



# Improved Genome Packaging Efficiency of Adeno-associated Virus Vectors Using Rep Hybrids

Mario Mietzsch,<sup>a</sup> Courtnee Eddington,<sup>a</sup> Ariana Jose,<sup>a</sup> Jane Hsi,<sup>a</sup> Paul Chipman,<sup>a</sup> Tom Henley,<sup>b</sup> Modassir Choudhry,<sup>b</sup> Robert McKenna,<sup>a</sup> Mavis Agbandje-McKenna<sup>a†</sup>

<sup>a</sup>Department of Biochemistry and Molecular Biology, Center for Structural Biology, McKnight Brain Institute, College of Medicine, University of Florida, Gainesville, Florida, USA

<sup>b</sup>Intima Bioscience, New York, New York, USA

**ABSTRACT** Recombinant adeno-associated viruses (rAAVs) are one of the most commonly used vectors for a variety of gene therapy applications. In the last 2 decades, research focused primarily on the characterization and isolation of new *cap*, genes resulting in hundreds of natural and engineered AAV capsid variants, while the *rep* gene, the other major AAV open reading frame, has been less studied. This is due to the fact that the *rep* gene from AAV serotype 2 (AAV2) enables the single-stranded DNA packaging of recombinant genomes into most AAV serotype and engineered capsids. However, a major by-product of all vector productions is empty AAV capsids, lacking the encapsidated vector genome, especially for non-AAV2 vectors. Despite the packaging process being considered the rate-limiting step for rAAV production, none of the *rep* genes from the other AAV serotypes have been characterized for their packaging efficiency. Thus, in this study AAV2 *rep* was replaced with the *rep* gene of a select number of AAV serotypes. However, this led to a lowering of capsid protein expression, relative to the standard AAV2-*rep* system. In further experiments the 3' end of the AAV2 *rep* gene was reintroduced to promote increased capsid expression and a series of chimeras between the different AAV Rep proteins were generated and characterized for their vector genome packaging ability. The utilization of these novel Rep hybrids increased the percentage of genome containing (full) capsids approximately 2- to -4-fold for all of the non-AAV2 serotypes tested. Thus, these Rep chimeras could revolutionize rAAV production.

**IMPORTANCE** A major by-product of all adeno-associated virus (AAV) vector production systems are “empty” capsids, void of the desired therapeutic gene, and thus do not provide any curative benefit for the treatment of the targeted disease. In fact, empty capsids can potentially elicit additional immune responses *in vivo* gene therapies if not removed by additional purification steps. Thus, there is a need to increase the genome packaging efficiency and reduce the number of empty capsids from AAV biologics. The novel Rep hybrids from different AAV serotypes described in this study are capable of reducing the percentage of empty capsids in all tested AAV serotypes and improve overall yields of genome-containing AAV capsids at the same time. They can likely be integrated easily into existing AAV manufacturing protocols to optimize the production of the generated AAV gene therapy products.

**KEYWORDS** adeno-associated virus, capsid, cryo-electron microscopy, empty and full capsids, genome, hybrid, packaging, RNA splicing, Rep protein

Adeno-associated viruses (AAVs) are not associated with any pathogenic effects but are widely studied, as they have been developed into one of the most promising gene therapy vectors for *in vivo* gene therapy for a wide variety of monogenetic diseases (1). To date, three AAV-vector-mediated gene therapies have gained approval for

**Citation** Mietzsch M, Eddington C, Jose A, Hsi J, Chipman P, Henley T, Choudhry M, McKenna R, Agbandje-McKenna M. 2021. Improved genome packaging efficiency of adeno-associated virus vectors using Rep hybrids. *J Virol* 95:e00773-21. <https://doi.org/10.1128/JVI.00773-21>.

**Editor** Rebecca Ellis Dutch, University of Kentucky College of Medicine

**Copyright** © 2021 Mietzsch et al. This is an open-access article distributed under the terms of the [Creative Commons Attribution 4.0 International license](https://creativecommons.org/licenses/by/4.0/).

Address correspondence to Robert McKenna, [rmckenna@ufl.edu](mailto:rmckenna@ufl.edu).

† This article is written in loving memory of Mavis Agbandje-McKenna.

**Received** 7 May 2021

**Accepted** 3 July 2021

**Accepted manuscript posted online** 21 July 2021

**Published** 9 September 2021

commercialization: Glybera, an AAV1 vector for the treatment of lipoprotein lipase deficiency (2); Luxturna, an AAV2 vector for the treatment of Leber's congenital amaurosis (3); and Zolgensma, an AAV9 vector for the treatment of spinal muscular atrophy type 1 (4). Common to all natural and engineered AAVs are their small, nonenveloped icosahedral capsids, ~260 Å in diameter, that contain a linear single-stranded DNA (ssDNA) genome (5). The wild-type AAVs have a genome size of approximately 4.7 kb (6). Both ends of the genome contain identical inverted terminal repeats (ITRs) of ~150 nucleotides (nt) that form T-shaped hairpin secondary structures and are important for genome replication and packaging (7). Between these ITRs, two open reading frames (ORFs) encode a series of replication (Rep) protein and virus proteins (VPs) (8), as well as two smaller accessory proteins: the assembly activating protein (AAP) and the membrane-associated accessory protein (MAAP) (9, 10). For recombinant AAV vectors, these viral ORFs are replaced with an approximately similar-sized therapeutic gene of interest and only the flanking *cis*-active ITRs are retained to allow packaging of the recombinant genomes into the capsids. Nevertheless, the Rep protein and VPs are also needed for vector manufacturing but can be supplied in *trans* (11, 12).

The Rep proteins play critical roles in the viral replication cycle as well as for vector production (8). The four Rep proteins—Rep78, Rep68, Rep52, and Rep40—are situated in the same ORF and are translated from transcripts generated by the p5 and p19 promoter (13). The larger Rep78/68 proteins are extended by ~224 amino acids at their N termini in comparison to Rep52/40, and the C terminus of Rep68/40 differs from those of Rep78/52 due to differential splicing (13). All Rep proteins share a central 305-amino-acid stretch that contains a helicase/ATPase domain (14) and nuclear localization signals (15). The N terminus of Rep78/68 contains a DNA-binding and endonuclease domain (16), whereas the C terminus of Rep78/52 is suggested to contain a zinc-finger domain (17). The large Rep proteins Rep78/68 are indispensable for genome replication and were shown to bind to a specific DNA sequence also present in the ITRs (18). The small Rep proteins Rep52/40 are essential for packaging of the genome into the capsids (19) formed by VP1, VP2, and VP3 that are encoded by the second ORF.

Similar to the Rep proteins, the VPs are located in one ORF and share a common C terminus. They are generated from transcripts of the p40 promoter by differential splicing and by the utilization of alternate start codons. VP1 is the largest capsid protein, with approximately 735 amino acids, and is encoded within the entire ORF. VP2 and VP3 are truncated at their N termini relative to VP1, with approximately 598 and 533 amino acids, respectively. The VPs are expressed and incorporated into the capsid in an approximate ratio of 1:1:10 (VP1:VP2:VP3) (20). The capsid assembly is assisted by AAP, which is encoded within the *cap* gene but situated in an alternate reading frame (9). Similarly, the second accessory protein, MAAP, is also encoded in an alternate reading frame within the *cap* gene and was suggested to have a role in capsid egress (10).

Packaging of the wild-type or recombinant vector genomes occurs in the nucleus into preformed "empty" VP capsids (21). Following replication of the genomes by the large Rep proteins, the generated ssDNA genomes are thought to be translocated into the capsid in 3'-to-5' direction by the helicase/ATPase domain of the small Rep proteins (19). The translocation of the ssDNA into the preformed capsid is proposed to occur through a channel at the 5-fold symmetry axis of the AAV capsids (5). Since genome packaging is considered to be the rate-limiting step, a significant number of capsids go unpacked and remain empty. For AAV-mediated gene therapy, these empty capsids are not desired since they do not provide any therapeutic benefit and may induce unwanted immune responses.

For the AAVs 13 primate AAV serotypes and numerous genotypes or engineered capsids have been described (22). The Rep and Capsid proteins of the AAV serotypes vary in amino acid sequence identity between ~52 to 99% (Table 1). The wide variety of available capsids is utilized in AAV-mediated gene therapy to package the transgene of choice into the capsid with the desired characteristics and tissue tropism. However, for the vast majority of manufactured AAV vectors the ITRs and *rep* gene of serotype 2

**TABLE 1** Comparison of the VP1 and Rep78 amino acid sequence identity of the AAV serotypes

AAV	AAV1	AAV2	AAV3	AAV4	AAV5	AAV6	AAV7	AAV8	AAV9	AAV10	AAV11	AAV12	AAV13		
AAV1		83	87	63	58	99	85	84	82	85	66	60	87	VP1 amino acid sequence identity (%)	
AAV2	87		88	60	57	83	82	83	82	84	63	60	88		
AAV3	89	89		63	58	87	85	86	84	86	65	61	94		
AAV4	90	90	93		53	63	63	63	62	63	81	78	65		
AAV5	58	58	58	58		58	58	58	57	57	53	52	58		
AAV6	99	87	89	89	58		85	84	82	85	66	60	87		
AAV7	98	88	90	90	58	98		88	82	88	67	62	85		
AAV8	95	85	87	88	57	95	96		85	93	65	62	85		
AAV9	NA		86	64	60	84									
AAV10	97	88	89	90	58	97	97	95	NA		66	61	86		
AAV11	97	88	89	89	58	97	97	95	NA	100		84	65		
AAV12	87	88	89	88	58	87	88	85	NA	88	88		60		
AAV13	89	90	93	99	58	89	90	88	NA	89	89	88			
Rep78 amino acid sequence identity (%)															

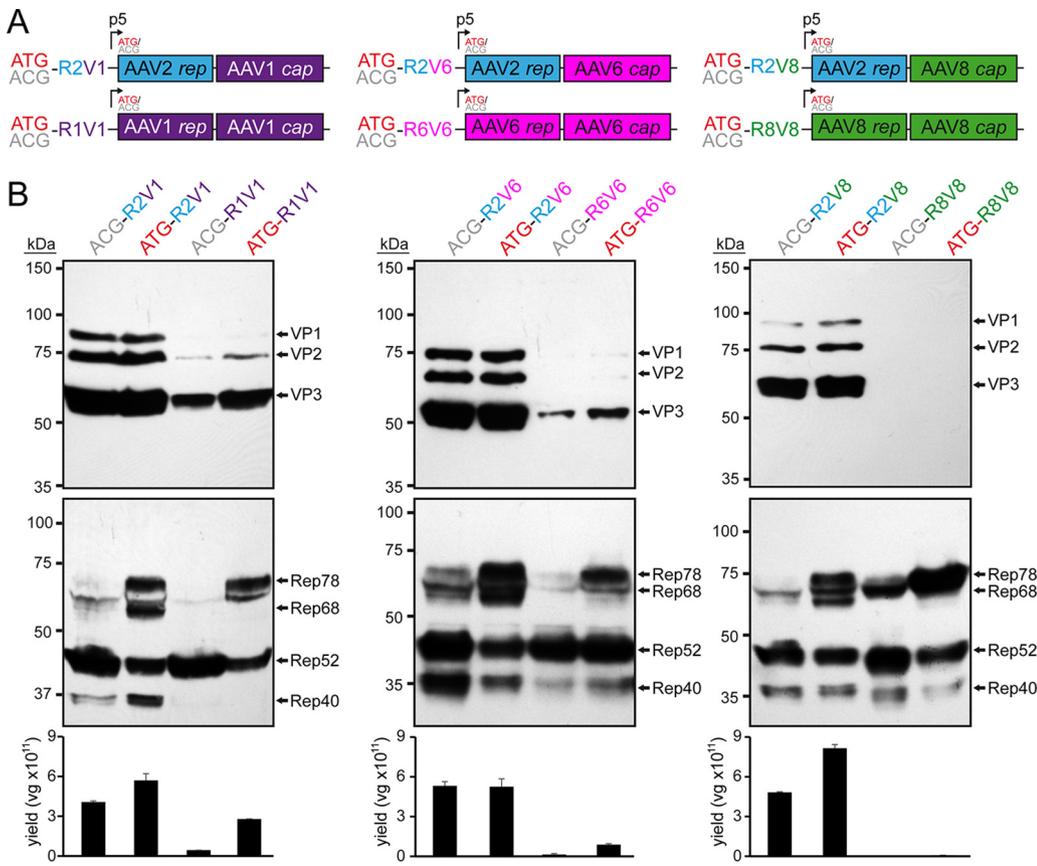
(AAV2) are utilized, and only the *cap* gene is changed, resulting in pseudotyped AAVs (23). Although the Rep proteins of AAV2 are able to package vector genomes in all AAV serotype capsids, their packaging efficiency is lower for some capsids compared to AAV2 (24). As a result, AAV vector preparations with non-AAV2 capsids contain a higher percentage of empty capsids.

Thus, in order to characterize the utilization of other *rep* genes in AAV vector productions, the AAV2 *rep* gene was replaced in this study with the *rep* gene of the same AAV serotype as the *cap* gene for a selection of different AAV serotypes. However, this substitution led to an overall lower VP expression for the analyzed AAV serotypes relative to the standard AAV2-*rep* system. Reverting the 3' end of the *rep* gene back to AAV2, as well as corrections in the DNA-binding domain of AAV8 Rep, restored high capsid expression. In addition, a higher packaging efficiency of vector genomes was observed. In a series of chimeras, primarily the DNA-binding and endonuclease domains of the AAV1 and AAV8 Rep proteins were shown to be responsible for improved vector genome packaging for AAV1 and AAV8 vectors. These observations were confirmed by enzyme-linked immunosorbent assay (ELISA), quantitative PCR (qPCR), alkaline gel electrophoresis, and cryo-electron microscopy (cryo-EM) of affinity-purified AAV vector preparations and show that the new Rep hybrids can increase the number of genome-containing capsids 2- to 4-fold in the case of AAV1 and AAV8 compared to AAV2-*rep* produced vectors. Furthermore, utilization of these hybrids for other AAVs, such as AAV6, AAV9, and AAVrh.10, also shows higher genome packaging efficiency. These observations indicate that the utilization of these Rep chimeras could revolutionize AAV vector production.

## RESULTS AND DISCUSSION

**Substitution of the standard AAV2 *rep* gene.** In the last 2 decades, a lot of research has focused on the characterization and isolation of new *cap* genes, resulting in hundreds of natural and engineered AAV capsid variants (22, 25–28). At the same time, the *rep* gene has been largely ignored due to the fact that the Rep proteins of AAV2 also package recombinant vector genomes into almost all AAV serotypes or variant capsids. A major by-product of all vector productions are empty AAV capsids, with even lower packaging efficiencies for non-AAV2 vectors (24). Compared to AAV2, the amino acid sequence identity of VP1 ranges from 82 to 88% for most AAV serotype capsids, with the exception of AAV4/5/11/12 (Table 1). For the foremost AAV serotypes, the amino acid sequence identity of Rep78 shows a similar range of 85 to 90%, indicating a potential coevolution of the Rep proteins with the capsid VPs, potentially due to optimal interaction of Rep proteins with the matched serotype capsids during genome packaging.

Thus, to understand the impact for the utilization of different AAV Rep proteins for AAV vector production, the AAV2 *rep* gene was replaced in the producer plasmids with



**FIG 1** Substitution of the AAV2 *rep* gene. (A) Overview of the utilized *rep-cap* constructs. AAV genes derived from AAV1 are indicated in purple, AAV2 in blue, AAV6 in pink, and AAV8 in green. The construct name indicates the utilized Rep78 start codon (ATG or ACG), as well as the origins of the *rep* (R) and *cap* (V) genes (e.g., R2V1: *rep* gene from AAV2 and *cap* gene from AAV1). (B) Analysis of Rep and VP expression by Western blotting after transfection of the constructs in HEK293 cells. The top blots were probed with MAb B1, and the lower blots were probed with MAb 1F. Individual VPs and Rep proteins are indicated.

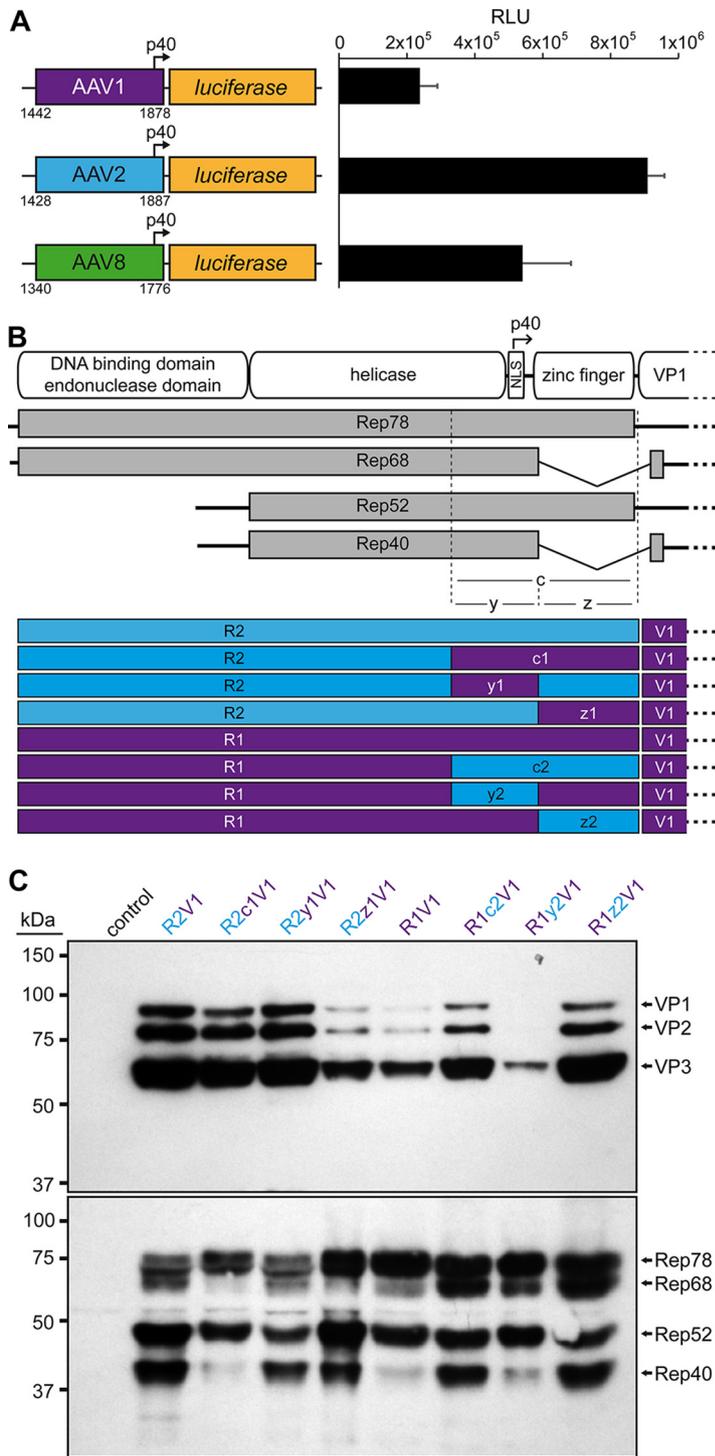
the *rep* gene of the AAV serotype identical to the *cap* gene for the serotypes AAV1, -6, and -8 (Fig. 1A). When the expression of the capsid proteins was analyzed, all of new constructs demonstrated low to no VP expression (in the case of AAV8) compared to the standard AAV2 *rep* constructs (Fig. 1B). The majority of AAV producer plasmids utilize AAV2 *rep* genes with the Rep78 start codon changed to an ACG to lower the expression of the large Rep proteins in favor of the small Rep proteins (29). This strategy was shown to improve the overall AAV yield (30). Thus, ATG and ACG constructs for each producer plasmid were generated. While the Rep78 start codon had no significant effect on VP expression for any construct, the ACG start codon reduced Rep78/68 expression in all constructs as anticipated (Fig. 1B), but in contrast to previous observations the overall yield of AAV vectors was not increased when using the ACG-Rep2 constructs (Fig. 1C). In fact, the yields of AAV1 and AAV6 vectors using the AAV1 and AAV6 *rep* constructs were higher with the Rep78 having an ATG start codon. Another observation for the AAV1 and AAV6 *rep* constructs was their lower expression levels of spliced Rep proteins, Rep68 and Rep40, compared to the AAV2 *rep* constructs. The similar behavior of the AAV1 and -6 constructs was expected since their *rep* genes differ only in 36 nt, resulting in only four amino acid differences or 99% sequence identity (Table 1). This level of similarity between AAV1 and -6 is comparable to VP1, with only six amino acid differences between the AAV serotypes.

**Reintroduction of AAV2 sequences at the 3' end of AAV1 *rep* gene rescues VP expression.** The low expression of the capsid proteins prevents efficient AAV vector production. Since the VPs are translated from transcripts generated from the p40

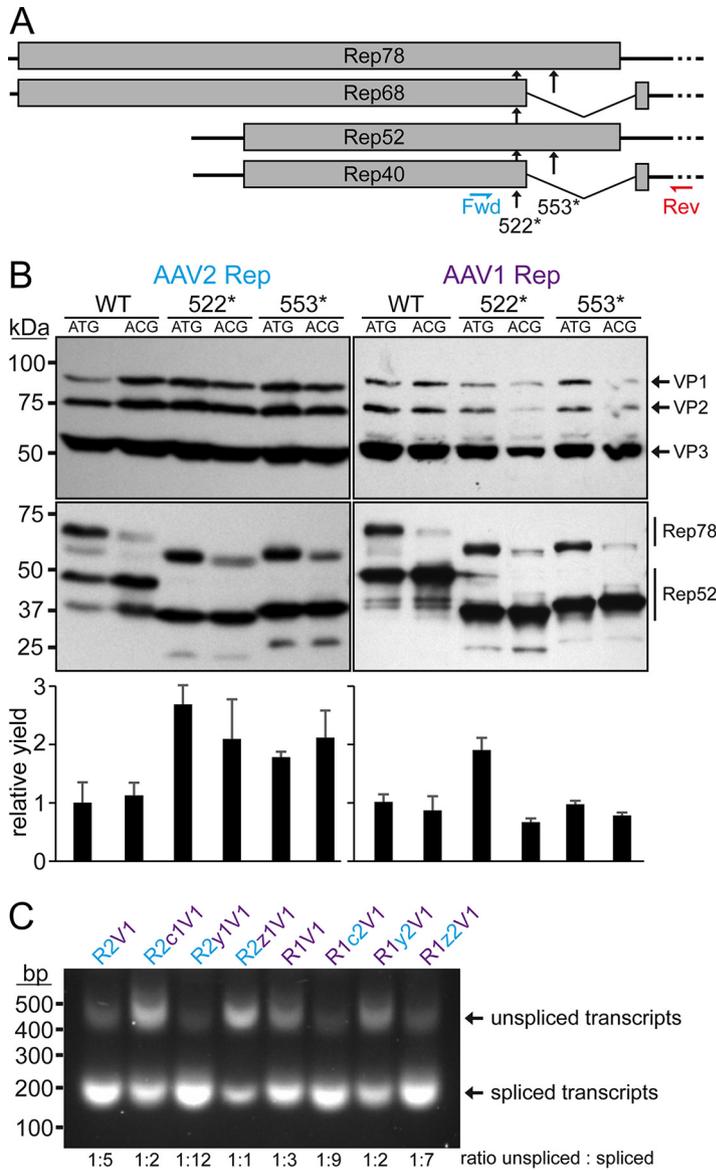
promoter, which is located in the 3' end of the *rep* gene, there was the possibility that the AAV1, -6, or -8 p40 promoter does not have the same activity as the AAV2 p40 promoter. Alternatively, the observed reduced expression of Rep68/40 (Fig. 1B) could indicate a disruption of mRNA splicing. The capsid proteins have been reported to be primarily translated from spliced p40 transcripts, and mutations of the splice donor site resulted in a strong reduction of VP expression (31, 32). However, the splice donor consensus sequences of AAV1, -6, and -8 are identical to that of AAV2 (32). In order to test the first possibility, AAV p40 promoter luciferase reporter constructs were generated for AAV1, -2, and -8, and their luciferase expressions were analyzed (Fig. 2A). Indeed, the AAV1 p40 promoter (AAV6 differs only by a single nucleotide) showed only about a quarter of the expression compared to AAV2. However, the AAV8 p40 promoter showed ~60% of the AAV2 p40-mediated expression (Fig. 2A), even though VP expression was lower for the AAV8 construct compared to AAV1 (Fig. 1B). Thus, there is the possibility of overlapping detrimental effects.

In order to rescue VP expression from the AAV1 *rep* constructs, a series of swaps of the 3' end of the *rep* gene (or the C terminus on protein level) for AAV1 and -2 were generated (Fig. 2B). While the selected region for the swaps do contain overlapping amino acids of the C terminus of the helicase domain, no amino acid differences are located in this region of the helicase between AAV1, -2, -6, and -8. These constructs were analyzed for their expression of the Rep and capsid proteins. As seen previously, the VP expression of the construct with the entire AAV1 *rep* (R1V1) was reduced, and less Rep68/40 was observed compared to the standard AAV2 *rep* construct (R2V1) (Fig. 2C). Substitution of the 3' end of AAV1 *rep* to AAV2 (R2c1V2) resulted only in a minor reduction of VP expression but showed less of Rep68/40. Vice versa, the introduction of 3' end of AAV2 *rep* in AAV1 increased VP expression and showed higher expression of the spliced Rep68/40 proteins (Fig. 2C). To further narrow down the region responsible for these effects the 3' end of the *rep* gene was divided into two parts. The first (y) region contains the p40 promoter and the splice donor site, whereas the second (z) region comprises the sequence encoding the zinc-finger domain (Fig. 2B). Swapping the y region to the respective sequences of the other AAV serotype resulted in little to no differences in VP and Rep expression compared to the constructs with the wild-type AAV1 or AAV2 *rep* gene (Fig. 2C). This observation was surprising given the differential p40 promoter activities (Fig. 2A). In contrast, swapping the zinc-finger domain sequences to the respective other AAV serotype resulted in significant differences of VP and Rep expression. The R2z1V1 construct showed a reduction of VP expression and a lower expression of Rep68/52 compared to Rep78/52 (Fig. 2C). Vice versa, R1z2V1 showed an increase of VP expression and an approximately equal expression of Rep68/52 to Rep78/52, unlike the R1V1 construct. Thus, it is possible that these observations are linked, particularly since the VPs are translated from spliced p40 transcripts and mutations of the splice sites result in a strong reduction of VP expression (31, 32). In the absence of nucleotide differences of the splice site in the AAV1 *rep* gene, the z region, which comprises the majority of the intron sequence, has to be responsible for the inefficient splicing and subsequently lower VP expression.

To analyze whether this effect is caused by the zinc-finger domain of the Rep proteins, stop codon mutations were introduced to the AAV2 or AAV1 *rep* gene at position 522 or 553, which are located prior to the AAV splice donor site or past the splice donor site, respectively (Fig. 3A). Truncating the Rep proteins at the C terminus by 100 or 69 amino acids generated Rep proteins either completely devoid of the zinc-finger domain or with a truncated version of it, respectively. The shifts of the size of the Rep proteins were visualized in Western blots from the lysates of the transfected producer cells (Fig. 3B). While a stop codon in position 522 eliminated the spliced version of Rep78 or Rep52, the stop codon in position 553 allowed the full generation of Rep68 and Rep40, which were indistinguishable on the Western blot since the size is very similar to the truncated Rep78 and Rep52 proteins. In addition, another Rep protein band was detected for the stop codon mutants that ran lower than the truncated Rep52 (Fig. 3B).



**FIG 2** Role of the 3' end of the rep gene in VP expression. (A) Analysis of the p40 promoter activity from AAV1, AAV2, and AAV8 driving luciferase expression in HEK293 cells 24 h posttransfection. The nucleotide positions of the fragments used are shown below the constructs based on the accession numbers: AF063497 (AAV1), AF043303 (AAV2), and AF513852 (AAV8). RLU, relative light units. (B) Schematic depiction of the Rep protein with its main domains. The approximate position of the p40 promoter in the rep gene is indicated. Below, the p5 and p19 transcripts are shown with the translated regions for Rep78, Rep68, Rep52, and Rep40. The C-terminal (c) ~250 amino acids have been subdivided into a N-terminal y region and a C-terminal z region. The constructs generated with all permutations between AAV1 and AAV2 involving these regions are depicted. rep gene fragments derived from AAV1 are colored in purple, and AAV2 are colored in blue. (C) Western blot analysis to determine Rep and VP expression of the constructs. The top blots were probed with MAb B1, and the lower blots were probed with MAb 1F. The individual VPs and Rep proteins are indicated.

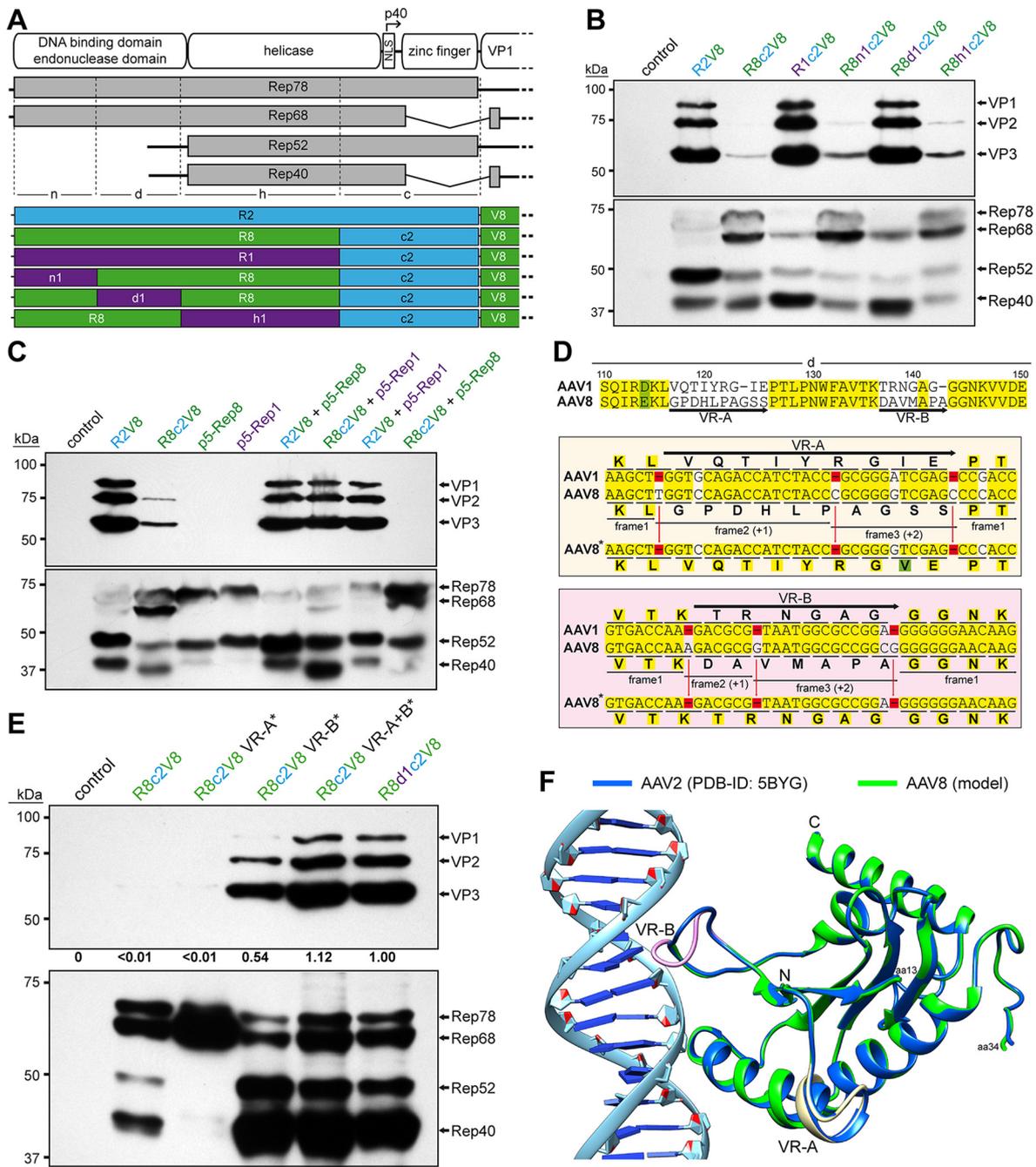


**FIG 3** Role of the 3' end of the *rep* gene in mRNA splicing. (A) Schematic depiction of the p5 and p19 transcripts resulting in the translation of Rep78, Rep68, Rep52, and Rep40. The locations of two introduced stop codons at amino acids 522 and 553 and the primer binding sites utilized for reverse transcription are shown. (B) Analysis of Rep protein and VP expression by Western blotting after transfection of HEK293 cells. The individual VPs and the approximate location of Rep78/52 or their truncated variants are indicated. The top blots were probed with MAb B1, and lower blots were probed with MAb 1F. Below, the yields of benzonase-resistant genomic AAV vectors as determined by qPCR are shown relative to ATG-R2V1 or ATG-R1V1, respectively. (C) Analysis of AAV splicing on RNAs extracted from transfected HEK293 cells with the indicated constructs. RNAs were DNase treated, reverse transcribed to cDNA, and amplified with the primers shown in panel A. The PCR products were run on agarose gels showing ~200- and ~500-bp bands corresponding to the expected sizes for spliced and unspliced mRNAs. The intensities of the bands were quantified using ImageJ and are shown as a ratio relative to the unspliced PCR product.

These bands likely represent truncated Rep40-like proteins that have previously been described for AAV5 (33). In their full form, these proteins likely run similar to Rep40 proteins. In contrast to these changes on the Rep proteins, VP expression and vector yield remained essentially the same (Fig. 3B), suggesting that the zinc-finger domain does not play a role in DNA packaging, reduced VP expression, and might even be dispensable for AAV vector production. Thus, the reduced VP expression is likely caused by the DNA sequence itself affecting mRNA splicing. In order to confirm this

hypothesis, RNA was extracted from the producer cells transfected with the series of 3' swaps of the *rep* genes described above and reverse-transcribed to cDNA, and PCRs were conducted to estimate the ratio of spliced to unspliced mRNAs. The Rep variants with low Rep68/40 and VP expression (Fig. 2C) showed a lower rate of mRNA splicing, with unspliced to spliced ratios of 1:1 to 1:3, whereas the Rep variants with high VP expression showed ratios of 1:5 to 1:12 (Fig. 3C). More research is needed to determine what DNA sequences are responsible for the reduction of splicing and VP expression. In the z region are a total of 64 nucleotide differences between AAV1 and AAV2. Fifty-seven of these nucleotide differences are also shared with AAV6 and AAV8 that showed a similar phenotype. For this study, all subsequent variants contained the 3' end of the AAV2 *rep* gene.

**Rescue of AAV8 VP expression by correction of the AAV8 Rep DNA-binding domain.** The substitution of the 3' end of the *rep* gene to the sequences of AAV2 rescued AAV1 VP expression, as shown above (Fig. 2C), and AAV6 VP expression. Thus, the same strategy was pursued in the case of AAV8 vector production using the AAV8 *rep* gene (Fig. 4A). However, unlike the AAV1 and AAV6 constructs, no significant AAV8 VP expression was observed with the R8c2V8 construct (Fig. 4B), comparable to the R8V8 construct, as seen previously (Fig. 1B). Furthermore, the inability to express VP proteins was not just restricted to AAV8, since AAV1 and AAV6 VP expression is also inhibited when the AAV8 *rep* gene was utilized (data not shown). Vice versa, AAV8 VP expression can be restored when using the AAV1 *rep* gene with the AAV2 3' end (R1c2V8) to a similar level compared to the standard AAV2 *rep* system (Fig. 4B). This observation was surprising given the 95% amino acid sequence identity between AAV1 Rep and AAV8 Rep (Table 1). In order to identify the region of the AAV8 *rep* gene responsible for this phenotype, *rep* was sectioned into three additional parts (n, d, and h) and substituted to the corresponding AAV1 sequences (Fig. 4A). Significant AAV8 VP expression was only achieved when the d region was substituted for the AAV1 sequences (R8d1c2V8, Fig. 4B), which is equivalent to the sequence encoding the DNA-binding domain. Further experiments showed that the lack of efficient AAV8 VP expression with the AAV8 *rep* gene can also be rescued by the addition of an AAV1 Rep expression construct (p5-Rep1) *in trans* (Fig. 4C). Conversely, an AAV8 Rep expression construct (p5-Rep8) does not repress AAV8 VP expression when using the standard AAV2 Rep system or enhance VP expression in the R8c2V8 construct. These observations point to a defect of the AAV8 Rep protein, that is associated with the DNA-binding domain. A sequence analysis of AAV8 to AAV1 Rep (and other AAV serotypes) in the affected region revealed two highly variable regions termed VR-A (residues 117 to 126, with 8 amino acid substitutions and 1 insertion) and VR-B (residues 137 to 143, with 5 amino acid substitutions and 1 insertion) (Fig. 4D). However, when the sequences were analyzed on the DNA sequence level, only 2 nucleotide exchanges and 3 insertions were found for VR-A and 1 nucleotide exchange and 3 insertions for VR-B, respectively. The fact that the individual nucleotide insertions of AAV8 are spread out results in a shift of the reading frame relative to AAV1 and causes the high amino acid variability despite little differences at the DNA level. Since both VRs contain 3 nucleotide insertions the ribosomes shift back into the original reading frame during translation (Fig. 4D). The occurrence of the 3 spaced-out nucleotide insertions is unlikely to happen naturally as they would need to be introduced simultaneously. Otherwise, the individual insertions would shift the reading frame and truncate the Rep proteins significantly. Thus, it is more likely that these insertions represent sequencing errors of the deposited AAV8 *rep* gene sequence. Therefore, the nucleotide sequences of VR-A and VR-B were corrected by removing the nucleotide insertions relative to AAV1, which resulted in a highly similar amino acid sequences with only a single amino acid substitution in VR-A and the identical amino acid sequence in VR-B (Fig. 4D). After transfection of these constructs, the AAV8 VP expression was analyzed. While the correction of VR-A alone did not improve VP expression, the construct with the corrected VR-B showed significant VP expression (Fig. 4E). This expression was further enhanced when both corrected regions were combined, resulting in a similar VP expression compared to the

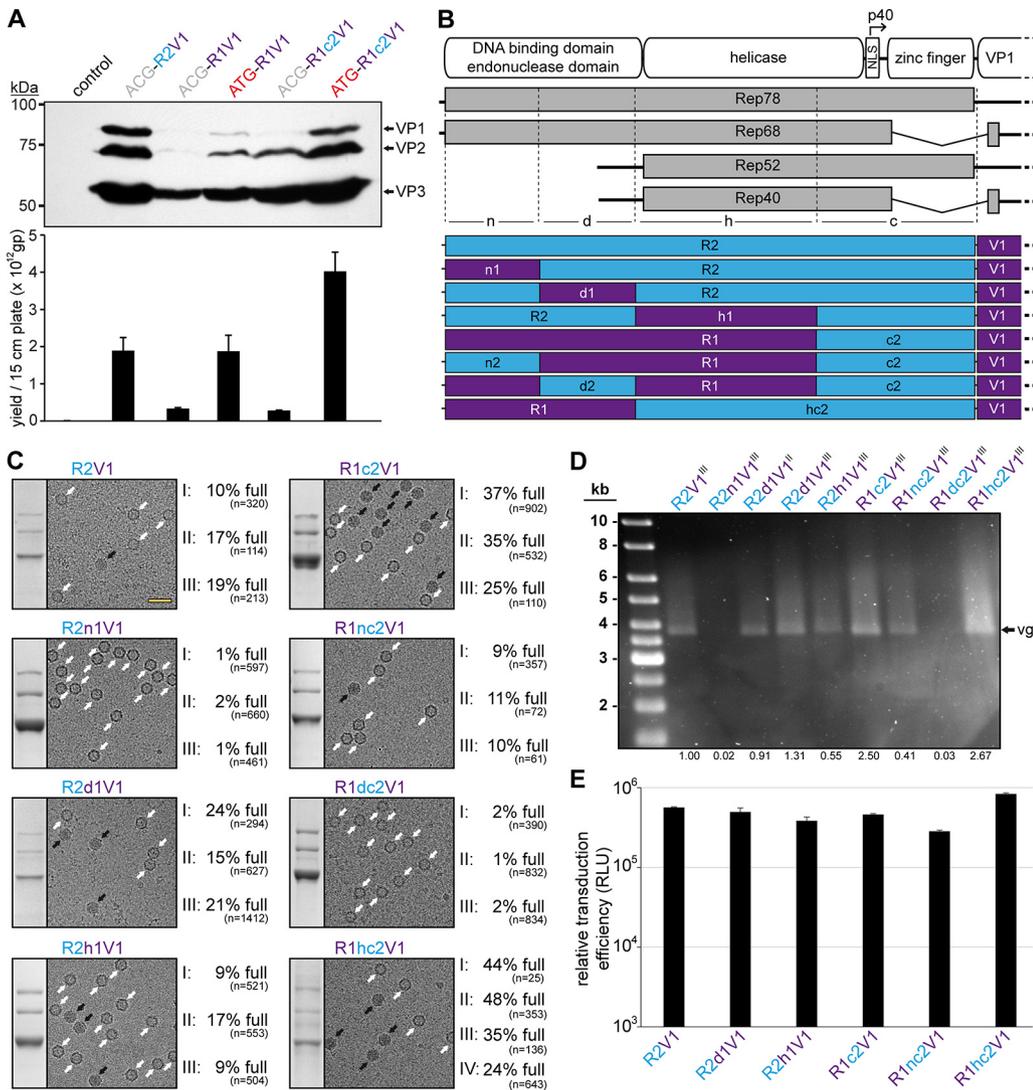


**FIG 4** Rescuing VP expression utilizing the AAV8 *rep* gene. (A) Schematic depiction of the Rep protein with its main domains. The approximate position of the p40 promoter in the *rep* gene is indicated. Below, the p5 and p19 transcripts are shown with the translated regions for Rep78, Rep68, Rep52, and Rep40. Rep has been subdivided into a N-terminal domain (n), the DNA-binding domain (d), the helicase domain (h), and a C-terminal domain (c). The generated constructs containing different domains from AAV1, AAV2, and AAV8 are depicted below. *rep* gene fragments derived from AAV1, AAV2, and AAV8 are colored in purple, blue, and green, respectively. (B) Western blot analysis to determine Rep and VP expression of the constructs shown in panel A. The top blots were probed with MAb B1, and lower blots were probed with MAb 1F. The individual VPs and Rep proteins are indicated. (C) Analysis as in panel B in presence or absence of cotransfected Rep constructs. (D) Amino acid sequence alignment of AAV1 and AAV8 in the DNA-binding domain (amino acids 110 to 150). Identical residues are highlighted in yellow, and amino acids with similar properties are highlighted in green. The two regions of significant amino acid difference, termed VR-A and VR-B, are indicated. Below, an analysis of these regions is depicted at the nucleotide level, with the encoded amino acids shown. Deletions in AAV1 relative to AAV8 are highlighted in red. The insertions in AAV8 potentially result in shifts to an alternative reading frame. Removal of these insertions (AAV8\*) results in an amino acid sequence similar to that for AAV1. (E) Analysis as in panel B utilizing the AAV8 Rep variants with the removed insertions. The intensity of the VP bands was quantified using ImageJ and normalized to R8d1c2V8. (F) Structural analysis of the DNA-binding site for AAV2-Rep (blue) and a superposed AAV8-Rep model (green) to AAVS1 dsDNA. The positions of VR-A and VR-B are indicated.

R8d1c2V8 construct. The rescue of AAV8 VP expression can be explained based on a previously determined X-ray crystallography structure of AAV2 Rep bound to its binding element within the AAVS1 sequence (34). A model generated for the uncorrected AAV8 Rep protein based on the AAV2 Rep structure placed VR-A distance from but VR-B directly in the major groove of the bound DNA probably preventing DNA binding (Fig. 4F). While VR-A does not directly bind to the DNA, it is part of the  $\alpha$ -helix D that contains residues at the N terminus of the helix involved with DNA recognition (34). The correction of VR-A may stabilize the formation of the  $\alpha$ -helix and enhance overall DNA-binding ability when combined with the corrected VR-B and thus further enhance VP expression. Since the Rep proteins have been described as an activator of AAV transcription (35), it is possible that the defective AAV8 Rep proteins are unable to transactivate the p40 promoter (36) needed for efficient VP expression. Another observation made following the correction of the nucleotide sequence of VR-A is the lack of expression of Rep52/40 unlike all the other related constructs, including the completely corrected construct (Fig. 4E). The lack of Rep52/40 expression indicated a loss of p19 promoter activity. Although the VR-A region is located near the equivalent AAV2 p19-promoter sequences, it does not align to the described promoter elements (37). However, previous studies exclusively characterized the AAV2 p19-promoter, and the other AAV serotypes may utilize different promoter elements.

**AAV1/2 Rep hybrids can enhance vector genome packaging efficiency into AAV1 capsids.** Initial analyses of lysates showed that vector genome titers were comparable to the standard AAV2 Rep system (R2V1) when the ATG-R1V1 construct was used despite displaying lower VP expression (Fig. 5A). Following the rescue of the VP expression with ATG-R1c2V1, the genome titer was higher than R2V1 but at a similar VP expression level. However, this enhancement was only observed for the ATG start codon constructs, whereas the ACG constructs showed very low genome titers (Fig. 5A).

Thus, the genome packaging efficiency of the Rep hybrids into AAV1 capsids was further investigated. For this purpose, the *rep* genes of AAV1 and AAV2 were divided into three additional regions (n, d, and h) similar to that described for AAV8 above to identify which region provides a potential benefit for genome packaging (Fig. 5B). For these regions, eight possible permutations between AAV1 and AAV2 were generated (R2, R2n1, R2d1, R2h1, R1c2, R1nc2, R1dc2, and R1hc2) and utilized for AAV1 vector production. Since the packaging efficiency was expected to vary slightly between different vector preparations, three independent vector preparations for each Rep variant were produced. The individual AAV1 vector preparations were purified also independently by AVB-affinity chromatography that indiscriminately binds empty and full capsids (38). Subsequently, each of the AAV1 vector preparations were analyzed by cryo-electron microscopy (cryo-EM), and the percentage of full capsids was determined (Fig. 5C). For the standard AAV2 Rep production system (R2V1), the percentage of full capsids in the three vector preparations ranged from approximately 10 to 20%, with a mean average of 15% (Table 2). In contrast, for AAV1 Rep with the AAV2 C terminus (R1c2V1) the percentage of full capsids in the three vector preparations is approximately doubled and ranged from 25 to 37%, with a mean average of 32%. Swaps of the N termini based on these constructs to the corresponding other AAV serotype was detrimental to both constructs. While the AAV1 N terminus in AAV2 Rep (R2n1V1) resulted in a vast majority of empty capsids (% full, 1 to 2%; average, ~1%), the AAV2 N terminus in AAV1 (R1nc2V1) reduced the percentage of full capsids to 9 to 11%, with a mean average of 10% (Fig. 5C and Table 2). Similarly, the AAV2 DNA-binding domain inserted into AAV1 Rep (R1dc2V1) also resulted in largely empty capsids (% full, 1 to 2%; average, ~2%). These results could indicate that the AAV2 DNA-binding domain is not compatible with the AAV1 N terminus, which are expressed by the R2n1V1 and R1dc2V1 constructs. Vice versa, the AAV1 DNA-binding domain is supported in AAV2 Rep and slightly enhanced packaging efficiency (% full, 15 to 24%; average, ~20%). Swapping the AAV1 helicase domain into AAV2 (R2h1V1) did not affect the packaging efficiency significantly (% full, 9 to 17%; average, ~12%) compared to R2V1. Lastly, the Rep hybrid with the best overall packaging efficiency contained the AAV2 helicase



**FIG 5** AAV1/2 Rep hybrids improve AAV1 genome packaging efficiency. (A) Analysis of VP expression by Western blotting and the AAV1 vector yield by qPCR after transfection of the constructs in HEK293 cells. The Western blot was probed with MAb B1. The individual VPs are indicated. (B) Schematic depiction of the Rep protein with its main domains. The approximate position of the p40 promoter in the *rep* gene is indicated. Below, the p5 and p19 transcripts are shown with the translated regions for Rep78, Rep68, Rep52, and Rep40. Rep has been subdivided into a N-terminal domain (n), the DNA-binding domain (d), the helicase domain (h), and a C-terminal domain (c). The generated constructs containing different domains from AAV1 and AAV2 are depicted below. *rep* gene fragments derived from AAV1 are colored in purple and those from AAV2 are colored in blue. (C) Analysis of the AVB-purified AAV1 vector preparations. Sections of SDS-PAGEs containing VP1, VP2, and VP3 and representative example cryo-EM micrographs are shown for each Rep hybrid. White arrows point to empty capsids (light appearance), and black arrows point to full capsids (dark appearance). The determined percentages of full capsids of at least three independently produced and purified AAV1 vector preparations are displayed with the total particle count of all micrographs collected for the individual sample. Scale bar (shown in R2V1 micrograph), 50 nm. (D) Alkaline gel electrophoresis of the AAV1 vector preparations. The capsid amount loaded is the same for all samples, based on the ELISA titer. The size of the packaged vector genome (vg) is ~3.9 kb. The intensities of the vector genome bands were quantified using ImageJ and normalized to R2V1. (E) Analysis of the transduction efficiency of the AAV1 vectors produced with different Rep hybrids in HEK293 cells.

within AAV1 Rep (R1hc2V1) with a percentage of full capsids in the four vector preparations ranging from 24 to 48%, with a mean average of 38% (Fig. 5D and Table 2). The approximate full-empty ratios for the different Rep hybrids were also confirmed by qPCR compared to the capsid titer determined by ELISA (data not shown). The packaged genomes were also visualized by performing an alkaline gel electrophoresis loading equal amounts of capsids (Fig. 5D). Thus, vector preparations with a higher percentage of full capsids will appear brighter such as for R1hc2V1, whereas no bands are

**TABLE 2** Summary of the quantification of the full capsids for AAV1

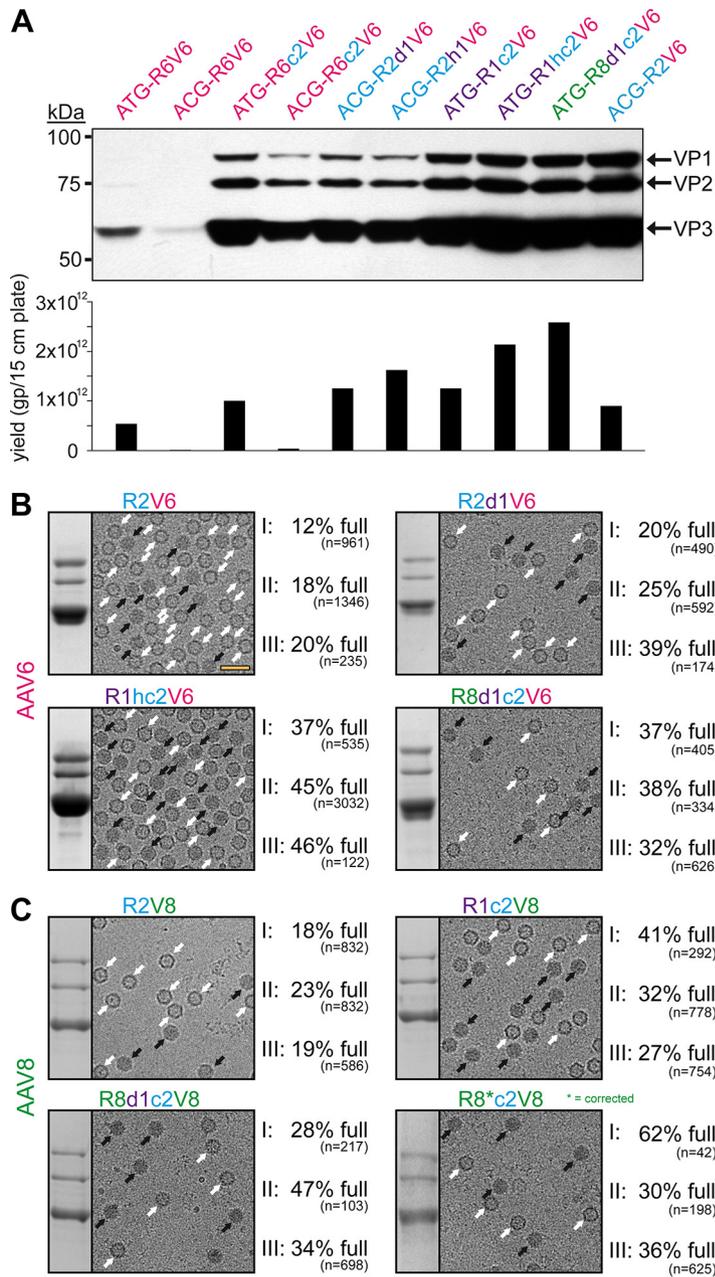
AAV serotype (capsid)	Utilized <i>rep</i> variant	Biological replicate	Percentage of "full" capsids	Mean (%)
AAV1	R2	I	10	15
		II	17	
		III	19	
	R2n1	I	1	1
		II	2	
		III	1	
	R2d1	I	24	20
		II	15	
		III	21	
	R2h1	I	9	12
		II	17	
		III	9	
	R1c2	I	37	32
		II	35	
		III	25	
	R1nc2	I	9	10
		II	11	
		III	10	
	R1dc2	I	2	2
		II	1	
III		2		
R1hc2	I	44	38	
	II	48		
	III	35		
	IV	24		

seen for largely empty capsids (R2n1V1 and R1dc2V1). Regardless of the utilized Rep hybrid during vector production, no significant differences of the transduction efficiency of the purified AAV vectors were observed (Fig. 5E). This was expected since the Rep proteins are not part of the final purified AAV vector preparations.

**The Rep hybrids also enhance vector genome packaging efficiency for other AAV serotypes.** A similar screen for the genome packaging efficiency, as done for AAV1 (Fig. 5A), was conducted for AAV6 utilizing AAV6 hybrids with the AAV2 C terminus. However, the AAV1 and AAV8 *rep* hybrids cloned upstream of the AAV6 *cap* gene appeared to exceed the packaging efficiency over the AAV6 Rep hybrids (Fig. 6A). Thus, the AAV1 and AAV8 Rep hybrids were utilized for AAV6 production and compared to the standard AAV2 Rep system (Fig. 6B). Similar to AAV1, the three independent AAV6 vector preparations were purified by AVB-affinity chromatography, and the percentage of full capsids was determined by cryo-EM imaging. For the standard AAV2 Rep system (R2V6), the percentage of full capsids ranged from 12 to 20% (average, 17%) (Table 3). In contrast, the percentage of full capsids is increased for each of the other tested Rep hybrids ranging from 20 to 39% (average, 28%) for R2d1V6, 37 to 46% (average, 43%) for R1hc2V6, and 32 to 38% (average, 36%) for R8d1c2V6 (Fig. 6B and Table 3). Due to the high sequence similarity of the AAV1 and AAV6 capsids (39), it was not surprising that the Rep hybrids improved packaging efficiency similarly.

For AAV8 vectors produced with AAV2 *rep*, the percentage of full capsids ranged from 18 to 23% (average, 20%) (Fig. 6C and Table 3). The utilization of the Rep hybrids that also rescued AAV8 VP expression (Fig. 4) improved the percentage of genome-containing capsids to 27 to 41% (average, 33%) for R1c2V8, 28 to 47% (average, 36%) for R8d1c2V8, and 30 to 62% (average, 43%) for the corrected R8c2V8. Following the correction of the AAV8 DNA-binding domain (Fig. 4D), the resulting Rep protein of the R8c2V8 construct varies only in two amino acids from the R8d1c2 Rep hybrid. Thus, a similar level of enhancement of genome packaging efficiency was expected.

For AAV9 and AAVrh.10 the *rep* genes or Rep protein sequences have not been isolated or deposited (22). Alternatively, the most promising Rep hybrids for AAV1, AAV2, and



**FIG 6** Rep hybrids enhance packaging efficiency for AAV6 and AAV8. (A) Analysis of VP expression by Western blotting and the AAV6 vector yield by qPCR after transfection of the constructs in HEK293 cells. The Western blot was probed with MAb B1. The individual VPs are indicated. (B) Analysis of the AVB-purified AAV6 vector preparations. Sections of SDS-PAGEs containing VP1, VP2, and VP3 and representative example cryo-EM micrographs are shown for each Rep variant. White arrows point to empty capsids (light appearance), and black arrows point to full capsids (dark appearance). The determined percentage of full capsids of three independently produced and purified AAV6 vector preparations are displayed with the total particle count of all micrographs collected for the individual sample. Scale bar (shown in R2V6 micrograph), 50 nm. (C) Analysis as in panel B for AAV8-Capture select affinity ligand purified AAV8 vectors.

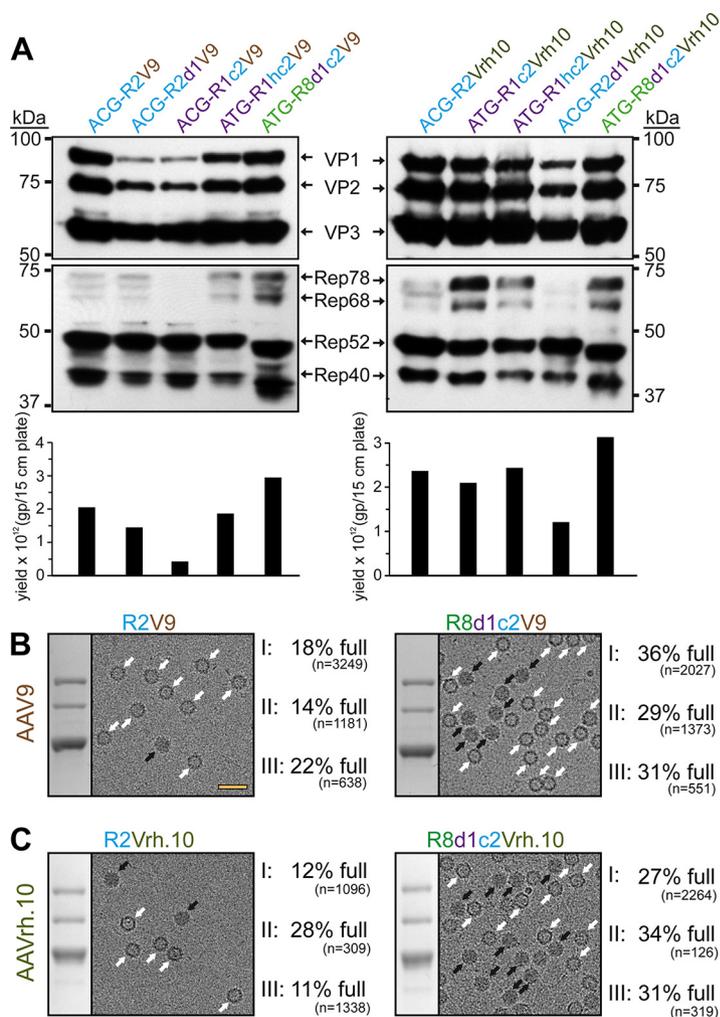
AAV8 were cloned upstream of the AAV9 or AAVrh.10 *cap* gene, and the yield of the constructs by qPCR was compared to the VP expression intensity by Western blotting (Fig. 7A). For both AAVs, the R8d1c2 Rep hybrid came ahead of the standard AAV2 Rep, and three independent AAV preparations were generated with either Rep variant as with the previous AAV serotypes for detailed analysis. For AAV9, the percentage of full capsids ranged from 14 to 22% (average, 18%) with R2V9, whereas genome-containing capsids

**TABLE 3** Summary of the quantification of the full capsids for AAV6, AAV8, AAV9, and AAVrh.10

AAV serotype (capsid)	Utilized <i>rep</i> variant	Biological replicate	Percentage of “full” capsids	Mean (%)
AAV6	R2	I	12	17
		II	18	
		III	20	
	R1hc2	I	37	43
		II	45	
		III	46	
	R2d1	I	20	28
		II	25	
		III	39	
R8d1c2	I	37	36	
	II	38		
	III	32		
AAV8	R2	I	18	20
		II	23	
		III	19	
	R8d1c2	I	28	36
		II	47	
		III	34	
	R1c2	I	41	33
		II	32	
		III	27	
R8*c2	I	62	43	
	II	30		
	III	36		
AAV9	R2	I	18	18
		II	14	
		III	22	
	R8d1c2	I	36	32
		II	29	
		III	31	
AAVrh.10	R2	I	12	17
		II	28	
		III	11	
	R8d1c2	I	27	31
		II	34	
		III	31	

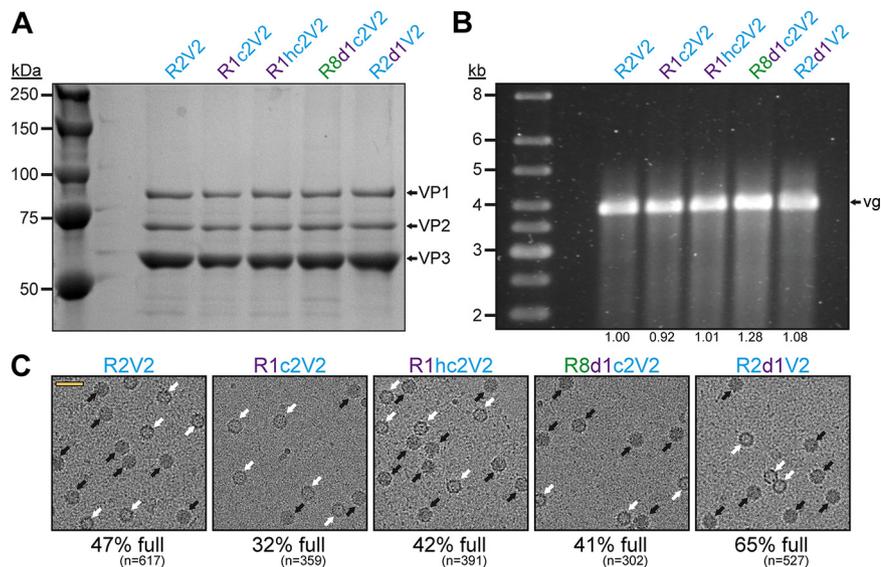
were increased to 29 to 36% (average, 32%) with R8d1c2V9 (Fig. 7B). For AAVrh.10, the results were similar with 11 to 28% (average, 17%) full capsids with R2Vrh.10 and 27 to 34% (average, 31%) full capsids with R8d1c2Vrh.10 (Fig. 7C and Table 3).

These results indicate that the new Rep hybrids may improve packaging regardless of which capsids are utilized since the amino acid sequence identity of these capsids vary from 82 to 99% (Table 1). In order to confirm this conclusion, the effects of the Rep hybrids were also analyzed on AAV2 vector production. The overall yields of capsids with the different Rep hybrids were comparable (Fig. 8A), and the vector genome cassettes were packaged equally (Fig. 8B). As previously mentioned, packaging of vector genomes into AAV2 capsids is generally more efficient (24). This was also observed using cryo-EM imaging with the R2V2 construct generating ~47% of full capsids (Fig. 8C). In contrast, the percentage of full capsids for the other AAV serotypes ranged from 12 to 20% when the AAV2 *rep* gene was used (Fig. 5C, Fig. 6B and C, and Fig. 7B and C). With the different Rep hybrids, the percentage of full AAV2 capsids ranged between 32 and 65% (Fig. 8C), which corresponds to a 0.7- to 1.4-fold decrease or increase relative to the wtAAV2 Rep construct, respectively. Since the portion of genome-containing capsids was high for AAV2 to start with, a further improvement of packaging was



**FIG 7** Rep hybrids also enhance packaging efficiency for AAV9 and AAVrh.10. (A) Analysis of VP and Rep expression by Western blotting and the vector yield by qPCR after transfection of the constructs in HEK293 cells. The Western blots were probed with MAb B1 or MAb 1F, respectively. The individual VPs and Rep proteins are indicated. (B) Analysis of the AAV9-Capture select affinity ligand purified AAV9 vector preparations. Sections of SDS-PAGEs containing VP1, VP2, and VP3 and representative example cryo-EM micrographs are shown for each Rep variant. White arrows point to empty capsids (light appearance), and black arrows point to full capsids (dark appearance). The determined percentage of full capsids of three independently produced and purified AAV9 vector preparations are displayed with the total particle count of all micrographs collected for the individual sample. Scale bar (shown in R2V9 micrograph), 50 nm. (C) Analysis as in panel B for AVB purified AAVrh.10 vectors.

not expected. On the other hand, utilizing the Rep hybrids such as R1c2 or R8d1c2 for AAV2 vector production, the percentage of full AAV2 capsids did not drop significantly either. This questioned the initial hypothesis that the Rep proteins of an AAV serotype might be adapted to its own capsid. Thus, there has to be an alternative explanation for the higher percentage of full capsids when using these Rep hybrids. The Rep hybrids that allowed the best packaging efficiency contained the ~240 N-terminal amino acids of either AAV1 or AAV8 Rep78/68, which contains the DNA-binding and endonuclease domain of the Rep proteins (16). This region is indispensable for genome replication (19, 40, 41). Thus, it is possible that the Rep hybrids replicate the vector genomes to higher copy numbers prior to their packaging. However, the copy number of vector genomes determined by qPCR after lysis of transfected cells without DNase treatment showed similar level of replicated genomes for all Rep variants (data not shown). The small Rep proteins (Rep52/40) have been suggested to be responsible for



**FIG 8** Effect of Rep hybrids on genome encapsidation of AAV2. (A) SDS-PAGE of the AVB-purified AAV2 vector preparations produced with different Rep variants. VP1, VP2, and VP3 are indicated. (B) Alkaline gel electrophoresis of the AAV2 vector preparations. The capsid amount loaded is the same for all samples. The size of the packaged vector genome (vg) is  $\sim 3.9$  kb. The intensities of the vector genome bands were quantified using ImageJ and normalized to R2V2. (C) Representative example cryo-EM micrographs are shown for each Rep variant. White arrows point to empty capsids (light appearance), and black arrows point to full capsids (dark appearance). The determined percentage of full capsids of the vector preparations are displayed with the total particle count of all micrographs collected for the individual sample. Scale bar (shown in R2V2 micrograph), 50 nm.

the encapsidation of the genomes into the capsids (19). However, the small Rep proteins are largely identical between Rep2 and the R1hc2 variant (Fig. 5D) except for a single amino acid substitution (K234R). Thus, it is likely that the large Rep proteins are also involved in the encapsidation process with the Rep hybrids improving genome packaging by a currently unknown mechanism. Finally, nucleotide sequence differences between the AAV serotype *rep* genes may also contribute to the observed increased packaging efficiencies by affecting promoter activities or mRNA splicing, resulting in differential expression ratios of the Rep proteins.

**Conclusions.** The vast majority of AAV production systems utilize AAV2 *rep* for vector manufacturing. The few exceptions reported previously include the utilization of the *rep* genes of AAV3 (42) and AAV4 (23) for production. In both cases, the vector yield was described to be higher, but no further investigation was conducted. In the present study, the *rep* genes of different AAV serotypes were analyzed and, while they were not directly usable “as is,” they were modified to generate hybrids supporting high capsid expression and improved packaging efficiency. The fact that the Rep hybrids enhanced packaging efficiency across multiple, widely used AAV serotypes in gene therapy trials indicates that they can be possibly used for almost all AAV capsids to improve overall vector yield and increase of the percentage of genome-containing capsids. Since the *rep* gene is often provided in *trans* for most AAV production systems, the new hybrid *rep* genes can easily replace the AAV2 *rep* constructs. Furthermore, during AAV vector purification the Rep proteins are removed, making the final product indistinguishable to the AAV2 *rep* system except for the higher content of genome-containing capsids. While the exact mechanism of the enhancement of packaging remains unclear, further research on the Rep proteins is needed in general. It will be also interesting to determine whether the Rep hybrids also improve packaging in insect cells for large scale AAV production setups that often suffer from high empty-to-full capsid ratios (43). Empty capsids are generally not desired in AAV vector preparations since they do not provide any curative benefit for the treatment of the targeted

disease and can reduce overall transduction efficiencies (44). Some purification protocols do not actively remove empty capsid, which could be problematic when administered to a patient with a given genome-containing vector dose because the empty particle titer can be 5- to 10-fold higher. These empty capsids can then potentially elicit additional immune responses *in vivo* gene therapies.

## MATERIALS AND METHODS

**Plasmids and cloning.** The *rep* genes of AAV1, AAV6, and AAV8 were synthesized by GenArt (Thermo Fisher) based on the deposited AAV serotype genomes; accession numbers: [AF063497](#) (AAV1), [AF028704](#) (AAV6), and [AF513852](#) (AAV8). These *rep* genes were inserted into plasmids containing the AAV2 *rep* gene ([AF043303](#)) and either the AAV1, -6, or -8 *cap* gene (R2V1, R2V6, or R2V8) by replacing the AAV2 *rep* gene to generate R1V1, R6V6, or R8V8, respectively. Site-directed mutagenesis was utilized by as previously described (45) to mutate the Rep78 start codon or to correct the AAV8 *rep* gene. In order to generate Rep hybrid proteins, the *rep* gene was subdivided into four regions and cloned using restriction sites that are conserved in equivalent positions of the AAV1, -2, -6, and -8 *rep* genes. Briefly, the N-terminal region (n) extends to NcoI (nt position 305 of the AAV1 *rep* gene, amino acid position 102 of AAV1 Rep), the DNA-binding domain (d) to BamHI (nt position 725, amino acid position 242), the helicase domain (h) to Sall (nt position 1108, amino acid position 370), and the C-terminal region consists of all the sequences past Sall. In the absence of usable restriction sites, Gibson assembly was performed using the NEB Gibson assembly master mix according to the manufacturer's instructions to further subdivide the C-terminal region at nt position 1590, amino acid position 530 (AAV1 numbering), termed the y and z regions. To determine the p40 promoter activity the nt 1442 to 1878 (AAV1, Sall-KpnI fragment), nt 1428 to 1887 (AAV2, Sall-HindIII fragment), and nt 1340 to 1776 (AAV8, Sall-KpnI fragment) from the accession numbers listed above were cloned upstream of a luciferase gene. The p40 promoter regions were selected very generously to contain all previously described p40 promoter elements (37, 46). Site-directed mutagenesis PCRs for the introduction of the stop codons to the 3' ends of the AAV1 and AAV2 *rep* genes and for the correction of the AAV8 *rep* gene were done as previously described (45).

**Cell culture.** HEK293 cells were maintained in Dulbecco modified Eagle medium supplemented with 10% heat-inactivated fetal calf serum with 100 U of penicillin/ml and 100  $\mu$ g of streptomycin at 37°C in 5% CO<sub>2</sub>.

**AAV production and purification.** Recombinant AAV vectors, with a packaged luciferase gene, were produced by triple transfection of HEK293 cells, utilizing pTR-UF3-Luciferase, pHelper (Stratagene), and a *rep-cap* plasmid containing either wild-type or hybrid *rep* genes. The transfected cells were harvested 72 h posttransfection as previously described (45). The cleared lysates containing AAV capsids were purified by using AVB Sepharose (Thermo Fisher) in the case of AAV1, -2, -6, and -rh.10 and by using POROS CaptureSelect AAV8 or AAV9 (Thermo Fisher) affinity chromatography in the case of AAV8 or -9, as previously described (38).

**SDS-PAGE and Western blot analysis.** The purity of the AAV preparations were confirmed by sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). For this purpose, the samples were incubated with 6 $\times$  Laemmle sample buffer (Bio-Rad) with 10%  $\beta$ -mercaptoethanol and boiled for 5 min at 100°C. The denatured proteins were applied to a 10% polyacrylamide gel and run at 120 V. The gel was washed three times with distilled water (dH<sub>2</sub>O) and stained with GelCode Blue Protein Safe stain (Invitrogen). In order to confirm and evaluate Rep and Cap expression from the new Rep hybrid plasmids, Western blot analyses were performed. For this purpose, the proteins were transferred to a nitrocellulose membrane following SDS-PAGE by electroblotting. The membrane was blocked in 6% milk in 1 $\times$  phosphate-buffered saline and probed with hybridoma supernatants containing either monoclonal antibody (MAb) B1, detecting VP1, VP2, and VP3 (47), or MAb 1F, detecting Rep78, Rep68, Rep52, and Rep40 (30). After incubation with a secondary antibody with a linked horseradish peroxidase, the proteins were visualized by applying Immobilon chemiluminescent substrate (Millipore) and detection on an X-ray film.

**RNA splicing analysis.** HEK293 cells were transfected as described above and harvested at 48 h posttransfection. The RNA was extracted utilizing a PureLink RNA minikit (Thermo Fisher) according to the kit's manual, and the RNA concentration was quantified. Then, 10  $\mu$ g of each sample was treated with a Turbo DNA-free kit (Thermo Fisher) prior to reverse transcription using ProtoScript II reverse transcriptase (NEB) with an AAV1 *cap* gene-specific primer (5'-AGCCATGTGGAATCGAATG-3'). The generated cDNA was amplified using the PfuUltra high-fidelity DNA polymerase AD (Agilent) with the following primer pair: p40-Fwd (5'-ATCGACGTGACGCGGAAG-3') and cap1-Rev (5'-CCTGTGTCGAGTCCGTTGAAG-3'). An elongation time of 30 s allowed the amplification of the spliced transcripts (201 or 228 bp) and unspliced transcripts (522 bp). The PCR products were run on agarose gels and imaged after SybrSafe staining.

**Quantification of AAV vectors.** Aliquots from the AAV vector preparations were digested with proteinase K to release the AAV vector genomes from the capsids. To this end, the samples were incubated in buffer containing 10 mM Tris (pH 8), 10 mM EDTA, and 1% SDS for 2 h at 56°C. The released DNA was purified utilizing a PureLink PCR purification kit (Thermo Fisher). The copy numbers of vector genome DNAs were determined by quantitative PCR using iQ SYBR Green Supermix (Bio-Rad, Hercules, CA). Primers specific for the luciferase gene of the vector genome were used (forward primer, 5'-GCAAAACGCTCCATCTTCC-3'; reverse primer, 5'-AGATCCACAACCTTCGTTCC-3').

**AAV capsid ELISA.** For the quantification of the physical capsid titer, AAV titration ELISA (Progen) was utilized. All the steps were performed in triplicate according to the protocol provided by the manufacturer. The colorimetric assay was analyzed by using a Synergy HT plate reader (BioTek).

**Alkaline gel electrophoresis.** For the alkaline gel electrophoresis, a 0.8% agarose gel in 1× TAE buffer (40 mM Tris, 20 mM acetic acid, 1 mM EDTA) was utilized. After solidification, the gel was equilibrated in 1× denaturing buffer (0.5 M NaOH, 50 mM EDTA) for 4 h. Prior to the loading, the samples were mixed with denaturing loading dye (final concentration: 1× Ficoll loading buffer, 1× denaturing buffer, 10% SDS). The agarose gel was run at low voltage overnight at 4°C. After the run, the gel was washed and neutralized in 1× TAE for 30 min and subsequently stained in a 0.02% SyBr-Gold solution in 1× TAE. The gel was imaged under UV light using a Bio-Rad GelDoc system.

**Analysis of transduction efficiency and promoter activity.** Purified AAV vectors with a packaged luciferase expression cassette were used to infect HEK293 cells at an multiplicity of infection (MOI) of 10<sup>5</sup>. After 48 h, the cells were lysed, and the luciferase activity was assayed using a luciferase assay kit (Promega) according to the manufacturer's protocol. Luminescence was measured on a Synergy HT plate reader (BioTek). In order to determine the p40 promoter activity in HEK293 cells, the AAV-p40-luciferase constructs were transfected, and the cells were incubated for 24 h. Subsequent steps were performed exactly as described above for the analysis of transduction efficiency.

**Cryo-electron microscopy imaging.** For each of the purified AAV capsids, 3.5 μl was applied to a glow-discharged Quantifoil copper grid with 2-nm continuous carbon support over holes (Quantifoil R 2/4 400 mesh), blotted, and vitrified using a Vitrobot Mark 4 (FEI) at 95% humidity and 4°C. Images were collected using an FEI Tecnai G2 F20-TWIN microscope (FEI) operated under low-dose conditions (200 kV, ~20 e<sup>-</sup>/Å<sup>2</sup>) on a GatanUltraScan 4000 CCD camera (Gatan). The numbers of empty and full capsids in these images were counted manually.

## ACKNOWLEDGMENTS

We thank the UF-ICBR Electron microscopy core for access to electron microscopes utilized for cryo-electron micrograph screening.

The University of Florida Research Foundation has filed and licensed patent applications based on the findings presented here.

## REFERENCES

- Rittié L, Athanopoulos T, Calero-Garcia M, Davies ML, Dow DJ, Howe SJ, Morrison A, Ricciardelli I, Saudemont A, Jespers L, Clay TM. 2019. The landscape of early clinical gene therapies outside of oncology. *Mol Ther* 27:1706–1717. <https://doi.org/10.1016/j.ymthe.2019.09.002>.
- Scott LJ. 2015. Alipogene tiparvovoc: a review of its use in adults with familial lipoprotein lipase deficiency. *Drugs* 75:175–182. <https://doi.org/10.1007/s40265-014-0339-9>.
- Lloyd A, Piglowska N, Ciulla T, Pitluck S, Johnson S, Buessing M, O'Connell T. 2019. Estimation of impact of RPE65-mediated inherited retinal disease on quality of life and the potential benefits of gene therapy. *Br J Ophthalmol* 103:1610–1614. <https://doi.org/10.1136/bjophthalmol-2018-313089>.
- Waldrop MA, Kolb SJ. 2019. Current treatment options in neurology-SMA therapeutics. *Curr Treat Options Neurol* 21:25. <https://doi.org/10.1007/s11940-019-0568-z>.
- Mietzsch M, Penzes JJ, Agbandje-McKenna M. 2019. Twenty-five years of structural parvovirology. *Viruses* 11:362. <https://doi.org/10.3390/v11040362>.
- Srivastava A, Lusby EW, Berns KI. 1983. Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J Virol* 45:555–564. <https://doi.org/10.1128/JVI.45.2.555-564.1983>.
- Bennett A, Mietzsch M, Agbandje-McKenna M. 2017. Understanding capsid assembly and genome packaging for adeno-associated viruses. *Future Virology* 12:283–297. <https://doi.org/10.2217/fvl-2017-0011>.
- Weitzman MD, Linden RM. 2011. Adeno-associated virus biology. *Methods Mol Biol* 807:1–23. [https://doi.org/10.1007/978-1-61779-370-7\\_1](https://doi.org/10.1007/978-1-61779-370-7_1).
- Sonntag F, Schmidt K, Kleinschmidt JA. 2010. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc Natl Acad Sci U S A* 107:10220–10225. <https://doi.org/10.1073/pnas.1001673107>.
- Ogden PJ, Kelsic ED, Sinai S, Church GM. 2019. Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science* 366:1139–1143. <https://doi.org/10.1126/science.aaw2900>.
- Gray JT, Zolotukhin S. 2011. Design and construction of functional AAV vectors. *Methods Mol Biol* 807:25–46. [https://doi.org/10.1007/978-1-61779-370-7\\_2](https://doi.org/10.1007/978-1-61779-370-7_2).
- Grieger JC, Samulski RJ. 2012. Adeno-associated virus vectorology, manufacturing, and clinical applications. *Methods Enzymol* 507:229–254. <https://doi.org/10.1016/B978-0-12-386509-0.00012-0>.
- Qiu J, Soderlund-Venermo M, Young NS. 2017. Human parvoviruses. *Clin Microbiol Rev* 30:43–113. <https://doi.org/10.1128/CMR.00040-16>.
- Smith RH, Kotin RM. 1998. The Rep52 gene product of adeno-associated virus is a DNA helicase with 3'-to-5' polarity. *J Virol* 72:4874–4881. <https://doi.org/10.1128/JVI.72.6.4874-4881.1998>.
- Cassell GD, Weitzman MD. 2004. Characterization of a nuclear localization signal in the C terminus of the adeno-associated virus Rep68/78 proteins. *Virology* 327:206–214. <https://doi.org/10.1016/j.virol.2004.06.034>.
- Im D-S, Muzyczka N. 1990. The AAV origin-binding protein Rep68 is an ATP-dependent site-specific endonuclease with helicase activity. *Cell* 61:447–457. [https://doi.org/10.1016/0092-8674\(90\)90526-k](https://doi.org/10.1016/0092-8674(90)90526-k).
- Di Pasquale G, Stacey SN. 1998. Adeno-associated virus Rep78 protein interacts with protein kinase A and its homolog PRKX and inhibits CREB-dependent transcriptional activation. *J Virol* 72:7916–7925. <https://doi.org/10.1128/JVI.72.10.7916-7925.1998>.
- Snyder RO, Im DS, Ni T, Xiao X, Samulski RJ, Muzyczka N. 1993. Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. *J Virol* 67:6096–6104. <https://doi.org/10.1128/JVI.67.10.6096-6104.1993>.
- King JA, Dubielzig R, Grimm D, Kleinschmidt JA. 2001. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J* 20:3282–3291. <https://doi.org/10.1093/emboj/20.12.3282>.
- Wörner TP, Bennett A, Habka S, Snijder J, Friese O, Powers T, Agbandje-McKenna M, Heck AJR. 2021. Adeno-associated virus capsid assembly is divergent and stochastic. *Nat Commun* 12:1642. <https://doi.org/10.1038/s41467-021-21935-5>.
- Yoon-Roberts M, Blouin AG, Bleker S, Kleinschmidt JA, Aggarwal AK, Escalante CR, Linden RM. 2004. Residues within the B' motif are critical for DNA binding by the superfamily 3 helicase Rep40 of adeno-associated virus type 2. *J Biol Chem* 279:50472–50481. <https://doi.org/10.1074/jbc.M403900200>.
- Gao G, Vandenbergh LH, Alvira MR, Lu Y, Calcedo R, Zhou X, Wilson JM. 2004. Clades of Adeno-associated viruses are widely disseminated in human tissues. *J Virol* 78:6381–6388. <https://doi.org/10.1128/JVI.78.12.6381-6388.2004>.
- Grimm D, Kay MA, Kleinschmidt JA. 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. [https://doi.org/10.1016/S1525-0016\(03\)00095-9](https://doi.org/10.1016/S1525-0016(03)00095-9).

24. Nass SA, Mattingly MA, Woodcock DA, Burnham BL, Ardinger JA, Osmond SE, Frederick AM, Scaria A, Cheng SH, O'Riordan CR. 2018. Universal method for the purification of recombinant AAV vectors of differing serotypes. *Mol Ther Methods Clin Dev* 9:33–46. <https://doi.org/10.1016/j.omtm.2017.12.004>.
25. Li C, Samulski RJ. 2020. Engineering adeno-associated virus vectors for gene therapy. *Nat Rev Genet* 21:255–272. <https://doi.org/10.1038/s41576-019-0205-4>.
26. Viney L, Bürckstümmer T, Eddington C, Mietzsch M, Choudhry M, Henley T, Agbandje-McKenna M. 2021. Adeno-associated virus (AAV) capsid chimeras with enhanced infectivity reveal a core element in the AAV genome critical for both cell transduction and capsid assembly. *J Virol* 95:e02023–20. <https://doi.org/10.1128/JVI.02023-20>.
27. Havlik LP, Simon KE, Smith JK, Klinc KA, Tse LV, Oh DK, Fanous MM, Meganck RM, Mietzsch M, Kleinschmidt J, Agbandje-McKenna M, Asokan A. 2020. Co-evolution of AAV capsid antigenicity and tropism through a structure-guided approach. *J Virol* 94:e00976–20. <https://doi.org/10.1128/JVI.00976-20>.
28. Tse LV, Klinc KA, Madigan VJ, Castellanos Rivera RM, Wells LF, Havlik LP, Smith JK, Agbandje-McKenna M, Asokan A. 2017. Structure-guided evolution of antigenically distinct adeno-associated virus variants for immune evasion. *Proc Natl Acad Sci U S A* 114:E4812–E4821. <https://doi.org/10.1073/pnas.1704766114>.
29. Xiao X, Li J, Samulski RJ. 1998. Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J Virol* 72:2224–2232. <https://doi.org/10.1128/JVI.72.3.2224-2232.1998>.
30. Li J, Samulski RJ, Xiao X. 1997. Role for highly regulated rep gene expression in adeno-associated virus vector production. *J Virol* 71:5236–5243. <https://doi.org/10.1128/JVI.71.7.5236-5243.1997>.
31. Farris KD, Pintel DJ. 2008. Improved splicing of adeno-associated viral (AAV) capsid protein-supplying pre-mRNAs leads to increased recombinant AAV vector production. *Hum Gene Ther* 19:1421–1427. <https://doi.org/10.1089/hum.2008.118>.
32. Stutika C, Gogol-Doring A, Botschen L, Mietzsch M, Weger S, Feldkamp M, Chen W, Heilbronn R. 2016. A comprehensive RNA sequencing analysis of the adeno-associated virus (AAV) Type 2 transcriptome reveals novel AAV transcripts, splice variants, and derived proteins. *J Virol* 90:1278–1289. <https://doi.org/10.1128/JVI.02750-15>.
33. Farris KD, Pintel DJ. 2010. Adeno-associated virus type 5 utilizes alternative translation initiation to encode a small Rep40-like protein. *J Virol* 84:1193–1197. <https://doi.org/10.1128/JVI.01961-09>.
34. Musayev FN, Zarate-Perez F, Bishop C, Burgner JW, II, Escalante CR. 2015. Structural insights into the assembly of the adeno-associated virus type 2 Rep68 protein on the integration site AAVS1. *J Biol Chem* 290:27487–27499. <https://doi.org/10.1074/jbc.M115.669960>.
35. Pereira DJ, McCarty DM, Muzyczka N. 1997. The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. *J Virol* 71:1079–1088. <https://doi.org/10.1128/JVI.71.2.1079-1088.1997>.
36. Pereira DJ, Muzyczka N. 1997. The adeno-associated virus type 2 p40 promoter requires a proximal Sp1 interaction and a p19 CArG-like element to facilitate Rep transactivation. *J Virol* 71:4300–4309. <https://doi.org/10.1128/JVI.71.6.4300-4309.1997>.
37. McCarty DM, Christensen M, Muzyczka N. 1991. Sequences required for coordinate induction of adeno-associated virus p19 and p40 promoters by Rep protein. *J Virol* 65:2936–2945. <https://doi.org/10.1128/JVI.65.6.2936-2945.1991>.
38. Mietzsch M, Smith JK, Yu JC, Banala V, Emmanuel SN, Jose A, Chipman P, Bhattacharya N, McKenna R, Agbandje-McKenna M. 2020. Characterization of AAV-specific affinity ligands: consequences for vector purification and development strategies. *Mol Ther Methods Clin Dev* 19:362–373. <https://doi.org/10.1016/j.omtm.2020.10.001>.
39. Ng R, Govindasamy L, Gurda BL, McKenna R, Kozyreva OG, Samulski RJ, Parent KN, Baker TS, Agbandje-McKenna M. 2010. Structural characterization of the dual glycan binding adeno-associated virus serotype 6. *J Virol* 84:12945–12957. <https://doi.org/10.1128/JVI.01235-10>.
40. Im DS, Muzyczka N. 1989. Factors that bind to adeno-associated virus terminal repeats. *J Virol* 63:3095–3104. <https://doi.org/10.1128/JVI.63.7.3095-3104.1989>.
41. Alex M, Weger S, Mietzsch M, Slanina H, Cathomen T, Heilbronn R. 2012. DNA-binding activity of adeno-associated virus Rep is required for inverted terminal repeat-dependent complex formation with herpes simplex virus ICP8. *J Virol* 86:2859–2863. <https://doi.org/10.1128/JVI.06364-11>.
42. Ling C, Yin Z, Li J, Zhang D, Aslanidi G, Srivastava A. 2016. Strategies to generate high-titer, high-potency recombinant AAV3 serotype vectors. *Mol Ther Methods Clin Dev* 3:16029. <https://doi.org/10.1038/mtm.2016.29>.
43. Wang L, Blouin V, Brument N, Bello-Roufai M, Francois A. 2011. Production and purification of recombinant adeno-associated vectors. *Methods Mol Biol* 807:361–404. [https://doi.org/10.1007/978-1-61779-370-7\\_16](https://doi.org/10.1007/978-1-61779-370-7_16).
44. Gao K, Li M, Zhong L, Su Q, Li J, Li S, He R, Zhang Y, Hendricks G, Wang J, Gao G. 2014. Empty virions in AAV8 vector preparations reduce transduction efficiency and may cause total viral particle dose-limiting side-effects. *Mol Ther Methods Clin Dev* 1:20139. <https://doi.org/10.1038/mtm.2013.9>.
45. Jose A, Mietzsch M, Smith JK, Kurian J, Chipman P, McKenna R, Chiorini J, Agbandje-McKenna M. 2019. High-resolution structural characterization of a new adeno-associated virus serotype 5 antibody epitope toward engineering antibody-resistant recombinant gene delivery vectors. *J Virol* 93:e01394–18. <https://doi.org/10.1128/JVI.01394-18>.
46. Weger S, Hammer E, Gonsior M, Stutika C, Heilbronn R. 2016. A regulatory element near the 3' end of the adeno-associated virus *rep* gene inhibits adenovirus replication in *cis* by means of p40 promoter-associated short transcripts. *J Virol* 90:3981–3993. <https://doi.org/10.1128/JVI.03120-15>.
47. Wobus CE, Hugle-Dorr B, Girod A, Petersen G, Hallek M, Kleinschmidt JA. 2000. Monoclonal antibodies against the adeno-associated virus type 2 (AAV-2) capsid: epitope mapping and identification of capsid domains involved in AAV-2-cell interaction and neutralization of AAV-2 infection. *J Virol* 74:9281–9293. <https://doi.org/10.1128/jvi.74.19.9281-9293.2000>.