*Sal*I. It was cleaved with *Bam*HI, ligated with *Kpn*I linker, and then cleaved with *Kpn*I to obtain λ -*Hind*III 564-*Kpn*I.

Construction of expression plasmids. The backbone plasmid used for the DNase I hypersensitivity tests (pEnh-bl1) and that used for the enhancer-blocking assays (pEnh-bl2) were constructed by insertion of the component DNA fragments into pGL3-basic (Invitrogen Corp.). CMV-Pro and CMV-En were inserted between the *Bgl*II and *Hind*III sites and at the *Bgl*II site of pGL3-basic, respectively. oriP/EBNA-1 was inserted between *Not*I and *Sal*I. Then cHS4-*Bam*HI/*Sal*I was inserted between *Bam*HI and *Sal*I. A synthetic linker with *Bgl*II-*Bss*III-*Bgl*II sites was inserted at the *Bgl*II site to obtain pEnh-bl1. The AAVS1 *Bam*HI/*Bam*HI fragment and β G-LCR were inserted at the *Bgl*II site of pEnh-bl1. AAVS1-EB/*Bss*HI, AAVS1-S/*Bss*HI, and λ -*Hind*III 594-*Bss*III were inserted at the *Bss*III site of pEnh-bl1. TK-RLuc was inserted between the *Kpn*I and *Sal*I sites of pEnh-bl1 to generate pEnh-bl2 plasmids.

The backbone plasmid used to test the position effects was pGL3-control (Invitrogen Corp.). AAVS1-S/*Bam*HI/*Sal*I, cHS4-*Bam*HI/*Sal*I, and λ -*Hind*III 564-*Bam*HI/*Sal*I were inserted between the *Bam*HI and *Sal*I sites of the pGL3-control, and then AAVS1-S/*Kpn*I, cHS4-*Kpn*I, and λ -*Hind*III 564-*Kpn*I were inserted to obtain pSV-Luc-AAVS1-S/S, pSV-Luc-cHS4, and pSV-Luc- λ , respectively.

DNase I hypersensitivity test. Normal HEK293 cells (3×10^7) were collected by low-speed centrifugation. HEK293 cells (2×10^6) transfected with 5 μ g of pEnh-bl1 plasmids were passaged at a split ratio of 1:4 at 48 h after the transfection, and at 120 h the cells were collected by a similar centrifugation. The cells were resuspended in phosphate-buffered saline and centrifuged for 10 min at 4°C. They were resuspended in TNM buffer (10 mM Tris-HCl [pH 7.5], 10 mM NaCl, 3 mM MgCl₂) at 5×10^6 cells per ml and incubated for 10 min on ice. The swollen cells were broken in a tight-fitting Dounce homogenizer. Intact nuclei were precipitated through a layer of 0.25 M sucrose in TNM buffer by centrifugation at $1,300 \times g$ for 10 min. The nuclei were resuspended in digestion buffer (15 mM Tris-HCl [pH 7.5], 15 mM NaCl, 3 mM MgCl₂, 60 mM KCl, 0.25 mM sucrose, 0.5 mM dithiothreitol, 1 mM phenylmethylsulfonyl fluoride) at 2×10^7 nuclei per ml. They were digested with various amounts of DNase I at 37°C for 10 min. Then, total DNA was extracted from the nuclei by digestion with proteinase, removal of denatured proteins, and precipitation with isopropanol using the SMITEST (Nippon Genetics Co. Ltd., Tokyo, Japan). The DNA sample was digested with the appropriate restriction enzymes, electrophoresed in an agarose gel, and transferred to a Hybond-N(+) membrane (Amersham Biosciences Corp., Piscataway, N.J.). The membrane was incubated with ³²P-labeled probe as indicated in the figure legends. Prehybridization and hybridization were done at 65°C for 12 h in hybridization solution (0.5 M NaHPO₄ [pH 7.2], 1 mM EDTA, 7% sodium dodecyl sulfate [SDS]) and were followed by washes in $2 \times$ SSC ($1 \times$ SSC is 0.15 M NaCl plus 0.015 M sodium citrate) containing 0.05% SDS at room temperature and in $0.1 \times$ SSC containing 0.1% SDS at 65°C.

Luciferase assay. To measure enhancer-blocking activity, HEK293 cells grown to 50% confluence in a 96-well plate were transfected with 100 ng of the reporter plasmid per well by using the Lipofectamine transfection reagent (Invitrogen Corp.). The cells were incubated with growth medium, passaged at a split ratio of 1:4 at 48 h after the transfection, and lysed at 120 h. The firefly and *Renilla* luciferase activities of the lysate were measured by using the Dual-Glo luciferase assay kit (Promega Corp.) and a TopCount microplate luminometer (Perkin-Elmer Life Sciences Inc., Boston, Mass.).

To measure transient expression of luciferase, HEK293 cells grown to 50% confluence in a 96-well plate were transfected with 100 ng of the reporter plasmid per well by using the Lipofectamine transfection reagent. At 48 h after transfection, firefly luciferase activities of cellular extracts were measured.

To measure the luciferase activity of HEK293 cell clones stably transformed with reporter plasmid, HEK293 cells were transfected with the reporter plasmid, which had been linearized by cleavage with *Ssp*I, together with a plasmid expressing the neomycin resistance gene, at 4:1 molar ratio, using Lipofectamine reagent. At 48 h later, G418 was added to the medium at 800 μ g/ml. The drug-resistant cell clones were picked at 2 weeks after the transfection and cultured with growth medium without G418. At 6, 9, and 15 weeks after the transfection, the firefly luciferase activity of extracts from each cell clone was measured.

Fluorescence microscopy. HeLa cells were transfected with the linearized plasmid expressing enhanced yellow fluorescent protein (EYFP) together with a plasmid expressing neomycin resistance gene at a 4:1 molar ratio, using Lipofectamine reagent (Invitrogen Corp.). At 48 h later, G418 was added to the medium at 600 μ g/ml. The drug-resistant cell colonies were picked 2 weeks after the transfection and cultured with growth medium without G418. At 4 and 6 weeks after the transfection, fluorescence of cells was detected under an inverted UV microscope (Olympus Co Ltd., Tokyo, Japan).

Electrophoretic mobility shift test. A DNA fragment encoding the CTCF DNA-binding domain (from amino acids 293 to 414) was synthesized by PCR

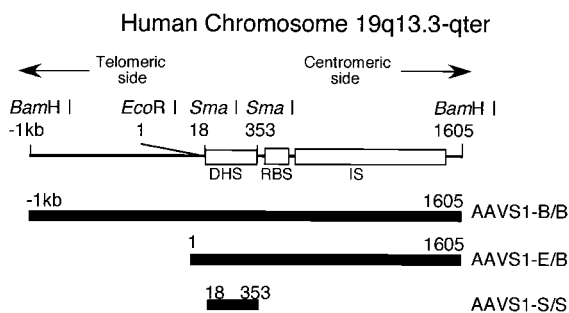


FIG. 1. Schematic representation of the region of AAVS1 analyzed. According to the sequencing by Kotin et al. (17), nucleotide numbering starts at the indicated *EcoRI* site. The Rep binding site (RBS) (30), the region where the recombination between viral and cell DNAs occurs (IS) (12, 15), and the region containing the DNase I hypersensitive site (DHS) (20) are indicated as open rectangles. The DNA fragments used in this study are indicated in the lower part of the figure. AAVS1-B/B, AAVS1-E/B, and AAVS1-S/S are DNA fragments isolated from AAVS1 by cleavage with *Bam*HI, *Eco*RI plus *Bam*HI, and *Sma*I, respectively.

with a sense primer (5' GAATCAATGAAGCCTCCAAAGCCAAC), an antisense primer (5' GAATTCAGATCTGGTTCAGCATTTCACCTG), and a HeLa cell cDNA. The PCR product was digested with *Eco*RI and inserted into the *Eco*RI site of pHM6 (Roche Diagnostics GmbH, Mannheim, Germany). Using the resultant plasmid, a CTCF DNA-binding domain having a hemagglutinin (HA) tag at the N terminus (CTCF-DB) was produced in vitro by using rabbit reticulocyte lysate (TNT coupled transcription/translation system; Promega Corp.).

AAVS1-*Sma*I/*Sma*I and *cHS4* fragments were dephosphorylated with calf intestinal alkaline phosphatase, and then the 5' ends of the fragments were labeled with [γ - 32 P]ATP by using Megalabel (Takara Co. Ltd.). The 32 P-labeled probe (50 fmol) was mixed with 1 μ l of in vitro-translated CTCF-DB or 1 μ l of poly(dI-dC) (50 μ g/ μ l) in a final volume of 20 μ l of binding-buffer (20 mM HEPES [pH 7.9], 150 mM KCl, 5 mM MgCl₂, 5% glycerol, 1 mM dithiothreitol, 0.5% Triton X-100). The mixture was incubated at room temperature for 30 min and electrophoresed on a 4% polyacrylamide gel. The gels were dried, and labeled probes were visualized by autoradiography.

RESULTS

DNase I hypersensitivity of AAVS1-*Bam*HI/*Bam*HI fragment cloned in an EBV-based eukaryotic episomal vector. A 2.6-kbp DNA fragment containing part of the AAV preintegration site was molecularly cloned from *Bam*HI-digested genomic DNA extracted from human lung fibroblasts. Comparison of the sequences of the fragment and AAVS1 (17) showed that the fragment contains DHS-S1, the Rep binding site, and part of the region where the recombination of cell and AAV DNAs occurs (Fig. 1). The 2.6-kbp fragment, designated AAVS1-B/B, was used to test the presence of an insulator in this study after it was confirmed that the fragment maintains an open chromatin structure when inserted into an expression plasmid.

DHS-S1 in AAVS1-B/B (Fig. 1) in an EBV-based plasmid that replicates as an episome in HEK293 cells was hypersensitive to digestion with DNase I. HEK293 cells were transfected with pEnh-bl1 (Fig. 2A) containing AAVS1-B/B and cultured with growth medium to allow the plasmid to replicate synchronously with cellular DNA replication (34). The cells were passaged at 48 h after the transfection, and nuclei were extracted at 120 h. The nuclei were digested with DNase I at the indicated concentrations for 10 min at 37°C. Then, total

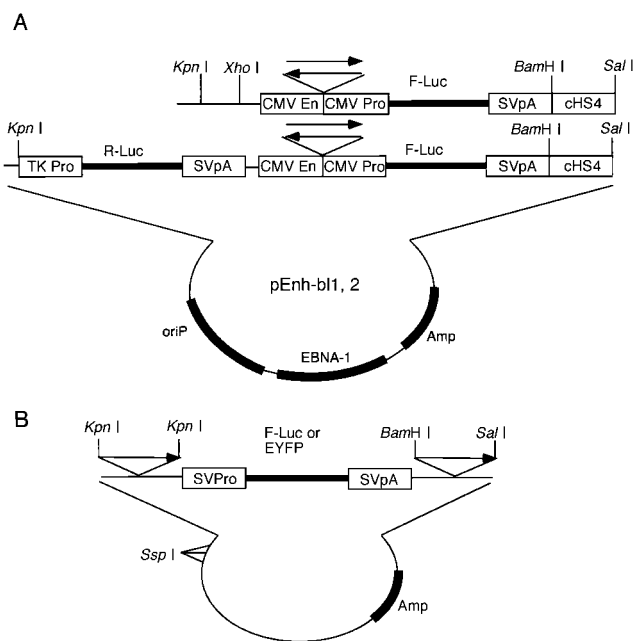


FIG. 2. Schematic representation of the plasmids used in this study. (A) Epstein-Barr virus (EBV) based episomal vectors. The plasmids contained oriP and the EBNA-1 expression unit to maintain the plasmids as episomal DNA in eukaryotic cells. The test DNA fragments were inserted between the CMV enhancer (CMV En) and CMV minimal promoter (CMV Pro). The direction from centromeric to telomeric side (CT) of the test DNA fragments is indicated by arrows. pEnh-bl1 was used for the DNase I hypersensitivity test, and pEnh-bl2, which contains the TK promoter-driven *Renilla* luciferase gene, was used for enhancer-blocking test. (B) Plasmid used to examine expression of firefly luciferase (F-Luc) or EYFP from the cassette integrated into cell DNA. The test DNA fragments were flanked by the expression cassette in the direction from centromeric to telomeric side (CT). The cassette was composed of SV40 early promoter (SVPro), the reporter gene (F-Luc or EYFP), and SV40 poly(A) signal (SVpA).

DNA was extracted and cleaved extensively with *Bam*HI and *Xho*I. Samples were electrophoresed in a 1% agarose gel, and the DNA fragments containing firefly luciferase gene were detected by Southern blotting using a 32 P-labeled probe hybridizing with firefly luciferase DNA (Fig. 3A, upper panel, and 3B, left panel). The nucleotide sequences of the insertion sites of the AAVS1-B/B fragment into the *Bgl*II site do not match the recognition sequence for *Bam*HI. Therefore, cleavage with *Bam*HI and *Xho*I generated a 5.0-kbp fragment containing the AAVS1-B/B fragment plus firefly luciferase DNA. The digestion with DNase I resulted in the emergence of a smear centered at 4.0 kbp, indicating that the DHS-S1 in the fragment in the plasmid was digested with DNase I.

For comparison, the sensitivity of cellular DHS-S1 to DNase I digestion was examined similarly. Nuclei of HEK293 cells were extracted and digested with DNase I. Then, total DNA was extracted, cleaved with *Bam*HI, and electrophoresed. The *Bam*HI fragment derived from cellular AAVS1 was detected with a probe hybridizing with the region from the *Bam*HI site (nt -1k) to the *Eco*RI site (nt 1) (Fig. 3A, lower panel). A smear centered at 1.5 kbp emerged after the digestion with DNase I. The results agree with previous data (18).

The DHS of the chicken β -globin gene, *cHS4*, in pEnh-bl1

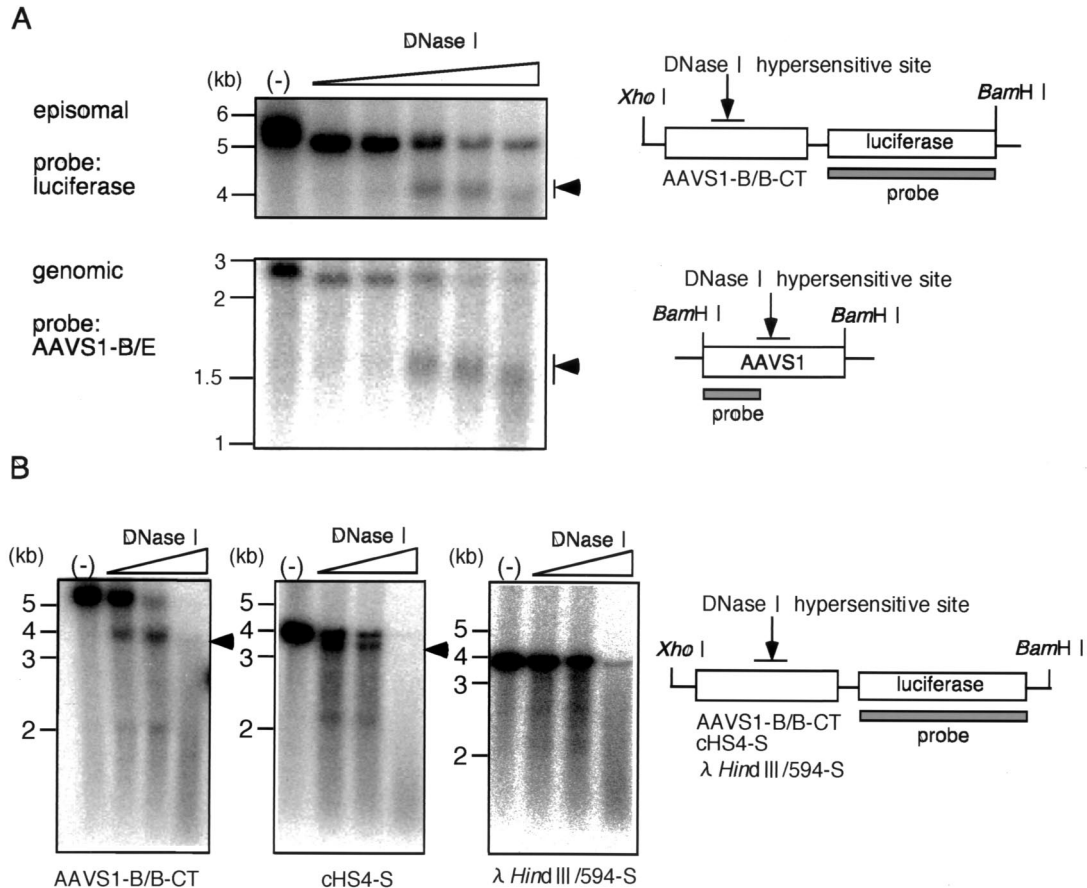


FIG. 3. DNase I hypersensitivity test. (A) DNase I hypersensitivity of DHS-S1 on the nuclear episomal plasmids replicated in HEK293 cells and in the AAVS1 locus of HEK293 cells. Nuclei were extracted from HEK293 cells transfected with pEnh-b11 containing the AAVS1-B/B fragment or from normal HEK293 cells. The nuclei were digested with 2.5, 5, 10, 20, and 40 U of DNase I per ml. Total DNA was isolated, digested with *Bam*HI plus *Xho*I (upper panel) or with *Bam*HI (lower panel), and electrophoresed in an agarose gel. After DNA samples were transferred to a nylon membrane, DNA fragments containing firefly luciferase gene (upper panel) or the AAVS1 region from the *Bam*HI site (nt -1k) to the DNase I hypersensitive site (lower panel) were detected by ³²P-labeled probes hybridizing with the DNA region indicated. (B) DNase I hypersensitivity of DHS-S1 and chicken β -globin insulator (cHS4) on the nuclear episomal plasmids replicated in HEK293 cells. Nuclei were extracted from HEK293 cells transfected with pEnh-b11 containing AAVS1-B/B, cHS4, and λ HindIII/594 fragments. The nuclei were digested with 10, 20, and 40 U of DNase I per ml. Total DNA was isolated, digested with *Bam*HI plus *Xho*I, and electrophoresed in an agarose gel. After DNA samples were transferred to a nylon membrane, DNA fragments containing firefly luciferase gene were detected by ³²P-labeled probes hybridizing to DNA carrying the luciferase gene.

was hypersensitive to DNase I digestion, but the 594-bp λ -HindIII fragment on pEnh-b11 was not hypersensitive (Fig. 3B, center and right panels). Thus, DNase I hypersensitivities of DHS-S1 and cHS4 were clearly retained on the EBV-based eukaryotic episomal plasmid, and the sensitivity of DHS-S1 to DNase I digestion seemed to be comparable to that of cHS4 (Fig. 3B, left and center panels). This strongly suggested that the chromatin conformation of DNA fragments on the EBV-based plasmid mimicked those in cellular chromosomes. The EBV vectors were used in the following enhancer-blocking assays.

Directional enhancer-blocking activity of AAVS1 fragments. Three AAVS1 fragments containing DHS-S1—AAVS1-B/B, AAVS1-E/B (nt 1 to 1605), and AAVS1-S/S (nt 18 to 353) (Fig. 1)—were subjected to the test for direction-dependent enhancer-blocking activities. The three AAVS1 fragments, cHS4, and a 594-bp λ -HindIII fragment were inserted in two directions between CMV-En and CMV-Pro of the pEnh-b12 plasmid,

which carries a *Renilla* luciferase expression unit upstream of CMV-Pro (Fig. 2A). The directions of the AAVS1 fragments are indicated as TC (from telomeric to centromeric side) or CT (from centromeric to telomeric side); the directions of cHS4 and the λ -HindIII 564-bp fragment are indicated as S (direction of transcription) or AS (direction against transcription). HEK293 cells were transfected with these plasmids and lysed at 120 h after the transfection. The firefly and *Renilla* luciferase activities of the lysate were measured, and to normalize plasmid copy numbers in the culture, a ratio of firefly luciferase activity to *Renilla* luciferase activity of each sample is presented (Fig. 4).

Comparison of the CT and TC directions showed that the three AAVS1 DNA fragments repressed the expression of firefly luciferase when placed in CT (Fig. 1), although the two longer fragments in TC, compared with AAVS1-S/S, apparently lowered the enhancer-promoter activity (Fig. 4). In these

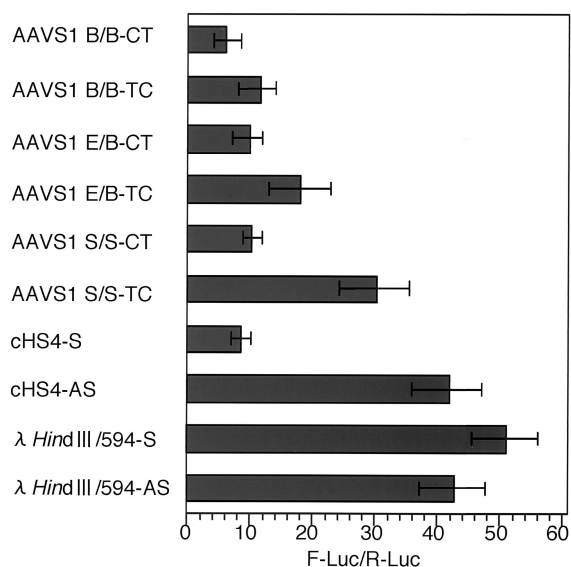


FIG. 4. The enhancer-blocking activity of AAVS1 fragments. HEK293 cells were transfected with pEnh-bl2 containing a test DNA fragment between the CMV enhancer and the CMV minimal promoter. At 120 h after the transfection, the firefly luciferase and *Renilla* luciferase activities of cell lysate were measured. To normalize plasmid copy numbers in the cells lysed, ratios of firefly luciferase activity to *Renilla* luciferase (F-Luc/R-Luc) activity were determined and are presented along with standard deviations of four independent experiments. The directions from the centromeric to telomeric side and from the telomeric to centromeric side of the test DNA fragments are indicated as CT and TC, respectively.

experiments, the insertion of cHS4 in the S direction showed the repression as previously reported (21) and insertion of the λ -HindIII 564-bp fragment in either direction did not affect the expression of firefly luciferase. The level of repression by AAVS1-S/S was comparable to that by cSH4. The data clearly indicate that, like cSH4, the AAVS1-S/S has a directional enhancer-blocking activity and that the activity is located in the region between two *Sma*I sites (nt 18 to 353), whose nucleotide sequence is identical to the corresponding sequence of AC010327 (GenBank).

Expression of a transgene-expressing cassette flanked by two AAVS1-S/S fragments. The property of an insulator to prevent the spread of heterochromatin was tested with AAVS1-S/S, and for this test a plasmid with an expression cassette for the firefly luciferase gene driven by the simian virus 40 (SV40) early promoter and enhancer was newly constructed (Fig. 2B). Three pairs of the AAVS1-S/S, cHS4, and λ -HindIII 564-bp fragment were each inserted at *Kpn*I site upstream of the promoter and between the *Bam*HI and *Sal*I sites downstream of the SV40 poly(A) site in the plasmid to generate pSV-Luc-AAVS1-S/S-CT, pSV-Luc-cHS4-S, and pSV-Luc- λ /HindIII 564-S, respectively. In these plasmids, the AAVS1-S/S or cHS4 fragments flanking the cassette were placed in the direction positive for the enhancer-blocking activities (Fig. 4). Levels of the transient expression of firefly luciferase were higher in HEK293 cells transfected with pSV-Luc-AAVS1-S/S-CT and pSV-Luc-cHS4-S than in those transfected with pSV-luc- λ /HindIII 564-S (Fig. 5A).

The AAVS1-S/S fragments flanking the cassette protected

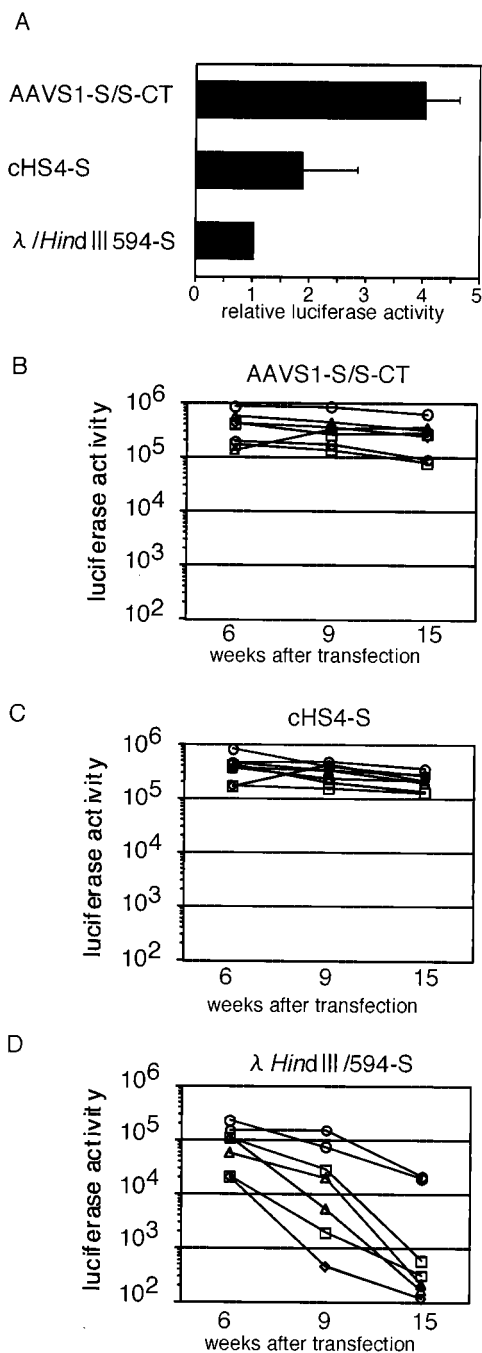


FIG. 5. Expression of firefly luciferase from the expression cassette shown in Fig. 2B. (A) Transient expression from the cassette flanked by the test DNA fragments. HEK293 cells were transfected with the linearized plasmid containing the test cassette. At 48 h after the transfection, the luciferase activity of cell lysate was measured. Results are presented as the mean and standard deviation of three independent experiments. (B to D) Expression from the cassette integrated in cell DNA. HEK293 cells were transfected with the linearized plasmid containing the cassette flanked by the test DNA fragments together with a plasmid expressing the G418 resistance gene at a molar ratio of 4:1. Cell colonies resistant to G418 were obtained by culturing the cells with growth medium containing G418 (800 mg/ml). Eight cell clones were randomly selected and cultured with growth medium without G418. Firefly luciferase activities of cell lysates were measured at 6, 9, and 15 weeks after the transfection.

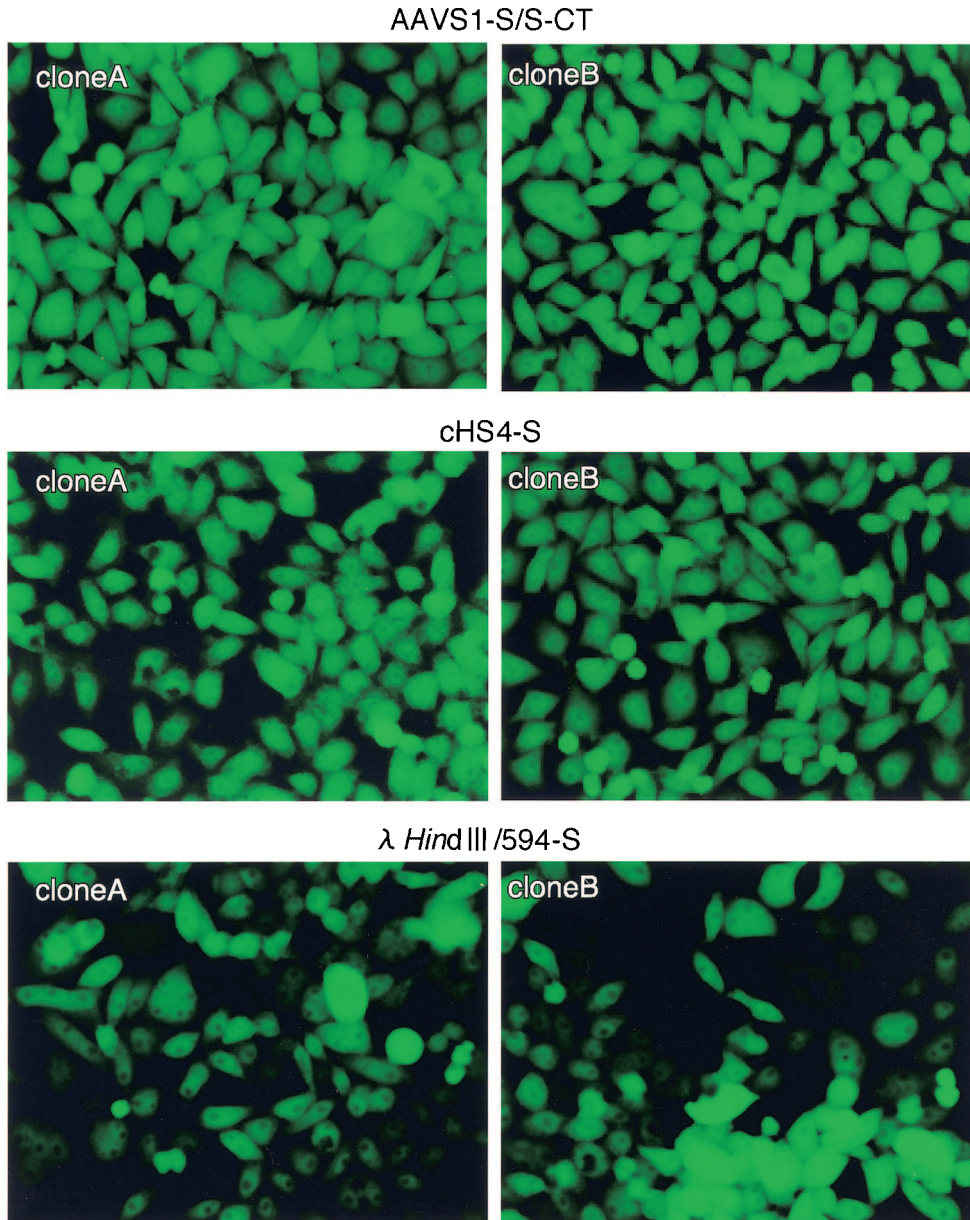


FIG. 6. Expression of EYFP in HeLa cells stably transformed with the expression cassette shown in Fig. 2B. HeLa cells were transfected with the linearized plasmids with the cassette flanked by the test DNA fragments together with a plasmid expressing the G418 resistance gene at a molar ratio of 4:1. Cell colonies resistant to G418 were obtained by culturing the cells with growth medium containing G418 (500 mg/ml). At 4 weeks after the transfection, fluorescence of the cell colonies was observed under the UV microscope. Typical photos of two cell clones (clones A and B) containing the test cassette are presented.

the expression of the reporter gene from chromosomal position effects. HEK293 cells were transfected with the expression plasmids together with a plasmid expressing the neomycin resistance gene. Eight (pSV-Luc-AAVS1-S/S-CT and pSV-Luc-cHS4-S) or seven (pSV-Luc- λ HindIII 564-S) drug-resistant cell clones were randomly selected at 2 weeks after the transfection and cultured with growth medium without G418. When the cells became confluent, the cultures were passaged at a split ratio of 1:4. Luciferase activities of cellular extracts were measured at 6, 9, and 15 weeks after the transfection (Fig. 5B to D). The luciferase activities in the clones with pSV-Luc- λ /

HindIII 564-S decreased; those in five clones decreased quickly, and those in two clones decreased less quickly (Fig. 5D). However, the luciferase activities in the clones with pSV-Luc-AAVS1-S/S-CT and pSV-Luc-cHS4-S were maintained at high levels (Fig. 5B and C).

To determine the levels of reporter gene expression in individual cells, the reporter gene in the plasmid was changed from the firefly luciferase gene to the EYFP gene to generate pSV-EYFP-AAVS1-S/S-CT, pSV-EYFP-cDHS4-S, and pSV-EYFP- λ HindIII 564-S. HeLa cells were transfected with the plasmids, which had been linearized by cleavage with *SspI*,

together with a plasmid expressing the neomycin resistance gene. Drug-resistant cell clones were randomly selected and cultured with growth medium without G418. When the cells became confluent, the cultures were passaged at a split ratio of 1:4. At 4 weeks after transfection, the fluorescence of the cells was observed under an inverted UV microscope (Fig. 6). Although the levels of fluorescence in the cells with pSV-EYFP- λ /*Hind*III 564-S varied from cell to cell, those in the cells with pSV-EYFP-AAVS1-S/S-CT and pSV-EYFP-cHS4-S were uniform and high.

Binding of CTCF with the AAVS1-S/S fragment. The AAVS1-S/S fragment was examined for its ability to bind CTCF, a protein that specifically binds to the chicken insulator cHS4. CTCF was produced by using a reticulocyte lysate in vitro (Fig. 7A). Despite the presence of several CTCF-binding motifs (CTCCC) within the fragment, a gel shift test did not detect the complex of CTCF and the ³²P-labeled fragment (Fig. 7B). Under the same experimental conditions, the complex of CTCF and ³²P-labeled cHS4 was detected as a broader band migrating more slowly than the uncomplexed DNA in electrophoresis (Fig. 7B). Addition of cold cHS4 DNA to the reaction mixture as a competitor eliminated the slowly moving part of the band, indicating that the experimental conditions were appropriate for detecting specific binding of CTCF with cHS4. The results suggest that there may be an unidentified protein that specifically binds to the AAVS1 insulator.

DISCUSSION

This study showed that an insulator is present in the region from nt 18 to 353 in AAVS1 or in the AAVS1-S/S fragment derived from the preintegration site. When the fragment was placed between the CMV minimal promoter and immediately early enhancer, the action of the enhancer was blocked in a direction-dependent manner. This directional enhancer-blocking activity is common to all of the insulators analyzed so far (7, 9). When the two fragments were placed to flank a transgene-expressing cassette, it continuously expressed the transgene irrespective of its integration site in cellular chromosomes. This action to buffer a gene from the repressing effects of heterochromatin is another property common to insulators. Thus, it was concluded that the newly identified AAVS1 insulator has typical insulator properties and is associated with DHS-S1 located within the region from nt 50 to 350 (20).

When insulator activity-negative DNA fragments, such as AAVS1 fragments with the TC orientation and λ /*Hind*III 564, were inserted between the enhancer and promoter of the episomal plasmids, the reporter activity in the transfected cells was apparently affected by the size of the inserted DNA; firefly luciferase activities from the insertion of *Bam*HI-*Bam*HI-TC (2.6 kbp) and *Bam*HI-*Eco*RI-TC (1.6 kbp) were lower than those from the insertion of *Sma*I-*Sma*I-TC (0.3 kb) and λ /*Hind*III564 (0.6 kbp) (Fig. 4). It is possible that the insertion of a relatively long DNA fragment between the enhancer and promoter in the episomal plasmid may hinder the enhancer-promoter activity by increasing the distance or introducing a repressive sequence between the two.

Although its role in normal cells is unclear, it is conceivable that the AAVS1 insulator has been playing significant roles in

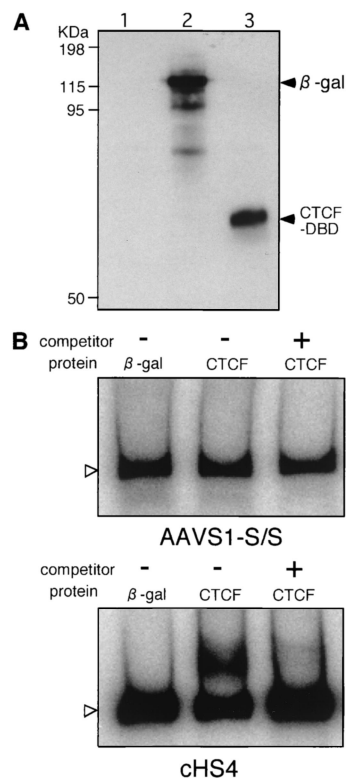


FIG. 7. Electrophoretic mobility shift test to examine the binding of CTCF to AAVS1-S/S. (A) Immunoblot detection of CTCF and β -galactosidase produced in vitro. HA-tagged CTCF-DNA-binding domain (CTCF-DB) and β -galactosidase were produced by using the rabbit reticulocyte system (TNT coupled transcription/translation system [Promega Corp.]). The reticulocyte lysate (0.3 μ l) containing synthesized CTCF-DB and β -galactosidase was electrophoresed in a 10% polyacrylamide gel, and the proteins were transferred onto Hybond-P membranes (Amersham Pharmacia Biotech). The membrane was incubated with rat monoclonal anti-HA antibody (clone 3F10; Roche Diagnostics GmbH) and then with peroxidase-conjugated goat anti-rat immunoglobulin G (Organon Teknica Corp., Durham, N.C.). Immunoreactive proteins were visualized by the ECL Plus chemiluminescence detection system (Amersham Pharmacia Biotech). (B) Electrophoretic mobility shift test. The AAVS1-S/S fragment and cHS4 fragment were labeled with ³²P by using T4 DNA kinase and used as probes. CTCF-DB and β -galactosidase produced in vitro were mixed with the probe. The probes complexing with CTCF-DB or β -galactosidase were separated by electrophoresis on a 4% polyacrylamide gel. Cold DNA fragments was added to the binding-reaction mixture at a 50-fold excess over the labeled probe for competition. Open arrowheads indicated free labeled probes.

the evolution and life cycle of AAV. The open chromatin structure in the preintegration site, caused by the AAVS1 insulator along with its probable counterpart somewhere in the distal side, may help the integration and rescue of AAV. For the integration, the open structure would allow AAV Rep proteins and viral DNA to have ready access to the Rep-binding site (nt 396 to 413) and the site of recombination, respectively, in the preintegration site. For the rescue, the open structure would allow the provirus to readily express its genes for replication on infection with a helper virus.

Because CTCF appeared not to bind to the AAVS1 insulator in the electrophoretic mobility shift test in this study, it is

possible that there is an unidentified binding protein for the insulator. The AAVS1 DNA of 3.5 kbp containing the insulator, when introduced into transgenic rats, is proficient for Rep-mediated site-specific integration of AAV (20). This suggests that the insulator must be active in rats and that its binding protein may be present in rat cells. Experiments to search for the binding protein are under way.

ACKNOWLEDGMENTS

We thank Kunito Yoshiike for critical reading of the manuscript. This work was supported by a grant-in-aid from the Ministry of Health and Welfare for the Second-Term Comprehensive 10-Year Strategy for Cancer Control and for the Research on Human Genome and Gene Therapy and a grant-in-aid for Scientific Research on Priority Areas (C) from the Ministry of Education, Culture, Sports, Science and Technology.

REFERENCES

1. Bell, A. C., A. G. West, and G. Felsenfeld. 1999. The protein CTCF is required for the enhancer blocking activity of vertebrate insulators. *Cell* **98**:387–396.
2. Berns, K. I., and R. A. Bohenzky. 1987. Adeno-associated viruses: an update. *Adv. Virus Res.* **32**:243–306.
3. Berns, K. I., and C. Giraud. 1996. Biology of adeno-associated virus. *Curr. Top. Microbiol. Immunol.* **218**:1–23.
4. Buller, R. M., J. E. Janik, E. D. Sebring, and J. A. Rose. 1981. Herpes simplex virus types 1 and 2 completely help adenovirus-associated virus replication. *J. Virol.* **40**:241–247.
5. Burgess-Beusse, B., C. Farrell, M. Gaszner, M. Litt, V. Mutskov, F. Recillas-Targa, M. Simpson, A. West, and G. Felsenfeld. 2002. The insulation of genes from external enhancers and silencing chromatin. *Proc. Natl. Acad. Sci. USA* **99**:16433–16437.
6. Carter, B. J., C. A. Laughlin, L. M. de la Maza, and M. Myers. 1979. Adeno-associated virus autointerference. *Virology* **92**:449–462.
7. Chen, S., and V. G. Corces. 2001. The Gypsy insulator of *Drosophila* affects chromatin structure in a directional manner. *Genetics* **159**:1649–1658.
8. Cheung, A. K., M. D. Hoggan, W. W. Hauswirth, and K. I. Berns. 1980. Integration of the adeno-associated virus genome into cellular DNA in latently infected human Detroit 6 cells. *J. Virol.* **33**:739–748.
9. Chung, J. H., A. C. Bell, and G. Felsenfeld. 1997. Characterization of the chicken β -globin insulator. *Proc. Natl. Acad. Sci. USA* **94**:575–580.
10. Georg-Fries, B., S. Biederlack, J. Wolf, and H. zur Hausen. 1984. Analysis of proteins, helper dependence, and seroepidemiology of a new human parvovirus. *Virology* **134**:64–71.
11. Giraud, C., E. Winocour, and K. I. Berns. 1994. Site-specific integration by adeno-associated virus is directed by a cellular DNA sequence. *Proc. Natl. Acad. Sci. USA* **91**:10039–10043.
12. Giraud, C., E. Winocour, and K. I. Berns. 1995. Recombinant junctions formed by site-specific integration of adeno-associated virus into an episome. *J. Virol.* **69**:6917–6924.
13. Gross, D. S., and W. T. Garrard. 1988. Nuclease hypersensitive sites in chromatin. *Annu. Rev. Biochem.* **57**:159–197.
14. Handa, H., K. Shiroki, and H. Shimojo. 1977. Establishment and characterization of KB cell lines latently infected with adeno-associated virus type 1. *Virology* **82**:84–92.
15. Huser, D., S. Weger, and R. Heilbronn. 2002. Kinetics and frequency of adeno-associated virus site-specific integration into human chromosome 19 monitored by quantitative real-time PCR. *J. Virol.* **76**:7554–7559.

16. Kalos, M., and R. E. K. Fournier. 1995. Position-independent transgene expression mediated by boundary elements from the apolipoprotein B chromatin domain. *Mol. Cell. Biol.* **15**:198–207.
17. Kotin, R. M., R. M. Linden, and K. I. Berns. 1992. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *EMBO J.* **11**:5071–5076.
18. Kotin, R. M., J. C. Menninger, D. C. Ward, and K. I. Berns. 1991. Mapping and direct visualization of a region-specific viral DNA integration site on chromosome 19q13-qter. *Genomics* **10**:831–834.
19. Kotin, R. M., M. Siniscalco, R. J. Samulski, X. Zhu, L. Hunter, C. A. Laughlin, S. McLaughlin, N. Muzyczka, M. Rocchi, and K. I. Berns. 1990. Site-specific integration by adeno-associated virus. *Proc. Natl. Acad. Sci. USA* **87**:2211–2215.
20. Lamartina, S., E. Sporeno, E. Fattori, and C. Toniatti. 2000. Characteristics of the adeno-associated virus preintegration site in human chromosome 19: open chromatin conformation and transcription-competent environment. *J. Virol.* **74**:7671–7677.
21. Li, Q., and G. Stamatoyannopoulos. 1994. Hypersensitive site 5 of the human β locus control region functions as a chromatin insulator. *Blood* **84**:1399–1401.
22. McPherson, R. A., L. J. Rosenthal, and J. A. Rose. 1985. Human cytomegalovirus completely helps adeno-associated virus replication. *Virology* **147**:217–222.
23. Mishra, L., and J. A. Rose. 1990. Adeno-associated virus DNA replication is induced by genes that are essential for HSV-1 DNA synthesis. *Virology* **179**:632–639.
24. Mori, S., M. Murakami, T. Takeuchi, T. Kozuka, and T. Kanda. 2002. Rescue of AAV by antibody-induced Fas-mediated apoptosis from viral DNA integrated in HeLa chromosome. *Virology* **301**:90–98.
25. Mutskov, V. J., C. M. Ferrell, P. A. Wade, A. P. Wolffe, and G. Felsenfeld. 2002. The barrier function of an insulator couples high histone acetylation levels with specific protection of promoter DNA from methylation. *Genes Dev.* **16**:1540–1554.
26. Pikaart, M. J., F. Recillas-Targa, and G. Felsenfeld. 1998. Loss of transcriptional activity of a transgene is accompanied by DNA methylation and histone deacetylation and is prevented by insulators. *Genes Dev.* **12**:2852–2862.
27. Richardson, W. D., and H. Westphal. 1981. A cascade of adenovirus early functions is required for expression of adeno-associated virus. *Cell* **27**:133–141.
28. Roeder, G. S. 1997. Meiotic chromosomes: it takes two to tango. *Genes Dev.* **11**:2600–2621.
29. Saitoh, N., A. C. Bell, F. Recilla-Targa, A. G. West, M. Simpson, M. Pikaart, and G. Felsenfeld. 2000. Structural and functional conservation at the boundaries of the chicken β -globin domain. *EMBO J.* **19**:2315–2322.
30. Surosky, R. T., M. Urabe, S. G. Godwin, S. A. McQuiston, G. J. Kurtzman, K. Ozawa, and G. Natsoulis. 1997. Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome. *J. Virol.* **71**:7951–7959.
31. Weindler, F. W., and R. Heilbronn. 1991. A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication. *J. Virol.* **65**:2476–2483.
32. West, M. H., J. P. Trempe, J. D. Tratschin, and B. J. Carter. 1987. Gene expression in adeno-associated virus vectors: the effects of chimeric mRNA structure, helper virus, and adenovirus VA1 RNA. *Virology* **160**:38–47.
33. Yakinoglu, A. O., R. Heilbronn, A. Burkle, J. R. Schlehofer, and H. zur Hausen. 1988. DNA amplification of adeno-associated virus as a response to cellular genotoxic stress. *Cancer Res.* **48**:3123–3129.
34. Yates, J. L., and N. Guan. 1991. Epstein-Barr virus-derived plasmids replicate only once per cell cycle and are not amplified after entry into cells. *J. Virol.* **65**:483–488.
35. Zaret, K. 1999. In vivo analysis of chromatin structure. 1999. *Methods Enzymol.* **304**:612–626.