



Adeno-associated virus capsid serotype identification: Analytical methods development and application

Kim Van Vliet^a, Yasmin Mohiuddin^a, Scott McClung^b, Veronique Blouin^c, Fabienne Rolling^c, Philippe Moullier^{a,c}, Mavis Agbandje-McKenna^d, Richard O. Snyder^{a,e,*}

^a Department of Molecular Genetics and Microbiology, University of Florida, College of Medicine, United States

^b Interdisciplinary Center for Biotechnology Research, Proteomics Core, University of Florida, United States

^c Laboratoire de Thérapie Génique, INSERM U649, CHU Hôtel Dieu, Nantes Cedex 1, France

^d Department of Biochemistry and Molecular Biology, University of Florida, College of Medicine, United States

^e Department of Pediatrics, University of Florida, College of Medicine, United States

ABSTRACT

Article history:

Received 27 October 2008

Received in revised form 11 March 2009

Accepted 16 March 2009

Available online 26 March 2009

Keywords:

Adeno-associated virus

AAV

Capsid

Retina

Serotype identity

Mass spectrometry

Mass spectrometry (MS) has been utilized to address the need for a rapid and reliable assay to confirm the capsid serotype identity of recombinant AAV gene transfer vectors. The differences in the primary amino acid sequence of AAV serotypes generate a unique set of fragments with different masses upon proteolytic digestion, and by comparing the fragment masses against common and custom databases, reliable capsid serotype identification is achieved. Highly homologous serotypes, such as AAV1, AAV2, and AAV8, can be distinguished from each other, as well as from less homologous serotypes such as AAV4, and AAV5. Furthermore, analysis of the MS data for wild-type AAV4 compared to an AAV4 capsid with a single amino acid mutation demonstrates the sensitivity of the method and validates the relevance of the method in the context of retinal gene transfer. With an expanding repertoire of AAV serotypes, physicochemical methods for capsid analysis, such as MS, are highly desirable and do not require product-specific analytical reagents such as monoclonal antibodies. A MS-based capsid identity test is suitable for cGMP lot release testing of rAAV gene transfer products and will help ensure patient protection.

© 2009 Elsevier B.V. All rights reserved.

1. Introduction

AAV is a small, single-stranded DNA virus that has a non-enveloped capsid composed of 60 subunits with $T=1$ icosahedral symmetry. The viral genome consists of two open reading frames, rep and cap, that are flanked by inverted terminal repeats (ITRs). The rep gene encodes proteins required for viral genome replication and packaging, while the cap gene encodes the structural proteins, VP1, VP2 and VP3, which differ in their N-terminal region and are present in the mature virion at a ratio of 1:1:10. The minimal viral sequences required in cis for viral DNA replication and packaging are the ITRs (McLaughlin et al., 1988; Xiao et al., 1997). To produce recombinant AAV (rAAV) vectors, the rep and cap genes are supplied in trans, and the therapeutic transgene expression cassette flanked by the ITRs is packaged into AAV capsids (Hermonat and Muzyczka, 1984). As a result, the ITRs are the only viral sequences expected in AAV vector virions.

There are twelve known serotypes of AAV (Schmidt et al., 2008), each with unique properties. Fig. 1 shows a phylogenetic tree of AAV1 through AAV12, with the recently isolated AAV12 shown closely related to AAV11 and AAV4. AAV2 was the first serotype to be used for gene transfer applications and has been utilized in the majority of AAV clinical trials to date. With the recent discovery of over 100 new AAV sequences (Chen et al., 2005; Gao et al., 2002, 2004; Mori et al., 2004; Rutledge et al., 1998; Schmidt et al., 2006), there is the potential for an expanding number of AAV serotypes that could be developed into gene transfer vectors in the future. AAV serotypes differ in tissue tropism, transduction efficiency, and antigenic reactivity. Studies have shown that neutralizing antibodies are generated against the AAV capsid which hinders subsequent re-administration (Hernandez et al., 1999; Moskalenko et al., 2000; Peden et al., 2004; Petry et al., 2008; Vandenberghe et al., 2006; Zaiss and Muruve, 2008). Pseudotyped rAAV vectors using different capsid serotypes are desirable for repeat administration as well as for patients that have preexisting circulating antibodies to AAV2. The rAAV vectors utilized to date in clinical trials have been well tolerated and have an excellent safety profile. Currently, AAV1, AAV2, AAV5, and AAV6 vectors are in clinical trials approved by the US Food and Drug Administration (Mueller and Flotte, 2008) for a variety of indications. Most recently, results from clinical tri-

* Corresponding author at: Department of Molecular Genetics and Microbiology, 1600 SW Archer Road, Gainesville, FL 32610 0266, United States. Tel.: +1 386 418 1642; fax: +1 352 392 4290.

E-mail address: rsnyder@cerhb.ufl.edu (R.O. Snyder).

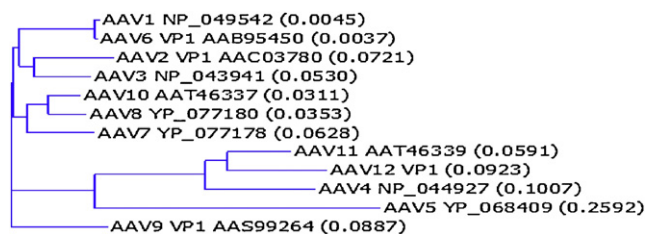


Fig. 1. Phylogenetic tree of AAV serotypes 1–12, depicting the evolutionary relationship among the serotypes.

als of the inherited blindness Leber Congenital Amaurosis (LCA) that utilize rAAV2-RPE65 vectors have been reported (Bainbridge et al., 2008; Cideciyan et al., 2008; Maguire et al., 2008; Hauswirth et al., 2008). rAAV serotype 4 vectors have also been shown to be efficient for retinal gene transfer and treating animal models of LCA (Le Meur et al., 2007), and are being developed for the clinic.

Currently, for rAAV vectors used in clinical trials, the vector genome is sequenced as an identity test needed for cGMP product release prior to patient administration; however, the number of existing methods is limited for confirming the capsid serotype of the rAAV vector lot that was manufactured. With the potential for vectors of several serotypes being utilized, reliable methods to verify the AAV capsid serotype identity are needed, and current methods may be inadequate. For example, a PCR-based assay using unique primers to the cap gene of different serotypes has been developed to distinguish the serotype of wild-type (wt) AAV2, AAV3B and AAV6 (Mitchell et al., 2006); however, for rAAV vectors the rep and cap genes have been removed and replaced with a therapeutic gene, thus this assay cannot be utilized. Grimm et al. (1999, 2003) developed a capture ELISA for AAV2 and AAV3 using the monoclonal antibody A20 and recently, new monoclonal antibodies (mAbs) were generated for serotypes 1/6, 4, and 5 (Kuck et al., 2007). These antibodies were shown to recognize a capsid-specific conformational epitope and do not cross react with the limited number of other serotypes tested; however, with an increasing number of AAV serotypes being developed for clinical application, antibodies that do not cross react may be difficult to obtain. Furthermore, specific capsid antibodies are not available for many serotypes, and utilizing mAbs is probably insufficient for verifying the identity of genetically modified AAV capsids or engineered capsids that are an ensemble of multiple serotypes (Muzyczka and Warrington, 2005). Previously, a method that relies on the proteolysis of intact AAV capsids for determining capsid serotype identity was described (Van Vliet et al., 2006). Since the current trend is toward customized AAV vectors, methods that rely on the inherent capsid composition are needed for analysis.

In the product manufacturing setting, the United States Pharmacopeia and European Pharmacopeia include Mass Spectrometry (MS) techniques for the identification and quantitation of compounds and impurities, including biotechnology derived products (European Pharmacopeia, 2002; United States Pharmacopeia, 2007). These product testing methods are utilized to determine if the product specifications are met prior to administration to humans. In the research setting, the application of MS has been shown to provide useful, robust, and rapid proteomics approaches to identify viral proteins produced during the course of infection for Marek's Disease Virus (MDV) (Liu et al., 2006). Advances in MS techniques have allowed detection of viruses that are human pathogens (Colquhoun et al., 2006) and have also been used to identify the viral proteins associated with a variety of viruses (Renesto et al., 2006; Resch et al., 2007; Kattenhorn et al., 2004; Swatkoski et al., 2007; Maaty et al., 2006). It was also demonstrated that MS could be used

to monitor hepatitis C virus (HCV) genotype 1a, which exists as a heterogeneous population of quasiespecies in patients (Yea et al., 2007). Furthermore, MS has been used to investigate dynamic capsid structural transitions that occur for viruses in solution, such as Cowpea chlorotic mottle virus (Speir et al., 2006). For AAV2, high resolution MS has been utilized to characterize capsid protein glycosylation (Murray et al., 2006); however, MS has not previously been utilized to determine the capsid serotype of rAAV gene transfer vectors.

In this work, we describe a method for AAV capsid serotype identity testing that addresses the need for a rapid and reliable assay for product lot release. We show that MS of individual AAV capsid proteins is a viable method for reliable serotype identification when fragment masses are compared against common and custom sequence databases. Utilizing these methods, highly homologous serotypes can be distinguished from each other, as well as from less homologous serotypes. Furthermore, MS is able to distinguish between the wtAAV4 capsid and an AAV4 capsid with a single point mutation even though the *in vivo* retinal transduction efficiencies showed little difference. MS will be useful for evaluating customized clinical gene therapy vectors with capsids that have been intentionally modified for purposes of tissue-specific targeting and higher transduction efficiencies. With the expanding number of AAV serotypes, physicochemical methods such as MS that do not rely on the availability of product-specific reagents such as mAbs are highly desirable as rAAV vector product release tests in the cGMP or cGMP setting.

2. Materials and methods

2.1. AAV1, 2, 5, 8 vectors

rAAV2-GFP vector virions were produced by transient transfection of HEK293 cells. Cells were harvested and lysed by 3 freeze/thaw cycles, cell debris was removed by centrifugation (3000 × *g* for 10 min), and deoxycholate (DOC) was added to the supernatant to a final concentration of 0.5%. Lysates were incubated for 30 min at 37 °C with 50 U/ml of Benzonase, and purified on sequential cesium chloride (CsCl) density gradients. Cesium fractions were dialyzed and heparin chromatography was performed as a concentration step (Zolotukhin et al., 2002). rAAV1-GFP, rAAV5-GFP and rAAV8-GFP vector virions were purified by freeze/thaw and CsCl gradients followed by Q-sepharose chromatography. Virions were dialyzed into 50 mM Tris-Cl pH 8 containing 100 mM NaCl using a 10,000 MWCO membrane (Pierce), aliquotted, and stored frozen at –20 °C or –80 °C. Infectivity of the AAV vectors was assayed on C12 cells in the presence of adenovirus as previously described (Clark et al., 1996).

2.2. AAV4 production

The wild-type AAV4: (pAAV₄₋₂ for rep4 cap4 and a part of the ITR) was provided by Jay Chiorini (Chiorini et al., 1997). The mutant AAV4 helper called pDP4r was provided by J.A. Kleinschmidt (Grimm et al., 2003). The amino acid mutation at position 544 (E instead of K) was identified by sequencing the cap gene (Genome-express, France), using the sequencing primer 1: 5'accgagtgactcagcagt 3', primer 2: 5'tcgagcagcagaaggccta3', primer 3: 5'gactggcagcgactcatcaa3', primer 4: 5'acgagcgcacagcactct3'. pDP4r was restored to the wild-type sequence using the following strategy: the cap gene from pAAV₄₋₂ was cloned by inserting a ClaI linker, and digesting with FseI and ClaI to isolate the wild-type cap gene. This fragment was used to replace the mutant cap gene sequence in pDP4r removed by FseI/ClaI digestion. The restored wild-type plasmid is called pDG4 and encodes K in position 544.

To produce rAAV4 particles, HEK293 cells were transfected with the helper plasmid pXX6 from Jude Samulski (Xiao et al., 1998), pAAV4-2, and the vector plasmid pAAV2-GFP. The AAV4-GFP particles were purified on 2 CsCl gradients and dialyzed in 1X dPBS. For in vivo studies, the mutant AAV4-GFP vector was produced by transient transfection of HEK293 cells (12.5 µg of pAAV2-GFP and 25 µg of pDP4r per 150 mm plate) using Calcium phosphate. Cells were harvested 60 h post transfection. rAAV particles were extracted in TNE buffer (20 mM Tris/100 mM NaCl/1 mM EDTA, pH 8.0) by 3 freeze/thaw cycles and lysed with DOC/Benzonase/MgCl₂, purified on two CsCl gradients and dialyzed into 50 mM Tris-HCl, pH 8.0, containing 100 mM NaCl as described in Van Vliet, 2006. The restored wild-type AAV4-GFP vector was produced by transient transfection of HEK293 cells with pDG4. Cells were extracted by 3 freeze/thaw cycles, purified on 2 CsCl gradients and dialyzed into 1X dPBS. For infectious titering, AAV4-GFP preparations were titered on COS cells, and rAAV4 vector genome titers were determined by dot-blot to be $\sim 4 \times 10^{11}$ vector genomes/ml (vg/ml). The ratio of vector genomes/transducing units is similar ($\sim 3 \times 10^3$) for the AAV4-GFP vectors produced by either pDP4r or pDG4 helper plasmids.

2.3. Subretinal injections

Subretinal injections were performed in rats as described (Duisit et al., 2002). Briefly, Wistar rats were anesthetized with an intramuscular injection of 50 mg/kg Ketamine and 6 mg/kg Xylazine. Subretinal injections were via a transscleral transchoroidal approach: the sclera and the choroid were punctured, a 33-gauge needle was then inserted in a tangential direction under an operating microscope and the vector was delivered into the subretinal space. Rats were cared for in accordance with the ARVO statement for the use of animals in ophthalmic and vision research.

2.4. In vivo GFP fluorescence imaging, retinal flat mount, and tissue sections

The sclera/choroid/RPE and neuroretinal flat mount was performed as previously described (Duisit et al., 2002). Briefly, RPE-choroid-sclera and neuroretinal flat mount of the enucleated eyes were fixed with 4% paraformaldehyde for 40 min at room temperature. After washing in PBS, the eyes were cut through the pars plana, and the anterior segment and the lens were removed. The eye cup was cut peripherally in four sections under an operating microscope and flattened. The neuroretina was detached from the RPE-choroid-sclera with forceps, both cell layers were mounted on a glass slide and were examined by fluorescence microscopy.

2.5. Protein analyses

2.5.1. SDS-PAGE

Samples were boiled for 3 min, centrifuged briefly and loaded on a 10% polyacrylamide gel (Bio-Rad, Hercules, CA). SDS-PAGE was performed at 125 V for 90 min (Laemmli, 1970). A 1% stock solution of Coomassie blue R-250 was prepared in dH₂O. Gels were stained in a solution consisting of 12.5% Stain stock, 10% acetic acid and 50% methanol for 1 h. Gels were destained in solution consisting of 10% acetic acid and 50% methanol overnight. AAV bands were cut from the gel and gel slices were stored in destain until the samples were prepared for mass spectrometry.

2.5.2. In-gel digestion

To prepare samples for mass spectrometry, gel slices were washed to remove SDS. Gel slices were washed with 50% acetonitrile (ACN) in dH₂O and vortexed 2 times for 15 min each.

Wash solution was removed and discarded. Gel slices were covered with neat acetonitrile until the gel was dehydrated. The liquid was removed after 5–10 min. The gel was rehydrated in 50 mM ammonium bicarbonate (ABC), pH 8.4 to cover the gel for 5 min (20–50 µl). An equal volume of acetonitrile was added to give a 1:1 ratio of acetonitrile to ABC and vortexed for 15 min. Wash solution was removed and the gel was dried down in a speedvac for 10–15 min. Next, reduction and alkylation of protein in the gel was performed. The gel piece was rehydrated in 100 µl of 45 mM Dithiothreitol (DTT) at 55 °C for 30 min. The gel was submerged in the reducing buffer. The buffer was discarded after incubation. Tubes were chilled to room temperature. The liquid was removed and replaced quickly with 100 µl of freshly made 100 mM iodoacetamide, and incubated in the dark for 30 min at room temperature. The buffer was removed. The gel was washed 3 times with 100 µl 50% ACN/50 mM ABC with agitation for 15 min each. Gel pieces were dried in a speedvac until they were completely dry, 10–15 min. Next, the proteins were digested with trypsin. Trypsin (Promega, Madison, WI) was prepared at a concentration of 12.5 ng/µl reconstituted in 50 mM ABC, pH 8.4, with 5 mM CaCl₂. Enzyme and buffer were kept cold in an ice bucket and the ratio of enzyme to protein was approximately 1:20. The dried gel piece was chilled on ice, and the gel pieces were covered with ice cold trypsin digestion buffer. Tubes were incubated overnight at 37 °C in a water bath. The reaction was stopped by addition of 5.0% glacial acetic acid to a final concentration of 0.5% acetic acid, and the supernatant was pipetted into a clean tube. Samples for digestion with chymotrypsin (Roche, Nutley, NJ) were prepared in the same fashion. Chymotrypsin was reconstituted in 1 mM HCl and the ratio of enzyme to protein was approximately 1:20.

2.5.3. Mass spectrometry

Capillary rpHPLC separation of protein digests was performed on a 15 cm × 75 µm i.d. PepMap C18 column (LC Packings, San Francisco, CA) in combination with an Ultimate Capillary HPLC System (LC Packings, San Francisco, CA) operated at a flow rate of 200 nl/min. Inline mass spectrometric analysis of the column eluate was accomplished by a hybrid quadrupole time-of-flight instrument (QSTAR XL, Applied Biosystems, Foster City, CA) equipped with a nanoelectrospray source.

Fragment ion data generated by Information Dependent Acquisition (IDA) via the QSTAR were searched against the NCBI nr sequence database using the Mascot Version 2.2.0 (Matrix Science, London, UK) database search engine. Probability-based MOWSE scores above the default significant value were considered for protein identification in addition to validation by manual interpretation of the tandem MS data.

2.5.4. Databases

Tandem mass spectra were extracted, charge state deconvoluted and deisotoped by Analyst version 1.1. All MS/MS samples were analyzed using Mascot (Matrix Science, London, UK; version 2.2.0) and X!Tandem (www.thegpm.org; version 2006.04.01.2). X!Tandem was set up to search a subset of the NCBI nr.20061201 database also assuming trypsin. Mascot was set up to search the NCBI nr.20070202 database (selected for viruses, unknown version, 375,509 entries) assuming the digestion enzyme trypsin. Mascot and X!Tandem were searched with a fragment ion mass tolerance of 0.30 Da and a parent ion tolerance of 0.30 Da. Iodoacetamide derivative of cysteine was specified in Mascot and X!Tandem as a fixed modification. S-carbamoylmethylcysteine cyclization of the N-terminus, deamidation of asparagine and glutamine and oxidation of methionine were specified in Mascot and X!Tandem as variable modifications. For samples digested with chymotrypsin, the protease chymotrypsin was selected and the database search was performed in the same fashion as the samples digested with trypsin.

A customized AAV database was built and consists of the amino acid sequences of VP1, VP2 and VP3 for the known AAV serotypes, AAV1 to AAV12. The customized database also includes the sequences for AAV4 K544E VP1, VP2 and VP3. Mascot was set to search the customized database, as described for the NCBI nr database. The database searches were repeated for all samples against the sequences in the customized database (see Supplemental Data).

2.5.5. Criteria for protein identification

Scaffold (version Scaffold-01.07.00, Proteome Software Inc., Portland, OR) was used to validate MS/MS-based peptide and protein identifications. Peptide identifications were accepted if they could be established at greater than 90.0% probability as specified by the Peptide Prophet algorithm (Keller et al., 2002). Protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least 1 identified peptide. Protein probabilities were assigned by the Protein Prophet algorithm (Nesvizhskii et al., 2003). Proteins that contained similar peptides and could not be differentiated based on MS/MS analysis alone were grouped to satisfy the principles of parsimony.

3. Results

A growing repertoire of AAV capsids is being evaluated in animals to determine empirically the tissue tropism of the different serotypes (reviewed by Van Vliet et al., 2008) that may be applicable for a given therapeutic indication, and desired target tissue and cell type. The degree of homology between the known AAV serotypes, AAV1 through AAV12 is provided in Table 1. To distinguish the various AAV capsids, a MS assay for AAV serotype identification was designed to incorporate speed and simplicity, and methods that have been accepted by the regulatory community (e.g. included in the pharmacopeia). The individual AAV viral proteins (VP1, 87 kDa; VP2, 72 kDa; and VP3, 62 kDa) are separated by SDS-PAGE, gels are stained, and MS is performed. The SDS-PAGE gel functions as a separation step for the three overlapping AAV capsid proteins that comprise the virion as well as impurities that may be present in the sample; therefore, samples of varying levels of purity can be analyzed. Following SDS-PAGE and staining, the individual AAV capsid bands are excised from the gel, digested with trypsin and the fragments are subjected to LC/MS/MS. Mass spectrometry data is analyzed using the software Mascot and X!Tandem. Mascot was set to search the NCBI database limited to viral proteins. The proteomics software Scaffold is used to confirm MS/MS data based peptide and protein identifications.

3.1. Analysis of known samples

Initially, the MS method was evaluated using VP1, VP2, and VP3 of three known AAV serotypes, AAV1, AAV2 and AAV5. Coomassie

stained viral protein bands were excised from the SDS-PAGE gel, trypsinized, and LC/MS/MS was performed. Peptide identifications were accepted if they could be established at greater than 90.0% probability as assigned by the Protein Prophet algorithm. AAV VP protein identifications were accepted if they could be established at greater than 95.0% probability and contained at least one identified peptide as assigned by the Protein Prophet algorithm. The results of LC/MS/MS were interpreted using the software Scaffold (Proteome Software Inc., Portland, OR) as shown in Fig. 2. The number of unique trypsin fragments is shown and the color of the box represents the probability. For example, Sample #1, which is AAV1 VP2, shows three fragments with identity to the capsid protein of AAV1 with over 95% probability. For Sample #6, which is AAV5 VP1, there are six fragments that match the capsid of AAV5. Although there are four fragments that suggest Sample #6 could be AAV2, based on a larger number of fragments matching AAV5, this sample was identified correctly as AAV5. Additionally, for Sample #7, AAV5 VP2, one of the fragments was identified as Rep40 protein [AAV2] using the Scaffold program that referenced the NCBI database. Upon closer inspection, the sequence of this peptide is present in all four of the Rep proteins (Rep78, Rep68, Rep52 and Rep40), thus this protein most likely represents Rep78, which is similar in size to AAV2 VP2 at 72 kDa and has been shown to be associated with the virion (Prasad and Trempe, 1995).

3.2. Analysis of unknown samples

Following validation of the identity of three known AAV serotypes, the assay was repeated as a blinded test to determine the reliability of MS for capsid serotype identification of unknown samples. rAAV samples were separated on SDS-PAGE gels and stained with Coomassie blue. After excising the bands from the gel, samples were assigned an alphabetic code and submitted for MS analysis without disclosing the identity of the samples. Because there is considerable overlap of the AAV viral proteins VP1, VP2, and VP3, and since VP3 is present in the largest amount in intact capsids, more VP3 protein is available and was used initially for the blinded study. Additionally, the N-terminus of VP3 of different serotypes was predicted to provide a unique fragment after trypsinization because of a greater degree of sequence heterogeneity. Mass spectrometry data was analyzed using Mascot and X!Tandem. Mascot was again set to search the NCBI database limited to viral proteins. Peptide identifications were accepted if they could be established at greater than 95.0% probability as specified by the Peptide Prophet algorithm using Scaffold software. AAV VP protein identifications were accepted if they could be established at greater than 99.0% probability and contained at least two identified peptides as assigned by the Protein Prophet algorithm. These software conditions are more stringent than those used for the known samples analyzed above and were increased to

Table 1
AAV serotype amino acid identity for AAV serotypes 1–12.

	AAV1	AAV2	AAV3	AAV4	AAV5	AAV6	AAV7	AAV8	AAV9	AAV10	AAV11	AAV12
AAV1	100%	83%	87%	64%	58%	99%	85%	84%	82%	85%	68%	63%
AAV2	83%	100%	87%	61%	58%	84%	83%	83%	82%	84%	64%	62%
AAV3	87%	87%	100%	63%	59%	87%	84%	86%	84%	86%	65%	62%
AAV4	64%	61%	63%	100%	54%	64%	65%	64%	64%	65%	82%	79%
AAV5	58%	58%	59%	54%	100%	58%	59%	58%	57%	58%	55%	54%
AAV6	99%	84%	87%	64%	58%	100%	85%	84%	82%	85%	68%	62%
AAV7	85%	83%	84%	65%	59%	85%	100%	88%	82%	88%	68%	63%
AAV8	84%	83%	86%	64%	58%	84%	88%	100%	85%	93%	67%	63%
AAV9	82%	82%	84%	64%	57%	82%	82%	85%	100%	86%	65%	62%
AAV10	85%	84%	86%	65%	58%	85%	88%	93%	86%	100%	68%	63%
AAV11	68%	64%	65%	82%	55%	68%	68%	67%	65%	68%	100%	85%
AAV12	63%	62%	62%	79%	54%	62%	63%	63%	62%	63%	85%	100%

Accession numbers for proteins used in the alignment are: AAV1 VP1 NP.049542, AAV2 VP1 AAC03780, AAV3 VP1 NP.043941, AAV4 VP1 NP.044927, AAV5 VP1 YP.068409, AAV6 VP1 AAB95450, AAV7 VP1 YP.077178, AAV8 VP1 YP.077180, AAV9 VP1 AAS99264, AAV10 AAT46337, AAV11 AAT46339, and AAV12 VP1 ABI16639.

Display Options: Number of Unique Peptides 95% min Protein; 1 min Peptides; 90% min Peptide

	Number of similar matches	Accession numbers	Protein molecular weight (AMU)	Sample 1 AAV1 VP2	Sample 2 AAV1 VP3	Sample 3 AAV2 VP1	Sample 4 AAV2 VP2	Sample 5 AAV2 VP3	Sample 6 AAV5 VP1	Sample 7 AAV5 VP2
Bio View: Identified Proteins (11) Chain A, The Atomic Structure Of AAV-2 capsid protein AAV5 [VP1] polymerase [Hendra virus] capsid protein VP1 [AAV hu.23] capsid protein VP1 [AAV1] capsid protein VP1 [AAV2] capsid protein [Avian AAV strain DA-1] Replicase polyprotein 1ab (pp1ab) Rep 40 protein [AAV2] tail tape measure protein [Bacillus clarkii bacteriophage BCJA1c] hypothetical protein PSSM2_284 [Cyanophage P-SSM2]										
	2	gi 110645922, gi 22219269	58698.1			21		25	4	
	1	gi 51593838	80407.2						6	9
	1	gi 29468608	257278	0	0	0	0	1	0	0
	1	gi 46487825	81756.6			9		0		
	1	gi 9632548	81358.5	3	2					
	1	gi 110645923	81927.5			1	1		4	
	1	gi 51949970	82410.8	1		1				
	2	gi 37999893, gi 9635157	748893		2					
	2	gi 110645919, gi 110645920	34946.7							1
	1	gi 56694917	131381							1
	1	gi 61806155	107603							1

Fig. 2. Mass spectrometry analysis of seven known samples of viral proteins from AAV serotypes 1, 2, and 5. Mass spectrometry data analyzed using the software program Scaffold shows the samples that were evaluated in columns on the top right, labeled Sample #1 through Sample #7. Proteins that were identified as a result of the database search, with their corresponding accession numbers, molecular mass, and taxonomy information are shown in rows 1 through 11. Data analysis was performed using the software Scaffold (Proteome Software Inc., Portland, OR). Scaffold results were analyzed using 95% minimum protein probability, 1 peptide minimum, and 90% minimum peptide probability. The database search was limited to the complete NCBI viral protein database for protein identification.

provide better confidence in assigning serotypes to the unknown samples.

Fig. 3A shows the Scaffold data analysis for seventeen unknown samples from AAV serotypes 1, 2, 4, 5 and 8 consisting of 15 VP3 samples and 2 VP2 samples (QQ, RR). Mass spectrometry data analyzed using the software program Scaffold enabled unambiguous serotype identification for 13/17 samples (76%) and the correct serotype assignment was made for 17/17 samples (100%) based on the serotype with the largest number of unique fragments. Fig. 3B shows MS analysis of twenty-four unknown samples of viral proteins, 12 VP1 samples and 12 VP2 samples, from AAV serotypes 1, 2, 4, 5, and 8. Table 2 lists the alphabetic codes assigned to each serotype sample prior to preparation and MS analysis. Mass spectrometry data analyzed using the software program Scaffold enabled unambiguous serotype identification for 21/24 samples (87.5%), and of the three other samples, no data was obtained for

two samples (SS, B), and one sample was ambiguous due to the same number of unique peptides matched to 3 different serotypes (Sample C). For each AAV serotype sample, separation by SDS-PAGE provides three samples for MS, namely VP1, VP2, and VP3. In this assay, VP3 alone is sufficient for AAV capsid serotype identification; however, analysis of VP1, VP2, and VP3 builds redundancy in the data set and additional confidence in the serotype assignment. Adenovirus DNA Binding Protein (DBP, 59 kDa) was identified in several VP3 samples (Fig. 3A) and originates from the helper plasmid used in the production of the rAAV vector stocks.

For protein identification by MS, the amino acid sequence of the protein tested must be present in the target database. The NCBI protein database (<http://www.ncbi.nlm.nih.gov/sites/entrez>) contains the capsid amino acid sequences of the known AAV serotypes. It also contains sequences from the Protein Data Bank (PDB) for AAV serotypes for which the X-ray crystal structures have been solved.

Table 2

Alphabetic codes for blinded AAV serotype samples for MS analysis shown in Fig. 3.

AAV1 VP1-A	AAV1 VP1-P	AAV2 VP1-H	
AAV1 VP2-AA	AAV1 VP2-PP	AAV2 VP2-HH	
AAV1 VP3-AAA	AAV1 VP3-PPP	AAV2 VP3-HHH	
AAV4 K544E VP1-L	AAV4 K544E VP1-M	AAV4 wt VP1-E	AAV4 wt VP1-K
AAV4 K544E VP2-LL	AAV4 K544E VP2-MM	AAV4 wt VP2-EE	AAV4 wt VP2-KK
AAV4 K544E VP3-LLL	AAV4 K544E VP3-MMM	AAV4 wt VP3-EEE	AAV4 wt VP3-KKK
AAV5 (pure) VP1-G	AAV5 (pure) VP1-Z	AAV5 (impure) VP1-Q	AAV5 (impure) VP1-R
AAV5 (pure) VP2-GG	AAV5 (pure) VP2-ZZ	AAV5 (impure) VP2-QQ	AAV5 (impure) VP2-RR
AAV5 (pure) VP3-GGG	AAV5 (pure) VP3-ZZZ	AAV5 (impure) VP3-QQQ	AAV5 (impure) VP3-RRR
AAV8 VP1-B	AAV8 VP1-S	AAV8 VP1-C	AAV8 VP1-X
AAV8 VP2-BB	AAV8 VP2-SS	AAV8 VP2-CC	AAV8 VP2-XX
AAV 8 VP2 small-1/2BB	AAV8VP2 small-1/2SS	AAV8 VP3-CCC	AAV8 VP3-XXX
AAV8 VP3-BBB	AAV8 VP3-SSS		

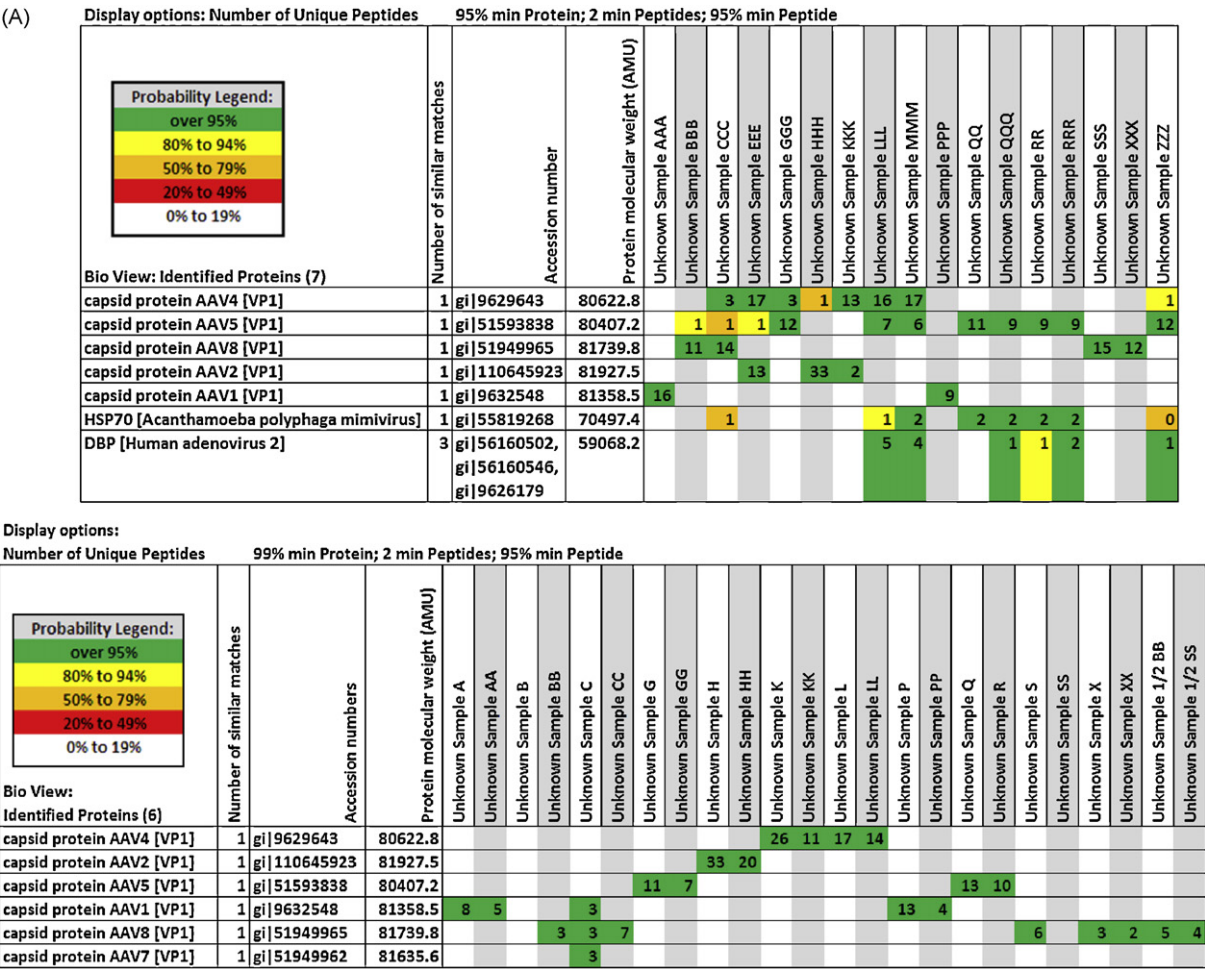


Fig. 3. (A) Mass spectrometry analysis of seventeen unknown samples of viral proteins from AAV serotypes 1, 2, 4, 5, and 8. The unknown samples are listed across the top of the table. The column on the left hand side are the identified proteins, the NCBI accession number, and the protein molecular mass. The numbers in the figure represent the number of unique trypsin fragments, and the color of the box represents the probability. Numbers in green boxes indicate the number of unique fragments with a probability of over 95% upon which the AAV serotype identification was based. (B) Mass spectrometry analysis of twenty-four unknown samples of viral proteins VP1 or VP2 from AAV serotypes 1, 2, 4, 5, and 8. Results were collected using 99% protein, a 2 peptide minimum, 95% peptide probability. The viral proteins, accession numbers, and protein mass are the result of the NCBI database search. The unknown samples are listed as their alphabetic code, and the values in the table are the number of unique trypsin fragments.

The PDB sequences are VP3, less approximately 12 N-terminal amino acids which were not resolved in the crystal structures. For most of the AAV serotypes, a search of the NCBI database of viral proteins most often matched the capsid protein VP1 because this sequence is usually the only sequence in the database, and is annotated to indicate the amino acids where the overlapping sequences of VP2 and VP3 begin. The exception is AAV2, where all three viral proteins have their own accession numbers, as well as the PDB sequence 1lp3. Therefore, out of necessity for MS analysis, a custom database was generated that includes separate files for VP1, VP2 and VP3 for the known serotypes (see [Supplemental Data](#)).

3.3. Analysis of an AAV4 capsid containing a mutation

AAV4 is being developed as a vector to treat Leber Congenital Amaurosis in humans ([Le Meur et al., 2007](#)). In anticipation of clinical trials, rAAV4-GFP vectors were produced using pDP4r, a helper plasmid harboring the AAV4 rep ORF and the AAV4 cap ORF ([Grimm et al., 2003](#)) to test the performance of the pDP4r helper plasmid. rAAV4-GFP preparations were evaluated in vivo in rats following subretinal injection. In advance of *E. coli* master cell banking for clinical manufacture, the cap and rep ORFs of pDP4r were sequenced and a mutation in the cap ORF was identified (K544E). The muta-

tion was repaired and new preparations of rAAV-GFP were made using the repaired pDP4r, renamed pDG4. The mutant and repaired vector preparations performed indistinguishably in vivo following subretinal injection in rats, as similar transduction patterns and efficiencies were observed with RPE cells being strongly transduced ([Fig. 4](#)). A different preparation of a rAAV4 vector was also produced for MS using the helper plasmid pAAV₄₋₂ that harbors the wtAAV4 capsid. MS was evaluated to determine if the method could distinguish the wtAAV4 capsid from the AAV4 capsid with the K544E amino acid mutation.

Generally for MS, the raw data is utilized to search a database to match theoretical proteolytic fragments that are based on the protease used, a protein's primary amino acid sequence, and the expected masses. If an amino acid change has occurred or there has been a post-translational modification, the proteolytic fragment will have a different mass than expected and will not match to the fragments in the database for the region harboring the mutation/modification. As a result, there will be a lack of fragment coverage in this region. Thus, if a proteolytic site is mutated, the resulting fragment mass will differ dramatically from the wild-type fragment, and if an amino acid mutation takes place between proteolytic sites, then the resulting fragment will have a different mass than the wild-type fragment. MS was performed on wtAAV4 and

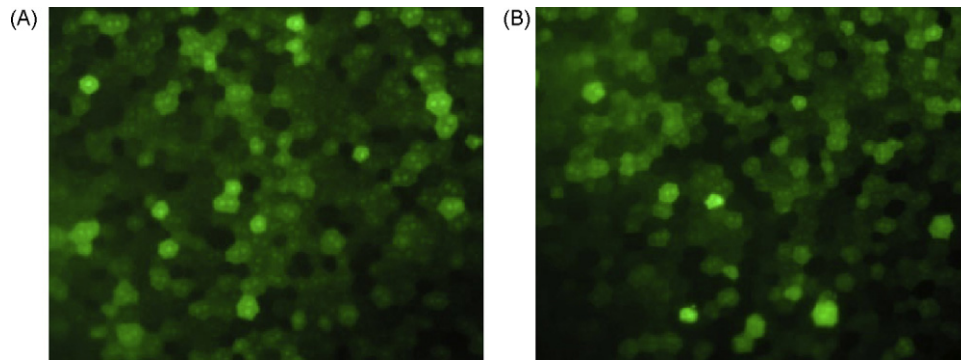


Fig. 4. Fluorescent imaging of the sclera/choroid/RPE flat mounts showing expression of GFP within the RPE following transduction with AAV4 K544E (A) and wild-type restored AAV4 (B) vectors.

AAV4 K544E capsid samples as described for capsid serotype identification. The samples were digested with trypsin, which cleaves at basic amino acid residues; therefore, as a result of the change from a basic to an acidic amino acid, it was predicted that the AAV4 K544E capsid would not be digested by trypsin at position 544 and the fragment generated would be larger. However, in the database search, the theoretical larger fragment is absent; therefore, the result obtained is a lack of coverage in this region. Coverage or the number of amino acids detected after MS analysis is rarely 100%, so the lack of coverage by itself is not sufficient to assume that a mutation has occurred. However, coverage was obtained resulting from trypsin cleavage at amino acid K544 for the wtAAV4 capsid “reference standard”, and for all of the AAV4 K544E capsid mutant VP protein samples, there was a lack of coverage in this region. This indicates that the corresponding fragment ionizes and is detectable for wtAAV4 when present, and the lack of coverage in this region for all samples of the AAV4 K544E mutant is strongly suggestive of a mutation in the capsid’s primary amino acid sequence.

To further verify that an amino acid change occurred, wtAAV4 and the AAV4 K544E capsid mutant proteins were separated by SDS-PAGE, stained with Coomassie blue and digested with chymotrypsin, which cleaves peptide bonds at the carboxyl side of Tyr, Trp, Phe, Leu, and Met amino acids. Fig. 5 shows the MS coverage for wtAAV4 compared with the AAV4 K544E capsid mutant after chymotrypsin digestion. Sequences highlighted in yellow represent peptides identified by MS. Green boxes indicate post-translational modifications identified during the MS data analysis: oxidized methionine, M (+16), and asparagine deamidation, N (+1).

The results shown are a compilation of two samples of VP3 for wtAAV4, and six samples of the AAV4 K544E capsid mutant consisting of two samples each of VP1, VP2 and VP3. Black boxes indicate two peptides identified in wtAAV4 that are not present in the AAV4 K544E mutant when digested with chymotrypsin. The two potential regions for mutations in the AAV4 capsid using chymotrypsin include the fragment ²⁵²KRLGESLQSN²⁶³ (VP1 numbering) and the fragment ⁵⁴¹AGPKQNGNTATVP⁵⁵⁶ (VP1 numbering). Mass spectrometry for the AAV4 K544E mutant treated with trypsin, sample MMM, had coverage for peptide ²⁴⁰TWVLPTYNHLYK²⁵², no coverage for Arg253, and coverage for peptide ²⁵⁴LGESLQSN²⁷⁷YNGFSTPWGYDFN²⁷⁷. This excludes the peptide 252–263 from the potential mutation sites, leaving the site from 541 to 556 as the region with the mutation. Fig. 6 shows the MS data for wtAAV4 and the AAV4 K544E mutant. As stated above, the custom database includes the amino acid sequences of VP1, VP2 and VP3 for the known AAV serotypes, and the mutant AAV4 capsid analyzed in this study (see Supplemental Data). In Fig. 6, of the 8 AAV4 samples analyzed, two proteins were identified: AAV4 VP1 as found in the NCBI database (gi9629643), and AAV4 VP3 which is identified by the unique accession number 43444344 assigned in the custom database for AAV4 VP3. The additional sequence files that were built into the custom database for VP sequences not found in the NCBI database were assigned unique accession numbers, in this case gi43444344.

A single sample digested with one protease (trypsin) was insufficient for identifying the region with the mutation in the AAV4 K544E capsid. A second set of samples digested with chymotrypsin

(A) wtVP3 Chymotrypsin [adeno-associated virus 4] 177/544 amino acids (32.5% coverage)				(B) K544E Chymotrypsin[adeno-associated virus 4] 235/544 amino acids (43.2% coverage)			
MSDDSEMRAA	AGGAAVEGGQ	GADGVGNASG	DWHCDSTWSE	MSDDSEMRAA	AGGAAVEGGQ	GADGVGNASG	DWHCDSTWSE
GHVTTTSTR	WVLPTYNHNL	YKRLGESLQS	NTYNGFSTPW	GHVTTTSTR	WVLPTYNHNL	YKRLGESLQS	NTYNGFSTPW
GYFDENRFHC	HFSPRDWQRL	INNNGWMRPK	AMRVKIFNIQ	GYFDENRFHC	HFSPRDWQRL	INNNGWMRPK	AMRVKIFNIQ
VKEVTTSNGE	TTVANLST	VQIFADSSYE	LPYVMDAGQE	VKEVTTSNGE	TTVANLST	VQIFADSSYE	LPYVMDAGQE
GSLLPFPNDV	FMVPQYGYCG	LVTGNTSQQQ	TDRNAFYCLE	GSLLPFPNDV	FMVPQYGYCG	LVTGNTSQQQ	TDRNAFYCLE
YFPSQMLRTG	NFEITYSFE	KVPFHSYAH	SQSLDRIMNP	YFPSQMLRTG	NFEITYSFE	KVPFHSYAH	SQSLDRIMNP
LIDQYLWGLQ	STTTGTTLNA	GTATTNFTKL	RPTNFSNFKK	LIDQYLWGLQ	STTTGTTLNA	GTATTNFTKL	RPTNFSNFKK
NWLPGPSIKQ	QGFSKTANQN	YKIPATGSDS	LIKYETHSTL	NWLPGPSIKQ	QGFSKTANQN	YKIPATGSDS	LIKYETHSTL
DGRWSALTPG	PPMATAGPAD	SKFSNSQLIF	AGPKQNGNTA	DGRWSALTPG	PPMATAGPAD	SKFSNSQLIF	AGPKQNGNTA
TVPGTLIFTS	EEELAATNAT	DTDMWGNLPG	GDQSNLNLP	TVPGTLIFTS	EEELAATNAT	DTDMWGNLPG	GDQSNLNLP
VDRLTALGAV	PGMVWQNRDI	YYQGPIWAKI	PHTDGHFHPS	VDRLTALGAV	PGMVWQNRDI	YYQGPIWAKI	PHTDGHFHPS
PLIGGFGLKH	PPPQIFIKNT	PVPANPATTF	SSTPVNSFIT	PLIGGFGLKH	PPPQIFIKNT	PVPANPATTF	SSTPVNSFIT
QYSTGQVSQ	IDWEIQKERS	KRWNEVQFT	SNYQQNSLL	QYSTGQVSQ	IDWEIQKERS	KRWNEVQFT	SNYQQNSLL
WAPDAAGKYT	EPRAIGTRYL	THHL		WAPDAAGKYT	EPRAIGTRYL	THHL	

Fig. 5. Mass spectrometry analysis of wtAAV4 and AAV4 K544E capsids. (A) Mass spectrometry coverage of wt AAV4 digested with chymotrypsin. (B) Mass spectrometry coverage of the AAV4 K544E mutant digested with chymotrypsin. When comparing the two sequences, there are two regions that show coverage for wt AAV4, but have a lack of coverage for the AAV4 K544E capsid mutant. These regions are shown in black boxes. Green boxes indicate post-translational modifications: oxidized methionine, M (+16), and asparagine deamidation, N (+1).

Display options: Number of Unique Peptides99% min Protein; 2 min Peptides; 95% min Peptide

Probability Legend: over 95% 80% to 94% 50% to 79% 20% to 49% 0% to 19%	Number of similar matches	Accession numbers	Protein molecular weight (AMU)	wt AAV4 VP3 Chymotrypsin	wt AAV4 VP3 Chymotrypsin	K544E AAV4 VP1 Chymotrypsin	K544E AAV4 VP2 Chymotrypsin	K544E AAV4 VP3 Chymotrypsin	K544E AAV4 VP1 Chymotrypsin	K544E AAV4 VP2 Chymotrypsin	K544E AAV4 VP3 Chymotrypsin
Bio View: Identified Proteins (2)											
capsid protein AAV4 [VP1]	1	gi 9629643	80622.8	11	13	12	9		5	9	15
capsid protein AAV4 [VP3]	2	gi 43444344,gi 4399334	59600.9					21			

Fig. 6. Mass spectrometry data for wt AAV4 and AAV4 K544E digested with chymotrypsin. Mass spectrometry of 2 samples of wild-type VP3 and 6 individual samples of AAV4 K544E mutant (2 samples each of VP1, VP2 and VP3) were digested with chymotrypsin. These were screened against the custom database (see Supplemental Data), and of these 8 AAV4 samples, two proteins were identified: AAV4 VP1 as found in the NCBI database (gi9629643), and AAV4 VP3, which is identified by the unique accession number 43444344 assigned in the custom database for AAV4 VP3.

also was insufficient to identify the region of the AAV4 capsid with the mutation; however, both data sets taken together and compared with wtAAV4 was sufficient for determining the region of the AAV4 capsid harboring the mutation. Thus, this assay provides the ability to distinguish between a wild-type capsid and a cap-

sid that harbors a single amino acid mutation. The region of the AAV4 K544E capsid mutation is shown in Fig. 7. The location of this mutation on the AAV capsid is a surface-associated residue that is not involved in subunit interactions that might affect capsid assembly or packaging, and given the similar performance in vivo in the retina, does not affect transduction (uptake, intracellular transport, and uncoating, Fig. 4). Therefore, there was no indication to suggest either during production, virus titering, or in vivo performance that this capsid harbored a mutation. Amino acid 544, based on the capsid alignment of serotypes 1 through 12, is not conserved among AAV serotypes. From the location of the mutation on the capsid surface, and the amino acid variability in this region among serotypes, as well as a glutamic acid at this position for AAV1, AAV3, and AAV6, it would be predicted that the change from K → E at amino acid 544 in AAV4 would be well tolerated, as was confirmed in vivo.

4. Discussion

rAAV gene transfer vectors derived from different AAV capsid serotypes are being evaluated in animals and humans (reviewed by Van Vliet et al., 2008). Except for the ITRs, rAAV vectors lack viral DNA sequences, thus there are few methods available that can be utilized for capsid serotype identification even though identity confirmation is critical for cGMP product lot release and clinical use. Prior to manufacturing, helper plasmids used in the manufacture of rAAV vectors are sequenced to ensure that the proper capsid serotype is produced, however since the cap gene is absent from the rAAV vector product, nucleic acid-based analyses (sequencing or PCR) cannot be used for capsid identity confirmation or to ensure that mutation has not occurred during rAAV manufacturing. The capsid proteins of different AAV serotypes separated by one-dimensional SDS-PAGE have slightly different mobilities and provide some limited information about the AAV serotype (Zolotukhin et al., 2002; Grimm et al., 2003), but to obtain information about the capsid serotype a comparison of several AAV samples need to be run in parallel with reference standards for comparison. Another approach is to use proteases and position-specific mAbs for peptide mapping and differentiation of the AAV serotypes (Van Vliet et al., 2006). An ideal assay for AAV serotype identification allows for a single sample to be identified on the basis of the inherent characteristics of the sample and without the need for product-specific reagents such as mAbs. Other desirable features of the assay include speed and simplicity, and methods that have been accepted by the

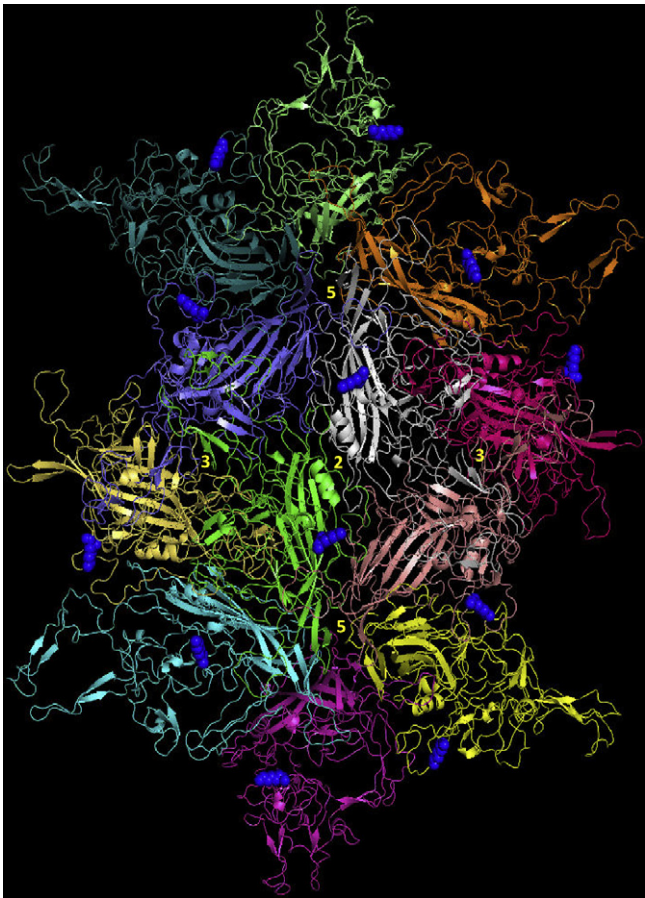


Fig. 7. Two AAV4 pentamers with two additional monomers (gold and dark pink) added for viewing of the 3-fold mounds. The two trimers in this image are dark pink, pink, white; gold, lime green, purple. Numbers indicate the 2-fold, 3-fold and 5-fold axis of symmetry, respectively. The location of the K544E mutation is shown as blue spheres.

regulatory community (e.g. included in the pharmacopeia). Based on these features, a MS method has been developed for determining AAV serotype identity.

Given the peptide coverage that was obtained, the MS analysis was capable of unambiguous serotype identification of different capsid serotypes (correct serotype assignment was made for 17 of 17 samples, Fig. 3A). In Fig. 3B, 21 of 24 samples (greater than 87%) were correctly identified; and of the three other samples, no data was obtained for two samples, and one sample was ambiguous due to the same number of unique peptides matched to 3 different serotypes. Furthermore, a wild-type capsid could be distinguished from a capsid that has a single amino acid mutation (Figs. 5 and 6) even though the *in vivo* performance of the two rAAV4 vectors was indistinguishable following subretinal injection in rats. The identification of Rep78 in Fig. 2 and Ad DBP in Fig. 3A demonstrates that in addition to identifying the serotype of the AAV capsid, impurities can be identified that are similar in size to the AAV viral proteins, which would not otherwise be distinguishable when assessing purity by one-dimensional SDS-PAGE. The samples used for these studies were of varying purity, as determined by SDS-PAGE (data not shown), to evaluate if the identity of the capsid serotype could be determined in the background of impurities, as is the case during *in-process* purification steps conducted prior to obtaining purified vector stocks. For mass spectrometry, a lightly stained blue band on a Coomassie stained gel is a sufficient amount of material; however, more material provides better coverage. Generally, 130 fmol (8 ng of a 62 kDa protein) represents the lower limit of detection by Coomassie staining. Since the lower limit required to assign an identity by MS on the QSTAR XL is 1–10 fmol of each peptide, and most of the AAV capsid samples were in the 0.5–5 µg range (8–81 pmol), an impurity of approximately 0.5% can be detected, and contamination by a different vector or a production mix-up should be detectable.

The next generation AAV vectors are being engineered with a variety of modifications of the capsid to attain specific tissue tropism, immune evasion, and better transduction. Site-directed mutagenesis studies have been performed to evaluate the functional significance of specific amino acids in various regions of the AAV capsid (Wu et al., 2000). These studies have provided the basis for an understanding of amino acids that are important for capsid assembly, viral entry into the host cell and infectivity, as well as the identification of regions of the AAV capsid that are amenable to modification. Structural studies of AAV capsids by X-ray crystallography and cryo-electron microscopy also provide for the rational design of modified AAV vectors (DiMattia et al., 2005; Govindasamy et al., 2006; Kronenberg et al., 2001; Nam et al., 2007; Padron et al., 2005; Walters et al., 2004; Xie et al., 2002, 2003). For AAV2, site-directed mutagenesis studies have identified the amino acid residues involved in heparin binding. Two of these, R585 and R588 are located in a surface loop at the 3-fold axis of symmetry (Kern et al., 2003; Opie et al., 2003). Alterations in this surface loop by insertion of an RGD integrin binding motif is able to ablate heparin binding and increase transduction of human lymphoblastic and chronic myelogenous leukemia cell lines, ovarian carcinoma cells in culture, and in an *in vivo* xenograft tumor model (Shi et al., 2006). Recently, insertion of atherosclerotic plaque targeting peptides into the AAV2 capsid at amino acid 587 re-targeted the vector (White et al., 2008). In addition to insertions into the AAV capsid for targeting purposes, studies of the functional domains on the AAV capsid by DNA shuffling and domain swapping also provide a better understanding of elements that are required in the AAV capsid for high level transduction in certain tissues, and elucidate the mechanism behind differences in serotype transduction efficiency (DiPrimio et al., 2008; Grimm et al., 2008; Shen et al., 2007). Single amino acid mutations are sufficient to influence titer, heparin binding, and tissue tropism in different AAV serotypes (Wu et al.,

2006). Mutation of specific amino acids on the AAV2 capsid surface, such as surface exposed tyrosines, has been shown to provide for high-efficiency transduction at lower doses of vector as a result of allowing the vectors to evade phosphorylation and ubiquitination, thus preventing proteasome-mediated degradation (Zhong et al., 2008). A directed evolution approach for high throughput selection of customized rAAV vectors with enhanced properties has been shown to generate vectors with altered affinities for heparin or the ability to evade neutralizing antibodies (Maheshri et al., 2006).

A customized AAV capsid database was assembled to support the MS analysis of samples evaluated in the present study. This database is more limited than the complete NCBI viral database, and as a result reduces the time required for protein identification. Additionally, future customized databases can include AAV capsid sequences that have been intentionally modified for various purposes, such as site specific targeting, increased transduction efficiency, or avoidance of an immune response upon subsequent re-administration. Since the next generation of rAAV vectors will likely consist of customized AAV capsids, a database of intentionally modified AAV vectors will allow the verification of AAV vectors with engineered modifications. For the capsid database, a computer file for each customized vector needs to be generated; however, since each modified helper plasmid construct (rep and cap expression) utilized in the cGMP manufacture of a customized AAV vector would have already been sequenced prior to manufacturing, the engineered capsid sequence can easily be derived from the helper plasmid sequence and added to the customized capsid database for MS analysis of the vector product.

New methods to verify the serotype of the AAV capsid incorporated into rAAV vector products will be valuable to ensure the safety of patients participating in gene transfer clinical trials. Using this assay, the serotype of rAAV samples with unknown identity can be determined using the current NCBI databases or customized databases that include VP1, VP2 and VP3 for the known AAV serotypes. Current methods using either ELISA or proteolytic fragmentation have their limitations for distinguishing between different serotype vectors, and between mutant and wtAAV capsids. MS is capable of AAV capsid identification and verification of capsid mutations because the method relies on the presence of specific proteolytic sites and the specific masses of the fragments that are generated. The use of MS for the identification and quantitation of components of the AAV capsid may be sufficient to meet compendial or other pharmaceutical requirements for product lot release when these products are evaluated in the clinic and eventually licensed.

Acknowledgements

We thank Nicholas Muzyczka and Sergei Zolotukhin for helpful comments; Jay Chiorini for providing wtAAV4; Nicole Brument for providing AAV8; Sophie Alvarez and Cindy Croft Bryant of the Interdisciplinary Center for Biotechnology Research (ICBR) for mass spectrometry; the Powell Gene Therapy Center Vector Core. RS is an inventor on patents related to recombinant AAV technology. RS owns equity in a gene therapy company that is commercializing AAV for gene therapy applications. To the extent that the work in this manuscript increases the value of these commercial holdings, RS has a conflict of interest. This project was funded by Association Francaise contre les Myopathies (AFM) Award 12263A (RS), AFM award 12263B (PM), HHMI Undergraduate Research Award and University Scholars Program Scholarship (YM), and NIH projects P01 HL59412 and P01 HL51811 to MA-M. This work was also performed under a Cooperative Agreement between INSERM, AFM, l'Etablissement Francais du Sang (EFS) and the Uni-

versity of Florida Center of Excellence for Regenerative Health Biotechnology.

Appendix A. Supplementary data

Supplementary data associated with this article can be found at doi:10.1016/j.jviromet.2009.03.020.

References

- Bainbridge, J.W., Smith, A.J., Barker, S.S., Robbie, S., Henderson, R., Balaggan, K., Viswanathan, A., Holder, G.E., Stockman, A., Tyler, N., Petersen-Jones, S., Bhat-tacharya, S.S., Thrasher, A.J., Fitzke, F.W., Carter, B.J., Rubin, G.S., Moore, A.T., Ali, R.R., 2008. Effect of gene therapy on visual function in Leber's congenital amaurosis. *N. Engl. J. Med.* 358, 2231–2239.
- Chen, C.L., Jensen, R.L., Schnepf, B.C., Connell, M.J., Shell, R., Sferra, T.J., Bartlett, J.S., Clark, K.R., Johnson, P.R., 2005. Molecular characterization of adeno-associated viruses infecting children. *J. Virol.* 79, 14781–14792.
- Chiorini, J.A., Yang, L., Liu, Y., Safer, B., Kotin, R.M., 1997. Cloning of adeno-associated virus type 4 (AAV4) and generation of recombinant AAV4 particles. *J. Virol.* 71, 6823–6833.
- Cideciyan, A.V., Aleman, T.S., Boye, S.L., Schwartz, S.B., Kaushal, S., Roman, A.J., Pang, J.-J., Sumaroka, A., Windsor, E.A.M., Wilson, J.M., Flotte, T.R., Fishman, G.A., Heon, E., Stone, E.M., Byrne, B.J., Jacobson, S.G., Hauswirth, W.W., 2008. Human gene therapy for RPE65 isomerase deficiency activates the retinoid cycle of vision but with slow rod kinetics. *Proc. Natl. Acad. Sci. U.S.A.* 105, 15112–15117.
- Clark, K.R., Voulgaropoulou, F., Johnson, P.R., 1996. A stable cell line carrying adenovirus-inducible rep and cap genes allows for infectivity titration of adeno-associated virus vectors. *Gene Ther.* 3, 1124–1132.
- Colquhoun, D.R., Schwab, K.J., Cole, R.N., Halden, R.U., 2006. Detection of norovirus capsid protein in authentic standards and in stool extracts by matrix-assisted laser desorption/ionization and nanospray mass spectrometry. *Appl. Environ. Microbiol.* 72, 2749–2755.
- DiMattia, M., Govindasamy, L., Levy, H.C., Gurda-Whitaker, B., Kalina, A., Kohlbrenner, E., Chiorini, J.A., McKenna, R., Muzyczka, N., Zolotukhin, S., Agbandje-McKenna, M., 2005. Production, purification, crystallization and preliminary X-ray structural studies of adeno-associated virus serotype 5. *Acta Crystallogr. Sect. F: Struct. Biol. Cryst. Commun.* 61, 917–921.
- DiPrimio, N., Asokan, A., Govindasamy, L., Agbandje-McKenna, M., Samulski, R.J., 2008. Surface loop dynamics in adeno-associated virus capsid assembly. *J. Virol.* 82, 5178–5189.
- Duisit, C., Conrath, H., Saleun, S., Folliot, S., Provost, N., Cosset, F.L., Sandrin, V., Moullier, P., Rolling, F., 2002. Five recombinant simian immunodeficiency virus pseudotypes lead to exclusive transduction of retinal pigmented epithelium in rat. *Mol. Ther.* 6, 446–454.
- European Pharmacopeia, 2002. Directorate for the Quality of Medicines of the Council of Europe. 4th ed., pp. 58–60 (2.2.43).
- Gao, G., Vandenberghe, L.H., Alvira, M.R., Lu, Y., Calcedo, R., Zhou, X., Wilson, J.M., 2004. Clades of adeno-associated viruses are widely disseminated in human tissues. *J. Virol.* 78, 6381–6388.
- Gao, G.P., Alvira, M.R., Wang, L., Calcedo, R., Johnston, J., Wilson, J.M., 2002. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc. Natl. Acad. Sci. U.S.A.* 99, 11854–11859.
- Govindasamy, L., Padron, E., McKenna, R., Muzyczka, N., Kaludov, N., Chiorini, J.A., Agbandje-McKenna, M., 2006. Structurally mapping the diverse phenotype of adeno-associated virus serotype 4. *J. Virol.* 80, 11556–11570.
- Grimm, D., Kern, A., Pawlita, M., Ferrari, F., Samulski, R., Kleinschmidt, J., 1999. Titration of AAV-2 particles via a novel capsid ELISA: packaging of genomes can limit production of recombinant AAV-2. *Gene Ther.* 6, 1322–1330.
- Grimm, D., Kay, M.A., Kleinschmidt, J.A., 2003. Helper virus-free, optically controllable, and two-plasmid-based production of adeno-associated virus vectors of serotypes 1 to 6. *Mol. Ther.* 7, 839–850.
- Grimm, D., Lee, J.S., Wang, L., Desai, T., Akache, B., Storm, T.A., Kay, M.A., 2008. In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *J. Virol.* 82, 5887–5911.
- Hauswirth, W., Aleman, T.S., Kaushal, S., Cideciyan, A.V., Schwartz, S.B., Wang, L., Conlon, T., Boye, S.L., Flotte, T.R., Byrne, B., Jacobson, S.G., 2008. Treatment of Leber congenital amaurosis due to RPE65 mutations by ocular subretinal injection of adeno-associated virus gene vector: short-term results of a Phase I trial. *Hum. Gene Ther.* 19, 979–990.
- Hermonat, P.L., Muzyczka, N., 1984. Use of adeno-associated virus as a mammalian DNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells. *Proc. Natl. Acad. Sci. U.S.A.* 81, 6466–6470.
- Hernandez, Y.J., Wang, J., Kearns, W.G., Loiler, S., Poirier, A., Flotte, T.R., 1999. Latent adeno-associated virus infection elicits humoral but not cell-mediated immune responses in a nonhuman primate model. *J. Virol.* 73, 8549–8558.
- Kattenhorn, L.M., Mills, R., Wagner, M., Lomsadze, A., Makeev, V., Borodovsky, M., Ploegh, H.L., Kessler, B.M., 2004. Identification of proteins associated with murine cytomegalovirus virions. *J. Virol.* 78, 11187–11197.
- Keller, A., Nesvizhskii, A.I., Kolker, E., Aebersold, R., 2002. Empirical statistical model to estimate the accuracy of peptide identifications made by MS/MS and database search. *Anal. Chem.* 74, 5383–5392.
- Kern, A., Schmidt, K., Leder, C., Muller, O.J., Wobus, C.E., Bettinger, K., Von der Lieth, C.W., King, J.A., Kleinschmidt, J.A., 2003. Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. *J. Virol.* 77, 11072–11081.
- Kronenberg, S., Kleinschmidt, J.A., Bottcher, B., 2001. Electron cryo-microscopy and image reconstruction of adeno-associated virus type 2 empty capsids. *EMBO Rep.* 2, 997–1002.
- Kuck, D., Kern, A., Kleinschmidt, J.A., 2007. Development of AAV serotype-specific ELISAs using novel monoclonal antibodies. *J. Virol. Methods* 140, 17–24.
- Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature* 227, 680–685.
- Le Meur, G., Stieger, K., Smith, A.J., Weber, M., Deschamps, J.Y., Nivard, D., Mendes-Madeira, A., Provost, N., Pereon, Y., Cherel, Y., Ali, R.R., Hamel, C., Moullier, P., Rolling, F., 2007. Restoration of vision in RPE65-deficient Briard dogs using an AAV serotype 4 vector that specifically targets the retinal pigmented epithelium. *Gene Ther.* 14, 292–303.
- Liu, H.C., Soderblom, E.J., Goshe, M.B., 2006. A mass spectrometry-based proteomic approach to study Marek's Disease Virus gene expression. *J. Virol. Methods* 135, 66–75.
- Maaty, W.S., Ortmann, A.C., Dlakic, M., Schulstad, K., Hilmer, J.K., Liepold, L., Weidenheft, B., Khayat, R., Douglas, T., Young, M.J., Bothner, B., 2006. Characterization of the archaeal thermophile *Sulfolobus turreted* icosahedral virus validates an evolutionary link among double-stranded DNA viruses from all domains of life. *J. Virol.* 80, 7625–7635.
- Maguire, A.M., Simonelli, F., Pierce, E.A., Pugh Jr., E.N., Mingozzi, F., Bencicelli, J., Banfi, S., Marshall, K.A., Testa, F., Surace, E.M., Rossi, S., Lyubarsky, A., Arruda, V.R., Konkle, B., Stone, E., Sun, J., Jacobs, J., Dell'Osso, L., Hertle, R., Ma, J.X., Redmond, T.M., Zhu, X., Hauck, B., Zelenia, O., Shindler, K.S., Maguire, M.G., Wright, J.F., Volpe, N.J., McDonnell, J.W., Auricchio, A., High, K.A., Bennett, J., 2008. Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N. Engl. J. Med.* 358, 2240–2248.
- Maheshri, N., Koerber, J.T., Kaspar, B.K., Schaffer, D.V., 2006. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat. Biotechnol.* 24, 198–204.
- McLaughlin, S.K., Collis, P., Hermonat, P.L., Muzyczka, N., 1988. Adeno-associated virus general transduction vectors: analysis of proviral structures. *J. Virol.* 62, 1963–1973.
- Mitchell, D.A., O'Donnell, J., Hare, J.T., Chapman, M.S., 2006. Serotype-specific detection of adeno-associated virus during laboratory preparation. *J. Virol. Methods* 136, 277–282.
- Mori, S., Wang, L., Takeuchi, T., Kanda, T., 2004. Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsid protein. *Virology* 330, 375–383.
- Moskalenko, M., Chen, L., van Roey, M., Donahue, B.A., Snyder, R.O., McArthur, J.G., Patel, S.D., 2000. Epitope mapping of human anti-adeno-associated virus type 2 neutralizing antibodies: implications for gene therapy and virus structure. *J. Virol.* 74, 1761–1766.
- Mueller, C., Flotte, T.R., 2008. Clinical gene therapy using recombinant adeno-associated virus vectors. *Gene Ther.* 15, 858–863.
- Murray, S., Nilsson, C.L., Hare, J.T., Emmett, M.R., Korostelev, A., Ongley, H., Marshall, A.G., Chapman, M.S., 2006. Characterization of the capsid protein glycosylation of adeno-associated virus type 2 by high-resolution mass spectrometry. *J. Virol.* 80, 6171–6176.
- Muzyczka, N., Warrington Jr., K.H., 2005. Custom adeno-associated virus capsids: the next generation of recombinant vectors with novel tropism. *Hum. Gene Ther.* 16, 408–416.
- Nam, H.J., Lane, M.D., Padron, E., Gurda, B., McKenna, R., Kohlbrenner, E., Aslanidi, C., Byrne, B., Muzyczka, N., Zolotukhin, S., Agbandje-McKenna, M., 2007. Structure of adeno-associated virus serotype 8, a gene therapy vector. *J. Virol.* 81, 12260–12271.
- Nesvizhskii, A.I., Keller, A., Kolker, E., Aebersold, R., 2003. A statistical model for identifying proteins by tandem mass spectrometry. *Anal. Chem.* 75, 4646–4658.
- Opie, S.R., Warrington Jr., K.H., Agbandje-McKenna, M., Zolotukhin, S., Muzyczka, N., 2003. Identification of amino acid residues in the capsid proteins of adeno-associated virus type 2 that contribute to heparan sulfate proteoglycan binding. *J. Virol.* 77, 6995–7006.
- Padron, E., Bowman, V., Kaludov, N., Govindasamy, L., Levy, H., Nick, P., McKenna, R., Muzyczka, N., Chiorini, J.A., Baker, T.S., Agbandje-McKenna, M., 2005. Structure of adeno-associated virus type 4. *J. Virol.* 79, 5047–5058.
- Peden, C.S., Burger, C., Muzyczka, N., Mandel, R.J., 2004. Circulating anti-wild-type adeno-associated virus type 2 (AAV2) antibodies inhibit recombinant AAV2 (rAAV2)-mediated, but not rAAV5-mediated, gene transfer in the brain. *J. Virol.* 78, 6344–6359.
- Petry, H., Brooks, A., Orme, A., Wang, P., Liu, P., Xie, J., Kretschmer, P., Qian, H.S., Hermiston, T.W., Harkins, R.N., 2008. Effect of viral dose on neutralizing antibody response and transgene expression after AAV1 vector re-administration in mice. *Gene Ther.* 15, 54–60.
- Prasad, K.M., Trempe, J.P., 1995. The adeno-associated virus Rep78 protein is covalently linked to viral DNA in a preformed virion. *Virology* 214, 360–370.
- Renesto, P., Abergel, C., Decloquement, P., Moinier, D., Azza, S., Ogata, H., Fourquet, P., Gorvel, J.P., Clavier, J.M., 2006. Mimivirus giant particles incorporate a large fraction of anonymous and unique gene products. *J. Virol.* 80, 11678–11685.
- Resch, W., Hixson, K., Moore, R.J., Lipton, M.S., Moss, B., 2007. Protein composition of the vaccinia virus mature virion. *Virology* 358, 233–247.
- Rutledge, E.A., Halbert, C.L., Russell, D.W., 1998. Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2. *J. Virol.* 72, 309–319.

- Schmidt, M., Grot, E., Cervenka, P., Wainer, S., Buck, C., Chiorini, J.A., 2006. Identification and characterization of novel adeno-associated virus isolates in ATCC virus stocks. *J. Virol.* 80, 5082–5085.
- Schmidt, M., Voutetakis, A., Afione, S., Zheng, C., Mandikian, D., Chiorini, J.A., 2008. Adeno-associated virus type 12 (AAV12): a novel AAV serotype with sialic acid- and heparan sulfate proteoglycan-independent transduction activity. *J. Virol.* 82, 1399–1406.
- Shen, X., Storm, T., Kay, M.A., 2007. Characterization of the relationship of AAV capsid domain swapping to liver transduction efficiency. *Mol. Ther.* 15, 1955–1962.
- Shi, W., Hemminki, A., Bartlett, J.S., 2006. Capsid modifications overcome low heterogeneous expression of heparan sulfate proteoglycan that limits AAV2-mediated gene transfer and therapeutic efficacy in human ovarian carcinoma. *Gynecol. Oncol.* 103, 1054–1062.
- Speir, J.A., Bothner, B., Qu, C., Willits, D.A., Young, M.J., Johnson, J.E., 2006. Enhanced local symmetry interactions globally stabilize a mutant virus capsid that maintains infectivity and capsid dynamics. *J. Virol.* 80, 3582–3591.
- Swatkoski, S., Russell, S., Edwards, N., Fenselau, C., 2007. Analysis of a model virus using residue-specific chemical cleavage and MALDI-TOF mass spectrometry. *Anal. Chem.* 79, 654–658.
- United States Pharmacopeia, 2007. The United States Pharmacopeial Convention, 30th ed., vol. 1 (736), pp. 301–305.
- Van Vliet, K., Blouin, V., Agbandje-McKenna, M., Snyder, R.O., 2006. Proteolytic mapping of the adeno-associated virus capsid. *Mol. Ther.* 14, 809–821.
- Van Vliet, K.M., Blouin, V., Brument, N., Agbandje-McKenna, M., Snyder, R.O., 2008. The role of the adeno-associated virus capsid in gene transfer. *Methods Mol. Biol.* 437, 51–91.
- Vandenbergh, L.H., Wang, L., Somanathan, S., Zhi, Y., Figueredo, J., Calcedo, R., Sanmiguel, J., Desai, R.A., Chen, C.S., Johnston, J., Grant, R.L., Gao, G., Wilson, J.M., 2006. Heparin binding directs activation of T cells against adeno-associated virus serotype 2 capsid. *Nat. Med.* 12, 967–971.
- Walters, R.W., Agbandje-McKenna, M., Bowman, V.D., Moninger, T.O., Olson, N.H., Seiler, M., Chiorini, J.A., Baker, T.S., Zabner, J., 2004. Structure of adeno-associated virus serotype 5. *J. Virol.* 78, 3361–3371.
- White, K., Buning, H., Kritiz, A., Janicki, H., McVey, J., Perabo, L., Murphy, G., Odenthal, M., Work, L.M., Hallek, M., Nicklin, S.A., Baker, A.H., 2008. Engineering adeno-associated virus 2 vectors for targeted gene delivery to atherosclerotic lesions. *Gene Ther.* 15, 443–451.
- Wu, P., Xiao, W., Conlon, T., Hughes, J., Agbandje-McKenna, M., Ferkol, T., Flotte, T., Muzyczka, N., 2000. Mutational analysis of the adeno-associated virus type 2 (AAV2) capsid gene and construction of AAV2 vectors with altered tropism. *J. Virol.* 74, 8635–8647.
- Wu, Z., Asokan, A., Grieger, J.C., Govindasamy, L., Agbandje-McKenna, M., Samulski, R.J., 2006. Single amino acid changes can influence titer, heparin binding, and tissue tropism in different adeno-associated virus serotypes. *J. Virol.* 80, 11393–11397.
- Xiao, X., Xiao, W., Li, J., Samulski, R.J., 1997. A novel 165-base-pair terminal repeat sequence is the sole cis requirement for the adeno-associated virus life cycle. *J. Virol.* 71, 941–948.
- Xiao, X., Li, J., Samulski, R.J., 1998. Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J. Virol.* 72, 2224–2232.
- Xie, Q., Bu, W., Bhatia, S., Hare, J., Somasundaram, T., Azzi, A., Chapman, M.S., 2002. The atomic structure of adeno-associated virus 2: three complete virus particles per asymmetric unit. *Acta Crystallogr. D: Biol. Crystallogr.* 59, 959–970.
- Xie, Q., Somasundaram, T., Bhatia, S., Bu, W., Chapman, M.S., 2003. Structure determination of adeno-associated virus 2: three complete virus particles per asymmetric unit. *Acta Crystallogr. D: Biol. Crystallogr.* 59, 959–970.
- Yea, C., Bukh, J., Ayers, M., Roberts, E., Krajden, M., Tellier, R., 2007. Monitoring of hepatitis C virus quasispecies in chronic infection by matrix-assisted laser desorption/ionization-time of flight mass spectrometry mutation detection. *J. Clin. Microbiol.* 45, 1053–1057.
- Zaiss, A.K., Muruve, D.A., 2008. Immunity to adeno-associated virus vectors in animals and humans: a continued challenge. *Gene Ther.* 15, 808–816.
- Zhong, L., Li, B., Mah, C.S., Govindasamy, L., Agbandje-McKenna, M., Cooper, M., Herzog, R.W., Zolotukhin, I., Warrington Jr., K.H., Weigel-Van Aken, K.A., Hobbs, J.A., Zolotukhin, S., Muzyczka, N., Srivastava, A., 2008. Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc. Natl. Acad. Sci. U.S.A.* 105, 7827–7832.
- Zolotukhin, S., Potter, M., Zolotukhin, I., Sakai, Y., Loiler, S., Fraites Jr., T.J., Chiodo, V.A., Phillipsberg, T., Muzyczka, N., Hauswirth, W.W., Flotte, T.R., Byrne, B.J., Snyder, R.O., 2002. Production and purification of serotype 1, 2, and 5 recombinant adeno-associated viral vectors. *Methods* 28, 158–167.

Adeno-associated Virus Capsid Serotype Identification: Analytical Methods Development and Application

Kim Van Vliet, Yasmin Mohiuddin, Scott McClung, Veronique Blouin, Fabienne Rolling,
Philippe Moullier, Mavis Agbandje-McKenna, and Richard O. Snyder

Supplemental Data

This customized database of the AAV capsid proteins was generated using the available capsid protein sequences (primarily VP1) found in the NCBI database. For the available NCBI sequences the NCBI accession number is provided as an active weblink. The VP2 and VP3 sequences generated (files with “KVV” that are custom files) are based on published results or predicted from the VP2 and VP3 start sites of AAV2. In some cases, redundant sequences are presented since different NCBI accession numbers are available for the same virus. In addition, the term “alternate” is included in the file to indicate alternative translation start sites, in the cases where initiation has not been confirmed. The AAV12 capsid sequences were not utilized for the MS analyses because the AAV12 sequence was not available at the time.

>gi|9632548|ref|NP_049542.1| capsid protein [adeno-associated virus 1]
MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPVEQS
PQEPDSSSGIGKTGQQPAKKRLNFGQTGDSESVDPDQPLGEPPATPAAVGPTTMASGGGAPMADNNEGADGVGNASG
NWHCDSTWLGDRVITTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCHFSRPDWQRLIN
NNWGRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLT
LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQYLYLNRQTQNSGSAQN
KDLLFSRGPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKDF
FPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDR
DVYLQGPWAKIPHTDGHFHPSPMLGGFGLKNPPPQILIKNTPVPANPPAEFSATKFASFITQYSTGQVSVEIEWELQKE
NSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|121212121|ref|KV_063497.1| major coat protein VP2 [Adeno-associated virus - 1]
MAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSESVDPDQPLGEPPATPAAVGPTTMASGGGAPMADN
NEGADGVGNASGNWHDSTWLGDRVITTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHC
HFSRPDWQRLINNNWGRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSDSEYQLPYVLGSAHQGCLPPFPA
DVFMIPQYGYLT LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQYLYL
NRTQNSGSAQNKDLLFSRGPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGT
AMASHKDDDEKFFPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHA
MGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLGGFGLKNPPPQILIKNTPVPANPPAEFSATKFASFITQYST
GQVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|131313131|ref|KV_063499.1| major coat protein VP3 [Adeno-associated virus - 1]
MASGGGAPMADNNEGADGVGNASGNWHDSTWLGDRVITTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYS
TPWGYFDFNRFHCHFSRPDWQRLINNNWGRPKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSDSEYQLPYV
LGSAGHQGCLPPFPADVFMIPQYGYLT LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLD
RLMNPLIDQYLYLNRQTQNSGSAQNKDLLFSRGPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASK
YNLNGRESIINPGTAMASHKDDDEKFFPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQS
SSTDPATGDVHAMGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLGGFGLKNPPPQILIKNTPVPANPPAEFS
ATKFASFITQYSTGQVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|[4689097](#)|gb|[AAD27757.1](#)| capsid protein [adeno-associated virus 1]
 MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLDKGEPVNAADAAALE
 HDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPVEQS
 PQEPDSSSGIGKTGQQPAKKRLNFGQTGDSESVDPQPLGEPPATPAAVGPTTMASGGGAPMADNNEGADGVGNASG
 NWHCDSTWLGDRVITTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCHFSRPRDWQRLIN
 NNWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLT
 LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQYLYLNRNTQNQSGSAQN
 KDLLFSRGSAPGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDDEKDF
 FPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHAMGALPGMVWQDR
 DVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKNPPPQILIKNTPVPANPPAEFSATKFASFITQYSTGQVSVEIEWELQKE
 NSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|7689541|gb|KVV14240.1| capsid protein VP2 [adeno-associated virus 1]
 MAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSESVDPQPLGEPPATPAAVGPTTMASGGGAPMADN
 NEGADGVGNASGNWHDSTWLGDRVITTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCH
 FFSRPRDWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSDEYQLPYVLGSAHQGCLPPFP
 ADVFMIPQYGYLTNNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLDRLMNPLIDQYLYL
 NRTQNQSGSAQNKDLLFSRGSAPGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGT
 AMASHKDDDEKDFPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSSSTDPATGDVHA
 MGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKNPPPQILIKNTPVPANPPAEFSATKFASFITQYST
 GQVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|3213524|gb|KVV89661.1| capsid protein VP3 [adeno-associated virus 1]
 MASGGGAPMADNNEGADGVGNASGNWHDSTWLGDRVITTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYST
 TPWGYFDFNRFHCHFSRPRDWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNDGVTTIANNLTSTVQVFSDEYQLPYV
 LGSAHQGCLPPFPADVFMIPQYGYLTNNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEEVPFHSSYAHSQSLD
 RLMNPLIDQYLYLNRNTQNQSGSAQNKDLLFSRGSAPGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASK
 YNLNGRESIINPGTAMASHKDDDEKDFPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNFQSS
 SSTDPATGDVHAMGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKNPPPQILIKNTPVPANPPAEFS
 ATKFASFITQYSTGQVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|[110645923](#)|ref|[YP_680426.1](#)| major coat protein VP1 [Adeno-associated virus - 2]
 MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLDKGEPVNEADAAALEH
 DKAYDRQLDSGDNPYLKYNHADADEFQERLKEDTSFGGNLGRAVFQAKKRVLEPLGLVEEPVKTAPGKKRPVEHSPV
 EPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGN
 WHCDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRPRDWQRLINNN
 WGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLT
 NNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRTNTPSGTTTQS
 RLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNGPAMASHKDDEEKFF
 PQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRD
 VYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFASFITQYSTGQVSVEIEWELQEN
 SKRWNPEIQYTSNYNKS NVNVDFTVDNNGVYSEPRPIGTRYLTRNL

>gi|[110645921](#)|ref|[YP_680427.1](#)| major coat protein VP2 [Adeno-associated virus - 2]
 MAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPQPLGQPPAAPSGLGTNTMATGSGAPMAD
 NNEGADGVGNSSGNWHDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCH
 FFSRPRDWQRLINNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFP
 ADVFMVPQYGYLTNNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLY
 YLSRTNTPSGTTTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSLVPNG
 PAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTNPVATEQYGSVSTNLQRGNRQAATADVNTQ
 GVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQILIKNTPVPANPSTTFSAAKFASFITQYSTGQ
 VSVEIEWELQKENS KRWNPEIQYTSNYNKS NVNVDFTVDNNGVYSEPRPIGTRYLTRNL

>gi|[110645922](#)|ref|[YP_680428.1](#)| major coat protein VP3 [Adeno-associated virus - 2]
 MATGSGAPMADNNEGADGVGNSSGNWHDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYST
 PWGYFDFNRFHCHFSRPRDWQRLINNNWGFPRKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVL

GSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDR
LMNPLIDQYLYLSRTNTPSGTTTTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYH
LNGRDSL VNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGN
RQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSA
KFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|[2906023](#)|gb|[AAC03780.1](#)| major coat protein VP1 [adeno-associated virus 2]

MAADGYLPDWLEDTLSEGIRQWWKLKPGPPPKPAERHKDDSRGLVLPGYKYLGPFGNGLDKGEPVNEADAAALEH
DKAYDRQLDSGDNPYLKYNHADA EFQERLKEDTSFGGNLGRAVFQAKKRVL EPLGLVEEPVK TAPGKKRPVEHSPV
EPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDQPLGQPPAAPSGLGTNTMATGSGAPMADNNEGADGVGNSSGN
WHCDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFS PRDWQRLINN
WGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTL
NNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDR LMNPLIDQYLYLSRTNTPSGTTTTQS
RLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNPGPAMASHKDDEEKFF
PQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLPGMVWQDRD
VYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAKFASFITQYSTGQVSVEIEWELQKEN
SKRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|[2906021](#)|gb|[AAC03778.1](#)| major coat protein VP2 [adeno-associated virus 2]

MAPGKKRPVEHSPVEPDSSSGTGKAGQQPARKRLNFGQTGDADSVDPDQPLGQPPAAPSGLGTNTMATGSGAPMAD
NNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFH
CHFS PRDWQRLINN WGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFP
ADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDR LMNPLIDQYLY
YLSRTNTPSGTTTTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNPG
PAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQ
GVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAKFASFITQYSTGQ
VSVEIEWELQKENS KRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|[2906022](#)|gb|[AAC03779.1](#)| major coat protein VP3 [adeno-associated virus 2]

MATGSGAPMADNNEGADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYST
PWGYFDFNRFHCHFS PRDWQRLINN WGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVL
GSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDR
LMNPLIDQYLYLSRTNTPSGTTTTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYH
LNGRDSL VNPGPAMASHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGN
RQAATADVNTQGVLPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSA
KFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|[22219269](#)|pdb|[1LP3](#)|A Chain A, The Atomic Structure Of Adeno-Associated Virus (Aav-2), A Vector For Human Gene Therapy

GADGVGNSSGNWHCDSTWMGDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFS
PRDWQRLINN WGFRPKRLNFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADV
MVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHSQSLDR LMNPLIDQYLYLSRT
NTPSGTTTTQSRLQFSQAGASDIRDQSRNWLPGPCYRQQRVSKTSADNNNSEYSWTGATKYHLNGRDSL VNPGPAMA
SHKDDEEKFFPQSGVLIFGKQGSEKTNVDIEKVMITDEEEIRTTNPVATEQYGSVSTNLQRGNRQAATADVNTQGVLP
GMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPSTTFSAKFASFITQYSTGQVSVE
IEWELQKENS KRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|[9628920](#)|ref|[NP_043941.1](#)| capsid protein [Adeno-associated virus 3]

MAADGYLPDWLEDNLSEGIREWALKPGVPQPKANQQHQDNRRGLVLPGYKYLGPFGNGLDKGEPVNEADAAALE
HDKAYDQQLKAGDNPYLKYNHADA EFQERLQEDTSFGGNLGRAVFQAKKRILEPLGLVEEAAKTAPGKKGA VDQS
PQEPDSSSGVGKSGKQARKRLNFGQTGDSESVDPDQPLGEPPAAPTSLGNTMASGGGAPMADNNEGADGVGNSSG
NWHCDSQWL GDRVITTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFS PRDWQRLINN
NWGFRPKKLSFKLFNIQVRGVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYLTL
LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHSQSLDR LMNPLIDQYLYLNRTQGTTSGTTN
QSRLFSQAGPQMSLQARNWLPGPCYRQQRVSKTANDNNNSNFPWTAASKYHLNGRDSL VNPGPAMASHKDDEE
KFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANLQSSNTAPTTGTVNHQ GALPGMVWQD

RDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQK
ENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|3332323|ref|KV_043943.1| VP2 capsid protein [Adeno-associated virus 3]
MAPGKKGAVDQSPQEPDSSSGVGKSGKQPKARKLNFGQTGDSESVDPQPPLGEPPAAPTSLGSNTMASGGGAPMAD
NNEGADGVGNSSGNWHCDSQWLGDRTVTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHC
HFSPRDWQRLINNNWGFPRPKLSFKLFNIQVRGVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPA
DVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYY
LNRTQGTTSGTTNQSRLLSQAGPQMSLQARNWLPGPCYRQRLSKTANDNNNSNFPWTAASKYHLNGRDSLVP
GPAMASHKDDEEKFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANNLQSSNTAPTTGTVN
HQQALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYS
TGQVSVEIEWELQKENS KRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|3333333|ref|KV_043942.1| VP3 capsid protein [Adeno-associated virus 3]
MASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGDRTVTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYST
PWGYFDFNRFHCHFSRPRDWQRLINNNWGFPRPKLSFKLFNIQVRGVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVL
GSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLD
RLMNPLIDQYLYYLNRTQGTTSGTTNQSRLLSQAGPQMSLQARNWLPGPCYRQRLSKTANDNNNSNFPWTAAS
KYHLNGRDSLVPNPGPAMASHKDDEEKFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANNL
QSSNTAPTTGTVNHQQALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQIMIKNTPVPANPPTT
FSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|[1408469](#)|gb|[AAC55049.1](#)| capsid protein [adeno-associated virus 3]
MAADGYLPDWLEDNLSEGIREWALKPGVPQPKANQQHQDNRRGLVLPGYKYLPGNGLDKGEPVNEADAAALE
HDKAYDQQLKAGDNPYLKYNHADAERLQEDTSFGGNLGRAVFQAKKRILEPLGLVEEAAKTAPGKKGAVDQS
PQEPDSSSGVGKSGKQPKARKLNFGQTGDSESVDPQPPLGEPPAAPTSLGSNTMASGGGAPMADNNEGADGVGNSSG
NWHCDSQWLGDRTVTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRPRDWQRLIN
NWGFPRPKLSFKLFNIQVRGVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVPQYGYL
TLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYYLNRTQGTTSGTTN
QSRLLSQAGPQMSLQARNWLPGPCYRQRLSKTANDNNNSNFPWTAASKYHLNGRDSLVPNPGPAMASHKDDEE
KFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANNLQSSNTAPTTGTVNHQQALPGMVWQD
RDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQK
ENSKRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|1333333|gb|KVV55050.1| capsid protein VP2 [adeno-associated virus 3]
MAPGKKGAVDQSPQEPDSSSGVGKSGKQPKARKLNFGQTGDSESVDPQPPLGEPPAAPTSLGSNTMASGGGAPMAD
NNEGADGVGNSSGNWHCDSQWLGDRTVTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHC
HFSPRDWQRLINNNWGFPRPKLSFKLFNIQVRGVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPA
DVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYY
LNRTQGTTSGTTNQSRLLSQAGPQMSLQARNWLPGPCYRQRLSKTANDNNNSNFPWTAASKYHLNGRDSLVP
GPAMASHKDDEEKFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANNLQSSNTAPTTGTVN
HQQALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYS
TGQVSVEIEWELQKENS KRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|1333313|gb|KVV55051.1| capsid protein VP3 [adeno-associated virus 3]
MASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGDRTVTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYST
PWGYFDFNRFHCHFSRPRDWQRLINNNWGFPRPKLSFKLFNIQVRGVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVL
GSAHQGCLPPFPADVFMVPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLD
RLMNPLIDQYLYYLNRTQGTTSGTTNQSRLLSQAGPQMSLQARNWLPGPCYRQRLSKTANDNNNSNFPWTAAS
KYHLNGRDSLVPNPGPAMASHKDDEEKFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANNL
QSSNTAPTTGTVNHQQALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPMLMGGFGLKHPPPQIMIKNTPVPANPPTT
FSPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYNKSNNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|[2766610](#)|gb|[AAB95452.1](#)| capsid protein VP1 [adeno-associated virus 3B]
MAADGYLPDWLEDNLSEGIREWALKPGVPQPKANQQHQDNRRGLVLPGYKYLPGNGLDKGEPVNEADAAALE
HDKAYDQQLKAGDNPYLKYNHADAERLQEDTSFGGNLGRAVFQAKKRILEPLGLVEEAAKTAPGKKRPVDQSP

QEPDSSSGVGKSGKQPARKRLNFGQTGDSESVDPDPQPLGEPPAAPTSLGSNTMASGGGAPMADNNEGADGVGNSSG
 NWHCDSQWLGDREVTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCHFSRDPWQRLINN
 NWGFRPKKLSFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMVMPQYGYLT
 LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYLNRQTGTTSGTTN
 QSRLFSQAGPQSMQLQARNWLPGPCYRQQRLSKTANDNNNSNFPWTAASKYHLNGRDSL VNP GPAMASHKDDEE
 KFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANLQSSNTAPTTRTVNDQGALPGMVWQD
 RDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYSTGQVSVEIEWELQK
 ENSKRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|2332333|gb|KVV95453.1| capsid protein VP2 [adeno-associated virus 3B]
 MAPGKKRPVDQSPQEPDSSSGVGKSGKQPARKRLNFGQTGDSESVDPDPQPLGEPPAAPTSLGSNTMASGGGAPMAD
 NNEGADGVGNSSGNWHCDSQWLGDREVTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYSTPWGYFDFNRFHCH
 HFSPRDWQRLINN NWGFRPKKLSFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVLGSAHQGCLPPFP
 ADVFMVMPQYGYLT LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLY
 LNRQTGTTSGTTNQSRLFSQAGPQSMQLQARNWLPGPCYRQQRLSKTANDNNNSNFPWTAASKYHLNGRDSL VNP
 GPAMASHKDDEEKFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANLQSSNTAPTTRTVN
 DQGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQIMIKNTPVPANPPTTFSPAKFASFITQYS
 TGQVSVEIEWELQKENS KRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|3332333|gb|KVV95454.1| capsid protein VP3 [adeno-associated virus 3B]
 MASGGGAPMADNNEGADGVGNSSGNWHCDSQWLGDREVTTSTRTWALPTYNNHLYKQISSQSGASNDNHYFGYST
 PWGYFDFNRFHCHFSRDPWQRLINN NWGFRPKKLSFKLFNIQVKEVTQNDGTTTIANNLTSTVQVFTDSEYQLPYVL
 GSAHQGCLPPFPADVFMVMPQYGYLT LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYTFEDVPFHSSYAHQSQSLD
 RLMNPLIDQYLYLNRQTGTTSGTTNQSRLFSQAGPQSMQLQARNWLPGPCYRQQRLSKTANDNNNSNFPWTAAS
 KYHLNGRDSL VNP GPAMASHKDDEEKFFPMHGNLIFGKEGTTASNAELDNVMITDEEEIRTTNPVATEQYGTVANL
 QSSNTAPTTRTVNDQGALPGMVWQDRDVYLQGPWAKIPHTDGHFHPSPLMGGFGLKHPPPQIMIKNTPVPANPPTTF
 SPAKFASFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYNKS VNVDFTVDTNGVYSEPRPIGTRYLTRNL

>gi|122920308|pdb|2G8G|A unnamed protein product [Adeno-associated virus - 4]
 GADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPPTYNNHLYKRLGESLQSN TYNGFSTPWGYFDFNRFHCHFSRDP
 WQRLINN NWGMRPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMV
 PQYGYCGLVTGNTSQQQTDRNAFYCLEYFPSQMLRTGNNFEITYSFEKVPFHSMYAHQSQSLDRLMNPLIDQYLWGLQ
 STTTGTTLNAGTATTNFTKL RPTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLIKYETHSTLDGRWSALTP
 GPPMATAGPADSKFSNSQLIFAGPKQNGNTATVPGTLIFTSEEELAA TNATD TDMWGNLPGGDQSNSNLPTVDRLTA
 LGAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSP LIGGFGLKHPPPQIFIKNTPVPANPATTFSSTPVNSFITQYSTGQ
 VSVQIDWEIQERSKRWNPEVQFTSNY GQQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|9629643|ref|NP_044927.1| capsid [adeno-associated virus 4]
 MTDGYLPDWLEDNLSEGVREWWALQPGAPKPKANQQHQDNARGLVLPGYKYLGPGNGLDKGEPVNAADAAALE
 HDKAYDQQLKAGDNPYLKYNHADA EFQQLQGDTSFGGNLGRAVFQAKKRVLEPLGLVEQAGETAPGKKRPLIESP
 QPDSSTGIGKKGKQPAKKKL VFEDETGAGDGPPEGSTSGAMSDDSEMRAAAGGA AVEGGQAGDGVGNASGDWH
 CDSTWSEGHVTTTSTRTWVLPPTYNNHLYKRLGESLQSN TYNGFSTPWGYFDFNRFHCHFSRDPWQRLINN NWGMRP
 KAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVMPQYGYCGLVTGNT
 SQQQTDRNAFYCLEYFPSQMLRTGNNFEITYSFEKVPFHSMYAHQSQSLDRLMNPLIDQYLWGLQSTTTGTTLNAGTA
 TTNFTKL RPTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLIKYETHSTLDGRWSALTPGPPMATAGPADS
 KFSNSQLIFAGPKQNGNTATVPGTLIFTSEEELAA TNATD TDMWGNLPGGDQSNSNLPTVDRLTALGAVPGMVWQN
 RDIYYQGPIWAKIPHTDGHFHPSP LIGGFGLKHPPPQIFIKNTPVPANPATTFSSTPVNSFITQYSTGQVSVQIDWEIQE
 RSKRWNPEVQFTSNY GQQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|4144414|ref|KV_044928.1| VP2 capsid [adeno-associated virus 4]
 MAPGKKRPLIESPQQPDSSTGIGKKGKQPAKKKL VFEDETGAGDGPPEGSTSGAMSDDSEMRAAAGGA AVEGGQGA
 DGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPPTYNNHLYKRLGESLQSN TYNGFSTPWGYFDFNRFHCHFSRDPW
 QRLINN NWGMRPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVMP
 QYGYCGLVTGNTSQQQTDRNAFYCLEYFPSQMLRTGNNFEITYSFEKVPFHSMYAHQSQSLDRLMNPLIDQYLWGLQS
 TTTGTTLNAGTATTNFTKL RPTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLIKYETHSTLDGRWSALTP
 PPMATAGPADSKFSNSQLIFAGPKQNGNTATVPGTLIFTSEEELAA TNATD TDMWGNLPGGDQSNSNLPTVDRLTAL

GAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPVPANPATTFSSTPVNSFITQYSTGQV
SVQIDWEIQKERSKRWNPEVQFTSNYGQQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|4344434|ref|KV_044930.1| VP3 capsid alternate [adeno-associated virus 4]

MSDDSEMRAAAGGAAVEGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNNTYN
GFSTPWGYFDNRFHCHFSRPRDWQRLINNNWGMMPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYEL
PYVMDAGQEGSLPPFPNDVFMVPQYGYCGLVTGNTSQQQTDNRNFAFYCLEYFSPQMLRTGNNFEITYSFEKVPFHSMY
AHSQSLDRLMNPLIDQYLWGLQSTTTGTTLNAGTATTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPAT
GSDSLIKYETHSTLDGRWSALTPGPPMATAGPADSKFSNSQLIFAGPKQNGNTATVPGLTIFTSEEELAAATNATDNDM
WGNLPGGDQSNNSNLPTVDRLTALGAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPV
PANPATTFSSTPVNSFITQYSTGQVSVQIDWEIQKERSKRWNPEVQFTSNYGQQNSLLWAPDAAGKYTEPRAIGTRYL
THHL

>gi|43444344|ref|KV_896744.1| VP3 capsid [adeno-associated virus 4]

MRAAAGGAAVEGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNNTYNGFSTPW
GYFDNRFHCHFSRPRDWQRLINNNWGMMPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDA
GQEGSLPPFPNDVFMVPQYGYCGLVTGNTSQQQTDNRNFAFYCLEYFSPQMLRTGNNFEITYSFEKVPFHSMY
AHSQSLDRLMNPLIDQYLWGLQSTTTGTTLNAGTATTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPAT
GSDSLIKYETHSTLDGRWSALTPGPPMATAGPADSKFSNSQLIFAGPKQNGNTATVPGLTIFTSEEELAAATNATDNDM
WGNLPGGDQSNNSNLPTVDRLTALGAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPV
PANPATTFSSSTPVNSFITQYSTGQVSVQIDWEIQKERSKRWNPEVQFTSNYGQQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|[2337940](#)|gb|[AAC58045.1](#)| capsid [adeno-associated virus 4]

MTDGYLPDWLEDNLSEGVREWWALQPGAPKPKANQQHQDNARGLVLPGYKYLPGNGLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLKYNHADAEEFQQRLQGDTSFGGNLGRAVFQAKKRVLEPLGLVEQAGETAPGKKRPLIESP
QQPDSSTGIGKKGKQPAKKKLVEFEDTGAGDGPPEGSTSGAMSDDSEMRAAAGGAAVEGGQGADGVGNASGDWH
CDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNNTYNGFSTPWGYFDNRFHCHFSRPRDWQRLINNNWGMMP
KAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVPQYGYCGLVTGNT
SQQQTDNRNFAFYCLEYFSPQMLRTGNNFEITYSFEKVPFHSMY AHSQSLDRLMNPLIDQYLWGLQSTTTGTTLNAGTA
TTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLIKYETHSTLDGRWSALTPGPPMATAGPADS
KFSNSQLIFAGPKQNGNTATVPGLTIFTSEEELAAATNATDNDM WGNLPGGDQSNNSNLPTVDRLTALGAVPGMVWQ
NRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPVPANPATTFSSTPVNSFITQYSTGQVSVQIDWEIQE
RSKRWNPEVQFTSNYGQQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|4222224|gb|KVV58046.1| VP2 capsid [adeno-associated virus 4]

MAPGKKRPLIESPQQPDSSTGIGKKGKQPAKKKLVEFEDTGAGDGPPEGSTSGAMSDDSEMRAAAGGAAVEGGQGA
DGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNNTYNGFSTPWGYFDNRFHCHFSRPRDW
QRLINNNWGMMPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVP
QYGYCGLVTGNTSQQQTDNRNFAFYCLEYFSPQMLRTGNNFEITYSFEKVPFHSMY AHSQSLDRLMNPLIDQYLWGLQ
STTTGTTLNAGTATTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLIKYETHSTLDGRWSALTPG
PPMATAGPADSKFSNSQLIFAGPKQNGNTATVPGLTIFTSEEELAAATNATDNDM WGNLPGGDQSNNSNLPTVDRLTAL
GAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPVPANPATTFSSTPVNSFITQYSTGQV
SVQIDWEIQKERSKRWNPEVQFTSNYGQQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|4333334|gb|KVV58047.1| capsid alternate VP3 [adeno-associated virus 4]

MSDDSEMRAAAGGAAVEGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNNTYN
GFSTPWGYFDNRFHCHFSRPRDWQRLINNNWGMMPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYEL
PYVMDAGQEGSLPPFPNDVFMVPQYGYCGLVTGNTSQQQTDNRNFAFYCLEYFSPQMLRTGNNFEITYSFEKVPFHSMY
AHSQSLDRLMNPLIDQYLWGLQSTTTGTTLNAGTATTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPAT
GSDSLIKYETHSTLDGRWSALTPGPPMATAGPADSKFSNSQLIFAGPKQNGNTATVPGLTIFTSEEELAAATNATDNDM
WGNLPGGDQSNNSNLPTVDRLTALGAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPV
PANPATTFSSTPVNSFITQYSTGQVSVQIDWEIQKERSKRWNPEVQFTSNYGQQNSLLWAPDAAGKYTEPRAIGTRYL
THHL

>gi|4399334|gb|KVV58447.1| capsid VP3 [adeno-associated virus 4]

MRAAAGGAAVEGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNTYNGFSTPW
GYFDNRFHCHFSRDPWQRLINNNWGMMPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDA
GQEGSLPPFPNDVFMVPQYGYCGLVTGNTSQQQTDRNAFYCLEYFPSQMLRTGNNFEITYSFEKVPFHSMYAHQSLS
DRLMNPLIDQYLWGLQSTTTGTTLNAGTATTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLI
KYETHSTLDGRWSALTPGPPMATAGPADSKFSNSQLIFAGPKQNGNTATVPGTLIFTSEEELAATNATDMDWGNLP
GGDQSNNSLPTVDRLTALGAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPVPANPAT
TFSSTPVNSFITQYSTGQVSVQIDWEIQKERSKRWNPEVQFTSNYGGQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|4544544|gb|KVV44544.1| Mutant R544E capsid [adeno-associated virus 4]

MTDGYLPDWLEDNLSEGVREWWALQPGAPKPKANQQHQDNARGLVLPGYKYLPGNGLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLKYNHADAEEFQQLQGDTSFGGNLGRAVFQAKKRVLEPLGLVEQAGETAPGKKRPLIESP
QQPDSSTGIGKKGKQPAKKKLVEFEDTGAGDGPPEGSTSGAMSDDSEMRAAAGGAAVEGGQGADGVGNASGDWH
CDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNTYNGFSTPWGYFDNRFHCHFSRDPWQRLINNNWGMMP
KAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVPQYGYCGLVTGNT
SQQQTDRNAFYCLEYFPSQMLRTGNNFEITYSFEKVPFHSMYAHQSLSLDRMLNPLIDQYLWGLQSTTTGTTLNAGTA
TTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLIKYETHSTLDGRWSALTPGPPMATAGPADS
KFSNSQLIFAGPEQNGNTATVPGTLIFTSEEELAATNATDMDWGNLPGGDQSNNSLPTVDRLTALGAVPGMVWQNR
DIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPVPANPATTFSSSTPVNSFITQYSTGQVSVQIDWEIQKERS
KRWNPEVQFTSNYGGQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|4544546|gb|KVV44546.1| Mutant R544E capsid VP2 [adeno-associated virus 4]

MAPGKKRPLIESPQQPDSSTGIGKKGKQPAKKKLVEFEDTGAGDGPPEGSTSGAMSDDSEMRAAAGGAAVEGGQGA
DGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNTYNGFSTPWGYFDNRFHCHFSRDPW
QRLINNNWGMMPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVP
QYGYCGLVTGNTSQQQTDRNAFYCLEYFPSQMLRTGNNFEITYSFEKVPFHSMYAHQSLSLDRMLNPLIDQYLWGLQS
TTTGTTLNAGTATTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLIKYETHSTLDGRWSALTPG
PPMATAGPADSKFSNSQLIFAGPEQNGNTATVPGTLIFTSEEELAATNATDMDWGNLPGGDQSNNSLPTVDRLTALG
AVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPVPANPATTFSSSTPVNSFITQYSTGQVS
VQIDWEIQKERSKRWNPEVQFTSNYGGQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|4544543|gb|KVV44543.1| Mutant R544E capsid alternate VP3 [adeno-associated virus 4]

MSDDSEMRAAAGGAAVEGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNTYN
GFSTPWGYFDNRFHCHFSRDPWQRLINNNWGMMPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYEL
PYVMDAGQEGSLPPFPNDVFMVPQYGYCGLVTGNTSQQQTDRNAFYCLEYFPSQMLRTGNNFEITYSFEKVPFHSMY
AHQSLSLDRMLNPLIDQYLWGLQSTTTGTTLNAGTATTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPAT
GSDSLIKYETHSTLDGRWSALTPGPPMATAGPADSKFSNSQLIFAGPEQNGNTATVPGTLIFTSEEELAATNATDMD
WGNLPGGDQSNNSLPTVDRLTALGAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPV
PANPATTFSSSTPVNSFITQYSTGQVSVQIDWEIQKERSKRWNPEVQFTSNYGGQNSLLWAPDAAGKYTEPRAIGTRYL
THHL

>gi|4578943|gb|KVV47743.1| Mutant R544E capsid VP3 [adeno-associated virus 4]

MRAAAGGAAVEGGQGADGVGNASGDWHCDSTWSEGHVTTTSTRTWVLPTYNNHLYKRLGESLQSNTYNGFSTPW
GYFDNRFHCHFSRDPWQRLINNNWGMMPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDA
GQEGSLPPFPNDVFMVPQYGYCGLVTGNTSQQQTDRNAFYCLEYFPSQMLRTGNNFEITYSFEKVPFHSMYAHQSLS
DRLMNPLIDQYLWGLQSTTTGTTLNAGTATTNFTKLRTNFSNFKKNWLPGPSIKQQGFSKTANQNYKIPATGSDSLI
KYETHSTLDGRWSALTPGPPMATAGPADSKFSNSQLIFAGPEQNGNTATVPGTLIFTSEEELAATNATDMDWGNLP
GGDQSNNSLPTVDRLTALGAVPGMVWQNRDIYYQGPIWAKIPHTDGHFHPSPPLIGGFGLKHPPPQIFIKNTPVPANPAT
TFSSTPVNSFITQYSTGQVSVQIDWEIQKERSKRWNPEVQFTSNYGGQNSLLWAPDAAGKYTEPRAIGTRYLTHHL

>gi|[51593838](#)|ref|[YP_068409.1](#)| capsid protein [Adeno-associated virus 5]

MSFVDHPPDWLEEVGEGLREFLGLAEGPPKPKNQHQDQARGLVLPGYNYLPGNGLDRGEPVNRADDEVAREHDI
SYNEQLEAGDNPYLKYNHADAEEFQEKLAADTSFGGNLGAVFQAKKRVLEPFGLVEEGAKTAPTGRIDDHFPKRK
KARTEEDSKPSTSSDAEAGPSGSQQLQIPAPASSLGADTMSAGGGPLGDNNQGADGVGNASGDWHCDSTWMGD
RVVTKSTRTWVLPSYNNHQRREIKSGSVDGSNANAYFGYSTPWGYFDNRFHSHWSRDPWQRLINNYWGFRPRSLR
VKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTE

RSSFFCLEYFPSKMLRTGNNFEFTYNFEEVPFHSSFAPSQNLFLKLANPLVDQYLRFVSTNNTGGVQFNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAATTNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIIVPGSVWMERDVYLQGPWAKIPETGAHFHSPAMGGFGLKHPPPMMLIKNTVPVGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|5151555|ref|KV_068405.1| capsid protein VP2 [Adeno-associated virus 5]
 MAPTGKRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVLPSTNNHQQYREIKSGSVSDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEEVPFHSSFAPSQNLFLKLANPLVDQYLRFVSTNNTGGVQFNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAATTNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIIVPGSVWMERDVYLQGPWAKIPETGAHFHSPAMGGFGLKHPPPMMLIKNTVPVGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|5551555|ref|KV_068406.1| capsid protein VP3 [Adeno-associated virus 5]
 MSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVLPSTNNHQQYREIKSGSVSDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEEVPFHSSFAPSQNLFLKLANPLVDQYLRFVSTNNTGGVQFNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAATTNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIIVPGSVWMERDVYLQGPWAKIPETGAHFHSPAMGGFGLKHPPPMMLIKNTVPVGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|4249658|gb|AAD13756.1| capsid protein [adeno-associated virus 5]
 MSFVDHPPDWLEEVGEGLREFLGLAAGPPKPKPNQHQDQARGLVLPGYNYLPGNGLDRGEPVNRADDEVAREHDI SYNEQLEAGDNPYLKYNHADAFAEFQEKLAADDTSFGGNLGAQVFAKKRVLEPFLVEEAKTAPTGRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVLPSTNNHQQYREIKSGSVSDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEEVPFHSSFAPSQNLFLKLANPLVDQYLRFVSTNNTGGVQFNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAATTNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIIVPGSVWMERDVYLQGPWAKIPETGAHFHSPAMGGFGLKHPPPMMLIKNTVPVGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|4245555|gb|KVV13757.1| capsid protein VP2 [adeno-associated virus 5]
 MAPTGKRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVLPSTNNHQQYREIKSGSVSDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEEVPFHSSFAPSQNLFLKLANPLVDQYLRFVSTNNTGGVQFNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAATTNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIIVPGSVWMERDVYLQGPWAKIPETGAHFHSPAMGGFGLKHPPPMMLIKNTVPVGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|4255555|gb|KVV13758.1| capsid protein VP3 [adeno-associated virus 5]
 MSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWVLPSTNNHQQYREIKSGSVSDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFPRSLRVKIFNIQVKEVTVQDSTTTIANNLTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTERSSFFCLEYFPSKMLRTGNNFEFTYNFEEVPFHSSFAPSQNLFLKLANPLVDQYLRFVSTNNTGGVQFNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAATTNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIIVPGSVWMERDVYLQGPWAKIPETGAHFHSPAMGGFGLKHPPPMMLIKNTVPVGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|[4160148](#)|emb|[CAA77024.1](#)| capsid protein [Adeno-associated virus - 5]
 MSFVDHPPDWLEEVGEGLEFLGLEAGPPKPKPNQQHQDQARGLVLPGYNYLPGNGLDRGEPVNRADDEVAREHDI
 SYNEQLEAGDNPYLKYNHADADEFQEKLAADDTSFGGNLGAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDHFPKRK
 KARTEEDSKPSTSSDAEAGPSGSQQLQIPAPASSLGADTMSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGD
 RVVTKSTRTWLPSYNNHQYREIKSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSRPDWQRLINNYWGFRPRSLR
 VKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTE
 RSSFFCLEYFPSKMLRTGNNFEFTYNFEEVVPFHSSFAPSQNLFLKLANPLVDQYLRFVSTNNTGGVQFNKNLAGRYAN
 TYKNWFPGPMGRTOGWNLGSGVNRASVSFAFATTNRMELEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPA
 NPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIIVPGSVWMERDVYLQGPWAKI
 PETGAHFHPSAMGGFGLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQY
 TNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|4160155|emb|KVV77025.1| capsid protein VP2 [Adeno-associated virus - 5]
 MAPTGKRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAPASSLGADTMSAGGGGGLGDNNQGADGVGN
 ASGDWHCDSTWMGDRVVTCKSTRTWLPSYNNHQYREIKSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSRPDW
 QRLINNYWGFRPRSLRVKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQ
 YGYATLNRDNTENPTERSFFCLEYFPSKMLRTGNNFEFTYNFEEVVPFHSSFAPSQNLFLKLANPLVDQYLRFVSTNNT
 GGVQFNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAFATTNRMELEGASYQVPPQPNGMTNNLQGS
 NTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMATNNQSSTTAPATGTYNLQEIIVPGSV
 WMERDVYLQGPWAKIPETGAHFHPSAMGGFGLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEW
 ELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|4165555|emb|KVV77555.1| capsid protein VP3 [Adeno-associated virus - 5]
 MSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGDRVVTCKSTRTWLPSYNNHQYREIKSGSVDGSNANAYFG
 YSTPWGYFDFNRFHSHWSRPDWQRLINNYWGFRPRSLRVKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDDYQLP
 YVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNTENPTERSFFCLEYFPSKMLRTGNNFEFTYNFEEVVPFHSSFAPSQ
 NLFKLANPLVDQYLRFVSTNNTGGVQFNKNLAGRYANTYKNWFPGPMGRTOGWNLGSGVNRASVSFAFATTNRME
 LEGASYQVPPQPNGMTNNLQGSNTYALENTMIFNSQPANPGTTATYLEGNMLITSESETQPVNRVAYNVGGQMAT
 NNQSSTTAPATGTYNLQEIIVPGSVWMERDVYLQGPWAKIPETGAHFHPSAMGGFGLKHPPPMMLIKNTPVPGNITS
 FSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDSTGEYRTTRPIGTRYLTRPL

>gi|[85070098](#)|gb|[ABC69726.1](#)| VP1 [Adeno-associated virus-Go.1]
 MSFVDHPPDWLEEVGEGLEFLGLEAGPPKPKPNQQHQDQARGLVLPGYNYLPGNGLDRGEPVNRADDEVAREHDI
 SYNEQLEAGDNPYLKYNHADADEFQEKLAADDTSFGGNLGAVFQAKKRVLEPFGLVEEGAKTAPTGKRIDDHFPKRK
 KARTEEDSKPSTSSDAEAGPSGSQQLQIPAPASSLGADTMSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGD
 RVVTKSTRTWLPSYNNHQYREIKSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSRPDWQRLINNYWGFRPRSLR
 VKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNGDNPT
 RSSFFCLEYFPSKMLRTGNNFEFTYSFEEVVPFHCSFAPSQNLFLKLANPLVDQYLRFVSTSATGAIQFQKNLAGRYAN
 YKNWFPGPMGRTOGWNTSSGSSTNRVSVNNFVSVMNRMLEGASYQVNPQPNGMTNTLQGSNRYALENTMIFNAQN
 ATPGTTSVYPEDNLLTSESETQPVNRVAYNTGGQMATNAQNATTAPTGTYNLQEVLPGSVWMERDVYLQGPWAKI
 PETGAHFHPSAMGGFGLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQ
 YTNNYNDPQFVDFAPDGSGEYRTTRAIGTRYLTRPL

>gi|85070755|gb|KVV69727.1| VP2 [Adeno-associated virus-Go.1]
 MAPTGKRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAPASSLGADTMSAGGGGGLGDNNQGADGVGN
 ASGDWHCDSTWMGDRVVTCKSTRTWLPSYNNHQYREIKSGSVDGSNANAYFGYSTPWGYFDFNRFHSHWSRPDW
 QRLINNYWGFRPRSLRVKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDDYQLPYVVGNGTEGCLPAFPQVFTLPQ
 YGYATLNRDNGDNPTERSFFCLEYFPSKMLRTGNNFEFTYSFEEVVPFHCSFAPSQNLFLKLANPLVDQYLRFVSTSAT
 GAIQFQKNLAGRYANTYKNWFPGPMGRTOGWNTSSGSSTNRVSVNNFVSVMNRMLEGASYQVNPQPNGMTNTLQGS
 NRYALENTMIFNAQNATPGTTSVYPEDNLLTSESETQPVNRVAYNTGGQMATNAQNATTAPTGTYNLQEVLPGS
 VWMERDVYLQGPWAKIPETGAHFHPSAMGGFGLKHPPPMMLIKNTPVPGNITSFSDVPVSSFITQYSTGQVTVEME
 WELKKENSKRWNPEIQYTNNYNDPQFVDFAPDGSGEYRTTRAIGTRYLTRPL

>gi|85070555|gb|KVV69728.1| VP3 [Adeno-associated virus-Go.1]
 MSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGDRVVTCKSTRTWLPSYNNHQYREIKSGSVDGSNANAYFG
 YSTPWGYFDFNRFHSHWSRPDWQRLINNYWGFRPRSLRVKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDDYQLP

YVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNGDNPTERRSSFFCLEYFPSKMLRTGNNFEFTYSFEEVPPHCSFAPSQ
NLFKLANPLVDQYLRFVSTSATGAIQFQKNLAGRYANTYKNWFPGPMGRTOGWNTSSGSSTNRVSVNNFVSNSRM
NLEGASYQVNPQPNMGMTNTLQGSNRYALENTMIFNAQNATPGTTSVYPEDNLLLTSESETQPVNRVAYNTGGQMAT
NAQNATTAPTGTYNLQEVLPGSVWMERDVYLQGPWAKIPETGAHFHPSAMGGFGLKHPPPMMLIKNTVPVGNIT
SFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDGSGEYRTTRAIGTRYLTRPL

>gi|[52630845](#)|gb|[AAU84890.1](#)| capsid protein [Caprine adeno-associated virus 1]

MSFVDHPPDWLEEVGEGLEFLGLEAGPPKPKPNQHQDQARGLVLPGYNYLPGNGLDREGPVNRADEVAREHDI
SYNEQLEAGDNPYLKYNHADAFAEFQEKLAADDTSFGGNLGKAVFQAKKRVLEPFGLVEEGAKTAPTGRIDDHFPKRK
KARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGD
RVVTKSTRTWLPSYNNHQQYREIKSGSVSDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFRPRSLR
VKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDYQLPYVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNGDNPT
RSSFFCLEYFPSKMLRTGNNFEFTYSFEEVPPHCSFAPSQNLFKLANPLVDQYLRFVSTSATGAIQFQKNLAGRYANT
YKNWFPGPMGRTOGWNTSSGSSTNRVSVNNFVSNSRMNLEGASYQVNPQPNMGMTNTLQGSNRYALENTMIFNAQN
ATPGTTSVYPEDNLLLTSESETQPVNRVAYNTGGQMATNAQNATTAPTGTYNLQEVLPGSVWMERDVYLQGPWA
KIPETGAHFHPSAMGGFGLKHPPPMMLIKNTVPVGNITSFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQ
YTNNYNDPQFVDFAPDGSGEYRTTRAIGTRYLTRPL

>gi|52630555|gb|KVV84893.1| capsid protein VP2 [Caprine adeno-associated virus 1]

MAPTGRIDDHFPKRKKARTEEDSKPSTSSDAEAGPSGSQQLQIPAQPASSLGADTMSAGGGGGLGDNNQGADGVGN
ASGDWHCDSTWMGDRVVTKSTRTWLPSYNNHQQYREIKSGSVSDGSNANAYFGYSTPWGYFDFNRFHSHWSPRDW
QRLINNYWGFRPRSLRVKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDYQLPYVVGNGTEGCLPAFPQVFTLPQ
YGYATLNRDNGDNPTERRSSFFCLEYFPSKMLRTGNNFEFTYSFEEVPPHCSFAPSQNLFKLANPLVDQYLRFVSTSAT
GAIQFQKNLAGRYANTYKNWFPGPMGRTOGWNTSSGSSTNRVSVNNFVSNSRMNLEGASYQVNPQPNMGMTNTLQGS
SNRYALENTMIFNAQNATPGTTSVYPEDNLLLTSESETQPVNRVAYNTGGQMATNAQNATTAPTGTYNLQEVLPGS
VWMERDVYLQGPWAKIPETGAHFHPSAMGGFGLKHPPPMMLIKNTVPVGNITSFSDVPVSSFITQYSTGQVTVEME
WELKKENSKRWNPEIQYTNNYNDPQFVDFAPDGSGEYRTTRAIGTRYLTRPL

>gi|52635555|gb|KVV84896.1| capsid protein VP3 [Caprine adeno-associated virus 1]

MSAGGGGGLGDNNQGADGVGNASGDWHCDSTWMGDRVVTKSTRTWLPSYNNHQQYREIKSGSVSDGSNANAYFG
YSTPWGYFDFNRFHSHWSPRDWQRLINNYWGFRPRSLRVKIFNIQVKEVTVQDSTTTIANNTSTVQVFTDDDYQLP
YVVGNGTEGCLPAFPQVFTLPQYGYATLNRDNGDNPTERRSSFFCLEYFPSKMLRTGNNFEFTYSFEEVPPHCSFAPSQ
NLFKLANPLVDQYLRFVSTSATGAIQFQKNLAGRYANTYKNWFPGPMGRTOGWNTSSGSSTNRVSVNNFVSNSRM
NLEGASYQVNPQPNMGMTNTLQGSNRYALENTMIFNAQNATPGTTSVYPEDNLLLTSESETQPVNRVAYNTGGQMAT
NAQNATTAPTGTYNLQEVLPGSVWMERDVYLQGPWAKIPETGAHFHPSAMGGFGLKHPPPMMLIKNTVPVGNIT
SFSDVPVSSFITQYSTGQVTVEMEWELKKENSKRWNPEIQYTNNYNDPQFVDFAPDGSGEYRTTRAIGTRYLTRPL

>gi|[2766607](#)|gb|[AAB95450.1](#)| capsid protein VP1 [adeno-associated virus 6]

MAADGYLPDWLEDNLSEGIREWWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFNGLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLRYNHADAFAEFQERLQEDTSFGGNLGRAVFAKAKKRVLEPFGLVEEGAKTAPGKKRPVEQSP
QEPDSSSGIGKTGQQPAKKRLNFGQTGDSESVDPDPQLGEPPATPAAVGPTTMASGGGAPMADNNEGADGVGNASG
NWHCDSTWLGDRVITSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCHFSRPRDWQRLIN
NNWGFRPKRLNFKLFNIQVKEVTTNDGVTIANNTSTVQVFSDEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLT
LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHQSOLDRLMNPLIDQYLYLNRTQNQSGSAQN
KDLLFSRGSAPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGTAMASHKDDKDKF
FPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNLQSSSTD PATGDVHVMGALPGMVWQDR
DVYLQGPWAKIPHTDGHFHSPLMGGFGLKHPPQPILIKNTVPANPPAEFSATKFASFITQYSTGQVSVEIEWELQKE
NSKRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|6766666|gb|KVV95456.1| capsid protein VP2 [adeno-associated virus 6]

MAPGKKRPVEQSPQEPDSSSGIGKTGQQPAKKRLNFGQTGDSESVDPDPQLGEPPATPAAVGPTTMASGGGAPMADN
NEGADGVGNASGNWHCDSTWLGDRVITSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYSTPWGYFDFNRFHCH
HFSRPRDWQRLINNNWGFRPKRLNFKLFNIQVKEVTTNDGVTIANNTSTVQVFSDEYQLPYVLGSAHQGCLPPFPA
DVFMIPQYGYLT LNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHQSOLDRLMNPLIDQYLYLN
NRTQNQSGSAQNKDLLFSRGSAPAGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASKYNLNGRESIINPGT
AMASHKDDKDKFFPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNLQSSSTD PATGDVHV

MGALPGMVWQDRDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPPAEFSATKFASFITQYST
GQVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|6666666|gb|KVV95666.1| capsid protein VP3 [adeno-associated virus 6]

MASGGGAPMADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISSASTGASNDNHYFGYS
TPWGYFDFNRFHCHFSRDPWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNDGVTIANNLTSTVQVFSSEYQLPYV
LGSAGHQCGLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFTFSYTFEDVPFHSSYAHQSGLD
RLMNPLIDQYLYLNRNTQNQSGSAQNKDLLFSRGSAPGMSVQPKNWLPGPCYRQQRVSKTKTDNNNSNFTWTGASK
YNLNGRESIINPGTAMASHKDDKDKFFPMMSGVMIFGKESAGASNTALDNVMITDEEEIKATNPVATERFGTVAVNLQS
SSTDPATGDVHVMGALPGMVWQDRDVYLQGPIWAKIPHTDGHFHPSPLMGGFGLKHPPPQILIKNTPVPANPPAEFS
ATKFASFITQYSTGQVSVEIEWELQKENS KRWNPEVQYTSNYAKSANVDFTVDNNGLYTEPRPIGTRYLTRPL

>gi|51949962|ref|YP_077178.1| capsid protein [Adeno-associated virus - 7]

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDNGRGLVLPGYKYLGPFGNGLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLPLGLVEEGAKTAPAKKRPEVPS
QRSPDSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPLGEPPAAPSSVSGSTVAAGGGAPMADNNEGADGVGNAS
GNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISSETAGSTNDNTYFGYSTPWGYFDFNRFHCHFSRDPWQRLIN
NNWGFPRKRLNFKLFNIQVKEVTTNDGVTIANNLTSTIQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLT
LNNGSQSVGRSSFYCLEYFPSQMLRTGNNFEFSYSFEDVPFHSSYAHQSGLDRLMNPLIDQYLYLARTQSNPGGTAG
NRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQNNNSNFAWTGATKYHLNGRNSLVNPGVAMATHKDDE
DRFFPSSGVLIFGKTGATNKTTLENVLMTNEEEIRPTNPVATEEYGIVSSNLQAANTAAQTQVVNNQGALPGMVWQN
RDVYLQGPIWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTPVPANPPEVFTPAKFASFITQYSTGQVSVEIEWELQK
ENSKRWNPFIQYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL

>gi|77949962|ref|KV_777177.1| capsid protein VP2 [Adeno-associated virus - 7]

MAPAKKRPEVPSQRSPDSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPLGEPPAAPSSVSGSTVAAGGGAPMAD
NNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISSETAGSTNDNTYFGYSTPWGYFDFNRFH
CHFSRDPWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNDGVTIANNLTSTIQVFSSEYQLPYVLGSAHQGLPPFP
ADVFMIPQYGYLTLNNGSQSVGRSSFYCLEYFPSQMLRTGNNFEFSYSFEDVPFHSSYAHQSGLDRLMNPLIDQYLYL
ARTQSNPGGTAGNRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQNNNSNFAWTGATKYHLNGRNSLVN
PGVAMATHKDDEDRFFPSSGVLIFGKTGATNKTTLENVLMTNEEEIRPTNPVATEEYGIVSSNLQAANTAAQTQVV
NNQGALPGMVWQNRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTPVPANPPEVFTPAKFASFITQYST
GQVSVEIEWELQKENS KRWNPEIQYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL

>gi|77749962|ref|KV_777777.1| capsid protein VP3 [Adeno-associated virus - 7]

MADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISSETAGSTNDNTYFGYSTPWGYFDF
NRFHCHFSRDPWQRLINNNWGFPRKRLNFKLFNIQVKEVTTNDGVTIANNLTSTIQVFSSEYQLPYVLGSAHQGL
PPFPADVFMIPQYGYLTLNNGSQSVGRSSFYCLEYFPSQMLRTGNNFEFSYSFEDVPFHSSYAHQSGLDRLMNPLIDQY
LYLARTQSNPGGTAGNRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQNNNSNFAWTGATKYHLNGRNS
LVNPGVAMATHKDDEDRFFPSSGVLIFGKTGATNKTTLENVLMTNEEEIRPTNPVATEEYGIVSSNLQAANTAAQTQV
VNNQGALPGMVWQNRDVYLQGPIWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTPVPANPPEVFTPAKFASFITQ
YSTGQVSVEIEWELQKENS KRWNPEIQYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL

>gi|22652861|gb|AAN03855.1|AF513851_2 capsid protein [adeno-associated virus 7]

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDNGRGLVLPGYKYLGPFGNGLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLPLGLVEEGAKTAPAKKRPEVPS
QRSPDSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPLGEPPAAPSSVSGSTVAAGGGAPMADNNEGADGVGNAS
GNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISSETAGSTNDNTYFGYSTPWGYFDFNRFHCHFSRDPWQRLIN
NNWGFPRKRLNFKLFNIQVKEVTTNDGVTIANNLTSTIQVFSSEYQLPYVLGSAHQGLPPFPADVFMIPQYGYLT
LNNGSQSVGRSSFYCLEYFPSQMLRTGNNFEFSYSFEDVPFHSSYAHQSGLDRLMNPLIDQYLYLARTQSNPGGTAG
NRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQNNNSNFAWTGATKYHLNGRNSLVNPGVAMATHKDDE
DRFFPSSGVLIFGKTGATNKTTLENVLMTNEEEIRPTNPVATEEYGIVSSNLQAANTAAQTQVVNNQGALPGMVWQN
RDVYLQGPIWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTPVPANPPEVFTPAKFASFITQYSTGQVSVEIEWELQK
ENSKRWNPFIQYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL

>gi|77772861|gb|KVV77855.1|KV713851_7 capsid protein VP2 [adeno-associated virus 7]
 MAPAKKRPVEPSPQRSPTSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPLGEPPAAPSSVSGTVAAGGGAPMAD
 NNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISSETAGSTNDNTYFGYSTPWGYFDFNRFH
 CHFSPRDWQRLINNNWGFPRPKLRFKLFNIQVKEVTNDGVTIANNTSTIQVFSDEYQLPYVLGSAHQGCLPPFP
 DVMIPQYGYLTLNNGSQSVGRSSFYCLEYFPSQMLRTGNNEFFSYSFEDVPFHSSYAHQSGLDRLMNPLIDQYLYL
 ARTQSNPGGTAGNRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQNNNSNFAWTGATKYHLNGRNSLVN
 GVAMATHKDDDRFFPSSGVLIFGKTGATNKTTLENVLMTNEEEIRPTNPVATEEYGIVSSNLQAANTAAQTQVNN
 QGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPMLMGGFGLKHPPPQILIKNTVPANPPEVFTPAKFASFITQYSTG
 QVSVEIEWELQKENS KRWNPEIQYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL

>gi|77772877|gb|KVV77877.1|KV713851_7 capsid protein VP3 [adeno-associated virus 7]
 MADNNEGADGVGNASGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISSETAGSTNDNTYFGYSTPWGYFDF
 NRFHCHFSPRDWQRLINNNWGFPRPKLRFKLFNIQVKEVTNDGVTIANNTSTIQVFSDEYQLPYVLGSAHQGCL
 PPFPADVMIPQYGYLTLNNGSQSVGRSSFYCLEYFPSQMLRTGNNEFFSYSFEDVPFHSSYAHQSGLDRLMNPLIDQY
 LYLARTQSNPGGTAGNRELQFYQGGPSTMAEQAKNWLPGPCFRQQRVSKTLDQNNNSNFAWTGATKYHLNGRNS
 LVNPGVAMATHKDDDRFFPSSGVLIFGKTGATNKTTLENVLMTNEEEIRPTNPVATEEYGIVSSNLQAANTAAQTQV
 VNNQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPMLMGGFGLKHPPPQILIKNTVPANPPEVFTPAKFASFITQ
 YSTGQVSVEIEWELQKENS KRWNPEIQYTSNFEKQTGVDFAVDSQGVYSEPRPIGTRYLTRNL

>gi|51949965|ref|YP_077180.1| capsid protein [Adeno-associated virus - 8]
 MAADGYLPDWLEDNLSEGIREWALKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLDKGEPVNAADAAALE
 HDKAYDQQLQAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPVEPSP
 QRSPTSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPLGEPPAAPSGVGPNTMAAGGGAPMADNNEGADGVGSSS
 GNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISNGTSGGATNDNTYFGYSTPWGYFDFNRFHCHFSPRDWQRL
 INNNWGFPRKLSFKLFNIQVKEVTQNEGTKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVMIPQYGYL
 TLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFTYTFEDVPFHSSYAHQSGLDRLMNPLIDQYLYLSRTQTGGTAN
 TQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNNSNFAWTAGTKYHLNGRNSLANPGIAMATHKDDEE
 RFFPSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGIVADNLQQQNTAPQIGTVNSQGALPGMVWQ
 RDVYLQGPWAKIPHTDGNFHPSPMLMGGFGLKHPPPQILIKNTVPADPPTTFNQSKLNSFITQYSTGQVSVEIEWELQ
 ENSKRWNPEIQYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

>gi|88949988|ref|KV_887188.1| capsid protein VP2 [Adeno-associated virus - 8]
 MAPGKKRPVEPSPQRSPTSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPLGEPPAAPSGVGPNTMAAGGGAPMAD
 NNEGADGVGSSSGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISNGTSGGATNDNTYFGYSTPWGYFDFNRFH
 HCHFSPRDWQRLINNNWGFPRKLSFKLFNIQVKEVTQNEGTKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFP
 ADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFTYTFEDVPFHSSYAHQSGLDRLMNPLIDQYLY
 LSRTQTGGTANTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNNSNFAWTAGTKYHLNGRNSLANP
 GIAMATHKDDEERFFPSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGIVADNLQQQNTAPQIGTVNS
 QGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPMLMGGFGLKHPPPQILIKNTVPADPPTTFNQSKLNSFITQYST
 GQVSVEIEWELQKENS KRWNPEIQYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

>gi|88889988|ref|KV_888888.1| capsid protein VP3 [Adeno-associated virus - 8]
 MAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISNGTSGGATNDNTYFGY
 STPWGYFDFNRFHCHFSPRDWQRLINNNWGFPRKLSFKLFNIQVKEVTQNEGTKTIANNLTSTIQVFTDSEYQLPYV
 LGSAGHQGCLPPFPADVMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFTYTFEDVPFHSSYAHQSGLD
 RLMNPLIDQYLYLSRTQTGGTANTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNNSNFAWTAGT
 KYHLNGRNSLANPGIAMATHKDDEERFFPSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGIVADNLQ
 QQNTAPQIGTVNSQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPMLMGGFGLKHPPPQILIKNTVPADPPTTFN
 QSKLNSFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

>gi|22652864|gb|AAN03857.1|AF513852_2 capsid protein [adeno-associated virus 8]
 MAADGYLPDWLEDNLSEGIREWALKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLDKGEPVNAADAAALE
 HDKAYDQQLQAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEGAKTAPGKKRPVEPSP
 QRSPTSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPLGEPPAAPSGVGPNTMAAGGGAPMADNNEGADGVGSSS
 GNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISNGTSGGATNDNTYFGYSTPWGYFDFNRFHCHFSPRDWQRL
 INNNWGFPRKLSFKLFNIQVKEVTQNEGTKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVMIPQYGYL

TLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFTYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYLSRTQTTGGTAN
TQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNNSNFAWTAGTKYHLNGRNSLANPGIAMATHKDDEE
RFFPSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGIVADNLQQQNTAPQIGTVNSQGALPGMVWQN
RDVYLQGPWAKIPHTDGNFHPSPMLGGFGLKHPPPQILIKNTVPADPPTTFNQSKLNSFITQYSTGQVSVEIEWELQK
ENSKRWNPEIQYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

>gi|82652884|gb|KVV88857.1|KV813852_8 capsid protein VP2[adeno-associated virus 8]

MAPGKKRPVEPSPQRSPDSSTGIGKKGQQPARKRLNFGQTGDSESVDPQPLGEPPAAPSGVGPNTMAAGGGAPMAD
NNEGADGVGSSSGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISNGTSGGATNDNTYFGYSTPWGYFDFNRF
HCHFSRPDWQRLINNNWGFRPKRLSFKLFNIQVKEVTQNEGKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFP
ADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFTYTFEDVPFHSSYAHSQSLDRLMNPLIDQYLYY
LSRTQTTGGTANTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNNSNFAWTAGTKYHLNGRNSLANP
GAMATHKDDEERFFPSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGIVADNLQQQNTAPQIGTVNS
QGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPMLGGFGLKHPPPQILIKNTVPADPPTTFNQSKLNSFITQYST
GQVSVEIEWELQKENS KRWNPEIQYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

>gi|82652888|gb|KVV88858.1|KV813852_8 capsid protein VP3[adeno-associated virus 8]

MAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRTTSTRTWALPTYNNHLYKQISNGTSGGATNDNTYFGY
STPWGYFDFNRFHCHFSRPDWQRLINNNWGFRPKRLSFKLFNIQVKEVTQNEGKTIANNLTSTIQVFTDSEYQLPYV
LGSAGHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFQFTYTFEDVPFHSSYAHSQSLD
RLMNPLIDQYLYLSRTQTTGGTANTQTLGFSQGGPNTMANQAKNWLPGPCYRQQRVSTTTGQNNNSNFAWTAGT
KYHLNGRNSLANPGIAMATHKDDEERFFPSNGILIFGKQNAARDNADYSDVMLTSEEEIKTTNPVATEEYGIVADNLQ
QQNTAPQIGTVNSQGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPMLGGFGLKHPPPQILIKNTVPADPPTTFN
QSKLNSFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYYKSTSVDFAVNTEGVYSEPRPIGTRYLTRNL

>gi|[46487805](#)|gb|[AAS99264.1](#)| capsid protein VP1 [Adeno-associated virus 9]

MAADGYLPDWLEDNLSEGIREWALKPGAPQPKANQQHQDNARGLVLPGYKYLPGNGLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLKYNHADAERLQKEDTSFGGNLGRAVFQAKKRLLLEPLGLVEEAAKTAPGKKRPVEQSP
QEPDSSAGIGKSGAQPAKKRLNFGQTGDTEVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADNNEGADGVGSSSGN
WHCDSQWLGDRTTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFHCHFSRPDWQRLINN
NWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSYQLPYVLGSAHEGCLPPFPADVFMIPQYGYLTL
NDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYLSKTINGSGQNQQT
LKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGPAMASHKEGEDRFFP
LSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNNQGILPGMVWQDRD
VYLQGPWAKIPHTDGNFHPSPMLGGFGMKHPPPQILIKNTVPADPPTAFNKDKLNSFITQYSTGQVSVEIEWELQKE
NSKRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL

>gi|99997805|gb|KVV99969.1| capsid protein VP2 [Adeno-associated virus 9]

MAPGKKRPVEQSPQEPDSSAGIGKSGAQPAKKRLNFGQTGDTEVPDPQPIGEPPAAPSGVGSLTMASGGGAPVADN
NEGADGVGSSSGNWHCDSQWLGDRTTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGYSTPWGYFDFNRFH
CHFSRPDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSYQLPYVLGSAHEGCLPPFP
ADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSLDRLMNPLIDQYLYY
LSKTINGSGQNQQTLLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSEFAWPGASSWALNGRNSLMNPGP
AMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQSAQAQAQTGWVQNN
QGILPGMVWQDRDVYLQGPWAKIPHTDGNFHPSPMLGGFGMKHPPPQILIKNTVPADPPTAFNKDKLNSFITQYST
GQVSVEIEWELQKENS KRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL

>gi|99997899|gb|KVV99999.1| capsid protein VP3 [Adeno-associated virus 9]

MASGGGAPVADNNEGADGVGSSSGNWHCDSQWLGDRTTSTRTWALPTYNNHLYKQISNSTSGGSSNDNAYFGY
STPWGYFDFNRFHCHFSRPDWQRLINNNWGFRPKRLNFKLFNIQVKEVTDNNGVKTIANNLTSTVQVFTDSYQLPY
VLGSAHEGCLPPFPADVFMIPQYGYLTLNDGSQAVGRSSFYCLEYFPSQMLRTGNNFQFSYEFENVPFHSSYAHSQSL
DRLMNPLIDQYLYLSKTINGSGQNQQTLLKFSVAGPSNMAVQGRNYIPGPSYRQQRVSTTVTQNNNSEFAWPGASSW
ALNGRNSLMNPGPAMASHKEGEDRFFPLSGSLIFGKQGTGRDNVDADKVMITNEEEIKTTNPVATESYGQVATNHQS
AQAQAQTGWVQNNQGILPGMVWQDRDVYLQGPWAKIPHTDGNFHPSPMLGGFGMKHPPPQILIKNTVPADPPTAF
NKDKLNSFITQYSTGQVSVEIEWELQKENS KRWNPEIQYTSNYYKSNNVEFAVNTEGVYSEPRPIGTRYLTRNL

>gi|[48728343](#)|gb|[AAT46337.1](#)| capsid protein [Adeno-associated virus 10]

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEAAKTAPGKKRPVEPSP
QRSPDSSTGIGKKKGQPAKKRLNFGQTGESESVDPQPIGEPPAGPSGLSGTMAAGGGAPMADNNEGADGVGSSSG
NWHCDSTWLGDRVITTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDFNRFHCHFSRWDWQRLIN
NNWGFPRKRLSFKLFNIQVKEVTQNEGKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFPADVFMIPQYGYLTL
NNGSQAVGRSSFYCLEYFPSQMLRTGNNFEFSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYLSRTQSTGGTQGTQ
QLLFSQAGPANMSAQAKNWLPGPCYRQQRVSTTSLSQNNNSNFAWTGATKYHLNGRDSLVPNGVAMATHKDDEERF
FPSSGVLMFQKQAGARDNVYSSVMLTSEEEIKTTNPVATEQYGVVADNLQQANTGPIVGNVNSQAGALPGMVWQN
RDVYLQGPWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTVPADPPTTFSQAKLASFITQYSTGQVSVEIEWELQK
ENSKRWNPFIQYTSNYYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL

>gi|10101010|gb|KVV11011.1| capsid protein VP2 [Adeno-associated virus 10]

MAPGKKRPVEPSPQRSPDSSTGIGKKKGQPAKKRLNFGQTGESESVDPQPIGEPPAGPSGLSGTMAAGGGAPMAD
NNEGADGVGSSSGNWHCDSTWLGDRVITTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGYSTPWGYFDFNRF
HCHFSRWDWQRLINNNWGFPRKRLSFKLFNIQVKEVTQNEGKTIANNLTSTIQVFTDSEYQLPYVLGSAHQGCLPPFP
ADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFEFSYTFEDVPFHSSYAHQSQSLDRLMNPLIDQYLYY
LSRTQSTGGTQGTQQLLFSQAGPANMSAQAKNWLPGPCYRQQRVSTTSLSQNNNSNFAWTGATKYHLNGRDSLVP
GVAMATHKDDEERFFPSSGVLMFQKQAGARDNVYSSVMLTSEEEIKTTNPVATEQYGVVADNLQQANTGPIVGNV
NSQAGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTVPADPPTTFSQAKLASFITQYS
TGQVSVEIEWELQKENSkrwnpeiQYTSNYYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL

>gi|12310123|gb|KVV12311.1| capsid protein VP3 [Adeno-associated virus 10]

MAAGGGAPMADNNEGADGVGSSSGNWHCDSTWLGDRVITTSTRTWALPTYNNHLYKQISNGTSGGSTNDNTYFGY
STPWGYFDFNRFHCHFSRWDWQRLINNNWGFPRKRLSFKLFNIQVKEVTQNEGKTIANNLTSTIQVFTDSEYQLPYV
LGSAHQGCLPPFPADVFMIPQYGYLTLNNGSQAVGRSSFYCLEYFPSQMLRTGNNFEFSYTFEDVPFHSSYAHQSQSLD
RLMNPLIDQYLYLSRTQSTGGTQGTQQLLFSQAGPANMSAQAKNWLPGPCYRQQRVSTTSLSQNNNSNFAWTGATK
YHLNGRDSLVPNGVAMATHKDDEERFFPSSGVLMFQKQAGARDNVYSSVMLTSEEEIKTTNPVATEQYGVVADNL
QQANTGPIVGNVNSQAGALPGMVWQNRDVYLQGPWAKIPHTDGNFHPSPLMGGFGLKHPPPQILIKNTVPADPPTTF
SQAKLASFITQYSTGQVSVEIEWELQKENSkrwnpeiQYTSNYYKSTNVDFAVNTEGTYSEPRPIGTRYLTRNL

>gi|[48728346](#)|gb|[AAT46339.1](#)| capsid protein [Adeno-associated virus 11]

MAADGYLPDWLEDNLSEGIREWDLKPGAPKPKANQQKQDDGRGLVLPGYKYLGPFGNGLDKGEPVNAADAAALE
HDKAYDQQLKAGDNPYLRYNHADADEFQERLQEDTSFGGNLGRAVFQAKKRVLEPLGLVEEAAKTAPGKKRPLESPQ
EPDSSSGIGKKKGKQPAKKRLNFEEDTGAGDGPPEGSDTSAMSSDIEMRAAPGGNAVDAGQGSQDGVGNASGDWHCDS
TWSEGVVTTTSTRTWVLPPTYNNHLYLRLGTTSSSNTYNGFSTPWGYFDFNRFHCHFSRWDWQRLINNNWGLRPKAM
RVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVPQYGYCGIVTGENQNQT
DRNAFYCLEYFPSQMLRTGNNFEMAYNFEKVPFHSMYAHQSQSLDRLMNPLLDQYLWHLQSTTSGETLNQGNAATTF
GKIRSGDFAFYRKNWLPGPCVKQQRFSKTASQNYKIPASGGNALLKYDTHYTLNNRWSNIAPGPPMATAGPSDGDFFS
NAQLIFPGPSVTGNTTTSANNLLFTSEEEIAATNPRDMDMFGQIADNNQNATTAPITGNVTAMGVLPGMVWQNRDIY
YQGPWAKIPHADGHFHPSPPLIGGFGLKHPPPQIFIKNTVPANPATTTAARVDSFITQYSTGQVAVQIEWEIEKERSK
RWNPVQFTSNYGNQSSMLWAPDTTGKYTEPRVIGSRYLTNHL

>gi|41111346|gb|KVV4111.1| capsid protein VP2 [Adeno-associated virus 11]

MAPGKKRPLESPQEPDSSSGIGKKKGKQPAKKRLNFEEDTGAGDGPPEGSDTSAMSSDIEMRAAPGGNAVDAGQGSQD
VGNASGDWHCDSTWSEGVVTTTSTRTWVLPPTYNNHLYLRLGTTSSSNTYNGFSTPWGYFDFNRFHCHFSRWDWQRL
INNNWGLRPKAMRVKIFNIQVKEVTTSNGETTVANNLTSTVQIFADSSYELPYVMDAGQEGSLPPFPNDVFMVPQYGY
YCGIVTGENQNQTDNAFYCLEYFPSQMLRTGNNFEMAYNFEKVPFHSMYAHQSQSLDRLMNPLLDQYLWHLQSTTS
GETLNQGNAATTFGKIRSGDFAFYRKNWLPGPCVKQQRFSKTASQNYKIPASGGNALLKYDTHYTLNNRWSNIAPGP
PMATAGPSDGDFFSNAQLIFPGPSVTGNTTTSANNLLFTSEEEIAATNPRDMDMFGQIADNNQNATTAPITGNVTAMGV
LPGMVWQNRDIYYQGPWAKIPHADGHFHPSPPLIGGFGLKHPPPQIFIKNTVPANPATTTAARVDSFITQYSTGQVA
VQIEWEIEKERSkrwnpevQFTSNYGNQSSMLWAPDTTGKYTEPRVIGSRYLTNHL

>gi|41111116|gb|KVV11119.1| capsid protein alternate VP3 [Adeno-associated virus 11]

MSSDIEMRAAPGGNAVDAGQGS DGVGNASGDWHCDSTWSEGVTTTSTRTWVLP TYNHLYLRLGTTSSSNTYNG
FSTPWGYFDNRFHCHFS PRDWQRLINNNWGLRPKAMRVKIFNIQVKEVTTSNGETT VANNTSTVQIFADSSYELPY
VMDAGQEGSLPPFPNDVFMVPQYGYCGIVTGENQNQTDRNAFYCLEYFPSQMLRTGNNFEMAYNFEKVPFHSMYA
HSQSLDRLMNPLLDQYLWHLQSTTSGETLNQGNAAATTFGKIRSGDFAFYRKNWLPGPCVKQQRFSKTASQNYKIPAS
GGNALLKYDTHYTLNNRWSNIAPGPPMATAGPSDGD FSNACLIFPGPSVTGNTTTSANNLLFTSEEEIAATNPRD TDM
FGQIADNNQNATTAPITGNVTAMGVLPGMVWQNRDIYYQGPIWAKIPHADGHFHPSP LIGGFGLKHPPPQIFIKNTPVP
ANPATTFTAARVDSFITQYSTGQVAVQIEWEIEKERSKRWNPEVQFTSNYGNQSSMLWAPD TTGKYTEPRVIGSRYL T
NHL

>gi|41168116|gb|KVV16819.1| capsid protein VP3 [Adeno-associated virus 11]

MRAAPGGNAVDAGQGS DGVGNASGDWHCDSTWSEGVTTTSTRTWVLP TYNHLYLRLGTTSSSNTYNGFSTPWG
YFDNRFHCHFS PRDWQRLINNNWGLRPKAMRVKIFNIQVKEVTTSNGETT VANNTSTVQIFADSSYELPYVMDAG
QEGSLPPFPNDVFMVPQYGYCGIVTGENQNQTDRNAFYCLEYFPSQMLRTGNNFEMAYNFEKVPFHSMYAHSQSLD
RLMNPLLDQYLWHLQSTTSGETLNQGNAAATTFGKIRSGDFAFYRKNWLPGPCVKQQRFSKTASQNYKIPASGGNALL
KYDTHYTLNNRWSNIAPGPPMATAGPSDGD FSNACLIFPGPSVTGNTTTSANNLLFTSEEEIAATNPRD TDMFGQIAD
NNQNATTAPITGNVTAMGVLPGMVWQNRDIYYQGPIWAKIPHADGHFHPSP LIGGFGLKHPPPQIFIKNTPVPANPAT
TFTAARVDSFITQYSTGQVAVQIEWEIEKERSKRWNPEVQFTSNYGNQSSMLWAPD TTGKYTEPRVIGSRYL TNHL

>gi|112379656|gb|ABI16639.1| capsid protein [Adeno-associated virus 12]

MAADGYLPDWLEDNLSEGI REWWALKPGAPQPKANQQHQDNGRGLVLP GYKYLGPFGNGLDKGEPVNEADAAALE
HDKAYDKQLEQGDNPYLKYNHADA EFQQLATDTSFGGNLGRAVFQAKKRILEPLGLVEEGVK TAPGKKRPLEKTP
NRPTNPDSGKAPAKKKQKDGE PADSARRTLDFEDSGAGDGPPEGSSSGEMSHDAEMRAAPGGNAVEAGQGADGVG
NASGDWHCDSTWSEGRVTTTSTRTWVLP TYNHLYLRIGTTANSNTYNGFSTPWGYFDNRFHCHFS PRDWQRLIN
NNWGLRPKSMRVKIFNIQVKEVTTSNGETT VANNTSTVQIFADSTYELPYVMDAGQEGSFPPFPNDVFMVPQYGYC
GVVTGKNQNQTDRNAFYCLEYFPSQMLRTGNNFEVSYQFEKVPFHSMYAHSQSLDRMMNPLLDQYLWHLQSTTTG
NSLNQGTATTTYGKITTGDFAYYRKNWLPGACIKQKQFSKNANQNYKIPASGGDALLKYDTHHTLNGRWSNMAPGP
PMATAGAGDSDFSNSQLIFAGPNPSGNTTSSNNLLFTSEEEIATTNPRD TDMFGQIADNNQNATTAPHIANLDAMGIV
PGMVWQNRDIYYQGPIWAKVPHTDGHFHPSP LMGGFGLKHPPPQIFIKNTPVPANPNTTFS AARINSFLTQYSTGQVA
VQIDWEIQKEHSKRWNPEVQFTSNYGTQNSMLWAPDNAGNYHELRAIGSRFLTHHL

AAV12 – Capsid Protein VP2

MAPGKKRPLEKTPNRPTNPDSGKAPAKKKQKDGE PADSARRTLDFEDSGAGDGPPEGSSSGEMSHDAEMRAAPGGN
AVEAGQGADGVGNASGDWHCDSTWSEGRVTTTSTRTWVLP TYNHLYLRIGTTANSNTYNGFSTPWGYFDNRFH
CHFS PRDWQRLINNNWGLRPKSMRVKIFNIQVKEVTTSNGETT VANNTSTVQIFADSTYELPYVMDAGQEGSFPPFP
NDVFMVPQYGYCGVVTGKNQNQTDRNAFYCLEYFPSQMLRTGNNFEVSYQFEKVPFHSMYAHSQSLDRMMNPLLD
QYLWHLQSTTTGNSLNQGTATTTYGKITTGDFAYYRKNWLPGACIKQKQFSKNANQNYKIPASGGDALLKYDTHHT
LNGRWSNMAPGPPMATAGAGDSDFSNSQLIFAGPNPSGNTTSSNNLLFTSEEEIATTNPRD TDMFGQIADNNQNATT
APHIANLDAMGIVPGMVWQNRDIYYQGPIWAKVPHTDGHFHPSP LMGGFGLKHPPPQIFIKNTPVPANPNTTFS AARI
NSFLTQYSTGQVAVQIDWEIQKEHSKRWNPEVQFTSNYGTQNSMLWAPDNAGNYHELRAIGSRFLTHHL

AAV12 – Capsid Protein VP3

MSHDAEMRAAPGGNAVEAGQGADGVGNASGDWHCDSTWSEGRVTTTSTRTWVLP TYNHLYLRIGTTANSNTYN
GFSTPWGYFDNRFHCHFS PRDWQRLINNNWGLRPKSMRVKIFNIQVKEVTTSNGETT VANNTSTVQIFADSTYELP
YVMDAGQEGSFPPFPNDVFMVPQYGYCGVVTGKNQNQTDRNAFYCLEYFPSQMLRTGNNFEVSYQFEKVPFHSMY
AHSQSLDRMMNPLLDQYLWHLQSTTTGNSLNQGTATTTYGKITTGDFAYYRKNWLPGACIKQKQFSKNANQNYKIP
ASGGDALLKYDTHHTLNGRWSNMAPGPPMATAGAGDSDFSNSQLIFAGPNPSGNTTSSNNLLFTSEEEIATTNPRD T
DMFGQIADNNQNATTAPHIANLDAMGIVPGMVWQNRDIYYQGPIWAKVPHTDGHFHPSP LMGGFGLKHPPPQIFIKN
TPVPANPNTTFS AARINSFLTQYSTGQVAVQIDWEIQKEHSKRWNPEVQFTSNYGTQNSMLWAPDNAGNYHELRAIG
SRFLTHHL