Cloning of an Avian Adeno-Associated Virus (AAAV) and Generation of Recombinant AAAV Particles

Ioannis Bossis and John A. Chiorini*

Gene Therapy and Therapeutics Branch, NIDCR, NIH, Bethesda, Maryland 20892

Received 2 December 2002/Accepted 28 March 2003

Recent studies have proposed that adeno-associated viruses (AAVs) are not evolutionarily linked to other mammalian autonomous parvoviruses but are more closely linked to the autonomous parvoviruses of birds. To better understand the relationship between primate and avian AAVs (AAAVs), we cloned and sequenced the genome of an AAAV (ATCC VR-865) and generated recombinant AAAV particles. The genome of AAAV is 4,694 nucleotides in length and has organization similar to that of other AAVs. The entire genome of AAAV displays 56 to 65% identity at the nucleotide level with the other known AAVs. The AAAV genome has inverted terminal repeats of 142 nucleotides, with the first 122 forming the characteristic T-shaped palindromic structure. The putative Rep-binding element consists of a tandem (GAGY)₄ repeat, and the putative terminal resolution site (trs), CCGGT/CG, contains a single nucleotide substitution relative to the AAV₂ trs. The Rep open reading frame of AAAV displays 50 to 54% identity at the amino acid level with the other AAVs, with most of the diversity clustered at the carboxyl and amino termini. Comparison of the capsid proteins of AAAV and the primate dependoviruses indicate that divergent regions are localized to surface-exposed loops. Despite these sequence differences, we were able to produce recombinant AAAV particles carrying a *lacZ* reporter gene by cotransfection in 293T cells and were able to examine transduction efficiency in both chicken primary cells and several cell lines. Our findings indicate that AAAV is the most divergent AAV described to date but maintains all the characteristics unique to the genera of dependovirus.

Adeno-associated viruses (AAVs) are small, nonpathogenic parvoviruses that require coinfection with a helper virus, such as adenovirus or herpesvirus, for productive infection (2). To date, eight AAV isolates (AAV types 1 to 8 [AAV1 to -8]) have been characterized and sequenced (2, 4, 19, 20, 25, 32, 51, 56), with AAV2 having been the most extensively studied.

AAV virions are approximately 20 to 25 nm in diameter and are composed of a mixture of assembled proteins (VPs) that encapsidate a linear ~4.7-kb single-stranded DNA (ssDNA) of plus or minus polarity (7, 43). The genome of AAVs is flanked by inverted terminal repeats (ITRs), which in the case of AAV₂ are 145 nucleotides in length. The ITR is organized as three interrupted palindromes that can fold in an energetically favored T-shaped hairpin structure, which can exist in two orientations, termed flip and flop (42). The ITRs serve as origin of replication and contain *cis*-acting elements required for rescue, integration, excision from cloning vectors, and packaging (41, 42, 49, 58).

The genetic map of the AAVs has been derived primarily from studies of AAV2 but is conserved in all serotypes (26, 27, 29, 36, 42, 45, 46, 58, 60, 64). Two major open reading frames (*rep* and *cap* ORFs) and three transcriptional active promoters (P_5 , P_{19} , and P_{40}) have been identified in the genome of AAV2. The P_5 and P_{19} promoters encode the nonstructural replication proteins Rep78 and Rep68 and Rep52 and Rep40, respectively. Due to differential splicing, Rep78 and Rep52 have different C termini from Rep68 and Rep40. Transcription initiation from two promoters results in Rep78 and Rep68 having different N termini from Rep52 and Rep40. The P_{40} promoter transcribes two alternatively spliced mRNAs. The major mRNA species encodes the major capsid protein VP3 from a conventional AUG codon and the minor capsid protein VP2 from an upstream in-frame ACG codon. The minor mRNA species encodes the entire *cap* ORF to produce the minor capsid protein VP1 (47). VP1, VP2, and VP3 are found in a ratio of 1 to 1 to 10, respectively, and this stoichiometry is generated by the high abundance of one of the mRNA species and the low translation efficiency from an ACG codon in the case of VP2 (14, 47, 55). Previous studies have indicated that VP2 and VP3 are sufficient for particle formation and accumulation of encapsidated ssDNA progeny, while VP1 is required for assembly of highly infectious particles (63, 64).

All four Rep proteins possess NTP binding activity, DNA helicase activity, and nuclear localization sequences; however, only Rep78and Rep68 possess DNA binding ability (33, 34, 66). Mutant AAVs defective for the synthesis of the small Rep proteins (Rep52 and Rep40) are able to replicate DNA, but no ssDNA progeny is encapsidated (16). The ability of Rep78 and Rep68 to bind and nick DNA in a sequence- and strand-specific manner inside the ITR is essential in every phase of the AAV life cycle, namely, DNA replication, AAV gene expression, rescue from the integrated state, and self-excision from cloning vectors (29, 35, 44). Nicking of the DNA within the ITR at the terminal resolution site (*trs*) requires binding of Rep78 and Rep68 proteins to a motif composed of tandem repeats of GAGY.

Among AAV serotypes, AAV1, AAV4, AAV7, and AAV8 are believed to be of simian origin, while AAV2, AAV3, and AAV5 are from humans. AAV6 was found in a human adenovirus preparation and is very similar to AAV1. AAVs have

^{*} Corresponding author. Mailing address: NIH 10/1N113, 10 Center Dr., MSC 1190, Bethesda, MD 20892-1190. Phone: (301) 496-4279. Fax: (301) 402-1228. E-mail: Jchiorini@dir.nidcr.nih.gov.

also been reported in other mammalian species including canines, bovines, ovines, and equines (8). An avian AAV (AAAV) was first isolated from the Olson strain of quail bronchitis adenovirus (68). It was later found that 50% of adenoviral field isolates from chickens in the United States and Ireland contained AAAVs serologically indistinguishable from the initial isolate (24). The AAAV was found to be 20 nm in diameter, was serologically distinct from AAV1-4, did not agglutinate erythrocytes from several species tested, and required adenovirus or herpesvirus for replication (5, 68). In addition, AAAV was found to inhibit replication of several avian adenoviruses and herpesviruses (5, 52, 53). Physicochemical studies revealed that the capsid of AAAV consists of three VP proteins similar to those of other AAVs. The buoyant density of AAAV in CsCl gradients (1.39 to 1.44 g/cm³) is similar to what has been reported for all AAVs (6, 30, 68). One study (30) also provided a limited restriction endonuclease map of AAAV.

The ability of AAV vectors to infect dividing and nondividing cells and establish long-term transgene expression and the lack of pathogenicity have made them attractive for use in gene therapy applications. Recent evidence has indicated lack of cross competition in binding experiments, suggesting that each AAV serotype may have a distinct mechanism of cell entry. Comparison of the *cap* ORFs from different serotypes has identified blocks of conserved and divergent sequences, with most of the latter residing on the exterior of the virion, thus explaining the altered tissue tropism among serotypes (19–21, 48, 56). Vectors based on new AAV serotypes may have different host ranges and different immunological properties, thus allowing for most efficient transduction in certain cell types. In addition, characterization of new serotypes will aid in identifying viral elements required for altered tissue tropism.

Serological studies have provided evidence of AAAV infection in humans (69). Six percent of an unselected adult population was found positive for antibody to AAAV by agar gel precipitation (AGP), and 15.6% was positive by virus neutralization. Fourteen percent of poultry workers (industry or research) were positive for AAAV antibody by AGP and 66% were positive by virus neutralization. In the same studies, no cross-reaction was noted by AGP when antiserum to AAAV was reacted against primate antigens of serotypes 1 to 4 or when antiserum to AAV serotypes 1 to 4 were reacted against AAAV antigen. In addition, antiserum prepared against primate AAV1-4 did not neutralize the AAAV. These results suggest that AAAV is a distinct serotype and that infections are not restricted to avian species but are found in the adult human population.

Based on the genome organization and sequence homology among insect densovirus, rodent parvovirus, and human dependovirus, it has been previously proposed that these viruses may have diverged from a common ancestor and evolved strictly in their hosts (3). However, the high sequence homology between avian autonomous parvovirus and primate AAVs and the epidemiological documentation of AAAV transmission to humans provide evidence for host-independent evolution of at least some parvovirus genera. To better understand the relationship between the avian and the primate AAVs, the complete viral genome of AAAV was cloned and sequenced and used to generate recombinant viral particles.

MATERIALS AND METHODS

Cell culture and virus propagation. 293T and COS cells were maintained in Iscove's modified Eagle medium and AMEM, respectively DF1 cells (spontaneously immortalized chicken embryonic fibroblasts), QNR cells (quail neuroretinal cells), A549 cells, and primary chicken embryonic fibroblasts (CEF) were maintained in Dulbecco's modified Eagle medium (DMEM) supplemented with 10% fetal bovine serum (FBS). Primary chicken embryonic kidney cells were maintained in β-mercaptoethanol supplemented with 10% fetal bovine serum (FBS). Primary chicken pituitary cells were maintained in DMEM supplemented with 5% horse serum. QT6 cells (quail fibrosarcoma) were maintained in Ham's F12K supplemented with 10% FBS. LMH cells (chicken hepatoma cells) were maintained in Waymouth's medium supplemented with 10% FBS. DT-90 (chicken lymphoblastoma) cells were maintained in DMEM supplemented with 15% FBS, 5% chicken serum, and 0.015% β-mercaptoethanol. Human primary fibroblasts were obtained from Clonetics and maintained in serum-free propertiary medium supplied by the manufacturer. AAAV (ATCC VR-865) was propagated in 10-day-old Spafas pathogen-free embryonated chicken eggs coinfected with the Phelps strain of fowl adenovirus type 1 (FAV1; ATCC VR-486). AAAV at 10⁴ to 10⁷ and FAV1 at 10⁵ infectious particles in saline were simultaneously injected into the chorioallantoic cavity of eggs and incubated for 96 h at 37°C. At the end of the incubation, allantoamniotic fluids were harvested and clarified by centrifugation at $6,000 \times g$ for 10 min.

Viral DNA isolation, cloning, and sequencing. Virus from infected clarified allantoamniotic fluids was precipitated by centrifugation at $100,000 \times g$ for 2 h. The supernatant was discharged, and the virus-containing pellet was resuspended in proteinase K digestion buffer (50 mM Tris [pH 8], 20 mM EDTA, 0.5% sodium dodecyl sulfate [SDS], 200 µg of proteinase K per ml) and incubated at 45°C for 2 h. Following a phenol-chloroform extraction and ethanol precipitation, the viral DNA was resuspended in Tris-EDTA buffer containing 0.1 M NaCl. The single-stranded viral DNA was annealed by heating to 95°C for 5 min followed by slow cooling to 65°C for 6 h. The annealed viral DNA was separated electrophoretically in 1% agarose gel, and the double-stranded AAAV DNA of approximately 4.7 kb was excised and purified with a gel extraction kit (Qiagen). The viral DNA was further processed to fill in the ends by treating with DNA polymerase (Klenow fragment) at 37°C for 15 min in the presence of deoxynucleoside triphosphates. The whole genome was then blunt end cloned in the pPCR-script cloning vector containing the lacZ gene, allowing blue-white screening of ampicillin-resistant colonies (Stratagene). Colonies that contained large inserts (4.7 kb) were initially screened by restriction digestion, and three clones were selected for sequencing. No sequence differences were found in these three clones. The sequence of the entire genome (except ITRs) was determined by using an ABI 373A automated sequencer and FS dye terminator chemistry (ABI). Due to the high degree of secondary structure, ITRs were sequenced by isothermal noncycling sequencing chemistry by using radiolabeled dCTP (Epicentre). One of the clones (pAAAV) that contained the entire consensus sequence of AAAV was further used to generate packaging and vector plasmids for construction of recombinant AAAV (rAAAV) virus. The complete DNA sequence of AAAV has been submitted to GenBank (accession number AY186198).

Sequence analysis. DNA and protein sequence alignments were performed by using the Clustal W multiple sequence alignment tool of the Biology Workbench web-based software (SDSC). Promoter, transcription initiation, and splice sites were predicted by using the Neural Network Promoter Prediction web-based software (BDGP). The presence of potential transcription binding sites was analyzed with the MaIInspector computer program (54). Putative motifs in the Rep proteins were identified with the BLIMPS program that searches for motifs in the Blocks protein database (28).

Southern blot hybridization. The ability of pAAAV to support self-excision, packaging, and generation of nuclease-resistant wild-type (wt) AAAV particles was examined. 293T cells seeded in 6-well plates were transfected by using calcium phosphate coprecipitation with pAAAV alone, pAAAV plus pAd12 (a helper plasmid containing the E2 and E4 ORFs and VA RNAs of Ad5), and pAAAV plus infection with Ad5. In addition, LMH cells seeded in gelatin-coated 6-well plates were similarly transfected with pAAAV alone or with pAAAV plus infection with FAV1. After 48 h, clarified lysates were prepared by using three freeze-thaw cycles and centrifugation at 3,800 × g for 20 min. The lysate (~100 μ l) was treated with 5 U of DNase for 2 h at 37°C to remove vector and unpackaged progeny. Subsequently, the solutions were adjusted to contain 20 mM EDTA (pH 8), 0.5% SDS, and 200 μ g proteinase K per ml and incubated at 45°C for 2 h. After one phenol-chloroform extraction, nucleic acids were precipitated with the addition of an equal volume of isopropanol, and the pellet was resuspended in 30 μ l of Tris-EDTA buffer containing 0.1 M NaCl. The

samples were heated to 95°C for 5 min, slowly cooled down to 65°C, and incubated for 5 h. After electrophoresis and blotting, the membrane was probed with a ³²P-labeled 1.2-kb *Bam*HI fragment of pAAAV.

Generation of recombinant AAAV particles. For production of recombinant particles, we generated and examined the efficiency of three different helper plasmids, pMA₃VRC, pCA₃VRC, and pA₃VRC, containing the AAAV rep and cap genes under control of a mouse mammary tumor virus (MMTV), cytomegalovirus (CMV), or native p5 promoter, respectively. For generation of pMA₃VRC, the rep and cap ORFs (nucleotides 243 to 4482) was produced by PCR with PFU polymerase (Stratagene) as specified by the manufacturer by using primers containing BstZ107 and NotI sites. The PCR products were digested with BstZ107 and NotI and ligated in a BstZ107/NotI fragment of pMMTV2.1 (18) containing an MMTV promoter and SV40 poly(A). For generation of pCA3VRC, the rep and cap ORFs (nucleotides 243 to 4482) were produced by PCR with PFU polymerase and blunt-end ligated in the pCMVscript (Stratagene) vector, which contains a CMV promoter and SV40 poly(A). For generation of pA3VRC, the rep and cap genes of AAAV including the p5 promoter and poly(A) signal (nucleotides 142 to 4516) was produced by PCR using PFU polymerase and blunt end ligated in pPCR-script. Orientation of inserts was verified by restriction digestion analysis, and final clones were confirmed by sequencing. For generation of the vector carrying the β-galactosidase gene flanked by AAAV ITRs, the plasmid pAAAV was digested with BsmBI (NEB). BsmBI does not cut in the plasmid backbone but cuts at positions 838, 1111, 2590, 4419, and 4530 of the AAAV genome. The resulting fragment that contained the plasmid backbone and 700 bp of AAAV genome flanked by ITRs was used to ligate a BsmBI-BsmI linker. The resulting plasmid was digested with Pml1 (cuts at nucleotide 146 of AAAV genome) and BsmI and used to ligate a BstZ107-BsmI fragment of pAAV2RnLacZ (18) that contains the β-galactosidase gene under the control of an RSV promoter and SV40 poly(A) tail. The resulting plasmid (pA3VRSVβ-Gal) was cotransfected with one of the helper plasmids described above and pAd12 in 293T cells plated in 150-cm dishes. Forty-eight hours posttransfection, cells were harvested and quantitated with a hemacytometer, and rAAAV was prepared by using standard CsCl gradient purification. The number of rAAAV genomes was estimated by using real-time quantitative PCR (QPCR) and expressed as nuclease-resistant particles per cell recovered after transfections. Titration of rAAAV was performed in exponentially growing CEF, DF-1, LMH, QNR, QT6, DT-90, 293T, COS, and primary embryonic chicken kidney cells and nondividing primary pituitary cells plated in 96-well plates and transduced with serial dilutions of recombinant virus for 48 h as previously described (20).

RESULTS

To obtain AAAV genomic DNA for cloning, a stock of AAAV was obtained from ATCC (VR-865) and coinfected with FAV1 in day 10 embryonated chicken eggs. Virus was concentrated after subjecting infected allantoamniotic fluids to high-speed centrifugation. Viral DNA was released by SDSproteinase K digestion and purified by gel electrophoresis after annealing the complementary single strands by heating the purified DNA to 95°C and slowly cooling to 65°C. Preliminary experiments indicated that 10⁵ infectious particles of FAV1 resulted in productive infection without killing the embryo prematurely. Coinfection with at least 10⁵ infectious particles of AAAV was required to detect viral DNA (\sim 4.7 kb) by ethidium bromide staining (data not shown). After recovery and end filling, the double-stranded AAAV genome was blunt end ligated and cloned into pPCR-script. Several clones that contained an insert of approximately 4.7 kb were initially screened by restriction digestion (data not shown) and all gave bands similar in size to those previously reported (30). We subsequently sequenced three of these clones and all gave identical sequences. One of the clones was randomly selected and used in subsequent analysis (pAAAV).

To verify that pAAAV can support self-excision, viral DNA replication, and packaging in mammalian and avian cells, we prepared viral lysates from 293T and LMH cells transfected



FIG. 1. Southern blot analysis of AAAV nuclease-resistant particles in 293T and LMH cells. (A) 293T cells were transfected with pAAAV alone (lane 3), pAAAV plus pAd12 (lane 2), and pAAAV plus infection with wt Ad (lane 1). (B) LMH cells were transfected with pAAAV alone (lane 2) or pAAAV plus infection with FAV1 (lane 1). Viral DNA was isolated as described in Materials and Methods and fractionated on agarose gel before Southern blot analysis with a ³²P-labeled pAAAV DNA.

with pAAAV and infected with wild-type (wt) Ad5 or FAV1, respectively. In addition, the ability of an Ad5 plasmid to provide helper functions was examined in 293T cells. Southern blot analysis showed encapsidated (nuclease-resistant particles) AAAV progeny in the presence of wt Ad5 or Ad helper plasmid in 293T cells and FAV1 in LMH cells but not in the absence of these (Fig. 1A and B). This result suggests that pAAAV can support rescue of AAAV in mammalian and avian cells in the presence of mammalian or avian adenoviral genes.

The AAAV ITR is composed of 142 nucleotides with the first 122 forming the characteristic T-shaped palindromic structure (Fig. 2), and it is 60 to 62% homologous with the ITRs of serotypes 2, 3, 4, and 6 and 48% homologous with AAV5. A tandem repeat of GAGY in the ITR, which serves as the binding element of Rep78 and Rep68 (RBE), is conserved between AAAV and the other AAVs (Fig. 3, 4). The *trs* rec-



FIG. 2. AAAV ITR. The sequence of the ITR is shown in the hairpin conformation. The putative Rep binding site is boxed, while the putative *trs* is underlined and the cleavage site is indicated by an arrow. RBE, Rep binding element.

	RBS RBS	
AAAV	$- \texttt{TGGC-CAGTTTC-CAAGACAG\underline{GCTCGCTCGCTCAC} \texttt{TCG} \texttt{GGCCGGGGCC}CCAAAGGGCCCCTAGCGACCGCTTCGCGGTCGCGGCCCGAGTGAGCGACCGACGACCGACGACCGACGACCGAC$	C 102
AAV2	${\tt TTGGC-CACTCCCCTCTCGCGCGCTCGCTCGCTCACTGAGGCCGGGCGACCAAAGGTCGCCCGACGCCCGGGCTTTGCCCGGGCGGCCTCAGTCAGCGAGCGACGACGACGACGACGACGACGACGACGACGA$	G 105
AAV4	$\label{eq:transformation} TTGGC-CACTCCCTCTATGCGCGCGCTCACTCACTCACTCGCCGGACGAGGAGGGCCGAGGGAGG$	G 105
AAV5	-IGGCACTCTCCCCCTGTCGCTCCCTCCCTGCCCGCTCGTTCGGCGG	G 119
	**** * * * * * * ****** *** ** * * * * *	
	TRS ▼ USF	
AAAV	IGENTIAL CONTRACT	G 150
AAVZ		G 182
7772		G 222
MAV J		A 213
	SD-1 E.F TATA	
AAAV		r 220
AAV2	TTAGGGAGGTCCTGTATTAGAGGTCACGTGAGTGTTTGCCACATTTTGCGACACCATGTGGTCACGCTGGGTATTTAAGCCCCGAGTGAGCACGC-AGGGTCTCC	т 289
AAV4	CGTAGCGGAGCGCATACCAAGCTGCCACGTCACAGCCACGTGGTCCGTTTGCGACAGTTTGCGACACCATGTGGTCAGGA-GGGTATATAACCGCGAGTGAGCCAGCGAGGAGGAGCTCC	т 341
AAV5	${\tt TAGTTAATGATTAACAGTCATGTGATGTGTTTTATCCA-ATAGGAAGAAAGCGCGCGTATGAGTTCCCGCGAGACTTCCG-GGGTATAAAAGACCGAGTGAACGAGCCCGCCGCCATTAAAAGACCGAGTGAACGAGCCCGCCGCCATTAAAAGACCGAGTGAACGAGCCCGCCGCCATTAAAAGACCGAGTGAACGAGCCCGCCGCCATTAAAAGACCGAGTGAACGAGCGCCGCCGCCATTAAAAGACCGAGTGAACGAGCGCCGCCGCCATTAAAAGACCGAGTGAACGAGCGCGCGC$	т 331
	* * * * * * * * * * * *	
	> ₽5	
	P5	
AAAV	ATACCGCCGCCCCCACC-GGCGACATGAGGTCATCTACCGACGTCATCGTTCAGCTGCCCAACGACCTCGACAGGTCAGGTACCTGGAATCTCCGACTGCTCGTCAACTG	A 331
AAV2	TIGAACCGCAACGTTTCAACCCCCA-CCCCCCATGCCGGGGTTTTACGACTTGATAACGTCCCCACCACCTTGACGCGCATCTGCCAACGTCTTGTGAACCGCCACCTTGCAACGCCCACCACCCAC	G 408
AAV4	THECE - CONSACT THE BACCAGCA-GCABCCCARTIC CONSTITUTION TO A DECOMPOSITION TO A CONSTITUTION OF A CONSTITUCION OF A CONST	G 459
MAVO		G 451
AAAV	ͲͳϨϹϾͲϹϾϲϾϨϪϨͳϾϾϨϹϾϒͳϨϨϹͲϾϨϨϾϾϨϪϾϨϾϹϾϨϪͲͲϔϾϾϨϹϾϨϪͲͲϔϾϾϨϾϾϨͲϾϨϪͲϾϫϪϾͲͲϾϪϨϲͲϾϲϪϲϲϤϲϫϲϫϫϫϫϲϲϲϫϫϲϫϫϫ	2 451
AAV2	TGGCCGAGAGAGTGGCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGCCGC	4 4JI
AAV4	TGGCCGAGAAGGAATGGGAGCTGCCGCCGGATTCTGACTGGACGAGGACGAGGACGCCCCGGACGGCGGCGGAGGACGGCGG	G 579
AAV5	TAACTGGTCAAATTTGGGAGCTGCCCCCCAGAGTCAGATTTAAATTTGACTCTGGTTGAACAGCCTCAGTGACGGTGGCTGATTGGAACAGCCCCCCCGCGTGTCCTGTACGAAGAACAC	T 571
		1 571
AAAV	TGGCCAAAGAACCGGACTTTCACTATTTATCCAACTGGAACAAGGTGAGGTGTTCTTTCATTTACACGTCCTGCTGGAAACGTGTTCCGTAAAGCCGATGGTACTCGGAACATATA	C 571
AAV2	TGAGTAAGGCCCCGGAGGCCCTTTTCTTGTGCAATTTGAGAAGGGAGAGAGCTACTTCCACATGCACGTGCTCGTGGAAACCACCGGGGTGAAATCCATGGTTTTGGGACGTTTCC	A 648
AAV4	TGAGTAAGGCCCCGGAGGCCCTCTTCTTGTCCAGTTCGAGAAGGGGGACAGCTACTTCCACCTGCACATCCTGGGGGGGCGCCGTCAAATCCATGGTGGGGGGGG	A 699
AAV5	${\tt TTTCCAAGCAGGAGTCCAAATTCTTTGTGCAGTTTGAAAAAGGGATCTGAATATTTTCATCTGCACACGCTTGTGGAGACCTCCGGCATCTCTTCCATGGTCCTCGGCCGCTACGTCGGAGTCCGACGCGCTACGTCGGAGTCCGGCCGCTACGTCGGAGTCCGGCGCGCGC$	A 688
	* ** * *** * *** * ** * ** * ** * ** * *	
	Ecc-1	
AAAV	GALATATICA CAAAAAAATIG GACTAAAGTCIACTGCGCCCA-CGAGCCTACGATGGAGGATGCGTGGGCCAAGACCAAAAATTCGGGGGCGCGAAACAAGGTCCGG	C 687
AAV2	GICASATICGCAAAAACIGATICAGAGATTACCGCGGGACGGACGACGACTICGCCGAAACIGGCCACAAACACCACAAAAGCGGCGGGGGGGG	T /6/
AAV4 88775		C 818
AAV J		1 001
		2 207
AAAV	ETS-1 P ₅₃ TATA GAGTCGTATATTCCCGCCTACCTGATCCCGAACGCGCAGGCGGGGGGGG	G 807
AAAV AAV2	ETS-1 P ₅₃ TATA P ₅₃ P19 GAGTGGTATATTCCCGCCTACCTGATCCCGAAACCAGCAGCGGGAGTGGGGGGGG	3 807 3 887 3 938
AAAV AAV2 AAV4 AAV5	$\label{eq:rest} \begin{array}{c} \mathbf{ETS-1} \\ \mathbf{Fs_3} \\ \mathbf{TATA} \\ \mathbf{TATA} \\ \mathbf{Fs_3} \\ \mathbf{TATA} \\ \mathbf{TATAA} \\ \mathbf{TATA} \\ \mathbf{TATAA} \\ \mathbf{TATAAA} \\ TATAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAAA$	3 807 3 887 3 938 2 921
AAAV AAV2 AAV4 AAV5	$ETS-1 \qquad P_{53} \qquad TATA \\ GAGTCGTATATTCCCGCCTACCTGATCCCGAACAGCGGACGGA$	G 807 G 887 G 938 C 921
AAAV AAV2 AAV4 AAV5	ETS-1 P ₅₃ TATA GAGTCGTATATTCCCGCCTACCTGATCCGAAACAGCAACGGAGTGGGCGGGGGGGG	3 807 3 887 3 938 3 921
AAAV AAV2 AAV4 AAV5	ETS-1 P ₅₃ TATA P ₅₃ TATA P ₅₃ CATA CATA P ₅₃ CATA CATA CAT	G 807 G 887 G 938 C 921
AAAV AAV2 AAV4 AAV5 AAAV	ETS-1 P53 TATA GAGTEGTATATTCCCGCCTACCTGATCCGAACAGCGAACGCGAAGTGCAGGGCGTGGACTACGTGCCCGAGTATATAAAAGCGCTGCTTGCACCGAGACGGCGGCGCGGCGCGGCGGCGGCGGCGGCGG	G 807 G 887 G 938 C 921 F 924
AAAV AAV2 AAV4 AAV5 AAAV AAAV2	ETS-1 P ₅₃ TATA GAGTCGTATATTCCCGCCTACTGATCCCGAAACAGCAACCGGAAGGGGGGGG	G 807 G 887 G 938 C 921 F 924 G 1007
AAAV AAV2 AAV4 AAV5 AAV5 AAAV AAV2 AAV4	ETS-1 P53 TATA GAGTCGTATATTCCCGCCTACCTGATCCGAAACAGCAACCGGAGTGGACGGGCGGG	G 807 G 887 G 938 C 921 F 924 G 1007 G 1058
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5	ETS-1 P53 TATA GAGTEGTATATTCCCGCCTACCTGATCCCGAACAGGAACCGGAGGTGGAGGGGGGGG	3 807 3 887 5 938 C 921 r 924 3 1007 G 1058 C 1038
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5	ETS-1 P ₅₃ TATA GAGTCGTATATTCCCGCCTACTGATCCCGAAACAGCGAACCGGAGGGGGGGG	3 807 3 887 3 938 2 921 5 924 3 1007 3 1058 2 1038
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5	ETS-1 P ₅₃ TATA GAGTCGTATATTCCCGCCTACCTGACCGAAACAGCGAACCGGAAGGGGGGGG	5 807 5 887 5 938 C 921 5 1007 6 1058 C 1038
AAAV AAV2 AAV4 AAV5 AAV5 AAV2 AAV4 AAV5	ETS-1 P53 IATA P53 BATA P53 BA	5 807 5 887 5 938 C 921 F 924 G 1007 G 1058 C 1038
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV2	ETS-1 P53 TATA GAGTCGTATATTCCCGCCTACCTGACCGGAACAGCGAACAGCGAACGGGGGGGG	G 807 G 887 G 938 C 921 F 924 G 1007 G 1058 C 1038 C 1044 F 1127
AAAV AAV2 AAV4 AAV5 AAV2 AAV2 AAV2 AAV2 AAV4 AAV5	ETS-1 P53 TATA GAGTGGTATATTCCCGCCTACCTGGCAGACGCGCAGACGCGCGAGGGGGGGG	G 807 G 887 G 938 C 921 F 924 G 1007 G 1058 C 1038 C 1044 F 1127 F 1178
AAAV AAV2 AAV4 AAV5	ETS-1 P53 TATA GAGTEGTATATTCCCGCCTACCTGAACAGGAACAGGGAACAGGGAACGGGGGTGGGCTGGGCTGGACTAATATAGAAGCGTGCTGCAGGACACAGGAACTGGGGGCGGGC	S 807 S 887 S 938 S 921 r 924 G 1007 G 1058 C 1038 C 1044 r 1127 r 1178 C 1158
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV2 AAV2 AAV2 AAV2 AAV2 AAV2 AAV2 AAV4 AAV5	ETS-1 P53 TATA GAGTCGTATATTCCCGCCTACCTGACCGGAACCGGAACGGGAGGGGGGGG	G 807 G 887 G 938 C 921 F 924 G 1007 G 1058 C 1044 F 1127 F 1178 C 1158
 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 	ETS-1 P53 TATA GAGTGGTATATTCCCGCCAACTGGCAACGGCAACGGGAACGGGGGGGG	G 807 G 887 G 938 C 921 G 1007 G 1058 C 1044 F 1127 F 1178 C 1158
AAAV AAV2 AAV4 AAV5 AAV2 AAV2 AAV2 AAV2 AAV4 AAV2 AAV4 AAV2 AAV4 AAV5	$ \begin{array}{c} \mathbf{FTS-1} & \mathbf{P_{53}} & \mathbf{TATA} \end{array} \right) \\ \mathbf{FTS-1} & \mathbf{P_{53}} & \mathbf{TATA} \\ \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} \\ \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} \\ \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} \\ \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} & \mathbf{FTS-1} \\ \mathbf{FTS-1} & $	G 807 G 887 G 938 C 921 G 1007 G 1058 C 1044 F 1127 F 1178 C 1158
АААV ААV2 ААV4 ААV5 АААV ААV2 ААV4 ААV5 АААV ААV2 ААV4 ААV5	ETS-1 P53 TATA GAGTEGTATATTCCCGCCTACCTGAACAGCGAACGGGAACGGGGGGGG	<pre>G 807 G 807 G 938 G 921 G 1007 G 1058 C 1038 C 1044 F 1127 F 1178 C 1158 C 1164</pre>
AAAV AAV2 AAV4 AAV5	ETS-1 P53 TATA GAGTGGTATATTCCGCCTACCGGAACAGCAACCGGAACGGGGGGGG	<pre>3 807 3 887 3 938 9 921 r 924 3 1007 3 1058 C 1044 r 1127 r 1178 C 1044 r 1127 r 1178 C 1164 C 1164</pre>
AAAV AAV2 AAV4 AAV5	ETS-1 P53 TATA GAGTEGTATATTCCCGCCTACCTGAACAGCGAACGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAACAGCGAGCGACAGCGGCACAACA	5 807 5 887 5 938 5 938 5 921 7 924 5 1007 5 1058 C 1044 r 1127 r 1178 C 1164 C 1164 C 1292 1292
 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV2 AAV4 AAV5 AAAV AAV5 AAAV AAV2 AAV4 AAV5 	ETS-1 P53 TATA GAGTCGTATATTCCCGCCTACCTGAACAGCGAACAGCGAACAGCGAACAGCGGACGGGGGGGG	3 807 3 887 3 938 2 921 r 924 r 924 r 924 r 1058 c 1058 c 1044 r 1127 r 1178 c 1164 c 11272 c 1292 c 1272
AAAV AAV2 AAV4 AAV5	ETS-1 P53 TATA GAGTCGTATATTCCCGCCTACCTGAACAGCAACAGCGAACGGGAGTGGGCGTGGGCTGGGCCGGCC	<pre>3 807 3 887 5 938 5 921 7 924 3 1007 3 1058 C 1044 r 1127 r 1178 C 1164 C 1164 C 1164 C 1164 C 1164 C 1177 C 1178 C 1177 C 11777 C 11777 C 11777 C 11777 C 11777 C 11</pre>
AAAV AAV2 AAV4 AAV5	ETS-1 P53 TATA GAGTCGTATATTCCCGCCAACGCGAACGCGAACGCGGACGGGGGGGG	3 807 3 887 3 924 3 1007 3 1058 C 1044 F 1127 F 1178 C 1164 C 1164 C 1241 C 1242 C 1272
АААV ААV2 ААV4 ААV5 АААV ААV2 ААV2 ААV4 ААV5 АААV ААV2 ААV4 ААV5 АААV ААV2 ААV4 ААV5	ETS-1 P53 TATA GAGTGGTATATTCCGGCAACTGGAACCGGAACTGGAACTGGAACTGGAACTGGAACTGGAACTGGAACTGGAACTGGAACTGGAACTGGAACTGGAACTGGAACGGTGGGGAGCGAACGGGTGGACTAACATGGACCGGTGATATATAGGGCCTGTTGAATCTCACGGAGCGTGAACAGGGTGGGACTAACATGGACGCTGATATATAGGGCCTGTTGAATCTCACGGAGCGTAAACGGGGGGGACTAACATGGACGCGCCCAATTCCACGGAGCGCAACGGGTGGACTAACGGGGGGGAGCAAACGGCGGGGAGCAAACGGCGGGGAGCAAACGGCGG	3 807 3 887 5 938 5 1007 5 1058 5 1038 5 1038 5 1044 5 1058 5 1044 5 1058 5 1038 5 1044 5 1158 5 1158 5 1158 5 1158 5 1158 5 1158 5 1158 5 1158 5 1124 5 1127 5 1127 5 1127 5 1127 5 1127 5 11284 5 11285 5 11285 5 11285 5 11285 5 11285 5 112
AAAV AAV2 AAV4 AAV5	ETG-1 P3.3 TATA GAGTCGTATATTCCCGCAAACGCGAAACGCGAAGCAGAGCGAGGGGGGGG	5 807 5 887 5 938 5 938 5 921 7 924 5 1058 5 10
AAAV AAV2 AAV4 AAV5 AAV4 AAV2 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV5 AAV4 AAV5 AAV5 AAV4 AAV5	ETS-1 P3.3 TATA GAGTGGTATATTCCGGCTAACCGAACGGCAAGGGGGGGGG	<pre>3 807 3 887 3 924 3 1007 3 1058 2 1044 3 1007 1 1127 1 1178 2 1164 2 1241 2 1272 4 272 4 1261 2 1241 2 1241 2 1242 4 1361 </pre>
AAAV AAV2 AAV4 AAV5	Product Pail	G 807 G 807 G 924 G 924 G 924 G 105 C 1044 G 105 C 1044 G 1127 C 1158 C 1164 C 1164 C 1292 C 1241 C 1272 F 1284 A 1361 C 1422 C 1424 C 1422 C 1424 C 1444 C 144
AAAV AAV2 AAV4 AAV5	ETG-1 Ps3 ETTA Ps3 ETTA GAGTCGTATATTCCCGCTACCTGAACGCGGAGCGGAGGGGTGGAGCTAACGGCGGGGTGAACGGGGTGGACGTAACGGGCGGG	G 807 G 807 G 938 G 938 C 921 F 924 G 1038 C 1038 C 1044 F 1127 F 1127 F 1178 C 1164 C 11412 C 1292 C 1292 C 1292 F 1284 A 1361 C 1412 C 1392
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV5 AAV5 AAV5 AAV5 AAV4 AAV5	Ergs-1 p-3 TATA	G 807 G 807 G 887 G 938 G 921 F 924 F 924 G 1058 C 1158 C 115
AAAV AAV2 AAV4 AAV5	ETS=1 P3 TATA P13 GAGTCGTATATTCCCGCCAACGAACCCGAACGCGAGTCGAGTGGCGGTGGACTAACTGGCGGGTGGACCGATTTAAAACGCGGCCCGAGAACTGCGGGGGGGAGACGCGGGTGAGACGGGGGGGAGACGGGGGGGG	G 807 G 807 G 924 G 924 G 921 C 924 G 105 C 1058 C 1058 C 1058 C 1044 C 1127 C 1158 C 1158 C 1164 C 1292 C 1272 C 1272 C 1292 C 1272 C 1292 C 1272 C 1292 C 1272 C 1292 C 1272 C 1292 C 1292 C 1272 C 1292 C 1272 C 1292 C 1292 C 1272 C 1292 C 129
AAAV AAV2 AAV4 AAV5	ETS =1 F33 TATA GAGTGGTATATTCCCGCCACTGATCCGAAAGCAGAGTGGGAGGGGGGGG	<pre>3 807 3 807 3 807 3 807 3 938 3 107 924 3 1007 1038 2 1038 2 1038 2 1038 2 1038 2 1044 f 1127 f 1178 c 1158 c 1164 c 1241 c 1242 c 1292 c 1292 c 1292 c 1292 c 1292 c 1292 c 1392 c 1392 c 1392 c 1402 c</pre>
AAAV AAV2 AAV4 AAV5 AAAV AAV5 AAV4 AAV5 AAAV AAV2 AAV4	ETS=1 P:3 TATA P:13 GAGTCGTATATTCCGCCAACGAACGCAACGGAAGTCGAGTGGGCGGGGGACAAATGGGGCTGGTTAACGGGTGGTGGACGGGGGCAGAACGGGGGGGAGCAAATGGGGGGGG	G 807 G 807 G 924 G 924 G 924 G 924 G 924 G 924 G 1058 G 1058 G 1058 G 1058 G 1058 G 1058 G 1058 G 1164 G 1292 G 1272 G 1292 G 1272 G 1292 G 1272 G 1292 G 1292
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV5 AAAV AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV4 AAV5 AAAV AAV4 AAV5 AAAV AAV5 AAAV AAV5	ETG-1 Ps [TATA] GATTCOTATATTCOCGCCTACTGATCCCCAAAGCCAGTGGGGTGGG	G 807 G 807 G 924 G 928 G 921 G 924 G 1058 G 1164 G 1164 G 1164 G 1164 G 1164 G 1164 G 1168 G 1189 G 11
AAAV AAV2 AAV4 AAV2 AAV4 AAV2 AAV4 AAV2 AAV4 AAV2 AAV4 AAV2 AAV4 AAV5 AAV4 AAV5 AAV4 AAV2 AAV4 AAV2 AAV4 AAV2 AAV4 AAV5 AAV4 AAV5 AAAV AAV4 AAV5	BTG-1 P:3 TATA GATTCOTATATTCCCGCCTACTGATCCCGAAACCCAGCTGGACTGAGCTGGCGTGGACTAACGTGCCCCGCGTATTAAAACCGGTGGCGTGCACATCCCCGACTGCGGGGGGGG	G 807 G 807 G 924 G 938 C 921 F 924 G 1058 C 1044 F 1277 F 11277 F 11277 F 11277 F 11278 C 1164 C 11412 C 1292 C 1292 C 1292 C 1292 G 1292 G 1491 F 1284 F 1284
AAAV AAV2 AAV4 AAV5	<pre>ETS-1</pre> <pre>Pis-1</pre> <pre> <pre>Pis-1</pre> <pre>Pis-1</pre></pre>	G 807 G 807 G 924 G 924 G 924 G 924 G 1058 C 1044 C 1044 C 1044 C 1044 C 1044 C 1127 C 1158 C 1164 C 1292 C 12
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAV5 AAV5 AAV5	ETTS-1 P, 3 TATA WATA ALCONG ALCON	G 807 G 807 G 924 G 924 G 1007 F 924 G 1058 G 1164 G 1127 G 1272 G 1272 G 1272 G 1322 G 1322 G 1322 G 1272 G 1322 G 1322 G 1322 G 1322 G 1272 G 1322 G 1404 G 1512 G 1
 AAAV AAV2 AAV4 AAV2 AAV2 AAV2 AAV2 AAV2 AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV 	ETG-1 P3 TATA CARTCOTTATTCCCCCCALACCCGAACCCGAACCCGAACCCGGCGGGGGGGG	G 807 G 807 G 924 G 938 C 921 F 924 G 1058 C 1044 F 1127 F 1127 F 1127 F 1127 F 1127 F 1127 F 1127 F 1127 F 1127 F 1284 G 1058 C 1164 C 1164 C 1164 C 1292 C 1292 G 1401 G 1401 G 1401 G 1401 G 1401 G 1402 G 1404 G 1452 C 1524
AAAV AAV2 AAV4 AAV5 AAV4	ETG-1 P3 ETG-1 P3 ETG-1 P3 ETG-1 P4 GAGTGGTATATTCCCCCCAATACCTGATCCCCCAAACCCGGGCGTGGCGTGGCGCGGATTATTAAAGCCGCCGGTTGACCCACGGGGGGGG	G 807 G 807 G 924 G 924 G 924 G 924 G 1058 C 1044 C 1044 C 1044 C 1044 C 1044 C 1127 G 1158 C 1164 C 1292 C 12
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV2 AAV2 AAV2 AAV5 AAV4 AAV5 AAV4	Ergs-1 p-3 TATA P-13 GAGTGGTTATTTCCGGCCACCGGAAACCCGGACCGCAAACGGGGGGGG	G 807 G 807 G 924 G 924 G 1007 C 924 G 1058 C 921 C 924 G 1058 C 1058 C 1058 C 1058 C 1164 C 1164 C 1164 C 1164 G 1127 C 1292 C 1292 C 1292 C 1392 G 1404 G 1461 G 1522 C 1524 C 1526 C 1526 C 1652 C 1526 C 1652 C 1526 C 1652 C 1526 C 1652 C 1556 C 1652 C 1556 C 1566 C 1556 C 155
AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV5 AAAV AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5	Control of the c	G 807 G 807 G 924 G 938 C 921 F 924 G 1058 C 1044 F 1127 F 1127 F 1127 F 1127 F 1127 F 1127 F 1127 F 1127 F 1127 F 1284 G 1058 C 1164 G 1401 G 1402 G 1652 G 1655 G 1655 G 1655 G 1655 G 1655 G 1655 G 1655 G 1655 G 16
AAAV AAV2 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV4 AAV5 AAV5 AAV5 AAV5 AAV5 AAV4 AAV5	Control of the c	G 807 G 807 G 887 G 938 G 921 G 924 G 1058 G 1058 G 1044 G 1044 G 1044 G 1044 G 122 G 1241 G 122 G 1241 G 1412 G 1412 G 1412 G 1412 G 1414 G 1401 G 1412 G 1414 G 1401 G 1412 G 1524 G 1632 G 1632

FIG. 3. Sequence of the AAAV genome. The genomes of AAAV, AAV2, AAV4, and AAV5 were aligned by using Clustal W. The sequences of the ITRs are presented in italics. The putative *trs* is indicated by a vertical arrow, and the putative Rep binding site is underlined. Proposed transcription factor binding sites and the polyadenylation signal are also underlined. Proposed transcription initiation sites of the p5, p19, and p40 promoters and splice donor and acceptor sites are indicated by horizontal arrows. Initiation and termination codons are presented in bold letters. USF, upstream factor.

AAAV AAV2 AAV4 GTGGATGGGAATTCCACGACCTTTGAACACCAGCAGCCGCCTGGAGGACCGCATGTTCAAATTTGAACTGACTAAGCGGCTCCCGCCAGATTTTGGCAAGATTACTAAGCAGGAAGTCAAG 1752 AAV5 TATA BC1-6 Ets-1 LATA CAATTCTTCAGGTGGTCTCCAGGATCACCTGACCCCGGACGCGACCCCAGGAAGGCGGAGGCGCAAAAGACCCGCCCCTTCCGGGGAAGGC<u>T-ATATAA</u>GCCCGACCAAA GACTTTTTCCGGTGGGCAAAGGATCACGTGGTTGAGGTGGAGCCAAGAATTCTACGTCAAAAGGGTGGAGCCAAGAAAAGACCCGCCCCCAGTGACGCGAGATATAAGTGGGGCCCAAACGG AAAV AAV2 AAV4 AAV5 Splice AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 Rep78 stop AAAV AAV2 AAV4 AAV5 Minor splice + Splice 🗲 Rep68 stop ∣ VP1 ٦ AAAV AAV2 AAV4 AAAGGAGT-AGTCTTTTG---TTGATCACCTCCCAGATTGGTTGGAAGAAG---TTGGTGAAGGGCTCGGA--------AGTTTTTGGGCCTTGAAGCGGGCCCCCCCACAA 2301 AAV5 CCTAAGGCAAATCAGCAAACTCAAGAATCTCTTGAAAAGGACGATTCGAGAGGGCTCCGTGTTCCCAGGCTACAATTATCTAGGCCCTTTCAACGGTCTAGATAAAGGACGACCCGTCAAC 2474 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV AAV2 AAV4 AAV5 AAAV CGAAGACGAACCTCGTTTGCCCGACACTTCTTCACAGACTCCCCAAGAAAAACAAGAAGACCTCGCAAGGAAAGACCTTCC-----GGCGGGGGCAGAAGATCCGGGCGAAGGCACCTCT 2825 AAV2 AAV4 AAV5 CGACCACTTTCC-AAAAAGAAAGAAGAAGCTC----GGACCGAAGAGAGACTCCAAGCCTTCCAC----CTCGTC-----AGACGCCGAAGCTGGACCCAGC----GGATCCCA 2738 VP3 AAAV AAV2 AAV4 AAVS **** * * * ** ***** ******* *** AAAV AAATTGGCATTGCGATTCCCAATGGCTGGAAAACGGAGTCGTCACTCGAACCACCCGGACCTGGGTCTTGCCCAGCTACAACCACCGTGTACAAACGAATCCAAGGACCCAGCGGGGG 3064 AAV2 AAV4 AAV5 AAAV CGACA---ACAACAACAAATTCTTTGGATTCAGCACCCCCTGGGGATACTTTGACTACAATCGATTCCACTGCCACTTTTCCCCGCGAGACTGGCAACGACTCATCAACAACAACTGGGG 3181 CTCGA---ACGACAATCACTACGTACAGCACCCCTTGGGGGTATTTTGACCTCCACAGAGTTCCACTGCCACTTTCACCACGGGCAAAGACTCATCAACAACGGGG 3117 C-----CAACACCCTACAACAGCGGATTCTCCACCCCCGGGGATACTTTGACTTCAACCGCGTTCCACCGCCACTTCCACCACGGGCGACTCGCCGCGCCCCCGGGGATACTTGGGG 314 AAV2 AAV4 CGGAAGCAACGCCAACGCCTACTTTGGATACAGCACCCCTGGGGGGTACTTTGACTTTAACCGCTTCCACAGCACCGCGGGGCCCACTGGCAAAGACTACTGGGG 3098 AAV5

AAAV AAV2 AAV4 AAV5	CATCCGTCCCAAAGCGATGCGCTTTAGACTCTTTAACATCCAGGTTAAAGAGGTCACGGTCCAAGACTTCAACACCACCATCGGCAACAACCTCACCAGTACGGTCCAGGTCCAGGTTTTGCGGA ATTCCGACCCAAGAGACTCCAACTTCAAGCTCTTTAACATTCAAGTCAAAGAGGTCACGGCGAGACGGTACGACGACGCGACACGGTGCCAATAACCTTACCAGCACGGTCAGGTCAGGTCAAGGAGGTCACGACGCGAGACAACGGTGGCTAATAACCTTACCAGCACGGTCAGGTCAGGTCAGGGGCGACACGCGGCGAAACACGGTGGCTAATAACCTTACCAGCACGGTCAGGTCAAGGAGGTCACGGCGAGACGACGACGACGCGAGACAACGGTGGCTAATAACCTTACCAGCACGGTCCAGGTCAAGGAGGTCACGGTGCGAGACGACGACGACGACGACGACGACGACAACGGTGGCTAATAACCTTCCACGTCCAGATTCAAGTCAAGAGGTCACGGTGCGAGGCGACGACCACGGTGGCTAACACCTCACCACCACCGCGCCAAGAACGTCACGGCGACCCACCACCGCGCCAACAACGCTCACCACCCCCCCC	3301 3237 3266 3218
AAAV AAV2 AAV4 AAV5	CAAGGACTACCAACTGCCGTACGTACGCATCGGCTACCGAAGGCACCTTCCCCCCGCGGTACTCCAGCGGATATCTACACGATCCCGCGCGTACGCGCGCG	3420 3355 3386 3336
AAAV AAV2 AAV4 AAV5	-AGGCGGTGGATCGTTCGGCCTTCTACTGTCTGGACTACTTTCCCTCAGACATGCTGCGGACAGGAAATAACTTTGAGTTTACTTAC	3532 3468 3506 3455
AAAV AAV2 AAV4 AAV5	CATGTTTGCCCACAACCAGACGCTAGACCGGCTGATGAATCCCCTCGTGGGTCAGTACCTCTGGGCTTTCAGCTCCGTCAGCCAAGCAGGCTCATCTGGACGAGCCATC CAGCTACGCTCACAGCCAGAGCCTGGACCGCTCCATGAATCCTCTCATCGACCAGTACCTGTATACTTGAGCAGAACAAACA	3641 3584 3624 3553
AAAV AAV2 AAV4 AAV5	CATTACTCGCGGGCGACTAAAACCAACATGGCGGCTCAATATAGGAACTGGTTACCTGGGCCTTTCTTCCGTGATCAGCAAATCTTTACGGGGCGCTAGCAACATCACCACTAAAAATAAC97C AGTTTTCTCAGGCCGGAGCG-AGTGACATTCGGGACCAGTCTAGGAACTGGCTTCCTGGACCCTGTTACCGCCAGCGAGTATCAAAGACATCGCGAACAACAACAACAACAGTGAATAC AACTTTACCAAGCTGCGGCCTACCAACTTTTCCAACTTTAAAAAGACTGGGTGCCGGGCCTTCAATCAA	3761 3703 3743 3672
AAAV AAV2 AAV4 AAV5	TTTAGCGTTTGGGAAAAAGGCAAGCAATGGGAACTCGACAATCGGACCAACTAATGCAGCCCGGTCCTGCGGCAGCGACCACCTTTAGCGGAGAACCTGACCGTCAAGCCAT TCGTGGACTGGAGCTACCAAGTACCACCTCAATGGCAGAGACTCTCTGGTGAATCCGGCCATGGCAAGCCACGAGAGAAAAGTTTTTCCTAG CACCGGGTCAGACAGTCTCATCAAATACGAGACGCACAGCACTC-TGGACGGAAGATGGAGTGCCCTGACCCGCCGGACCTCCAATGGCCACGGCTGGACCTCGGGACAGCAAGAAAGTTTTTCCCTGG GCCTTCGCCAGGACCAATAGGATGGAGCTCGAGGGGGGGGGG	3874 3805 3860 3782
AAAV AAV2 AAV4 AAV5	GCAAAACACGCTGGCTTTTAGCAGGACCGTCTACGATCAAACGACCGCCACGACCGATCGTAACCAGATACTCATCACCAAAGACGAAATCAGACCCACCACCACCGGTCGGT	3994 3918 3977 3902
AAAV AAV2 AAV4 AAV5	CGCGTGGGGGGGGGGGTCCCCCGGCACCACCACCAGCCGGGGGCGGCGCGCGC	4114 4037 4096 4021
AAAV AAV2 AAV4 AAV5	AGGGACCCATTTGG-CCAAAATTCCCGGACACTGACAATCACTTCCATCCGTCCCCGGCTATTGGGCGGTTTGGCTGCAAGCATCCCCCCCC	4233 4157 4216 4141
AAAV AAV2 AAV4 AAV5	CTGCCAACCCTTCGGAAACGTTCCAGACGGCCAAAGTGGCCTCCTTCATCAACCAGTACTCGACCGGACA-GTGCACCGTCGAAATCTTTTGGGAACTCAAGAAGGAAACCTCCAAGGGC CTGCCAATCCTTCGACCACCTTCAGTGCGGCCAAAGTTGCTTCCTTC	4352 4277 4335 4257
aaav aav2 aav4 aav5	TGGAACCCCGAAATCCAGTTCACCTCCAACTTTGGCAACGCGGCCGA-CATCCAGTTGCCGTCTCCGACACGGGATCCTATTCCGAACCTCGTCCCATCGGTACCCGTTACCTTA TGGAATCCCGAAATTCAGTACACTTCCAACTACAACAAGTCTGTTAATCGTGGACTTA-CCGTGGATACTAATGGCGTGTATTCAGAGCCTCGCCCATTGGCACCAGAATACCAG TGGAACCCCGAGGTCCAGTTTACCTCCAACTACAACAACTCTCTGTTGTGGGCTCC-CGATGCGGCTGGGAAATACACTGAGCCTGGGCTATCGGGCCCCATTGGCACCAGAACTACCAGAGATCCAGGACACCGATACCAGAACTCCCGCTACCCCGACAGCACCCGGGGAAATACACTGAGCCTGAGCCTATCGGAACCCCGCTACCGCTA TGGAACCCCAGAGATCCAGTACACAAAAACAACTACAACGA-CCCCCAGTTTGTGGGCTCC-CGATGCGGCCGGGAAATACACGAGACCCACAGGCTATCGGAACCCCGATACCTTA *****	. 4467 . 4392 . 4450 . 4372
AAAV AAV2 AAV4 AAV5	VP-stop Poly-A CCAAACCTCTGTAA	4550 4506 4569 4470
AAAV AAV2 AAV4 AAV5	CCAGTTTCCAAGACAGGCTCGCTCGCTCACTCGGGCGGG- 	; 4617 · 4594 · 4686 ; 4559
AAAV AAV2 AAV4 AAV5	GCCCCAAAGGGGCCCCTAGCGGCTTCGCGGGTCGCGGCCCGAGTGAGCGAGCGAGCCAGCCGACCATCTTG-GAAACTG-GCCA- 4694 GCGACCAAAGGTCGCCCGGCCCGGG-CTTTGCCCGGGCGGGCCGCAGTGAGCGAGCGAGCGAGCGAGCGAGGGGGGGGGG	

FIG. 3—Continued.

ognition motif of serotypes 2, 3, 4, and 6 (CCGGT'TG) is highly homologous with that of the putative AAAV *trs* (CCG GT'CG) and weakly homologous with the AAV5 *trs* site (ACG GT'GT). In addition, the spacing between the RBE and the putative *trs* is similar to that found in other serotypes, a characteristic that has been shown to be essential for Rep activity (12).

It has been proposed that a potential inverted repeat flanking the core *trs* sequence of AAV serotypes might be required for Rep *trs* nicking (11). Such an inverted repeat is not found around the AAAV *trs* sequence. This observation may indicate that avian Rep nicking does not require any secondary structure around the core *trs* element. Methylation interference experiments have indicated the importance of the CTTTG motif found at the tip of one palindrome in AAV2 Rep binding (9, 57). Most of this motif is conserved in AAAV ITR (CT-TCG) and only one T residue is changed to C. Interestingly, the AAV4 ITR has a similar substitution in this motif (CTCTG). Thus, regardless of the overall nucleotide sequence homology, the secondary structure and the elements required for viral replication are conserved in the AAAV ITR.

The entire AAAV genome (Fig. 3) is 4,694 nucleotides in length and has a similar organization to that of other AAVs. It has two inverted terminal repeats and two distinct ORFs. The entire genome of AAAV displays 56 to 65% identity at the nucleotide level with the other known AAVs. The p5 promoter region of AAAV is much shorter and shows some divergence from homologous regions of other AAV serotypes. Core regulatory elements such as the TATAA box and Ebox/USF are conserved; however, YY1 and Rep binding sites are not present. This suggests that AAAV gene expression might be regulated differently from that of other AAVs. The p19 promoter, the p40 promoter, and poly(A) can also be identified in the AAAV genome by homology to those in primate AAV serotypes. Based on the general organization and sequence, these elements are highly conserved.

Clustal W protein sequence alignment indicates that the left ORF of AAAV is 46 to 54% identical and equally divergent from that of the primate AAVs and the goose parvovirus (GPV) Rep ORF (Fig. 4A) and only 18 to 22% identical with the Rep ORF of other mammalian autonomous parvovirus (data not shown). In comparison, the Rep ORFs of isolates 1, 2, 3, 4, 6, 7, and 8 show greater than 90% similarity and are approximately 67 to 70% identical with that of the AAV5 Rep ORF. The central region of the AAAV Rep ORF (amino acids [aa] 322 to 470), which is present in all Rep proteins, displays the greatest identity (82%) with the same region of the other AAVs and the GPV. This region of the Rep proteins is necessary for ATPase and helicase activity and contains an ATPbinding site (aa 334 to 349) and a divalent cation binding site at amino acid residue 421 (44, 61, 65). The amino terminus (aa 1 to 251) shows 42 to 45% similarity between AAAV and the other AAVs. This region of the Rep78 and Rep68 proteins is required for DNA binding and trs endonuclease activities (22, 50). A tyrosine residue at 155 is homologous to the Tyr156 in AAV2 that functions as the active nucleophile in the trs endonuclease site (22, 62). The active site is assembled by the spatial convergence of a divalent metal ion that is tetrahedrally coordinated by Asp24, Glu83, His90, and His92. In addition, Glu6 is required for the correct orientation of the two activesite imidazoles from His90 and His92 (31). All of these amino acid residues are strictly conserved among AAV serotypes, including AAAV. Furthermore, a helix region important for Rep multimerization (aa 159 to 179) is also conserved in AAAV. The carboxyl terminal portion (aa 490 to 662) of the unspliced AAAV Rep proteins appears highly divergent, displaying less than 15% homology with the primate serotypes. However, a characteristic zinc finger motif was identified by using the BLIMPS algorithm. This feature is conserved in all AAV serotypes.

The right ORF of AAAV, which encodes the three viral capsid proteins, is approximately 54 to 57% identical to the capsid ORF of the other AAVs and the GPV (Fig. 4B). It has been previously reported (6) that the AAAV capsid proteins VP1, VP2, and VP3 have apparent molecular masses of 92, 69, and 61 kDa, respectively, as determined by SDS-polyacryl-amide gel electrophoresis. The calculated molecular masses based on amino acid composition for VP1, VP2, and VP3 are 83, 67, and 60 kDa. We also subjected purified AAAV virions to SDS-polyacrylamide gel electrophoresis and found that they have molecular masses of 91, 68, and 60 kDa (data not shown). As with the primate AAVs and the goose and duck autonomous parvovirus, the AAAV cap gene contains two ATG initiator codons, one for VP1 and the other for VP3. The unusual ACG initiator codon for VP2 is also conserved in AAAV.

Clustal W alignment of the VP ORFs indicated the presence of conserved and divergent regions. The N terminus of VP1 (aa 1 to 143), which is required for particle formation, is relatively conserved among AAAV, AAV2, AAV4, AAV5, and GPV. However, the start sites for VP2 and VP3 are found in a divergent region. Based on the published three-dimensional structure of the canine parvovirus and comparisons of parvovirus capsid sequences (15, 67), most of the divergent regions among AAAV, AAV2, AAV4, AAV5, and GPV are located on the exterior of the virus, thus suggesting different uptake mechanisms and altered tissue tropism.

In the present study, we constructed recombinant AAAV particles containing the gene for nucleus-localized β-galactosidase. Virus was produced as previously described (19, 20) by constructing a vector plasmid containing the β-galactosidase gene under the control of a Rous sarcoma virus promoter flanked by AAAV ITRs (pA3Vβ-Gal, Fig. 5A) and a helper plasmid containing the AAAV rep and cap genes. Virus was isolated from 293T cell lysates by CsCl banding, and the distribution of recombinant virus across the gradient was determined by QPCR analysis of gradient fractions. The majority of packaged genomes were found in fractions with a density of 1.42 g/cm³, which is similar to that of wt AAAV. We also examined the yield of rAAAV when using helper plasmids with the rep gene under the control of three different promoters, CMV, MMTV, or the native P5 promoter (Fig. 5A). The different helper plasmids (pCA3VRC, pMA3VRC, and pA3VRC) were cotransfected with pA3Vβ-Gal and an adenovirus helper plasmid in 293T cells, and rAAAV was purified from the three different clarified viral lysates by using CsCl gradients. The number of rAAAV genomes was determined by QPCR. In three independent trials, the yield of rAAAV was 5-fold and 15-fold greater when using the stronger CMV promoter than the yield with the MMTV and the native P5 promoter, respectively (Fig. 5A). This finding with rAAAV is in

A		₽5			*			*	* *		
	AAAV		MRSYYEV	IVQLPNDVESQVPC	GIS D SFVNW	ITSREWTLPEDADWDLDQVDQV	QLTLGDKIQREIRTHWO	GTMAKEPDFHYFIQL E (GEVFFHLHVLLETCSVK	PMVLGRYIR	110
	AAV2		.PGF.I	VIKV.S.LDGHL.		VAEKEP.S.MNLIE.A	PVAE.L.,DFL.E.H	RRVS.A.EALF.V.F.H	SYMVTG;	SFLS	110
	AAV4		.PGFI	VLKV.S.LDEHL.	s.	VAEKEP.S.MNLIE.A	PVAE.LFLVE.H	RRVS.A.EALF.V.F.H	DSYI.VVG	SVVS	110
	AAV5		.ATF	RV.FEHL	D.	V.GQI.EPES.LN.TL.E.P	VA.R.R.VFLYE.1	NKFS.Q-ESKF.V.F.H	.SEYT.VSGIS	svs	109
	GP	MALSRPLQIS	SDKF	.IR.SS.IDQD	.L.LNE.	LSTGV.EPTGI.NMEH.NLP	MVAEKNIFIQR.N	NQFNQD-ETDF.FH	S.SEYICCIAQGN.R	SFMS	117
						 Dimerizati 	on domain		D.	10	
	0 D D J	HTOOKTVCKV	WCARGED	WEDGCIARENT.	ROCANIEVOA	POVIDAVI TOKOODEVOMAMUM	VORVIVACIUDRIDACI	ADI HEFENOVCOCKEN	ים זיגם ייעז אייק א זו זיגר ייעז אייק א זו	VEVMET UDW	220
	AAV2	O.RE.LIORT	RGIEPT	LPNWFAR.G	A.GVD	.C N I T I	MEQ.LSNLTE.KR	VAO. LTHVSOT. EONF	ENONPNSD TRSKT.	AG.	230
	AAV4	O.KE.L.TRI	RGVEPO	LPNWFAR.GZ	AGVD	DCN.L.TL	MDOS., NLAE.KR.	VAO.LTHVSOT.EON	ENONPNSD IRSKT.	AG.	230
	AAV5	Q.RAQL.KV.	FQGIEPQ	IN.WVAIV.K	vd	SGLVL	LD.,KL.A.NL.E.KR	.VAQFLA.SSQRS-Q.A	ASQREFSAD. IKSKT.	QKAN.	226
	GP	Q.KDS.IRD.	.EGKQIK	IP.WFAIR	QTVT	AALHKLF	M.LFTA.A.CLQK.QE	.LDAFQ.SDLAAPI	PDPQ.STV.,LISN.AA	.N.SN	232
									ATP bin	nding	
	AAAV	LVEKGITTEK	.EWLLENR	ESFRSFQASSNSAF	RQIKTALQG	AIQEMLLTKTAEDYLVGKDPVS	DDDIRQNRIYKILELN	HYDPAYVGSILVGWCQF	KWGKRNTLWLF <u>GHATTG</u>	KTNIAEAIA	348
	AAVZ	D	Q. TQ. DQ	A.IIN.AR	5ADN	GKI.SPQU	-Е	JQ.AA.VFLAT.	.r	• • • • • • • • • •	348
	22/25	н с		VI. NGTO DO		TWIC GV CON-	-E. SSRM.C.		SEN V V D		340
	GP	.I.MS	0T	YT.SNN.	V.AEN	.RA	-L.,TK.,V.O.,KM.N	N.N.O.ICVKF	EFNAIY.P		350
								•			
	AAAV	HAVPFYGCVN	WTNENFP	'FNDCVEKMIIWWE	EGKMTAKVV	ETAKAILGGSRVRVDQKCKASV	PIEPTPVIITSNTNMC	YVIDGNTTTFEHKQPLE	DRMFKLELLTRLPDDFG	KVTKQEVRQ	468
	AAV2		• • • • • • •	DV		.SKS.A	Q.DV	ASQ,	2F. TR. DH	KD	468
	AAV4		• • • • • • •	Dv		.5	Q.DV	A	2FTKEH	KD	400
	CP CP			D L	N	S A C	יי ה יי ה	VIVI S M RT	E OTV SHK EDS	TCK E	470
	01						C	arv	- Classe damain	.10.10.11	110
				NLS		NLS			y ringer domain		
	AAAV	FFRWSQDHLI	PVIPEFL	VRKAESRKRP	A	PSGEGYISPT <u>KRP</u> ALAEQQQAS	ESADPVP	-TRYRIK <u>CSKHCGMDKN</u>	ILFPCQICESMNRDINIC	AIHKTTDCK	567
	AAV2	AKVV	Έ.ΕΗ.Υ	.K.GGAK		DADISE.KRVRESVA.PST.	DAEASINYA	ADQNR.VNL.	RQRQNS	FT.GQKL	569
	AAV4	ASV.	E.THY	GGA		.NDADISE.KRACPSVA.PST.	DAEADYA	ADQNR.VNL.	RQRQNVD	FT.GVMA	569
	AAV5	A.AKVNQV	THK	C. PRELAGTKGAEKS	SL <u>KRP</u> LGDV	TNTSYKSLEKRARLSFVPETPR	S.DVT.DPAPLRPLNW	NSDCDY.AQF.NI	SNK.DEYLGK.G.	IC.NV.H.Q	584
	GP	K.AN.N.V	vsĸ	TN.QTNLP	K	.VP.RANE.EEP.KIWAPPTRE	LEELLRASI	PELFSSVAPIPVTPQNS	SPE.KRSRNNYQVRCALH	TYDNSM.VF	571
				Zinc finger dom	nain		Zinc finger doma	in ,			
	AAAV	ECFPDYGDKD	DVELPPC	TEHNVSRCYQCHSC	GELYRVTSD	SDEKPAPESDEGTEPSYAP <u>CTI</u>	HHLMGKSHGLVTCAACH	RLKNSTLHDDLDDGDLH	Q 662		
	AAV2	VS			V.E	.QVSVVKK-AYQKL.Y.	IVP-D-A.TI	D.VVCIF.	. 621		
	AAV4	VS			V.E	.QVSVVRKRTYQKL.P.	.,1.,RAP-E,A,S.,I	S.AVCDM.	. 623		
	GP	I.MECEKAN-				FP.FOPLG.N.CDEHGW	YDCATCKELKNEL EI-	EHVFE AEN	. 627		
						11.11.91.201111.022410	1 Doni Chi Dhanda Di				
в	νααα	MSLISDATEDM	I.ERLWKK	CUNA A A DEVHLES	CPPRPKANC	OTOESLEKDDSEGLVE	PGYNYLGPENGLDKGE	ρυνιβάτια α άτ. εμπικά γι	TETRICHNDAREANETD	RRFORRIK	114
В	AAAV AAV2	MSLISDAIPDW	LERLVKK	(GVNAAADFYHLES)	GPPRPKANÇ	QTQESLEKDDSRGLVF	PGYNYLGPFNGLDKGE	PVNEADAAALEHDKAYI	DLEIKDGHNPYFEYNEAD	RRFQERLK	114 105
в	AAAV AAV2 AAV4	MSLISDAIPDW .A-ADGYL .TDGYL	VLERLVKK	(GVNAAADFYHLES) (LSEGIRQWWK.KP ILSEGVREWWA.OP	GPPRPKANÇ PPAE .A.K	QTQESLEKDDSRGLVF RHK	PGYNYLGPFNGLDKGE	PVNEADAAALEHDKAYI	DLEIKDGHNPYFEYNEAD RQLDS.DLKH OOL.A.DLKH	PRRFQERLK AE AEOO	114 105 104
В	AAAV AAV2 AAV4 AAV5	MSLISDAIPDW .A-ADGYL .TDGYL FV.HP	<pre>%LERLVKKDTDNE</pre>	(GVNAAADFYHLES) ILSEGIRQWWK.KP ILSEGVREWWA.QP IVGEGLRE.LG.A	GPPRPKANQ PPAE .A.K KP	QTQESLEKDDSRGLVF RHK	PGYNYLGPFNGLDKGE KG	PVNEADAAALEHDKAYI	DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQL.A.DLKH	RRFQERLK AEQ.Q AEK.A	114 105 104 104
в	AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL FV.HP TFL.SFEE.	<pre>%LERLVKKDTDNE YE</pre>	(GVNAAADFYHLES) TLSEGIRQWWK.KP ILSEGVREWWA.QP IVGEGLRE.LG.A TSWRN.KA	GPPRPKANQ PPAE .A.K KP .A.QP	QTQESLEKDDSRGLVF RHKNAI .HQAI S.SVSPDR.PER.NN.F.I	PGYNYLGPFNGLDKGE KG GR 	PVNEADAAALEHDKAYI A	DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQL.A.DLKH QQL.A.DLKH QQL.A.DIKF.H	PRRFQERLK AEQQ AEK.A QD.IDS.Q	114 105 104 104 113
в	AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL FV.HP TFL.SFEE.	VLERLVKK DT DN E Y	(GVNAAADFYHLES) (LSEGIRQWWK.KP VLSEGVREWWA.QP SVGEGLRE.LG.A TSWRN.KA	GPPRPKANQ PPAE .A.K KP .A.QP	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NN.FI	PGYNYLGPFNGLDKGE. KGR. GR. KGP	PVNEADAAALEHDKAYI A	DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQL.A.DLKH NEQLEA.DLKH QQL.A.DIKF.H	PRFQERLK AEQ.Q AEK.A QD.IDS.Q	114 105 104 104 113
в	AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL FV.HP TFL.SFEE.	VLERLVKK	KGVNAAADFYHLES LSEGIRQWWK.KP USEGVREWNA.QP SVGEGLRE.LGA TSWRN.KA	GPPRPKANQ PPAE .A.K KP .A.QP VP2	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NN.FI	PGYNYLGPFNGLDKGE, KG GP	PVNEADAAALEHDKAYI	DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQL.A.DLKH NEQLEA.DLKH QQL.A.DIKF.H VP3	PRFQERLK AEQ.Q AEK.A QD.IDS.Q	114 105 104 104 113
в	AAAV AAV2 AAV4 AAV5 GP AAAV	MSLISDAIPDW .A-ADGYL DGYL FV.HP .TFL.SFEE.	VLERLVKK DI E YE Y	KGVNAAADFYHLES LSEGIRQWWK.KP USEGVREWWA.QP SVGEGLRE.LGA TSWRN.KA <u>CRVLEPFGLVEDS-I</u> L EPV	GPPRPKANQ PPAE .A.K KP .A.QP VP2 KTAPTGDKF GKREF	QTQESLEKDDSRGLVF RHKNAI HQAI S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI 	DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQL.A.DLKH QQL.A.DLKH QQL.A.DIKF.H VP3 SSIMAEGGGGPVGDAGQG NFT T.S.A.MA.NNE	RRFQERLK AEQ.Q AEK.A QD.IDS.Q	114 105 104 104 113 231 225
В	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF	VLERLVKK DI E YE Y (AIFQAKK 	KGVNAAADFYHLES(LLSEGURQWWK.KP JLSEGVREWWA.QP JVGEGLRE_LG_A TSWRN.KA <u>RVLEPFGLVEDS-1</u> EPV LOAG	GPPRPKANQ PPAE .A.K .K.P .A.QP VP2 KTAPTGDKF GKRFE EGKRFE	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEMSPVEPDSSTGGKAQQPP LI.SPOOPDSSTGGKAQQP	PGVNYLGPFNGLDKGE: KG GR.P. KG.LOPGEGT: K.LOPGOTG.ADSVPI KKLVFEDETGAGD.P	PVNEADAAALEHDKAYI R.EV.RIS.I K.SVS SSNAGAAAPASSVG	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H QQL.A.DLK.H QQL.A.DLK.H QQL.A.DIK.H VP3 SSIMAEGGGGPVGDAGQG NNTT.S.A.MA.NNE. -E.RAAA.AAVEG.	PRRFQERLK AE AEQQ AEK.A QD.IDS.Q SADGVGNSS	114 105 104 104 113 231 225 219
В	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF GF	VLERLVKK DI E YE YE YE YE YE Y V	KGVNAAADFYHLES(FLSEGIRQWWK KP JLSEGVREWWA QP JUSEGURE LG .A TSWRN.KA KRVLEPFGLVEDS-1 LEPV LEQX	GPPRPKANQ PPAE .A.K .K.P .A.QP VP2 KTAPTGDKF GKKRF EGKKRF	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NNF.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPP ULI.SPQQPDSSTGIGKKGQQPP DDFFPK.KKARTEEDS.PSTSS	PGYNYLGPFNGLDKGE KGP KGP RKERPSOGAEDPGEGT: RKLINFGQTG.ADSVPI KKKLVFEDETGAGD.P. 	PVNEADAAALEHDKAYI REV.RIS.I KSV. SSNAGAAAPASSVG DPQPLGQAPSGLG? PEGSTSGAM.DDS DPASSLGADTM.A	DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQLA.DLKH NEQLEA.DLKH QQL.A.DIKF.H VP3 SSIMAEGGGGPVGDAGQG INTT.S.A.MA.NNE. E.RAAA.AAVEG	PRRFQERLK AE AEQ.Q AEK.A QD.IDS.Q BADGVGNSS A.	114 105 104 104 113 231 225 219 215
В	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL FV.HP .TFL.SFEE. DDTSFG <u>GNLG8</u> EF GF	VLERLVKK DI E .YE .YE .YE .YE .YE .YE .YE .YE .YE .YE .YE	KGVNAAADFYHLESK FLSEGIRQWWK KP USEGUREWA QP SVGEGLRE LG. A T SWRN KA REVLEPFGLVEDS-1 L. EPV L. QAG EQG EQG 	GPPRPKANQ PPAE .A.K .A.QP VP2 KTAPTGDKF GKKRF EGKKRF NAKKNY	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NNF.I KGEDEPRLPDTSSQTPKKNKKF VEMSPVEPDSSSGTGKAGQQP LI.SPQ0PDSSTGIGKKGKQPP DDHFPK.KKATEEDS.PSFS GKLTDHPVV	PGYNYLGPFNGLDKGE KGP KGP RKERPSGGAEDPGEGT: RKRLNFGQTG.ADSVPI KKRLVFEDETGAGD.P JDA.AGPSGSQQLQIPAI KLTEEVSAGGGSSAVQ	PVNEADAAALEHDKAYI REV.RIS.I K.SV. SSNAGAAAPASSVG DPQPLGQAPSGLG PEGSTSGAM.DDS QPASSLGADTM.A DGG.T.EGTEPVAI	CLEIKDGHNPYFEYNEAD RQLDS.DLKH NEQLBA.DLKH QQL.A.DIK.H QQL.A.DIK.F.H VP3 SIMAEGGGGPVGDAGQG TNTT.S.A.MA.NNE. E.RAAAAAVEG L.NN. A.EAM.SGG	RRFQERLK AE AEQ.Q AEK.A QD.IDS.Q SADGVGNSS A. A.	114 105 104 113 231 225 219 215 221
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV2 AAV4 AAV5 GP	MSLISDAIPDM .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EF GF	VLERLVKK DI E .Y	(GVNAAADFYHLES(LLSEGUREWMA, QP SVGEGLRE, LG, A T SWRN, KA (RVLEPFGLVEDS-1 	GPPRPKANÇ PPAE .A.K .KP .A.QP VP2 KTAPTGDKF GKKRF EGKKRF KRI NAKKNI	QTQESLEKDDSRGLVF RHKNAI HQAI S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQPP LI.SPQQPDSSTGIGKKGQP DDHFPK.KKARTEEDS.PSTSS GKLTDHYPVV	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI 	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H VEQLEA.DLK.H QQL.A.DIK.H QQL.A.DIKF.H VP3 SSIMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. L.NN.A. A.EAM.SSG.	PRRFQERLK AE, AE,Q.,Q AE,K.A QD.IDS.Q SADGVGNSS ,A. ,A.	114 105 104 113 231 225 219 215 221
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV	MSLISDAIPDØ .A-ADGVL .TDGVL FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EFG GFG Q. GNWHCDSQWLE	VLERLVKK DI E YE YE XIFQAKK V.V V.V V.V NGVVTRT	KGVNAAADFYHLES(LLSEGURGWWK KP USEGURELG.A T.SWRN.KA (RVLEPFGLVEDS-1 L.EPV L.QAG GAL I.PVI TRTWVLPSYNNL)	GPPRPKANQ PPAE .A.K .A.QP VP2 KTAPTGDKF GKKRF EGKKRF NAKKNT YKRIQGPS-	QTQESLEKDDSRGLVF RHKNAI HQAI S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPA LI.SPQ2PDSSTGIGKKGKQPA DDHFPK.KKARTEEDS.PSTSS GKLTDHYPVV	PGYNYLGPFNGLDKGE: 	PVNEADAAALEHDKAYI R. EV.RISI K. SVS SSNAGAAAPASSVG DPQPLGQAPSGLG PEGSTSGAM.DDS QPASSLGADTM.A DGG.T.EGTEPVA NWGIRPKAMRFRLFNI(DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQL.A.DLKH QQL.A.DLKH QQL.A.DIKH QQL.A.DIKF.H VP3 SIMAEGGGGPVGDAGQG FNTT.S.A.MA.NNE. E.RAAAAAVEG LNN. .EAMSSG.	PRRFQERLK AE AEQ.Q AEKA QD.IDS.Q ADGVGNSS A. A. A. A. A.	114 105 104 113 231 225 219 215 221 350
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV2	MSLISDAIPDW A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF G Q GNWHCDSQWLE 	VLERLVKK DI DN F YF YF YF YF X VV V V V V V V V V V V.	KGVNAAADFYHLES(FLSEGURQWWK.KP VLSEGVREWWA.QP SVGECLRE.LG.A T.SWRN.KA (RVLEPFGLVEDS-1	GPPRPKANQ PPAE .A.K .K.P .A.QP VP2 KTAPTGDKF GKKRF EGKKRF EGKKRF MKRI NAKKNT YKRIQGPS- QSQ	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEMSPVEPDSSTGIGKAGQ2PP LI.SPQQPDSSTGIGKAGQ2PA DHFPK.KKARTEEDS.PSTSS GKLTDHYPV	PGYNYLGPFNGLDKGE K.G., K.G., RKERPSGGAEDPGEGT RKERPSGGAEDPGEGT RKELNFGQTG.ADSVP KKKLVFEDETGAGD.P. DA.AGP5GSQQLQTPA KLTEEVSAGGSSAVQ FHCHFSPRDWQRLINN	PVNEADAAALEHDKAYI R.EV.RIS.I K.SVSSNAGAAAPASSVGS DPQPLAQAPSGLG PEGSTSGAM.DDS QFASSLGADTM.A DGG.T.EGTEPVAI NWGIRPKAMRFRLFNI 	DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQL.A.DLKH QQL.A.DLKH QQL.A.DIKH QQL.A.DIKF.H VP3 SSIMAEGGGGPVGDAGQG NTT.S.A.MA.NNE. 	PRRFQERLK AEA AEKA QD.IDS.Q SADGVGNSS A. A. A. A. A. A.	114 105 104 113 231 225 219 215 221 350 344 225
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV2 AAV4 AAV2	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EF GF GF G.WHCDSQWLE T.MC .D.T.T.S.	VLERLVKK DN DN Y YE Y	KGVNAAADFYHLESK LLSEOUTBWWA, QP SVGEGLRE, LG., A T., SWRN, KA (RVLEPFGLVEDS-1 , L., EPV , L., EPV , L., QAG , PVI YTRTWVLPSYNNHL: 3., A.T., 3.,, T., 3.,, T., 3.,, O	GPPRPKANC P. PAE A.K K. P. A.Q. P. A.Q. P. VP2 KTAPTGDKF GKKRF GKKRF KKNT YKRIQGPS- Q.SSQ 	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NNF.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQ2PF LI.SPQ0PDSSTGIGKKGKQPF DDHFPK.KKATEEDS.PSTSG GKLTDHYPVV	PGYNYLGPFNGLDKGE K. G K RKERPSOGAEDPGEGT: RKERPSOGAEDPGEGT: RKELNFOQTG.ADSVPJ KKKLVFEDETGAGD.P: DA.AGFSGSQQLQTPA KLTEEVSAGGSSAVQ: FHCHFSPRDWQRLINNI S.W	PVNEADAAALEHDKAYI 	DLE IKDGHNPYFEYNEAD RQLDS.DLK.H VQLA.DLK.H QQLA.DLK.H QQLA.DIKF.H VP3 SSIMAEGGGGPVGDAGQG NT.T.S.A.MA.NNE. E.RAAA.AAVEG E.RAAA.AAVEG L.NN A.EAM.SSG. VKEVTVQDFNTTIGNNL ST A.S.	PRRFQERLK AE AEK.A QD.IDS.Q SADGVGNSS A. AA. AA. A. A. A. A. A. A.	114 105 104 113 231 225 219 215 221 350 344 335 335
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDM .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EF GF GF GF GF GF GNWHCDSQWLE .T.MC .DT.MC	VLERLVKK DI DN E .Y .Y .Y .Y	(GVNAAADFYHLES(LLSEGUREWWA, QP SVGEGLRE, LG, A T SWRN, KA (RVLEPFGLVEDS-1 	GPPRPKANC P. PAB .A.K .A.Q.P. VP2 KTAPTGDKF GKKRF GKKF GKKF GKKF GKKF GKKF GKKF GKKF GKKF GKKF GKF GKKF GKKF GKF GF GKF GF GKF GF GF GF GF GF GKF GF	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPP 'LI.SPQQPDSSTGIGKKGQPP DDHFPK.KRATEBDS.PSTSG 'GKLTDHYPVV	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI 	DLE IKDGHNPYFEYNEAD RQLDS.DLK.H VQLA.DLK.H QQLA.DIK.H QQLA.DIKF.H VP3 STIMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. L.NN.A.EAM.SSG. VVEVTVQDFNTTIGNNL 	PRRFQERLK AEQ.Q AEKA QD.IDS.Q SADGVGNSS A. AEA. A. T. T.	114 105 104 113 231 225 219 215 221 350 344 335 335 340
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EF GF GF GF GF GF GNWHCDSQWLE DT.MC .DT.MC	VLERLVKK DI DN YP Y Y X.V. V V V SNGVVTRT SDR.I.TS DPRKS ;.T.I.K.	(GVNAAADFYHLES(CLSEGIRQWWK.KP ULSEGVIEWMA.QP BVGEGLRE.LG.A T.SWRN.KA (RVLEPFGLVEDS-1 L.QAG	GPPRPKANC P. PAB .A.K .A.Q. P. VP2 VP2 VP2 VP2 VP2 KTAPTODKF GKKRF GKKRF KRT NAKKWT YKRIQOPS- Q.SSQ LGE RE.KSG.V. .A.TSGT-	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPA LI.SPQQPDSSTGIGKKKQPA DDHFPK.KARTEEDS.PSTSG GKLTDHYPVV	PGYNYLGPFNGLDKGE K. G. R. G. R. K. G. R. RKERPSGGAEDPGEGT: RKERPSGGAEDPGEGT: KKLVFEDETGAGD. P DA. AGPSGSQQLQIPA KLTEEVSAGGSSAVQ: FHCHFSPRDWQRLINN S.W.	PVNEADAAALEHDKAYI R.EV.RISI SSNAGAAAPASSVG DPQPLGQAPSGLG PEGST5GAM.DDS QPASSLGADTM.A DGG.T.EGTEPVA NWGIRPKAMRFRLFNI F.RLN.K MVKI Y.F.RSL.VKIV	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H HQLA.DLK.H. QQLA.DLK.H. QQLA.DIKF.H. VP3 STMAEGGGPVGDAGQG TNT.T.S.A.MA.NNE -E.RAAA.AAVEG -E.RAAA.AAVEG. -E.RAAA.AAVEG. L.NN. A.EAM.SSG. VKEVTVQDFNTTIGNNL .QN.GT.A. .ST.A. .T.QTK.A.	PRRFQERLK AEQ.Q AEKA QD.IDS.Q SADGVGNSS A. AA. A.	114 105 104 104 113 231 225 219 215 221 350 344 335 340
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV	MSLISDAIPDW A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF GF .Q. GNWHCDSQWLE 	VLERLVKK DT DK YP Y Y Y Y V V V	KGVNAAADFYHLES(CLSEGIRQWWK KP ULSEGVREWWA QP VUEGELRE.LGA TSWRN.KA (RVLEPFGLVEDS-1	GPPRPKANC 	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NNF.I KGEDEPRLPDTSSQTPKNKKK UEHSPVEPDSSGTGKAGQOPA LI.SPQQPDSSTGIGKKGKQPA DDHFPK.KKARTEEDS.PSTSS GKLTDHYPVV	PGVNYLGPFNGLDKGE: KGP RKERPSGGAEDPGEGT: RKRLNFGQTG.ADSVPI KKKLVFEDETGAGD.P DA.AGPSGSQQLQIPAK KLTEEVSAGGSSAVQI PHCHFSPRDWQRLINNI 	PVNEADAAALEHDKAYI 	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H., QQL.A.DLK.H., QQL.A.DIK.H., QQL.A.DIK.H., QQL.A.DIK.F.H., VP3 SSIMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. E.RAAA.AVEG L.NN. A.EAVEG. VKEVTVQDFNTTIGNNL 	PRRFQERLK AEQ.Q AEQ.Q AEQ.Q SADGVGNSS A. A. A. A. A. A. A. TSTVQVFA T T T T	114 105 104 113 2231 225 219 215 221 350 344 335 335 3340
в	AAAV AAV2 AAV4 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EF GF GF GF GT.MC DNWHCDSQWLE T.MC DT.S. .DT.MC DLT.S. MC	VLERLVKK DI PI VF YF YF YF YF V V V V V SQL.I.TS GH.T.TS DRKS DRKS SATEGTF HQ.CL	KGVNAAADFYHLES(CLSEGURGWWK KP ULSECVREWMA, QP SVGEGLRE, LG, A T SWRN, KA QEVLEPFGLVEDS-1 L	GPPRPKANC P.PAB .A.K .A.Q.P. VP2 KTAPTGDKF EGKKRE EGKKRE KRIN NAKKWI YKRIQGPS- LCE .RE.KSG.V. A.TSGT- .A.TSGT- ATSGT- ATSGT- 	QTQESLEKDDSRGLVF RHKNAI HQAI S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQP LI.SPQPDSSTGTGKAGQPP DDHFPK.KKATEEDS.PSTSG GKLTDHYPV	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI 	DLE IKDGHNPYFEYNEAD RQLDS.DLK.H VQLA.DLK.H QQLA.DLK.H QQLA.DIKF.H VP3 SSIMAEGGGGPVGDAGQG NT.T.S.A.MA.NNE. E.RAAA.AAVEG E.RAAA.AAVEG E.RAAA.AAVEG L.NN A.EAM.SSG. VKEVTVQDFNTTIGNNL 	PRRFQERLK AE AEQ.Q AEK.A QD.IDS.Q SADGVGNSS A. A. A. A. A. T TT TT TT T T T TT T	114 105 104 104 113 225 219 215 221 350 344 335 335 340 464 460
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2	MSLISDAIPDN .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EF GF	VLERLVKK D1 P1 YF Y Y Y Y Y S S S S	(GVNAAADFYHLES(CLSEGIRQWWK.KP ULSEVUREWA.QP EVGEGLRE.LG.A T.SWRN.KA (RVLEPFGLVEDS-1 L.QAGI	GPPRPKANC P. PAB .A.K .A.Q.P. VP2 KTAPTGDKF GKKRF EGKKRF EGKKRF KRT NAKKNT YKRIQGPS- GKRT GKTSG- GKTSG- GYCTLNYNN G,VTG.	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQOPP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQPDSSTGIGKKGQOP LI.SPQQTSSTGIGKKGQOP LI.SPQQQT.S.S.S.S.S.C.S.C.S.C.S.C.S.C.S.C.S.C.S.	PGVNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI A	DLE IKDGHNPY FEYNEAD RQLDS.DLK.H VQLA.DLK.H QQLA.DIK.H VP3 STMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. -E.RAAA.AAVEG -E.RAAA.AAVEG -E.RAAA.AAVEG QVKEVTVQDFNTTIGNNL 	ARRFQERLK AE AEQ AEKA QD.IDS.Q ADGVGNSS A. AA. AA. ATSTVQVFA T TT T T T T T T T T T T T T T	114 105 104 113 225 219 215 221 350 344 335 335 340 464 460
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV5 GP AAAV AAV5 AAV2 AAV4 AAV5	MSLISDAIPDW A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGH</u> EF GF .Q. GINWHCDSQWLE T.MC .DT.MC .DMC DKDYQLPYVLC .SS.EMI DV.	VLERLVKK DI DN YP Y Y XAIFQAKK V V V V V V V SNGVVTRI SDR.I.TS GH.T.TS JDRKS ; T.I.K. SATEGTFI JAGQ.SL NGCL	KGVNAAADFYHLES(CLSEGUREWMA, QP ULSEGVREWMA, QP EVGEGLRE, LG, .A TSWRN, KA (RVLEPFGLVEDS-1 .LEPV .LEPV	GPPRPKANC P. PAE .A.K .A.Q.P. VP2 KTAPTODKF GKKRF EGKKRF EGKKRF EGKKRF KRT NAKKNT YKRIQGPS- Q.SSQ Q.SSQ CES RE.KSG.V A.TSGT- GYCTLNYNN G.VTG G.VTG G.VTG G.VTG	QTQESLEKDDSRGLVF RHK	PGVNYLGPFNGLDKGE: K.GG G.R.P K.G.P K.G.P K.G.P K.SO K.SO 	PVNEADAAALEHDKAYI 	DLEIKDGHNPYFEYNEAD RQLDS.DLKH QQL.A.DLKH QQL.A.DIKH QQL.A.DIKF.H VP3 SIMAEGGGGFVGDAGQG FNTT.S.A.MA.NNE. L.NN. -E.RAAA.AAVEG L.NN. .EAM.SSG. VKEVTVQDFNTTIGNNL .ST.A .ST.A .ST.A .ST.A .TQTK.A .YL.RTNPSG I,YL.RTNPSG I,YR.V.TNN .YR.V.TNN	PRRFQERLK AE A AE AE	114 105 104 113 225 219 215 221 350 344 335 335 340 464 460 455
в	AAAV AAV2 AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGR</u> EF G Q GNWHCDSQWLE T.MG .DT.S DT.MG SS.EMI DV. DEH	VLERLVKK DI DI P VY Y Y Y V V V V	KGVNAAADFYHLESK KGVNAAADFYHLESK ULSEOUREWWA.QP SVGEGLRE.LGA TSWRN.KA <u>(RVLEPFGLVEDS-1)</u> LEPV LL.EPV J FUL QAG S GA TTTTWVLPSYNNHL S	GPPRPKANC 	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VENSPVEPDSSSGTGKAQQPA DDHFPK.KKARTEEDS.PGTS GGDNNNKFFGFSTPWGYFDYNF .AS.D.HY.Y.F. -SLQS.TYN.F. D.S.A.AY.Y.F. SQ.A.VQYA.Y.F. EAVDRSAFYCLDYFFSDMI MGSQAVG.S.E.Q. TSQQQT.N.E.Q.	PGVNYLGPFNGLDKGE: KGP RKERPSGGAEDPGEGT: RKRLNFGQTG.ADSVPI KKKLVFEDTGAGD.P DA.AGPSGSQQLQIPAK KLTEEVSAGGSSAVQI PHCHFSPRDWQRLINNI 	PVNEADAAALEHDKAY 	DLE IKDGHNPYFEYNEAD RQLDS.DLK.H., QQL.A.DLK.H., QQL.A.DIK.H., QQL.A.DIK.H., QQL.A.DIK.F.H., VP3 SSIMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. E.RAAA.AAVEG. L.NN. A.EAVEG. QVKEVTVQDFNTTIGNNL 	PRRFQERLK AEQ.Q AEQ.Q AEQ.Q AEQ.Q SADGVGNSS A. A. A. T T T T T T T T T T T T T T	1114 105 104 104 113 225 219 215 221 350 344 335 340 464 460 455 447 453
в	AAAV AAV2 AAV4 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGR</u> EF GF GF GF GT.MC DT.MC DT.S .DT.MC DKDYQLPYVLC .SEMI DV. .DEH	VLERLVKK DN PI VF YF YF YF YF V.V.V.V.V.V.V.V.V.V.V.V.V.V.V.V.V.V.V.	KGVNAAADFYHLES(CLSEOIRQWWK KP ULSEOURBWA QP SVGEGLRE LG. A T. SWRN KA QULEPFGLVEDS-1 L . EPV L . QAG	GPPRPKANC P. PAB .A.K .A.Q.P VP2 KTAPTGDKF EGKKRE EGKKRE EGKKRE KRIN VKRIQGPS- LCB .RE.KSG.V. A.TSGT-	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NNF.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPP LI.SPQPDSSTGTGKKGKQPP DDHFPK.KKATTEEDS.PSTSG GKLTDHYPV	PGYNYLGPFNGLDKGE K.G. G.R. K.G.P RKERPSGGAEDPGEGT RKENPGQTG.ADSVP KKLVFEDETGAGD.P. KKLVFEDETGAGD.P. IDA.AGPSGSQQLQIPA KLTEEVSAGGGSSAVQ FHCHFSPRDWQRLINN 	PVNEADAAALEHDKAYI 	DLE IKDGHNPYFEYNEAD RQLDS.DLK.H VQLA.D.LK.H QQLA.D.LK.H QQLA.D.IK.H VP3 STMAEGGGGPVGDAGQG NT.T.S.A.MA.NNE. E.RAAA.AAVEG E.RAAA.AAVEG NN E.RAAA.AAVEG L.NN A.EAM.SSG. VKEVTVQDFNTTIGNNL 	PRRFQERLK AE AEQ.Q AEKA QD.IDS.Q SADGVGNSS A. A. A. A. T TT 	1114 105 104 104 113 225 219 215 221 350 344 335 340 464 460 4455 447 453
B	AAAV AAV2 GP AAAV AAV2 AAV2 AAV4 AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EFE GF GF GF GF GT.MC DT.S. .DT.MC DT.MC D MC DKDYQLPYVLC .SEMI DV. .DEH	VLERLVK# D1 D1 D1 N N 	(GVNAAADFYHLES((CLSEGIRQWWK, KP (LSEQUREWMA, QP BVGEGLRE, LG, A TSWRN, KA (RVLEPFGLVEDS-1 LEPV LQAGI	GPPRPKANC P. PAB .A.K .A.Q.P. VP2 KMAPTGDKF GKKRF EGKKRF EGKKRF MAKKNT YKRIQOPS- Q.SSQ- LGE LGE LGE A.TSGT- GYCTLNYNN LTLNRI LTLNRI MHT.C ASNITKNNV	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NNF.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSTGGKAGQ2PA LI.SPQQPDSSTGIGKKGQ2PA LI.SPQQPDSSTGIGKKGQ2PA LI.SPQQPDSSTGIGKKGQ2PA LI.SPQQPDSSTGIGKKGQ2PA CHAPPESSTGGKAGQ2PA GGDNNKFFGFSTPWGYFDYNF GGDNNKFFGFSTPWGYFDYNF GGDNNKFFGFSTPWGYFDYNF .S.A.AY.Y.F.F. SQ.A.VQYA.Y.F.F. SQ.A.VQYA.Y.F.F. SQ.A.VQYA.Y.F. NGGQAVG.S.E.Q. TSQQQT.N.E.Q. NTENPTE.S.F.E.K. NGARFN.E.Q. TSVWE-KGKQWELDWRTNLMQF	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI R. EV.R. IS.I SSNAGAAAPASSVG5 DPOPLQQAPSGLG PEGSTSGAM.DDS DFASSLGADTM.A DGG.T.EGTEPVAI NWGIRPKAMRFRLFNI(F.R.N.K MVKI Y.F.RSL.VKIV FHSMFAHNQTLDRLMNI Y.S.S Y.S.S S.D.N.FK.A. 	DLE IKDGHNPY FEYNEAD RQLDS.DLK.H VQLA.DLK.H QQLA.DIK.H QQLA.DIKF.H VP3 ST MAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. L.NN A.EAM.SSG. VKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKEVTVQDFNTIGNNL .QNKA.	PRRFQERLK AEX.A AEX.A QD.IDS.Q SADGVGNSS A.	114 105 104 104 113 225 221 350 344 335 340 460 455 447 453
B	AAAV AAV2 AAV4 AAV5 GP AAV4 AAV2 AAV4 AAV5 GP AAV4 AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV2	MSLISDAIPDM .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EF. GF. GF. .Q. .Q. .Q. .GNWHCDSQWLE 	VLERLVKK DI DN YDR Y Y X.V V. V. V. V. V. V. V. V. V. V. V. V	(GVNAAADFYHLES(CLSEGIRQWWK.KP ULSEGVIREWA.QP EVGEGLRE.LG.A T.SWRN.KA (RVLEPFGLVEDS-1 L.OPU CLQAGI	GPPRPKANC P. PAB .A.K .A.Q.P. .A.Q.P. .VP2 KMAPTODKF GKKRF GKRF GKKF GKKF 	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPF LI.SPQQPDSSTGIGKKGQPP DDHFPK.KARTEEDS.PSTSS GKLTDHYPVV	PGVNYLGPFNGLDKGE KG RK.GP RKERPSGAEDPGEGT: RKRLNFGQTG.ADSVPI KKRLVFEDETGAGD.P DA.AGPSGSQQLQIPA KLTEEVSAGGSSAVQI PHCHFSPRDWQRLINN S.W RTGNNFEFTYTFEDVPI T.S I.S.K. 	PVNEADAAALEHDKAYI R. EV.RISI K. SVS SSNAGAAAPASSVG DPQPLGQ. APSGLG PEGSTSGAM.DDS QPASSLGADTM.A DGG.T.EGTEPVAJ NWGIRPKAMRFRLFNI F. RSL.VKI MVKI Y.F. RSL.VKIV FLSMFAHNQTLDRLMNI Y.S.S S.PS.N.FK.A. S.D MUNTLAFSRTVYDQTP PQSGVLIFGKQGSEK	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H QQLA.DLK.H. QQLA.DIK.H. QQLA.DIKF.H. VP3 SIMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE L.NN. A.EAM.SSG VKEVTVQDFNTTIGNNL QN.GT.A. TSNGE.VA T.QTK.A. T.QTK.A. YL.RTNTPSG YL.RTNTPSG YR.V.TNN N.NE N.NE	PRRFQERLK AEQ.Q AEX QD.IDS.Q SADGVGNSS A.	114 105 104 104 113 231 225 229 215 221 350 344 4355 340 453 578 570
в	AAAV AAV2 AAV4 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5	MSLISDAIPDW A-ADGYL FV.HP FV.HP .TFL.SFEE. DDTSFG <u>GNLGR</u> EF GF Q GNWHCDSQWLE T.MG T.MG T.MG T.MG T.MG T.MG T.MG T.MG T.MG 	VLERLVKK DI DI PI VF Y V V V V V V V	KGVNAAADFYHLESK KULSEGUIRQWWA.QP ULSEGUIRDWA.QP SVGEGLRE.LG.A ILSEGUIRDUIS LL.CHURGUIRDUIS GENERAL LL.CHURGUIRDUIS GENERAL LL.CHURGUIRDUIS GENERAL LL.CHURGUIRDUIS GENERAL LL.CHURGUIRDUIS GENERAL LL.CHURGUIRDUIS LL.CHURGUIRDUIS	GPPRPKANC P. PAE .A.K .K. P .K.Q. P VP2 KTAPTGDKF GKKRI EGKKRI EGKKRI KTAPTGDKF Q.SSQ LGE RE.KSG.V. ATLNRI 	QTQESLEKDDSRGLVF RHK	PGVNYLGPFNGLDKGE: KGP RKERPSGGAEDPGEGT: RKELNFQUTG.ADSVPI KKRLVFEDETGAGD.P DA.AGPSGSQQLQIPAI KLTEEVSAGGSSAVQI FHCHFSPRDWQRLINNI S.W RTGNNFEFTYTFEDVPI T.SK 	PVNEADAAALEHDKAYI R. EV.RIS.I K. SV. SSNAGAAAPASSVG DPQPLQQAPSGLG: PEGST5GAM.DDS QPASSLGADTM.A DGG.T.EGTEPVA DGG.T.EGTEPVA NWGIRPKAMRFRLFNI (F.RSL.VKIVKI MVKIVKIVKIVKI HSLK.KI.V FHSMFAHNQTLDRLMNI Y.S.S S.PS.N.FK.A. S.D MONTLAFSRTVYDQTTM PQSGVLIFGKQGSEK SNSQ.I.AGFKQNGN	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H., QQL.A.DLK.H., QQL.A.DLK.H., QQL.A.DIK.H., QQL.A.DIK.F.H., VP3 SIMAEGGGGFVGDAGQG FNT.T.S.A.MA.NNE. L.NN. L.NN. A.EANESG VKEVTVQDFNTTIGNNL 	PRRFQERLK AEQ.Q AEQ.Q.Q AEQ.Q.IDS.Q AA. A. A. A. A. A. A.	114 105 104 104 113 231 225 229 215 221 350 344 460 455 345 4457 453
в	AAAV AAV2 AAV4 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV2 AAV4 AAV2 GP AAAV AAV2 AAV4 AAV2 AAV4 AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF GF GF G GNWHCDSQWLE T.MC DT.MC DT.MC DT.MC DY.MC SE DY.MC PLOUPYVLC SE DV. DEH .DEH	VLERLVKK DI DI PI YF Y	KGVNAAADFYHLESG CLSEOIRGWWK KP ULSEOUREWA.QP SVGEGLRE.LG.A T.SWRN.KA QULEPFGLVEDS-1 L.OAG L.OAG I.L.PPF L.OAG FPFPADIYTIPQY	GPPRPKANC P. PAB .A.K .A.Q.P. VP2 KTAPTGDKF EGKKRE EGKKRE EGKKRE KRIN VKRIQGPS- Q.SQ LCB LCB LCB A.TSGT-	QTQESLEKDDSRGLVF RHKRA	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI A	DLE IKDGHNPYFEYNEAD RQLDS.DLK.H VQLA.DLK.H QQLA.DLK.H QQLA.DIKF.H VP3 STMAEGGGGPVGDAGQG VTT.T.S.A.MA.NNE. E.RAAA.AAVEG L.NN A.EAM.SSG. VKEVTVQDFNTTIGNNL 	ARFQERLK AE AEQ.Q AEKA QD.IDS.Q ADGVGNSS A. A. A. A. A. T TTTQS-R. LNAGTATT TGGV TGGV TGGV TGGV TGGV TGGV TGGV 	114 105 104 104 113 231 225 219 225 221 350 344 335 335 344 455 570 570 569 569 569
в	AAAV AAV2 GP AAAV AAV2 AAV4 AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EF GF GF GF GT.MC DT.MC DT.MC DT.MC D MC DMC D MC DV. D MC DV. V. V. V. V. V. V. V. V. 	VLERLVK# DI DI DI NE NE 	KGVNAAADFYHLES(CLSEGURGWWK, KP ULSEGUREWWA, QP SVGEGLRE, LG, A TSWRN, KA (RVLEPFGLVEDS-1 LCPV LCPV	GPPRPKANC P. PAB K.P.P. K.P.P. VP2 KMAPTGDKF GKKRF EGKKRF EGKKRF MAKKNT YKRIQGPS- Q.SSQ LGE LGE LGE A.TSGT- GVCTLNVNN A.TSGT- GVCTLNVNN A.TSGT- SADNITKNNV SADNITKNV SADNITKNNV SADNITKNV S	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPA LI.SPQQPDSSTGIGKKGKQPA DDHFFK.KKARTEEDS.PSTSS GKLTDHYPVV	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI A R. EV.R. IS. SSNAGAAAPASSVG5 DPOPLQQ. APSGLG PEGSTSGAM.DDS PASSLGADTM.A DGG.T.EGTEPVAI NWGIRPKAMRFRLFNI MVKI Y.F.RSL.VKI HSLK.KI.V FHSMFAHNQTLDRLMNI .SY.S.S .Y.S.S S.D. .Y.S.S.N.FK.A. S.D. MDNTLAFSRTVYDQTP FQSGVLIFGKQGSEK SNSQ.I.AGPKQNGN. ENTMIFN.QPANGG. A.I.GTAKDP.RSG52	DLE IKDGHNPY FEYNEAD RQLDS.DLK.H QQLA.D.LK.H QQLA.D.IK.H QQLA.D.IK.H QQLA.D.IK.H VP3 STMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. L.NN.A.E. L.NN.A.E. L.NN.A.E. L.NN.SG L.NN.SG L.NN.SG L.NN.SG 	PRRFQERLK AEX.A AEX.A QD.IDS.Q SADGVGNSS A. T T T T T T T T T	114 105 104 103 231 225 219 225 2219 221 330 344 335 334 460 455 447 455 340 578 5578 5578
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW A-ADGYL FV.HP FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF GF G Q GNWHCDSQWLE T.MG T.MG MG DKDYQLPYVLG .SS.EMI DV. DEHV. DEH. MYSRATKTNMA QFNKLAGRY. QFNKNLAGRY. QFKLV.GAYC Heparit	VLERLVKK DN DY F 	KGVNAAADFYHLESK CLSEGIRQWWK.KP ULSEGVIEWWA.QP SVGEGLRE.LG.A T.SWRN.KA CRVLEPFGLVEDS L.EPY L.EPY L.EPY J.A.PQY S.A.T.SWRN.KA CRVLEPFGLVEDS L.EPY L.EPY GENERS GR FPFPADIYTIPQY N.VFMV N.VFMV	GPPRPKANC P. PAE .A.K. P. PAE K. P. P. VP2 KTAPTGDKF EGKKRI EGKKRI EGKKRI EGKKRI EGKKRI CKRIW YKRIQOPS- Q.SSQ LCE RE.KSG.V. ATLNRI ATLNRI ATLNRI ATLNRI ATLNRI ATLNRI SGVNRASVE TGGTDNYAN	QTQESLEKDDSRGLVF RHK	PGVNYLGPFNGLDKGE: K.GG K.G.R. K.G.R. K.G.R. K.G K.G K.G RKENPEGGAEDPGEGT: 	PVNEADAAALEHDKAYI R. EV.RISI K. SV SSNAGAAAPASSVG DPQPLGQAPSGLG: PEGST5GAM.DDS QPASSLGADTM.A DGG.T.EGTEPVA2 NWGIRPKAMRFRLENI(F. RSL.VKI MVKI Y.F. RSL.VKIV HSLK.KIV FHSMFAHNQTLDRLMNI Y.S.S Y.S.S Y.S.S S.PS.N.FK.A. S.PS.N.FK.A. S.D PNTLAFSRTVYDQTE FQSGVLIFGKQGSK SNSQ.I.AGPKQNSN. ENTMIFN.QPANPG.A. GIAKDP.RSGE	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H., QQL.A.DLK.H., QQL.A.DIK.H., QQL.A.DIKF.H., VP3 SIMAEGGGGPVGDAQQG INT.T.S.A.MA.NNE. L.NN.A. 	PRRFQERLK AE AE AE AE QD.IDS.Q ADGVGNSS A. A. A. A. TSTVQVFA T TT T T T T T TTTCS-R. LNAGTATT TCGV D.S.NA TSVGIDA P.ATEQ ATDT.M V.R.AYNV GWKP	114 105 104 113 231 225 221 350 3340 455 335 464 453 340 578 5570 5570 5569 560
в	AAAV AAV2 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF G G GNWHCDSQWLE 	VLERLVKK DN PH F YF YF YF YF YF YF YF X.V CAIFQAKK V.V SAUGUTRI SOUCH	KGVNAAADFYHLES(CLSEGIRQWWK KP LLSEGVREWA.QP SVGEGLRE.LG.A T.SWRN.KA CLVLEPFGLVEDS-1 L.EQVREWA.QP L.OAGI L.OAGI TRTWVLPSYNNHL S.A.T.SWRN.HL S.A.T.SWRNHL S.A.T.S.V.SV L.OAGI TPFPADIYTIPQY L.A.PQVFL. A.S.V.AL. PEFFERDQUIFTGG S.V.AL. PGPFFRDQUIFTGG S.V.AL. PGPFFRDQUIFTGG S.K.L., RVRMY S Y.MNNQGALPGMVW(GPPRPKANC P. PAB .A.K .A.Q.P. VP2 KTAPTGDKF EGKKRE EGKKRE KRI NAKSUT YKRIQGPS- LCB .RE.KSG.V. A.TSGT- A.TSGT- A.TSGT- A.TSGT- A.TSGT- SADNINKST, MHT.Q SQUNRASVE TGGTDNYAN QNRDIYPTG	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEMSPVEPDSSSGTGKAGQQPA LI.SPQQPDSSTGTGKKGXQPA GGDNNNKFFGFSTPWGYFDYNF .AS.D.HY.Y.F SQ.STYNF. D.S.A.AY.Y.F.F. D.S.A.AY.Y.F.F. EAVDRSAFYCLDYFPSDMI NGSQAVG.S.E.Q. TSQQQT.N.E.Q. TSQQQT.N.E.Q. TSQQQT.N.E.Q. TSQQQT.S.F.E.K. NGARFNE.Q. TSQDSILKYETHSTLDGRWSAI AFATTNRKELEGASYQVPPC WNI.S-N.NKVN.KD.QY.L.	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI A	DLE IKDGHNPYFEYNEAD RQLDS.DLK.H VQLA.DLK.H QQLA.DIK.H QQLA.DIKF.H VP3 STMAEGGGGPVGDAGQG NT.T.S.A.MA.NNE. 	WRRFQERLK AE AE AE AE AE QD.IDS.Q SADGVGNSS	114 105 104 113 231 225 221 350 342 464 465 578 578 578 569 560 567 698
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EFE GFG GFG GT.MC DT.MC DT.MC DT.MC DMC DKDYQLPYVLC .SS.EMI DV. DEHV. DEHV. QFNKLAGRY. QFKK.V.GAYC Heparit MGAVPTNNQSI Y.S.S.L.RC	VLERLVK# DN PH F YF Y	KGVNAAADFYHLES(CLSEGUREWWA, QP ULSEQUREWWA, QP SVGEGLRE, LG, A TSWRN, KA KRVLEPFGLVEDS-J LCPV LCPV LCPV CPV CPV	GPPRPKANC P. PAB .A.K .A.Q.P. PAB K.P.P. VP2 KMAPTGDKF GKKRF EGKKRF EGKKRF MAKKNT VKRIQOPS- LGE LGE LGE LGE A.TSGT- GVCTLNYNN A.TSGT- GVCTLNYNN MHT.C SADNINSEY SGVNRASV5 TGGTDNYAN QNRDIYPTC V.LQ.	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP SKTGA.AY.Y.F. SLQS.TYN.F. EAVDRSAFYCLDYFPSDMI NGSQAVG.S.E.Q. TSQQQT.N.E.Q. TSQQQT.N.E.Q. TSQQQT.N.E.Q. FSVWE-KGKQWELDWRTNLMQF SWTGASTRFHSTLORGBSAI AFATTNRMELEGASYQVPPC WNI.S-N.NKVN.KD.QY.L. THLAKIPDTDNHFHPSPLIGRF	PGYNYLGPFNGLDKGE: 	PVNEADAAALEHDKAYI A R. EV. R. IS. SSNAGAAAPASSVG5 DPOPLQQAPSGLG PEGSTSGAM.DDS PASSLGADTM.A DGG.T.EGTEPVAI NWGIRPKAMRFRLFNI(F.R.N.K MVKI Y.F.RSL.VKI HSLK.KI.V FHSMFAHNQTLDRLMNI .SY.S.S S.S S.S.N.FK.A. S.S.N.FK.A. S.S.N.FK.A. S.S.N.FK.A. S.S.S.I.GERQNGN. ENTMIFN.QPANPG A.I.GIAKDP.RSGE PANPSETFQTAKVASF: 	DLE IKDGHNPYFEYNEAD RQLDS.DLK.H QQLA.D.LK.H QQLA.D.IK.H QQLA.D.IK.H QQLA.D.IK.H QQLA.D.IK.H VP3 SIMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. 	PRRFQERLK AEX.A AEX.A QD.IDS.Q SADGVGNSS A. T T T T T T T T T	114 105 104 113 225 2219 215 221 350 3355 3355 221 464 455 570 560 560 560 560
в	AAAV AAV2 AAV3 GP AAV4 AAV5 AAV4 AAV5 AAV4 AAV4 AAV4 AAV4 AAV4 AAV4 AAV4 AAV4	MSLISDAIPDW A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFGGNLGH EF GF .Q. GINWHCDSQWLE T.MC .DT.MC .DT.MC MC DKDYQLPYVLC .SS.EMI DV. DEH .SS.EMI DV. DEH .SS.EMI DV. DEH MGAUPTNNOGI Y.S.S.L.RC MGAUPTNNOGI Y.S.S.L.RC	VLERLVKK DN DN PF Y S.T.I.K. SATEGTF MG Y Y NPKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK NDKK ND	(GVNAAADFYHLES(CLSEGIRQ/WWK.KP ULSECUREWWA.QP EVGEGLRE.LG.A T.SWRN.KA EVGEGLRE.LG.A T.SWRN.KA KRVLEPFGLVEDS-1 L.QAGI	GPPRPKANC P. PAB .A.K .A.Q.P. PAB K.P.D. .K.Q.P. .VP2 KTAPTODKF GKKRF EGKKRF EGKKRF GKF 	QTQESLEKDDSRGLVF RHKNAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQQPDSSTGIGKKGQPP 'LI.SPQPDSSTGIGKKGQPP 'LI.SPQPDSSTGIGKKGQPP 'LI.SPQPDSSTGIGKKGQPP 'LI.SPQPDSSTGIGKKGQPP 'LI.SPQPDSSTGIGKKGQPP 'LI.SPQPDSSTGIGKKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQPP 'LI.SPQPDSSTGIGKGQP 'LI.SPQPDSSTGIGKGQP 'LI.SPQPDSSTGIGKGQP 'LI.SPQPDSSTGIGKGQP 'LI.SPQPDSSTGIGKGQP 'LI.SPQPDSSTGIGKGQP 'LI.SPQPDSSTGIGKGQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGIGKQQP 'LI.SPQPDSSTGI	PGVNYLGPFNGLDKGE KG RK.G RKERPSGAEDPGEGT: RKERPSGAEDPGEGT: RKKLVFEDETGAG.P DA.AGPSGSQQLQIPA KKLVEDETGAG. RKENFEFTYFEDVP 	PVNEADAAALEHDKAYI A	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H QQLA.DLK.H QQLA.DIK.H QQLA.DIKF.H VP3 SIMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. 	PRRFQERLK AEQ.Q AEX.A QD.IDS.Q SADGVGNSS A.	114 105 104 113 231 225 221 3500 335 335 335 335 335 340 460 4455 4453 578 5578 5570 560 567 569 560 567
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGR</u> EF GF GF GF GF GF GF GF GF GF GF GF GF GF GF GF MGAUCDSQWLE T.MG F MGAUCDSQWLE F 	VLERLVKK DI 	KGVNAAADFYHLESG KGVNAAADFYHLESG LLSEOUREWWA.QP BUSEOUREWWA.QP SUCCEVEWA.QP SUCCEVEWA.QP SUCCEVEWA.QP L.SCVREWA.QP CRVLEPFGLVEDS-J L.SCVREWA.QP GA I.SVREWA GA I.SVREWA PTTTWVLPSUNHLY A.T.S. GA I.VFMV. A.PQVFL. A.S.V.AL. PEOPFFREDQUETEGG J.A.PQVFL. A.S.V.AL. PEOPFFREDQUETEGG S.M.G.T.GWRLS K.L.RVRAY B VAVNNQGALPGMVWQ NYLLEIN.EIN.S.Y	GPPRPKANC P. PAB .A.K .A.Q.P. PAB .K.P.PTGDKF KTAPTGDKF EGKKRE EGKKRE KRI N.AKKWT YKRIQGPS- LCB RE.KSG.VG A.TSGT- GYCTLNYNN A.TSGT- GYCTLNYNN TLTL. A.TSGT- GYCTLNYNN TLTL. A.TSGT- GYCTLNYNN TLTL. A.TSGT- GYCTLNYNN TLTL. A.TSGT- GYCTLNYNN TLTL. A.TSGT- GYCTLNYNN TLTL. A.TSGT- GYCTLNYNN A.TSGT- GYCTLNYNN A.TSGT- GYCTLNYNN A.TSGT- GYCTLNYNN A.TSGT- GYCTLNYNN A.TSGT- 	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEMSPVEPDSSSTGKAGQQPA LI.SPQQPDSSTGGKAGQQPA DHFPK.KKATTEDS.PSTSS GGDNNKFFGFSTPWGYFDYNF .AS.D.HY.Y.F SQ.S.YYN.F D.S.A.AY.Y.F.F. SQ.A.VQYA.Y.F. EAVDRSAFYCLDYFPSDMI NGSQAVG.S.E.Q. TSQQQT.N.E.Q. TSQQQT.N.E.Q. TSQQQT.N.E.Q. TSQQQT.N.E.Q. TSQDSLIKYETHSTLDGRWSAI AFATTNRKELEGASYQVPPC WNI.S-N.NKVN.KD.QY.L. THLAKIPDTDNHFHPSPLIGRR PIWH.G.MG	PGVNYLGPFNGLDKGE: KGP KGP KGP KGP KGP KGP KGP KGP 	PVNEADAAALEHDKAYI A	DLE IKDGHNPYFEYNEAD RQLDS D LK. H VQL A. D LK. H VQL A. D IK. H VQL A. D IKF. H VY3 SSIMAEGGGGPVGDAGOG NT. T. S. A. MA. NNE. 	PRRFQERLK AE AE AE AE QD.IDS.Q ADGVGNSS A. AA. A. A. TSTVQVFA T TTTQS-R. ITT TTTGS-R. LNAGTATT TCGV D.S.NA TTSVGIDA P.ATEQ ATDT.M V.R.AYNV GWKP	114 105 104 113 231 225 221 350 344 460 455 578 570 569 567 698 690 679 275 205 205 205 205 205 205 205 20
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF G G G GNWHCDSQWLE 	VLERLVKK DN DN (AIFQAKK VF YF YF YF YF YF (AIFQAKK V X.V SATEGATE SATEGTF NGCL SATEGTF NGCL NGCL NGCL NGCL NGCL NFKK NT.K.F TMG NTKK.F TMG YTPGTRA NRQAAT. SNLP.VU T.APATG T.APATG	KGVNAAADFYHLES(CLSEOIRQWWK KP ULSEOUREWA.QP SVGEGLRE.LG.A T.SWRN.KA (RVLEPFGLVEDS-) L.E.CVREWA.QP L.OAG	GPPRPKANC P. PAB .A.K .A.Q.P. PAB .K.P.PAB .K.P.PAB .A.K.P. .K.PAB GKKR GKKR GKKR GKKR GKKR GKKR GKKR GKKR GKKR GKKR GKR 	QTQESLEKDDSRGLVF RHK	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI 	DLE IKDGHNPY FEYNEAD RQLDS.DLK.H QQL.A.DLK.H QQL.A.DIK.H QQL.A.DIK.H VP3 STMAEGGGGPVGDAGQG VTT.T.S.A.MA.NNE. 	PRRFQERLK AE AEQ.Q AEK.A QD.IDS.Q ADGVGNSS A. A. A. A. T TTTQVFA T T.IT -GSSGRAL TTTQS-R. LNAGTATT TGGV D.S.NA TNSVGIDA P.ATEQ ATDT.M V.R.AYNV G.WKP	114 105 104 113 225 229 215 221 350 3350 460 4457 4453 5788 570 560 567 560 567 698 699 687
в	AAAV AAV2 AAV5 GP AAV2 AAV4 AAV5 GP AAV4 AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAV2 AAV2 AAV4 AAV5 GP AAV2 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLGK</u> EF GF GF GT.MC DT.MC DT.MC DT.MC DMC DKDYQLPYVLC .SEMI DV. DEHV. DEHV. DEHV. DEHV. GYK.V.GAYC HEPATI MGAVPTNNQSI Y.S.S.L.RC NL.GGD.N	VLERLVK# DN DI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI 	KGVNAAADFYHLESG CLSEGUREWWA.QP VLSEGUREWWA.QP SUCGELRE.LG.A T.SWRN.KA KRVLEPFGLVEDS-1 L.OPU L.OPU L.OPU GAG FPFPADIYTIPQY	GPPRPKANC P. PAB .A.K .A.Q.P. PAB .K.P.P VP2 KTAPTGDKF GKKRT GKKRT GKKRT GKKRT KRI NAKKNT VKRIQGPS- Q.SSQ. 	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP DDHFFK.KKARTEEDS.PSTSS GGLNNNKFFGFSTPWGYFDYNF -SLQS.TYNF. D.S.A.AY.YF. D.S.A.AY.Y.F. EAVDRSAFYCLDYFPSDMI NGSQAVG.S.E.Q. TSQQQT.N.E.Q. TSQQQT.N.E.Q. FSVWE-KGKQWELDNFTNLMQF SWTGAFWLNGRDSLN AFATTNRMELEGASYQVPPC WNI.S-N.NKVN.KD.QY.L. THLAKIPDTDNHFHPSPLIGRF PIWH.GMG	PGYNYLGPFNGLDKGE: 	PVNEADAAALEHDKAYI A R. EV. R. IS. SSNAGAAAPASSVG5 DPOPLQQAPSGLG PEGSTSGAM.DDS DGG.T.EGTEPVA DGG.T.EGTEPVA NWGIRPKAMRFRLFNI F.RN.K MVKI Y.F.RSL.VKI HSLK.KI.V FHSMFAHNQTLDRLMNI S.S S.S S.S MDNTLAFSRTVYDQTE FOSGVLIFGKQNGN. ENTMIFN.QPANPG. A.I.GIAKDP.RSGE PANPSETFQTAKVASF D.PVEYVHQ.WN.Y	DLE IKDGHNPY FEYNEAD RQLDS.DLK.H QQLA.D.LK.H QQLA.D.LK.H QQLA.D.IK.H QQLA.D.IK.H VF3 STMAEGGGGPVGDAGQG STMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. 	PRRFQERLK AEX.A AEX.A QD.IDS.Q SADGVGNSS A. T T T T T T	114 105 104 113 225 221 225 221 350 455 3355 340 455 570 550 560 560 567 698 699 689 689 689
В	AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP AAAV AAV2 AAV4 AAV5 GP	MSLISDAIPDW .A-ADGYL .TDGYL .FV.HP .TFL.SFEE. DDTSFG <u>GNLG¥</u> EFE GF GF GF GT.MC DT.MC DT.MC D	VLERLVK# DI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI IDI 	KGVNAAADFYHLES(CLSEGUREWWA, QP VLSEVUREWWA, QP SVGEGLRE, LG, A TSWRN, KA KRVLEPFGLVEDS-J LCPU LCPU LCPU CPU CPU CPU CPU CPU	GPPRPKANC P. PAB .A.K .A.Q.P. PAB K.P.P. VP2 KWAPTGDKF GKKRF EGKKRF EGKKRF MAKKNT VKRIQOPS- LGE 	QTQESLEKDDSRGLVF RHKQAI .HQAI .S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEMSPVEPDSSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP LI.SPQQPDSSTGIGKKGQPP GGDNNKFFGFSTPWGYFDYNF GGDNNKFFGFSTPWGYFDYNF .S.J.A.Y.Y.F. SLQS.TYN.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y.F. .S.Q.A.VQYA.Y. .S.Q.A.VQYA.Y. .S.Q.A.VQYA.Y. .S.Q.A.VQYA.Y. .S.Q.A.VQYA.Y. .S.J.	PGYNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI 	DLE IKDGHNPY FEYNEAD RQLDS.DLK.H QQLA.D.LK.H QQLA.D.LK.H QQLA.D.IK.H QQLA.D.IK.H QQLA.D.IK.H VF3 SIMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. 	PRRFQERLK AEX.A AEX.A QD.IDS.Q SADGVGNSS A. T T T T T T T	114 105 104 113 225 221 225 221 350 464 455 578 578 570 560 560 560 560 689 687
в	AAAV AAV2 AAV4 AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV AAV5 GP AAAV5 GP AAAV5 GP AAV5 GP AAV5 GP AAV5 GP AAV5 GP AAV4 AAV5 GP AAAV AAV4 AAV4	MSLISDAIPDW A-ADGYPL .TDGYL .FV.HP .TFL.SFEE. DDTSFGGNLGH EF GF .Q GNWHCDSQWLE T.MC .DT.MC .DT.MC .DMC DKDYQLPYVLC .SS.EMI DV. DEH MGAVPUNNOSI Y.S.S.L.RC .NIL.GGD.N G.QKK.V.GAYC Hepari Y.S.S.L.RC .NIL.GGD.N G.QKL.E.MI PEIQFTSNFGM Y. YN	VLERLVK# DI 	(GVNAAADFYHLES(CLSEGUREWWA.QP ULSEQUREWWA.QP BVGEGLRE.LG.A TSWRN.KA (RVLEPFGLVEDS-) LPPV LPPV LPVI LPVI GAGI	GPPRPKANC P. PAB .A.K .A.Q.P. PAB K.P.P. VP2 KTAPTGDKF GKKRF EGKKRF EGKKRF 	QTQESLEKDDSRGLVF RHKNAI HQAI S.SVSPDR.PER.NN.F.I KGEDEPRLPDTSSQTPKKNKKF VEHSPVEPDSSSGTGKAGQQPP LI.SPQQPDSSTGIGKKGKQPP DDHFPK.KKARTEEDS.PSTSS GKLTDHYPVV	PGVNYLGPFNGLDKGE 	PVNEADAAALEHDKAYI 	DLEIKDGHNPYFEYNEAD RQLDS.DLK.H QQL.A.DLK.H QQL.A.DIK.H QQL.A.DIKF.H VP3 STMAEGGGGPVGDAGQG INT.T.S.A.MA.NNE. 	PRRFQERLK AEQ.Q AEKA QD.IDS.Q SADGVGNSS A.	114 105 104 113 221 225 221 350 335 335 335 335 335 464 455 578 570 560 560 560 560 689 687 687

AAV5Y.N.YNDPQFVD...PDS..E.RTT. GP

FIG. 4. Comparisons of rep and cap ORFs. The rep and cap ORFs of AAAV, AAV2, AAV4, AAV5, and GPV (GP) were aligned by using Clustal W. Identical amino acids are indicated by a dot. Dashes indicate gaps in the sequence added by the alignment program. (A) Horizontal arrows indicate the initiator codon of the p5 and p19 Rep proteins. The Rep endonuclease site established by Tyr155 and the tetrahedrally coordinated Asp24, Glu83, His90, and His92 are presented in bold letters and are overlined by an asterisk. The region important for Rep multimerization, the ATP binding site, and the basic amino acids of the nuclear localization signal are underlined. The zinc finger motifs in the carboxy terminus are underlined and the coordinating cystine and histidine residues are indicated by dots. (B) The theoretical initiator codons of VP2 and VP3 are indicated in bold letters. Regions that have been proposed to be on the surface of AAV2 are underlined and divergent regions are boxed. The heparin binding region in the capsid of AAV2 is also indicated.



FIG. 5. Vector constructs for the generation of recombinant AAAV and transduction of chicken fibroblasts. (A) Wild-type AAAV, vector plasmid (pA3Vβ-Gal), and production yields of rAAAV using helper plasmids providing the *rep* gene under the control of CMV, MMTV, or the native P5 promoter. The helper plasmids pCA3VRC, pMA3VRC, and pA3VRC were individually cotransfected with pA3Vβ-Gal and an adenovirus helper plasmid in 293T cells, and rAAAV was produced as described in Materials and Methods. The number of rAAAV genomes produced in each group was determined by quantitative PCR and is expressed as DNase-resistant particles per cell (DRP/cell). (B) Relative transduction efficiency of primary chicken embryonic fibroblasts (CEF) and immortalized chicken embryonic fibroblasts (DF1) with equal particles of rAAAV expressing LacZ.

contrast to previous work with AAV2, which demonstrated that the use of a CMV promoter inhibited the production of rAAV2 (39).

In preliminary studies, we observed that the addition of detergents during virus purification affected infectivity. To better understand the effect of detergents, we prepared rAAAV in the presence of the following conditions: 0.5% deoxycholate, 0.5% 3-[(3-chloamidopropyl)-dimethylammonio]-1-propanesulfonate (CHAPS), 0.5% octylglucoside, or no detergent, respectively. The virus from the four groups was purified by using CsCl gradients, and rAAAV genomes were quantitated by using quantitative PCR. No effect was observed on the yield of viral particles or density of rAAAV in the four preparations. After dialysis against phosphate-buffered saline, transduction efficiency was measured by titration on CEF cells. Addition of OCG or CHAPS had no significant effect on transduction efficiency. However, deoxycholate, which is a stronger ionic detergent, reduced transduction efficiency almost 10-fold (data not shown).

Tissue tropism of rAAAV was determined in CEF, DF1, LMH, DT-90, QNR, QT6, 293T, COS, primary chicken embryonic kidney cells, primary chicken pituitary cells, and pri-

mary human fibroblasts and compared with that of rAAV2, rAAV4, and rAAV5 (Table 1). Transduction efficiency of rAAV was 10- to 300-fold higher in avian cells than in rAAV2, rAAV4, and rAAV5. In contrast, transduction of the mammalian cells in the panel by rAAAV was almost absent. This observation suggests that AAAV is using a different uptake or transduction mechanism compared with the primate AAVs. Interestingly, rAAAV exhibited ~15-fold higher transduction efficiency in primary chicken embryonic fibroblasts than did the immortalized embryonic fibroblasts (Fig. 5B).

DISCUSSION

Although the molecular and biological properties of AAAV were largely unknown, serological and immunological data have indicated that AAAV is distinct from the primate AAV (68, 69). That evidence prompted us to isolate, clone, and sequence an infectious clone of AAAV and construct a recombinant vector.

Previous studies have indicated difficulties in directly cloning full-length infectious clones of AAVs. These difficulties have been attributed to the genetic instability of parvoviral inverted

TABLE 1. Titers for rAAAV, rAAV2, rAAV4, and rAAV5 expressing LacZ in avian and mammalian cell lines and primary cells

	Transducing units per 10 ⁶ genomes ^a					
Cell type	rAAAV	rAAV ₂	$rAAV_4$	rAAV ₅		
CEF	$7,140 \pm 380$	25 ± 3.5	84 ± 6.3	58 ± 5.7		
DF-1	530 ± 35	8 ± 0.9	45 ± 4.7	60 ± 6.1		
LMH	$2,380 \pm 145$	230 ± 25	34 ± 5.6	40 ± 4.9		
DT-90	ND	ND	ND	ND		
QNR	$1,260 \pm 90$	176 ± 18	42 ± 5.2	185 ± 26		
QT6	930 ± 62	112 ± 21	23 ± 3.8	33 ± 5		
Chicken primary embryonic kidney cells	$8,080 \pm 560$	422 ± 46	350 ± 40	235 ± 38		
Chicken primary pituitary cells	$4,640 \pm 375$	144 ± 17	70 ± 12	91 ± 8.4		
293T	ND	$4,500 \pm 355$	$3,130 \pm 270$	684 ± 57		
COS	5 ± 0.7	$6,920 \pm 420$	$3,550 \pm 165$	592 ± 53		
A549	ND	$2,190 \pm 315$	$1,360 \pm 140$	26 ± 4.3		
Human primary fibroblasts	ND	$1,990 \pm 170$	$1,130 \pm 145$	292 ± 31		

^{*a*} Transductions were performed as described in Materials and Methods, and efficiency is expressed as transducing units per 10^6 recombinant particles. Numbers represent mean \pm standard deviation from four independent assays. ND, none detected.

terminal repeats. For that reason, investigators have used *recBC* bacterial strains (10, 70) or low-copy-number plasmids (56, 59) or constructed the full genome from cloned subfragments (20, 37, 48). Surprisingly, in the present study we did not encounter any difficulties in directly cloning the full AAAV genome in a medium- to high-copy-number pUC18 derivative plasmid. This may indicate a higher genetic stability of the AAAV ITRs than that of the primate isolates.

The nucleotide sequence of AAAV is 56 to 65% identical with the other known AAVs and contains all the structural components and genetic elements that characterize the family of AAVs. These elements include the ITRs, promoters, ORFs, transcription start and stop sites, and intron splice junctions. Particularly, the ITRs of all serotypes (including AAAV) are similar in length and symmetry, contain a conserved *rep*-binding site, and retain the ability to form the characteristic hairpin structure. Previous studies have demonstrated that the ability to form the terminal hairpin structure is important for AAV replication (38). This observation is further supported by the conservation of this structure between AAAV and the primate AAVs.

The high degree of conservation of the rep ORF between the primate AAVs indicates the importance of this gene to the life cycle of the virus. The rep ORF of AAAV is significantly more divergent than it is in other serotypes; however, the core region (aa 322 to 470) containing the ATPase and helicase activity is highly conserved (82% identity). This region is highly conserved in all vertebrate parvovirus, both autonomous and dependovirus, indicating the region's importance in the parvovirus life cycle. The N-terminal region of rep has been shown to be important for DNA binding; however, the exact amino acids involved are not known. The rep-binding site in the ITRs is highly conserved among AAVs including AAAV. Therefore, it is anticipated that the motif in the N-terminal region of rep involved in DNA binding must also be conserved. The Nterminal region of AAAV (aa 1 to 251) only shows 43% similarity with that of the other serotypes. Thus, the low degree of homology may help in identifying the conserved motif involved in DNA binding.

The carboxyl terminal of the unspliced Rep proteins encodes a zinc finger motif, and it is conserved among the AAV serotypes. The function of this region is largely unknown; however, previous studies have indicated involvement in transactivation (23). In addition, this region has been shown to be important for interaction with the cellular kinases PrKX and PKA, causing inhibition of kinase activity (17, 23). The carboxyl terminus of AAAV *rep* is highly divergent, displaying less than 15% homology. This fact may explain the increased titers of rAAAV obtained when using a helper plasmid that drives expression of the *rep* gene from a CMV promoter.

The predicted amino acid sequence of the capsid proteins indicates several regions with significant variation between serotypes (Fig. 4B). Several differences in the capsid proteins (aa 450 to 613) lie in regions that have been proposed to be on the exterior surface. These regions may play a role in serotype-specific properties such as antigenicity and/or binding to specific cellular receptors. However, not all of the changes are confined to the proposed exterior regions (aa 152 to 221), and they may also be important for other unique properties of AAAV versus the primate AAVs.

Transduction efficiency in avian and mammalian cells was very distinct between rAAAV and rAAV2, rAAV4, and rAAV5. The only difference in our four recombinant constructs is the presence of serotype-specific ITRs flanking the expression cassette [RSV- β -Gal-SV40poly(A)] and the serotype-specific capsid. Although it is possible that each serotypespecific ITR could interact differently with host-specific intracellular factors, it is more likely that transduction efficiencies are affected by the presence of distinct cell surface receptors. These data, combined with the extensive divergence of the *cap* ORF, suggest that AAAVs utilize a different uptake mechanism from those utilized by other serotypes.

The original hypothesis of a host-dependent evolution of parvovirus (3) is in contrast to the high sequence homology reported between the goose and Muscovy duck autonomous parvovirus and the primate AAVs (13, 40, 70). This observation raised the possibility of horizontal transfer of parvovirus between different species during evolution. The AAAV might be the missing link between the avian autonomous parvovirus and the AAVs or may constitute a distinct branch in the evolution of dependovirus. The AAAV genome (both *rep* and *cap* genes) is equally divergent between the avian autonomous parvovirus and AAVs. However, the structure and function of the AAAV ITR are very similar to those of AAVs. Thus, the proposed classification of parvoviruses based on the properties of the ITRs (1) gains further merit.

REFERENCES

- Astell, C. R. 1990. Terminal hairpins of parvovirus genomes and their role in DNA replication, p. 59–79. *In* P. Tijssen (ed.), Handbook of parvoviruses. CRC Press, Boca Raton, Fla.
- Atchison, R. W., B. C. Casto, and W. Hammon. 1965. Adenovirus-associated defective virus particles. Science 149:754–756.
- Bando, H., J. Kusuda, T. Gojobori, T. Maruyama, and S. Kawase. 1987. Organization and nucleotide sequence of a densovirus genome imply a host-dependent evolution of the parvoviruses. J. Virol. 61:553–560.
- Bantel-Schall, U., and H. zur Hausen. 1984. Characterization of the DNA of a defective human parvovirus isolated from a genital site. Virology 134:52– 63.
- Bauer, H. J., and G. Monreal. 1986. Herpesviruses provide helper functions for avian adeno-associated parvovirus. J. Gen. Virol. 67:181–185.
- Bauer, H. J., R. Schneider, H. R. Gelderblom, R. Lurz, V. Friehmelt, and G. Monreal. 1991. Biological and physicochemical characterization of the major (1.40) and minor (1.45) component of infectious avian adeno-associated virus. Arch. Virol. 120:123–133.
- Berns, K. I., and S. Adler. 1972. Separation of two types of adeno-associated virus particles containing complementary polynucleotide chains. J. Virol. 9:394–396.
- Berns, K. I., M. Bergoin, M. Bloom, M. Lederman, N. Muzyczka, G. Siegl, J. Tal, and P. Tattersall. 1994. *Parvoviridae*. Vth report of International Committee on Taxonomy of Viruses, p. 166. *In* F. A. Murphy, C. M. Faquet, D. H. L. Bishop, S. A. Ghabrial, A. W. Jarvis, G. P. Martelli, M. A. Mayo, and M. D. Summers (ed.), Virus taxonomy. Springer-Verlag, Vienna, Austria.
- Bohenzky, R. A., R. B. LeFebvre, and K. I. Berns. 1988. Sequence and symmetry requirements within the internal palindromic sequences of the adeno-associated virus terminal repeat. Virology 166:316–327.
- Boissy, R., and C. R. Astell. 1985. An Escherichia coli recBCsbcBrecF host permits the deletion-resistant propagation of plasmid clones containing the 5'-terminal palindrome of minute virus of mice. Gene 35:179–185.
- Brister, J. R., and N. Muzyczka. 1999. Rep-mediated nicking of the adenoassociated virus origin requires two biochemical activities, DNA helicase activity and transesterification. J. Virol. 73:9325–9336.
- Brister, J. R., and N. Muzyczka. 2000. Mechanism of Rep-mediated adenoassociated virus origin nicking. J. Virol. 74:7762–7771.
- Brown, K. E., S. W. Green, and N. S. Young. 1995. Goose parvovirus—an autonomous member of the dependovirus genus? Virology 210:283–291.
- Buller, R. M. L., and J. A. Rose. 1978. Characterization of adenovirusassociated virus-induced polypeptides in KB cells. J. Virol. 25:331–338.
- Chapman, M. S., and M. G. Rossmann. 1993. Structure, sequence and function correlations among parvoviruses. Virology 194:491–508.
- Chejanovsky, N., and B. J. Carter. 1989. Mutagenesis of an AUG codon in the adeno-associated virus rep gene: effects on viral DNA replication. Virology 173:120–128.
- Chiorini, J. A., B. Zimmermann, L. Yang, R. H. Smith, A. Ahearn, F. Herberg, and R. M. Kotin. 1998. Inhibition of PrKX, a novel protein kinase, and the cAMP-dependent protein kinase, PKA, by the regulatory proteins of adeno-associated virus type 2. Mol. Cell. Biol. 18:5921–5929.
- Chiorini, J. A., C. M. Wendtner, E. Urcelay, B. Safer, M. Hallek, and R. M. Kotin. 1995. High-efficiency transfer of the T cell co-stimulatory molecule B7–2 to lymphoid cells using high-titer recombinant adeno-associated virus vectors. Hum. Gene Ther. 6:1531–1541.
- Chiorini, J. A., F. Kim, L. Yang, and R. M. Kotin. 1999. Cloning and characterization of adeno-associated virus type 5. J. Virol. 73:1309–1319.
- Chiorini, J. A., L. Yang, Y. Liu, B. Safer, and R. M. Kotin. 1997. Cloning of adeno-associated virus type 4 (AAV4) and generation of recombinant AAV4 particles. J. Virol. 71:6823–6833.
- 21. Davidson, B. L., C. S. Stein, J. A. Heth, I. Martins, R. M. Kotin, T. A. Derksen, J. Zabner, A. Ghodsi, and J. A. Chiorini. 2000. Recombinant adeno-associated virus type 2, 4, and 5 vectors: transduction of variant cell types and regions in the mammalian central nervous system. Proc. Natl. Acad. Sci. USA 97:3428–3432.
- Davis, M. D., J. Wu, and R. A. Owens. 2000. Mutational analysis of adenoassociated virus type 2 Rep68 protein endonuclease activity on partially single-stranded substrates. J. Virol. 74:2936–2942.
- Di Pasquale, G., and S. N. Stacey. 1998. Adeno-associated virus Rep78 protein interacts with protein kinase A and its homolog PRKX and inhibits CREB-dependent transcriptional activation. J. Virol. 72:7916–7925.
- El Mishad, A. M., K. J. McCormick, V. J. Yates, and J. J. Trentin. 1975. Detection and serological identification of adeno-associated virus in avian adenovirus stocks. Infect. Immun. 11:287–299.
- Gao, G. P., M. R. Alvira, L. Wang, R. Calcedo, J. Johnston, and J. M. Wilson. 2002. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. Proc. Natl. Acad. Sci. USA 99:11854–11859.

- Green, M. R., and R. G. Roeder. 1980. Definition of a novel promoter for the major adenovirus-associated virus mRNA. Cell 22:231–242.
- Green, M. R., and R. G. Roeder. 1980. Transcripts of the adeno-associated virus genome: mapping of the major RNAs. J. Virol. 36:79–92.
- Henikoff, S., J. G. Henikoff, and S. Pietrokovski. 1999. Blocks+: a nonredundant database of protein alignment blocks derived from multiple compilations. Bioinformatics 15:471–479.
- Hermonat, P. L., M. A. Labow, R. Wright, K. I. Berns, and N. Muzyczka. 1984. Genetics of adeno-associated virus: isolation and preliminary characterization of adeno-associated virus type 2 mutants. J. Virol. 51:329–339.
- Hess, M., G. Paul, S. Kling, and G. Monreal. 1995. Molecular characterization of two strains of the avian adeno-associated virus (AAAV). Arch. Virol. 140:591–598.
- Hickman, A. B., D. R. Ronning, R. M. Kotin, and F. Dyda. 2002. Structural unity among viral origin binding proteins: crystal structure of the nuclease domain of adeno-associated virus Rep. Mol. Cell 10:327–337.
- Hoggan, M. D., N. R. Blacklow, and W. P. Rowe. 1966. Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. Proc. Natl. Acad. Sci. USA 55:1467–1474.
- Im, D. S., and N. Muzyczka. 1992. Partial purification of adeno-associated virus Rep78, Rep52, and Rep40 and their biochemical characterization. J. Virol. 66:1119–1128.
- Kleinschmidt, J. A., M. Mohler, F. W. Weindler, and R. Heilbronn. 1995. Sequence elements of the adeno-associated virus rep gene required for suppression of herpes-simplex-virus-induced DNA amplification. Virology 206:254–262.
- Labow, M. A., P. L. Hermonat, and K. I. Berns. 1986. Positive and negative autoregulation of the adeno-associated virus type 2 genome. J. Virol. 60: 251–258.
- Laughlin, C. A., H. Westphal, and B. J. Carter. 1979. Spliced adenovirusassociated virus RNA. Proc. Natl. Acad. Sci. USA 76:5567–5571.
- Laughlin, C. A., J. D. Tratschin, H. Coon, and B. J. Carter. 1983. Cloning of infectious adeno-associated virus genomes in bacterial plasmids. Gene 23: 65–73.
- LeFebvre, R. B., S. Riva, and K. I. Berns. 1984. Conformation takes precedence over sequence in adeno-associated virus DNA replication. Mol. Cell. Biol. 4:1416–1419.
- Li, J., R. J. Samulski, and X. Xiao. 1997. Role for highly regulated *rep* gene expression in adeno-associated virus vector production. J. Virol. 71:5236– 5243.
- Lukashov, V. V., and J. Goudsmit. 2001. Evolutionary relationships among parvoviruses: virus-host coevolution among autonomous primate parvoviruses and links between adeno-associated and avian parvoviruses. J. Virol. 75:2729–2740.
- Lusby, E., K. H. Fife, and K. I. Berns. 1980. Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. J. Virol. 34: 402–409.
- Lusby, E., R. Bohenzky, and K. I. Berns. 1981. Inverted terminal repetition in adeno-associated virus DNA: independence of the orientation at either end of the genome. J. Virol. 37:1083–1086.
- Mayor, H. D., K. Torikai, J. L. Melnick, and M. Mandel. 1969. Plus and minus single-stranded DNA separately encapsidated in adeno-associated satellite virions. Science 166:1280–1282.
- McCarty, D. M., T. H. Ni, and N. Muzyczka. 1992. Analysis of mutations in adeno-associated virus Rep protein in vivo and in vitro. J. Virol. 66:4050– 4057.
- McLaughlin, S. K., P. Collis, P. L. Hermonat, and N. Muzyczka. 1988. Adeno-associated virus general transduction vectors: analysis of proviral structures. J. Virol. 62:1963–1973.
- Mendelson, E., J. P. Trempe, and B. J. Carter. 1986. Identification of the trans-acting Rep proteins of adeno-associated virus by antibodies to a synthetic oligopeptide. J. Virol. 60:823–832.
- Muralidhar, S., S. P. Becerra, and J. A. Rose. 1994. Site-directed mutagenesis of adeno-associated virus type 2 structural protein initiation codons: effects on regulation of synthesis and biological activity. J. Virol. 68:170–176.
- Muramatsu, S. I., H. Mizukami, N. S. Young, and K. E. Brown. 1996. Nucleotide sequence and generation of an infectious clone of adeno-associated virus 3. Virology 221:208–217.
- Nahreini, P., S. H. Larsen, and A. Srivastava. 1992. Cloning and integration of DNA fragments in human cells via the inverted terminal repeats of the adeno-associated virus 2 genome. Gene 119:265–272.
- Owens, R. A., M. D. Weitzman, S. R. Kyostio, and B. J. Carter. 1993. Identification of a DNA-binding domain in the amino terminus of adenoassociated virus Rep proteins. J. Virol. 67:997–1005.
- Parks, W. P., J. L. Melnick., R. Rongey, and H. D. Mayor. 1967. Physical assay and growth cycle studies of a defective adeno-satellite virus. J. Virol. 1:171–180.
- Pronovost, A. D., V. J. Yates, and D. E. Fry. 1978. Effect of avian adenoassociated virus on pathogenicity of Tipton virus in chicks. Avian Dis. 22: 354–357.
- 53. Pronovost, A. D., V. J. Yates, and D. E. Fry. 1979. Inhibition and enhance-

ment of avian adenovirus plaque production by heavy and light avian adenovirus-associated viral particles. Am. J. Vet. Res. **40**:549–551.

- 54. Quandt, K., K. Frech, H. Karas, E. Wingender, and T. Werner. 1995. MatInd and MatInspector: new fast and versatile tools for detection of consensus matches in nucleotide sequence data. Nucleic Acids Res. 23:4878–4884.
- Rose, J. A., J. V. Maizel, Jr., J. K. Inman, and A. J. Shatkin. 1971. Structural proteins of adenovirus-associated viruses. J. Virol. 8:766–770.
- Rutledge, E. A., C. L. Halbert, and D. W. Russell. 1998. Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2. J. Virol. 72:309–319.
- Ryan, J. H., S. Zolotukhin, and N. Muzyczka. 1996. Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats. J. Virol. 70:1542–1553.
- Samulski, R. J., A. Srivastava, K. I. Berns, and N. Muzyczka. 1983. Rescue of adeno-associated virus from recombinant plasmids: gene correction within the terminal repeats of AAV. Cell 33:135–143.
- Samulski, R. J., K. I. Berns, M. Tan, and N. Muzyczka. 1982. Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. Proc. Natl. Acad. Sci. USA 79:2077–2081.
- 60. Senapathy, P., J. D. Tratschin, and B. J. Carter. 1984. Replication of adenoassociated virus DNA. Complementation of naturally occurring rep 2 mutants by a wild-type genome or an ori 2 mutant and correction of terminal palindrome deletions. J. Mol. Biol. 179:1–20.
- Smith, R. H., and R. M. Kotin. 1998. The Rep52 gene product of adenoassociated virus is a DNA helicase with 3'-to-5' polarity. J. Virol. 72:4874– 4881.
- 62. Smith, R. H., and R. M. Kotin. 2000. An adeno-associated virus (AAV)

initiator protein, Rep78, catalyzes the cleavage and ligation of singlestranded AAV ori DNA. J. Virol. **74**:3122–3129.

- Smuda, J. W., and B. J. Carter. 1991. Adeno-associated viruses having nonsense mutations in the capsid genes: growth in mammalian cells containing an inducible amber suppressor. Virology 184:310–318.
- 64. Tratschin, J. D., I. L. Miller, and B. J. Carter. 1984. Genetic analysis of adeno-associated virus: properties of deletion mutants constructed in vitro and evidence for an adeno-associated virus replication function. J. Virol. 51:611–619.
- Walker, S. L., R. S. Wonderling, and R. A. Owens. 1997. Mutational analysis of the adeno-associated virus Rep68 protein: identification of critical residues necessary for site-specific endonuclease activity. J. Virol. 71:2722–2730.
- Wonderling, R. S., S. R. M. Kyostio, and R. A. Owens. 1995. A maltosebinding protein/adeno-associated virus rep68 fusion protein has DNA-RNA helicase and ATPase activities. J. Virol. 69:3542–3548.
- 67. Xie, Q., W. Bu, S. Bhatai, J. Hare, T. Somasundaram, A. Azzi, and M. S. Chapman. 2002. The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. Proc. Natl. Acad. Sci. USA 99:10405–10410.
- Yates, V. J., A. M. el-Mishad, K. J. McCormick, and J. J. Trentin. 1973. Isolation and characterization of an avian adenovirus-associated virus. Infect. Immun. 7:973–980.
- Yates, V. J., G. J. Dawson, and A. D. Pronovost. 1981. Serologic evidence of avian adeno-associated virus infection in an unselected human population and among poultry workers. Am. J. Epidemiol. 113:542–545.
- Zadori, Z., R. Stefancsik, T. Rauch, and J. Kisary. 1995. Analysis of the complete nucleotide sequences of goose and muscovy duck parvoviruses indicates common ancestral origin with adeno-associated virus 2. Virology 212:562–573.