

# A robust and flexible baculovirus-insect cell system for AAV vector production with improved yield, capsid ratios and potency

Yoko Marwidi,<sup>1</sup> Hoang-Oanh B. Nguyen,<sup>1</sup> David Santos,<sup>1</sup> Tenzin Wangzor,<sup>1</sup> Sumita Bhardwaj,<sup>1</sup> Gabriel Ernie,<sup>1</sup> Gregg Prawdzik,<sup>1</sup> Garrett Lew,<sup>1</sup> David Shivak,<sup>1</sup> Michael Trias,<sup>1</sup> Jada Padilla,<sup>1</sup> Hung Tran,<sup>1</sup> Kathleen Meyer,<sup>1</sup> Richard Surosky,<sup>1</sup> and Alex Michael Ward<sup>1</sup>

<sup>1</sup>Sangamo Therapeutics, 501 Canal Boulevard, Richmond, CA 94804, USA

**Manufacturing of adeno-associated viruses (AAV) for gene and cell therapy applications has increased significantly and spurred development of improved mammalian and insect cell-based production systems. We developed a baculovirus-based insect cell production system—the SGMO Helper—with a novel gene architecture and greater flexibility to modulate the expression level and content of individual Rep and Cap proteins. In addition, we incorporated modifications to the AAV6 capsid sequence that improves yield, capsid integrity, and potency. Production of recombinant AAV 6 (rAAV6) using the SGMO Helper had improved yields compared to the Bac-RepCap helper from the Kotin lab. SGMO Helper-derived rAAV6 is resistant to a previously described proteolytic cleavage unique to baculovirus-insect cell production systems and has improved capsid ratios and potency, *in vitro* and *in vivo*, compared with rAAV6 produced using Bac-RepCap. Next-generation sequencing sequence analysis demonstrated that the SGMO Helper is stable over six serial passages and rAAV6 capsids contain comparable amounts of non-vector genome DNA as rAAV6 produced using Bac-RepCap. AAV production using the SGMO Helper is scalable using bioreactors and has improved yield, capsid ratio, and *in vitro* potency. Our studies demonstrate that the SGMO Helper is an improved platform for AAV manufacturing to enable delivery of cutting-edge gene and cell therapies.**

## INTRODUCTION

Adeno-associated virus (AAV) is a member of the family Parvoviridae, genus *Dependoparvovirus*, and is composed of a non-enveloped capsid containing an approximately 4.7 kb single-stranded DNA genome.<sup>1</sup> AAV is a promising delivery modality for gene and cell therapies, as demonstrated by the 2017 Food and Drug Administration approval of Luxturna, to treat retinal dystrophy, and the 2019 approval for Zolgensma, to treat spinal muscular atrophy (<https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/approved-cellular-and-gene-therapy-products>). AAV is a non-pathogenic, replication-dependent virus with low cytotoxicity and the ability to infect both dividing and non-dividing cells,

thus avoiding some of the challenges associated with retroviral or lentiviral delivery.<sup>2</sup>

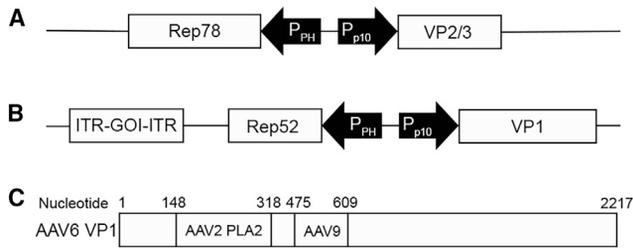
AAV requires helper functions to replicate, typically provided by either adenovirus coinfection or a subset of adenovirus genes, though AAV replication can also be supported by Herpesviruses, Papillomaviruses, and Bocaviruses.<sup>3</sup> AAV expresses two major sets of proteins, the enzymatic Rep proteins and the structural Cap proteins. Wild-type AAV uses a combination of internal promoters, RNA splicing, and suboptimal start codons that allow leaky ribosome scanning to control expression of the Rep (Rep78, Rep68, Rep52, and Rep40) and Cap proteins (VP1, VP2, and VP3).<sup>1</sup> In addition, the *Rep* and *Cap* genes contain overlapping, in-frame sequences, where the *Rep52* sequence is contained within the *Rep78* sequence and the *VP2* and *VP3* sequences are contained within the *VP1* sequence.<sup>1</sup> In the process of developing the HEK-293 production system, it was determined that Rep78 and Rep52 are sufficient to generate recombinant AAV (rAAV) and that AAV can be pseudotyped using Rep proteins and vector genomes containing the inverted terminal repeats (ITRs) from the same serotype (i.e., AAV2 Rep and AAV2 ITRs).<sup>1</sup> The mammalian HEK-293 production system is one of the most widely used technologies for the generation of rAAV at laboratory scale and has been optimized for transfection of HEK-293 cells with plasmids encoding the AAV Rep and Cap proteins, the vector genome of interest, and adenoviral helper genes.<sup>3</sup> The need for scaled-up production for larger clinical trials and commercialization of rAAV products led to the development of baculovirus-insect (Sf9) cell systems for rAAV production (reviewed in Joshi et al.<sup>4</sup>). The Bac-RepCap system developed by the Kotin lab contains the AAV2 Rep (*Rep78* and *Rep52*) gene and the AAV Cap gene (*VP1*, *VP2* and *VP3*) on a single baculovirus.<sup>5</sup> The Rep protein is expressed from a single mRNA, where the *Rep78* open reading frame (ORF) contains a suboptimal start codon that allows leaky ribosomal scanning

Received 9 February 2024; accepted 28 February 2024;  
<https://doi.org/10.1016/j.omtm.2024.101228>.

**Correspondence:** Alex Michael Ward, PhD, Sangamo Therapeutics, 501 Canal Boulevard, Richmond, CA 94804, USA.

**E-mail:** [award@sangamo.com](mailto:award@sangamo.com)





**Figure 1. Schematic of Rep/Cap helpers and VP1 sequence modifications** (A) The Rep78 and VP2/3 ORFs were cloned into the pFastBacDual vector under the control of the polyhedrin ( $P_{PH}$ ) or p10 ( $P_{p10}$ ) promoters, respectively. (B) The Rep52 and VP1 ORFs were cloned into the pFastBacDual vector under the control of the polyhedrin ( $P_{PH}$ ) or p10 ( $P_{p10}$ ) promoters, respectively. The ITR-containing GOI was cloned adjacent to the Rep52 ORF. (C) The AAV6 VP1 sequence was modified to contain the AAV2 phospholipase A2 (PLA2) domain and a short fragment of VP1 from AAV9 at the indicated nucleotide coordinates.

through to the strong *Rep52* start codon. The Cap proteins are also expressed from a single mRNA, where the *VP1* and *VP2* ORFs contain suboptimal start codons, allowing leaky ribosomal scanning through to the strong *VP3* start codon. To produce rAAV, an additional baculovirus containing the vector genome of interest is used to coinfect Sf9 cells with Bac-RepCap and the resulting rAAV is purified away from the infected cells and baculovirus.

While the baculovirus-Sf9 cell system has been successfully utilized for AAV production at multiple scales, it was observed that rAAV generated in insect cells has reduced VP1 content (VP1:VP2:VP3 ratio is approximately 1:1:30 to 1:1:60 compared with 1:1:10 in wild-type AAV) as well as reduced potency compared with rAAV generated in mammalian cells.<sup>6–11</sup> In addition, a proteolytic cleavage site was identified in AAV8 VP1 that is unique to the Sf9-baculovirus cell system and appears to be the result of cleavage by the baculovirus-encoded v-Cathepsin gene, resulting in the incorporation of an aberrant capsid protein into the AAV capsid.<sup>12</sup> A similar cleavage has been observed with AAV1 and AAV6 produced using the Sf9-baculovirus cell system, demonstrating that multiple serotypes are susceptible to cleavage by v-Cathepsin (A.M.W., unpublished observation). Since VP1 is required for AAV infectivity, we hypothesized that increasing incorporation of VP1 into capsids and removing the proteolytic cleavage site to preserve full-length VP1 would lead to increased rAAV6 potency.<sup>13–17</sup>

In order to increase the flexibility and utility of the helper system, we generated a two-baculovirus system, which we named “the SGMO Helper.” This includes a *Rep* gene from serotype 2 and a *Cap* gene derived from serotype 6, where *Rep78* and *VP2/3* are incorporated into the first baculovirus; and *Rep52*, *VP1*, and the vector genome of interest are incorporated into the second baculovirus. In addition to separating the *Rep* and *Cap* genes, we incorporated improvements that (1) increased VP1 incorporation compared with Bac-RepCap-derived rAAV6, (2) prevented the proteolytic cleavage observed with Bac-RepCap-derived rAAV6, and (3) increased yield of

rAAV6 derived from the helper system compared with Bac-RepCap. We demonstrate that the SGMO Helper system is stable over multiple passages, and rAAV6 generated using the SGMO Helper has increased potency *in vitro* and *in vivo* compared with rAAV6 generated using Bac-RepCap. The results presented here show that the SGMO Helper is an improved system for scalable production of rAAV in Sf9 cells.

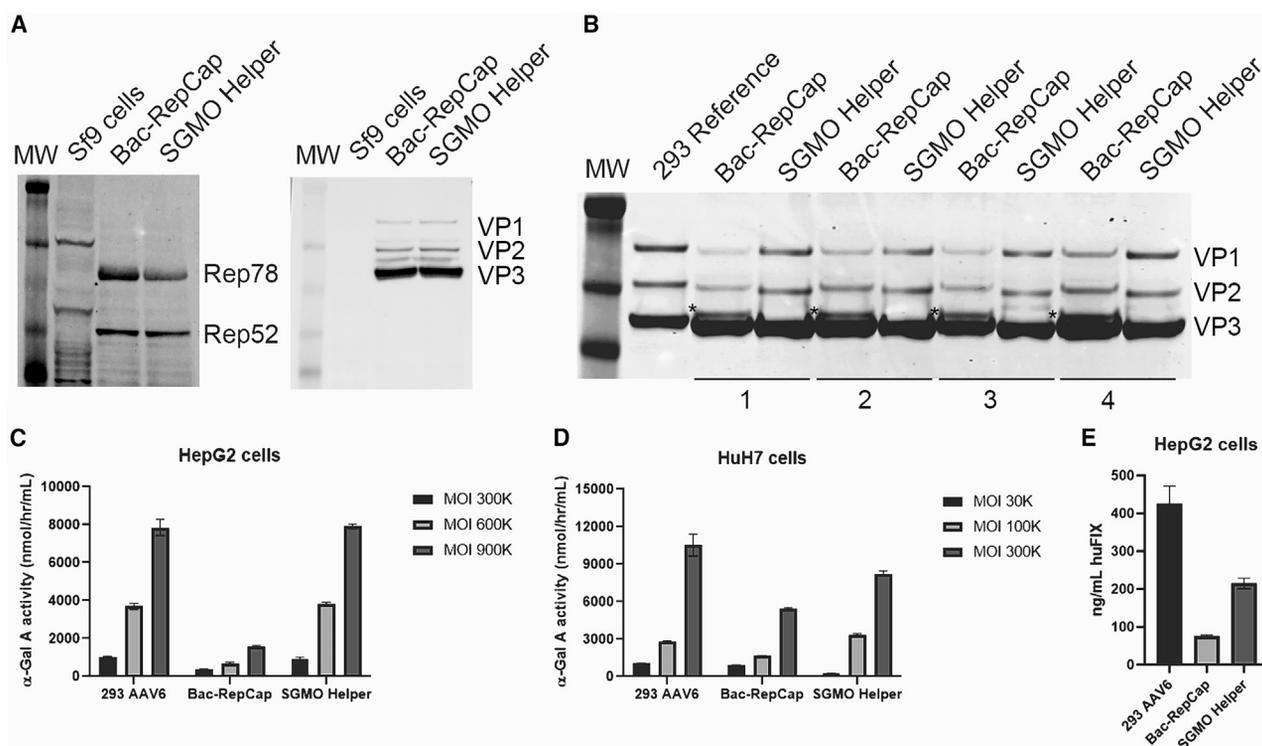
## RESULTS

### Cloning of the AAV6 SGMO Helper

Wild-type and *Rep/Cap* plasmids used to produce rAAV contain in-frame overlapping coding sequences, where the smaller *Rep* sequence (*Rep40/52*) is embedded within the larger *Rep* sequence (*Rep68/78*) and the smaller *Cap* sequences (*VP2* and *VP3*) are embedded within the larger *Cap* sequence (*VP1*).<sup>1</sup> In order to separate the *Rep* and *Cap* ORFs, the coding sequence for *Rep52* and *Rep78*, as well as the *VP1* and *VP2/3* sequences, were cloned into two vectors. The pFastBac Dual Expression Vector was purchased from ThermoFisher and the *Rep78* and *VP2/3* ORFs were cloned into the vector under the control of the baculovirus polyhedrin and p10 promoters, respectively (Figure 1A). The *Rep52* and *VP1* ORFs were cloned into a separate pFastBac Dual Expression Vector, under the control of the baculovirus polyhedrin and p10 promoters, respectively (Figure 1B). In the *Rep52/VP1* vector, an SbfI restriction site was added downstream of the HSV thymidine kinase pA signal by site-directed mutagenesis for the purpose of adding a vector genome cassette. Cloning was performed as described in the [materials and methods](#).

Several improvements were incorporated into the primary nucleotide sequence of the unique *VP1/VP2* region that improve yield, potency, and capsid integrity. The AAV2 Phospholipase A2 (PLA2) domain was reported to improve the capsid ratio and potency of rAAV when transplanted into other serotypes.<sup>9,18</sup> These studies utilized sequences from the start of *VP1* to the end of the PLA2 domain, so it is not possible to differentiate the impact of the primary nucleotide sequence (including the start codon and Kozak context) from the impact of the PLA2 domain alone. In our studies, we determined that a minimal AAV2 PLA2 domain transplanted into the wild-type AAV6 VP1 sequence (corresponding to nucleotides 148 to 318, Figure 1C and resulting in five amino acid changes to the VP1u region) increased the yield of rAAV6 produced in Sf9 cells, but had no impact on the potency of the rAAV6 (Table S1; Figure S1).

In addition to the AAV2 PLA2 domain transplant, we and others observed a proteolytic cleavage event unique to the Sf9-baculovirus production system that resulted in the cleavage of rAAV capsid proteins from serotypes 1, 6, and 8, whereas several serotypes are resistant to this cleavage, such as AAV2, 3B, 5, and 9.<sup>12</sup> The major cleavage product is prominent when purified capsids are imaged on Coomassie-stained SDS-PAGE gels and runs at a slightly larger size than the VP3 band (see Figure 2B, Bac-RepCap, asterisked band; Figure S2, asterisked band). We identified the cleavage site in rAAV6 by performing Edman degradation sequencing to identify the N-terminal amino acid residues present in the major cleavage



**Figure 2. Testing of AAV6 produced using either Bac-RepCap or the SGMO Helper**

(A) Western blot analysis of cell lysates collected at day 6 were harvested using an antibody specific to either Rep or Cap. Naive Sf9 cells were run as a negative control for antibody recognition. The individual Rep (Rep78 and Rep52) and Cap (VP1, VP2, and VP3) proteins are indicated in their respective panels. (B) SDS-PAGE and Coomassie blue stained gel of purified AAV6 from four independent production runs. The individual Cap (VP1, VP2, and VP3) proteins are indicated on the panel. The proteolytic cleavage fragment observed with AAV6 produced using Bac-RepCap is indicated by an asterisk. AAV8 produced using the 293 production system (293 Reference) was also included for comparison. The capsid ratio for each production run is reported in Table 1. (C) *In vitro* potency of AAV6 produced using 293 cells or Sf9 cells (Bac-RepCap and SGMO Helper) on HepG2 cells. HepG2 cells were transduced using the indicated MOI and, after 5 days, cell supernatants were assayed for  $\alpha$ -Gal A activity. The nmol/h/mL  $\alpha$ -Gal A activity was calculated based on fluorescent activity using a standard curve. Each virus sample was assessed for potency in triplicate, with each of the replicates being assayed in duplicate, and the standard deviation of the six data points was used to plot error bars. (D) *In vitro* potency of AAV6 produced using 293 cells or Sf9 cells (Bac-RepCap and SGMO Helper) on HuH7 cells. HuH7 cells were transduced using the indicated MOI and, after 5 days, cell supernatants were assayed for  $\alpha$ -Gal A activity. The nmol/h/mL  $\alpha$ -Gal A activity was calculated based on fluorescent activity using a standard curve. Each virus sample was assessed for potency in triplicate, with each of the replicates being assayed in duplicate, and the standard deviation of the six data points was used to plot error bars. (E) *In vitro* potency of AAV6 produced using 293 cells or Sf9 cells (Bac-RepCap and SGMO Helper) on HepG2 cells. HepG2 cells were transduced using an MOI of  $1 \times 10^6$  and, after 5 days, cell supernatants were assayed for huFIX expression. The ng/mL huFIX quantity was calculated using a standard curve. Each virus sample was assessed for potency in triplicate, with each of the replicates being assayed in duplicate, and the standard deviation of the three data points was used to plot error bars.

product. Similar to the results with AAV8 reported by Galibert et al., the residues in the N-terminal portion of the major cleavage fragment mapped to the unique sequence shared by VP1 and VP2.<sup>12</sup> However, the specific amino acids in the N-terminal portion of the AAV6 major cleavage fragment were EPPATP (also see electropherograms in Figure S3), in contrast to the AAV8 cleavage site residues of EPPAA. In order to prevent proteolytic cleavage of rAAV6, we transplanted a short fragment from AAV9, which is resistant to the cleavage, into the unique region of the AAV6 VP1 and VP2 proteins (corresponding to nucleotides 475 to 609, Figure 1C). This resulted in a 10-amino acid change to the VP1 and VP2 proteins that prevented proteolytic cleavage and increased the potency of rAAV6 capsids (see Figure 2B, SGMO Helper; Figure S2, Bac-RepCap and AAV9 transplant).

#### AAV production using the AAV6 SGMO Helper

We tested rAAV6 production using the SGMO Helpers by coinfecting naive Sf9 cells with baculovirus-infected insect cells (BIICs) generated using each of the helper constructs (*Rep78-VP2/3* and *Rep52-VP1*) at a cell ratio of 1:1:10,000. The vector genome used to generate rAAV6 contains the GLA cDNA, which encodes  $\alpha$ -Galactosidase A ( $\alpha$ -Gal A), under the control of a liver-specific promoter. For comparison, the AAV6 Bac-RepCap developed in the Kotin lab was used to produce rAAV6 in parallel production flasks.<sup>5</sup> After 6 days of culturing, cell pellets were collected and processed for western blot analysis and purification on cesium chloride density gradients as described in the materials and methods. Western blot analysis using an antibody to AAV Rep proteins showed that Rep52 and Rep78 were expressed in cells infected with either Bac-RepCap or the SGMO

**Table 1. Purified AAV titer (vg/mL), purified AAV yield (vg), harvest yield (vg/mL culture), and VP1:VP2:VP3 ratio for AAV lots used in this study**

Experiment	Sample	Purified AAV titer (vg/mL)	Purified AAV yield (vg)	Harvest yield (vg/mL culture)	VP1	VP2	VP3
1	Bac-RepCap	2.23E+13	2.45E+13	1.23E+11	1	2.1	45
	SGMO Helper	5.99E+13	8.99E+13	4.50E+11	1	1.1	21
2	Bac-RepCap	2.25E+13	2.70E+13	1.35E+11	1	2.1	60
	SGMO Helper	7.97E+13	1.20E+14	6.00E+11	1	1.3	24
3	Bac-RepCap	9.46E+12	9.46E+12	4.73E+10	1	2.1	67
	SGMO Helper	6.21E+13	8.69E+13	4.35E+11	1	1	19
4	Bac-RepCap	1.83E+13	1.65E+13	8.25E+10	1	1.8	60
	SGMO Helper	8.40E+13	1.09E+14	5.45E+11	1	0.8	17
Serial passaging	SGMO Helper P.0	1.84E+13	1.29E+13	6.45E+10	1	0.4	16
	SGMO Helper P.3	1.05E+12	9.45E+11	4.73E+09	1	0.8	18
	SGMO Helper P.6	2.70E+12	3.24E+12	1.62E+10	1	0.4	17
Mid-scale production	Bac-RepCap	3.56E+13	3.98E+14	2.65E+11	1	1.7	82
	SGMO Helper	4.85E+13	9.89E+14	6.59E+11	1	1.2	24
Mid-scale production for long-read sequencing	Bac-RepCap	9.89E+13	8.41E+14	5.61E+11	1	1.8	99
	Bac-RepCap	5.31E+13	3.98E+14	2.65E+11	1	1.4	88
	SGMO Helper	2.75E+14	3.38E+15	2.25E+12	1	0.7	34
<i>In vivo</i> study	SGMO Helper	1.02E+14	9.89E+14	6.59E+11	1	1.0	24
	Bac-RepCap	1.37E+13	3.01E+13	1.51E+11	1	1.5	59
	SGMO Helper	3.47E+13	1.15E+14	5.75E+11	1	0.9	27

Helper at 6 days post-infection (Figure 2A). Western blot analysis using an antibody to the AAV Cap proteins further demonstrated that VP1, VP2, and VP3 were expressed at 6 days post-infection (Figure 2A). We observed additional bands between the VP proteins that may process impurities or degradation products common to both production platforms (Figure 2A).

Vg titers of the purified rAAV6 were measured by qPCR and the harvest and total yields were calculated for each production flask. In four independent production runs, the SGMO Helper showed a 3- to 8-fold higher harvest and purified yield of rAAV6 with respect to Bac-RepCap (Table 1). The increase in yield can be attributed to either the presence of the AAV2 PLA2 domain, as observed previously, a fundamental property of the SGMO Helper gene architecture, or both. We tested rAAV6 production using a Bac-RepCap containing the AAV2 PLA2 domain and found that, similar to our observation with the SGMO Helper, the presence of the PLA2 domain increased the harvest and purified yield (Table S1), demonstrating that the AAV2 PLA2 domain alone is sufficient to increase rAAV6 yield. In order to determine whether this gene architecture was generalizable to other serotypes, similar constructs were generated using the Cap sequences from serotypes 1, 2, 3B, and 9. Using the same process described for rAAV6, we determined that the SGMO Helper produced intact AAV capsids with harvest yields ranging from 7.93E+9 to 3.95E+10 vg/mL culture (Table S2).

Analysis of Coomassie-stained SDS-PAGE gels confirmed the presence of VP1, VP2, and VP3 in purified capsids and that the AAV9

transplant in the SGMO Helper sequence eliminated the proteolytic cleavage event observed with unmodified rAAV6 produced in Sf9 cells using Bac-RepCap (Figure 2B). Densitometry analysis of the Coomassie-stained SDS-PAGE gels was used to determine the capsid ratio of the rAAV6 produced using Bac-RepCap and the SGMO Helper. Bac-RepCap-produced rAAV6 had an average capsid ratio of approximately 1:2.9:80, whereas SGMO Helper-produced rAAV6 had an average capsid ratio of approximately 1:1.3:23, demonstrating that more VP1 is incorporated into rAAV6 produced using the SGMO Helper compared with rAAV6 produced using Bac-RepCap (Table 1). A similar analysis was performed with 12 additional production runs and showed an average increase in VP1 incorporation with the SGMO Helper (1:1:26) compared with the Bac-RepCap (1:2:77) (Table S3). We performed a one-way ANOVA with Tukey's multiple comparisons on all 16 production runs, comparing capsid ratios of rAAV6 produced using the Bac-RepCap and SGMO Helper, and showed that the difference in VP2 incorporation was not significantly different ( $p = 0.9949$ ), whereas the difference in VP3 incorporation was significantly different ( $p < 0.0001$ ) (Table S3). The increased incorporation of VP1 into capsids using the SGMO Helper can be attributed to the ACG start codon present in the VP1 ORF, which results in increased translation efficiency compared with the TTG start codon used by the Kotin lab to generate the Bac-RepCap AAV6 Cap ORF (reviewed in Kears and Wilusz<sup>19</sup>), as well as the presence of the AAV9 transplant, which prevents cleavage of VP1. Indeed, this is demonstrated by the stepwise increase in VP1 incorporation using Bac-RepCap containing the ACG start codon alone, the AAV9 transplant alone, or a combination of the two modifications

**Table 2. NGS analysis of encapsidated DNA**

Target aligned against	Filter step	Bac-RepCap	SGMO Helper	Bac-RepCap	SGMO Helper
Input vector genome	0	94.5	94.1	98.7	92.3
Kanamycin resistance gene	1	0.008	0.000	0.000	0.000
Gentamicin resistance gene	2	0.004	0.000	0.000	0.000
Bacmid with F9 cDNA vector genome	3	1.998	0.758	0.163	0.820
pRep2Cap1	4	0.023	0.000	0.010	0.011
pRep2Cap2	5	0.000	0.000	0.000	0.000
pRep2Cap5	6	0.000	0.000	0.000	0.000
pRep2Cap6	7	0.002	0.000	0.000	0.000
pRep2Cap8	8	0.000	0.000	0.000	0.000
pRep2Cap8_Kan	9	0.000	0.000	0.000	0.000
pRep2Cap9	10	0.000	0.000	0.000	0.000
pRepCap3B	11	0.000	0.008	0.000	0.000
Sf9 genome (FDA)	12	0.076	0.032	0.010	0.063
Transfer vector with F9 cDNA vector genome	13	0.002	0.000	0.000	0.000
Empty transfer vector	14	0.000	0.000	0.000	0.000
pMON7124 (transposase plasmid)	15	0.008	0.072	0.004	0.000
Plant rhabdovirus genome	16	0.000	0.000	0.000	0.000
phiX	17	0.000	0.000	0.000	0.000
Total mapped reads:		96.654	95.000	98.893	93.165
Percentage of mapped reads that are vector genome:		97.806	99.085	99.811	99.041
Percentage of unmapped reads:		3.346	5.000	1.107	6.835

Capsid-associated DNA was sequenced and mapped back to the vector genome or other potential contaminating sequences in sequential order, as indicated by the Filter Step. Values are presented as percentage of total sequences that mapped to a particular target. The percentage of mapped and unmapped reads is reported at the bottom of the table and the percentage of the total mapped reads that match the vector genome is also reported.

(Figure S4; Table S4). These results contrast from those reported in Kohlbrenner et al., where the entire VP1 unique sequence from AAV5 was used to replace the VP1 unique sequence in AAV8, and was shown to increase potency.<sup>9</sup> Our studies indicate that the AAV2 PLA2 domain alone is sufficient to improve yields, but not VP1 incorporation or potency (Table S4; Figure S1), whereas the ACG start codon and AAV9 transplant, which prevents proteolytic cleavage of VP1, do drive the increased VP1 incorporation and potency (Table S4; Figure S2).

A similar analysis of capsid ratios was performed using rAAV generated using AAV1, 2, 3B, and 9 SGMO Helpers. The AAV1, 2, 3B, and 9 SGMO Helpers produced rAAV with capsid ratios ranging from 1:1:14 to 1:4:66 (Table S5), demonstrating that the SGMO Helper gene architecture is generalizable to multiple AAV serotypes.

Since packaging of non-vector genome sequences is a concern with both 293 and Sf9 AAV production systems,<sup>7</sup> we extracted DNA from Bac-RepCap and SGMO Helper-derived rAAV6 preps and performed next-generation sequence (NGS) analysis of the encapsidated DNA. Since AAV preps are treated with Benzonase during the purification process and DNase I prior to DNA extraction, we expect that sequence reads only originate from encapsidated DNA. To identify

contaminating DNA sequences, we compared the sequence reads with a panel of expected contaminating DNAs, using a workflow where individual sequence reads are sequentially compared with reference sequences and binned according to whether the sequence matches the reference sequence. Mapped reads are removed from the analysis once assigned to a reference sequence, which could impact the results. We compared two independent AAV production runs using Bac-RepCap and the SGMO Helper. We found the highest percentage of mapped reads matched the vector genome used for rAAV6 production (98% with Bac-RepCap-derived rAAV6 and 99% with SGMO Helper-derived rAAV6, Table 2, Percentage of mapped reads that are vector genome). The most common contaminating DNA was found to be bacmid or baculovirus sequences (0.8%–2% with Bac-RepCap-derived rAAV6 and 0.2%–0.8% with SGMO Helper-derived rAAV6, Table 2). Sequence reads that map to the AAV1 RepCap plasmid (pRep2Cap1) may result from the high sequence homology between AAV1 and AAV6 capsids or the bioinformatic workflow incorrectly mapping AAV6 sequences to the AAV1 sequence. Between 1.1% and 6.8% of reads were not able to be mapped to a reference sequence, either due to poor sequence quality or because the reference sequence is not present in the bioinformatic workflow (Table 2, Percentage of unmapped reads). In order to confirm the homogeneity of the packaged vector genomes, as suggested by the

sequencing analysis, we performed alkaline gel electrophoresis on the lots used in the contaminating DNA experiment and compared them qualitatively with rAAV6 produced with the same vector genome using the 293 cells. We found there was a prominent band corresponding to the expected vector genome size for all samples (Figure S5). This analysis demonstrates that the SGMO Helper packages AAV vector genomes with the same fidelity as a previously described Sf9 baculovirus AAV production system.

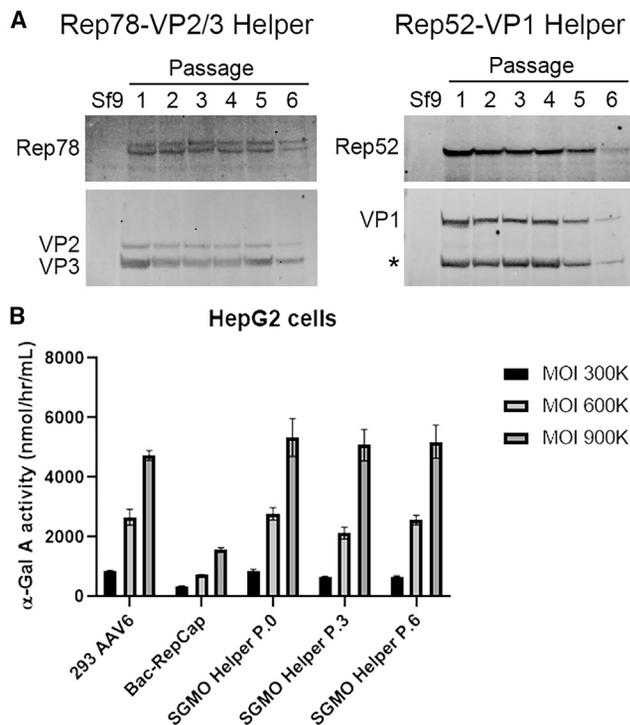
We tested the potency of rAAV6 generated using the 293 production system, Bac-RepCap, and the SGMO Helper using two different transgenes and cell lines. We first tested the material described above by transducing HepG2 or HuH-7 cells with rAAV6 carrying a vector genome encoding  $\alpha$ -galactosidase A ( $\alpha$ -Gal A) under the control of a liver-specific promoter using three different multiplicities of infection (MOIs). This gene is clinically relevant since patients diagnosed with Fabry disease carry mutations that reduce their ability to produce the  $\alpha$ -Gal A enzyme. Since HepG2 and HuH-7 cells have different levels of transducibility, an MOI of 300K, 600K, and 900K was used with HepG2-based potency testing, whereas an MOI of 30K, 100K, and 300K was used with HuH-7-based potency testing. Five days post-transduction, the tissue culture supernatants were collected and assayed for  $\alpha$ -Gal A activity, as described in the [materials and methods](#) section. We found that, with HepG2 cells, rAAV6 generated using the 293 production system and the SGMO Helper produced similar amounts of  $\alpha$ -Gal A activity, whereas rAAV6 produced using Bac-RepCap produced approximately 4-fold less  $\alpha$ -Gal A activity (Figure 2C). The relative potency with respect to the 293 AAV6 reference sample is summarized for all *in vitro* experiments in Table S6. A similar trend was observed when the same material was tested on HuH-7 cells, though the potency of SGMO Helper-derived rAAV6 at the lowest and highest MOIs (30K and 300K) was lower with respect to the 293-derived rAAV6 (Figure 2D; Table S6). The overall difference in potency in HuH-7 cells with 293- or SGMO Helper-derived rAAV6 compared with Bac-RepCap-derived rAAV6 was less than the difference observed with HepG2 cells. However, in both cell lines using the same vector genome, an increase in potency was observed with rAAV6 produced using the SGMO Helper. To extend this analysis to a second transgene, we produced rAAV6 using the 293 production system, Bac-RepCap, and SGMO Helper using a vector genome carrying the human F9 cDNA under the control of a liver-specific promoter, which is a common gene therapy strategy to correct the genetic deficiency present in Hemophilia B patients. We tested the potency of the rAAV6 on HepG2 cells by transducing with an MOI of 1E6 and assayed for huFIX expression by ELISA. We found that rAAV6 produced using the SGMO Helper had 2- to 3-fold higher huFIX expression compared with rAAV6 produced using Bac-RepCap, though the SGMO Helper AAV6 still had 2-fold lower potency compared with 293 AAV6 (Figure 2E; Table S6). Based on the DNA contaminant analysis, there is no fundamental difference between rAAV6 produced using Bac-RepCap and the SGMO Helper. However, the amount of VP1 content in an AAV capsid directly impacts the potency of an AAV vector, and VP1 incorporation with SGMO Helper-derived rAAV6 was greater than VP1 incorporation

with Bac-RepCap-derived rAAV6.<sup>20,21</sup> Along with the elimination of proteolytic cleavage observed with SGMO Helper-derived rAAV6, thus increasing the amount of intact VP1 present in the capsid, the increase in yield and potency demonstrates the SGMO Helper is an improved system for AAV production in Sf9 cells.

#### Stability of Rep and Cap expression from the SGMO Helper during serial passaging

Baculoviruses are known to accumulate gene deletions and generate increasing amounts of defective interfering particles after serial passaging.<sup>22</sup> Serial passaging of baculoviruses containing the *Rep* and/or *Cap* genes has been shown to lead to reduced protein expression and rAAV yields as the passage number increases.<sup>9</sup> Other groups have also observed a passage effect that is either the result of loss of functional baculovirus particles or instability of the *Rep* and *Cap* genes inserted into the Tn7 transposase sites.<sup>18</sup> To assess the stability of the SGMO Helper we performed serial passaging of the individual baculoviruses and assayed for expression of the AAV Rep and Cap proteins by western blot. Briefly, fresh baculovirus was generated by bacmid transfection of Sf9 cells and the infectious titer was determined using an infectious titer assay (see [materials and methods](#)). The fresh baculovirus was used to generate passage 0 BIICs by inoculation of naive Sf9 cells with an MOI of 0.1 and were cryopreserved once the cell diameter increased by 2  $\mu$ m.<sup>23</sup> Cryopreserved BIICs were thawed, diluted, and a 1/100th volume was used to inoculate naive Sf9 cultures. Every 3 days, cultures were sampled for analysis and a 1/100th volume was used to inoculate a new naive Sf9 culture. This process was continued for six passages total and samples were analyzed by NGS for genetic stability and by western blot for AAV Rep and Cap expression. In addition, cell-free culture media from passages 3 and 6 were titered and used to generate BIICs for testing in AAV production.

Cell pellets from each passage were lysed using RIPA buffer, clarified by centrifugation, and analyzed by western blot for the expression of the Rep and Cap proteins. The *Rep78-VP2/3* baculovirus expressed Rep78, VP2, and VP3 throughout all six passages, though a modest reduction in expression levels can be seen at later passages (Figure 3A). The *Rep52-VP1* baculovirus expressed Rep52 and VP1 throughout all six passages and also saw a reduction in expression at later passages (Figure 3A). However, in the case of the *Rep52-VP1* baculovirus, the western blot for Cap proteins also revealed the presence of an additional band corresponding to the approximate size of VP3 (\*, Figure 3A). While it cannot be ruled out that the band is the result of proteolytic cleavage in Sf9 cells, it is likely the result of translation from an in-frame ATG codon occurring 21 nucleotides downstream of the normal VP3 start codon. This truncated VP3 protein has been observed in other publications and a similar mutation strategy prevented expression of the truncated VP3.<sup>24,25</sup> The use of the downstream start codon is not unique to the Sf9-baculovirus production system, since the materials used in these studies were derived from mammalian cell AAV production.<sup>24,25</sup> To confirm the source of the truncated band, we mutated the downstream ATG to a GTG in the *Rep52-VP1* Helper to give rise to the *Rep52-VP1* Helper (GTG). We transfected the bacmids into Sf9 cells and analyzed Cap expression after 72 h by western



**Figure 3. Serial passaging of the SGMO Helper**

(A) Western blot analysis of cell lysates collected 72 h post-infection for the indicated passage of each individual helper (Rep78-VP2/3 or Rep52-VP1). Cell lysates were analyzed using an antibody specific to either Rep or Cap. Naive Sf9 cells were run as a negative control for antibody recognition. The individual Rep (Rep78 or Rep52) and Cap (VP1, VP2, or VP3) bands are indicated in each panel. An additional band resulting from internal translation initiation from the VP1 ORF is indicated by an asterisk. (B) *In vitro* potency of AAV6 produced using 293 cells or Sf9 cells (Bac-RepCap and SGMO Helper) on HepG2 cells. SGMO Helper baculovirus from passages zero, three, and six were used to generate baculovirus-infected insect cells (BIICs) and the BIICs were used to initiate AAV production in naive Sf9 cells. HepG2 cells were transduced using the indicated MOI and, after 5 days, cell supernatants were assayed for  $\alpha$ -Gal A activity. The nmol/h/mL  $\alpha$ -Gal A activity was calculated based on fluorescent activity using a standard curve. Each virus sample was assessed for potency in triplicate, with each of the replicates being assayed in duplicate, and the standard deviation of the six data points was used to plot error bars.

blot analysis. For comparison, we also analyzed a 72 h time point from an AAV production culture containing both the Rep78-VP2/3 and Rep52-VP1 Helpers (SGMO Helper). Western blot analysis confirmed that the additional band observed during serial passaging was not present in the Rep52-VP1 Helper (GTG) sample, confirming that the additional band is the result of translation initiation from the downstream ATG codon present in the Rep52-VP1 Helper (Figure S6, VP3 or \*). Furthermore, serial passaging of the Rep78-VP2/3 and Rep52-VP1 Helpers demonstrates that Rep and Cap are expressed over multiple passages, but declines during passages five and six, consistent with observations made with other helper systems.<sup>5,9</sup>

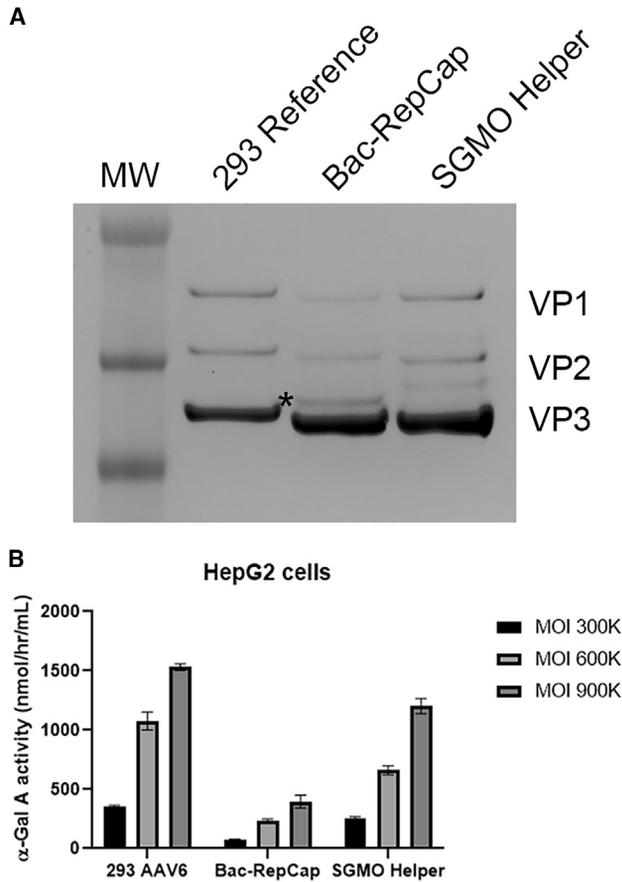
We analyzed the sequence of the baculovirus in cell-free supernatants from each passage to determine whether mutations accumulated dur-

ing serial passaging. Cell-free supernatants from each passage were extracted for total DNA and prepared for Nextera sequencing. Mapping of the individual sequencing reads back to the original bacmid sequence was used to identify single nucleotide polymorphisms or deletions for each passage. No sequence variants were detected in sample NGS paired-end reads using pilon (v1.23, using flags -variant and mindepth 10), indicating that the baculoviruses did not accumulate mutations over the course of six serial passages.<sup>26</sup>

In order to test the productivity and potency of the serial passaged baculoviruses, we generated BIICs from passage three and six and used them to initiate AAV production in naive Sf9 cells. Briefly, the baculovirus titer for passages three and six cell-free supernatants was determined and used to initiate BIIC cultures as described in the materials and methods section. The BIICs were used to initiate AAV production in naive Sf9 cells, harvested 6 days post-infection and purified on CsCl gradients. The vg titer was determined by qPCR and the purified rAAV6 was used to transduce HepG2 cells at an MOI of 300K, 600K, and 900K. After 5 days, the cell supernatants were assayed for  $\alpha$ -Gal A activity and compared with parallel transductions with rAAV6 derived from 293 cells or Sf9 cells using passage zero Bac-RepCap or SGMO Helper BIICs. Compared with the rAAV6 yield using passage zero BIICs, the rAAV6 yield using passage three and passage six BIICs was reduced 4- to 13-fold (Table 1). This suggests that, despite a decline in Rep and Cap expression, the Rep and Cap genes remained intact and there was sufficient Rep and Cap expression to support rAAV production, even at higher passages. In spite of the reduced yield, the material was tested for potency on HepG2 cells using  $\alpha$ -Gal A expression. We found that rAAV6 from 293 cells or produced in Sf9 cells using the SGMO Helper, had similar potency, whereas rAAV6 produced in Sf9 cells using Bac-RepCap had reduced potency (Figure 3B; Table S6). The potency of SGMO Helper-derived rAAV6 was similar using passages zero, three, and six BIICs, suggesting that, while the yield of rAAV6 declines over the course of serial passaging, the potency of the rAAV6 did not change over the course of serial passaging. Furthermore, the decline in Rep/Cap expression and rAAV6 yield is consistent with results using other helper systems, thus is not a unique feature of the SGMO Helper.<sup>5,9</sup> These results demonstrate that, over six passages, the SGMO Helper expresses the full complement of Rep and Cap proteins, is a genetically stable Sf9-baculovirus production system, and is capable of generating rAAV6 with similar potency up to passage six.

#### Scaled-up rAAV6 manufacturing using the SGMO Helper

One challenge to developing new AAV manufacturing platforms is the inability to predict performance during scale-up to larger volumes. Our previous work utilized 0.4-L culture volumes in 1-L shake flasks for upstream production and CsCl density gradients for downstream purification (small-scale process), but production at clinical scale is typically performed in bioreactors and utilize column chromatography for purification. To model the performance of the SGMO Helper at larger scale, we produced rAAV6 carrying the GLA transgene in 3 L volumes using Bac-RepCap and the SGMO Helper in a bioreactor system and purified the resulting material using column



**Figure 4. Testing of AAV6 produced at 3L scale using either Bac-RepCap or the SGMO Helper**

(A) SDS-PAGE and Coomassie blue stained gel of purified AAV6 from 3L production run. The individual Cap (VP1, VP2, and VP3) proteins are indicated on the panel. The proteolytic cleavage fragment observed with AAV6 produced using Bac-RepCap is indicated by an asterisk. AAV8 produced using the 293 production system (293 Reference) was also included for comparison. The capsid ratio for each production run is reported in Table 1. (B) *In vitro* potency of AAV6 produced using 293 cells or at 3L scale in Sf9 cells (Bac-RepCap and SGMO Helper) on HepG2 cells. HepG2 cells were transduced using the indicated MOI and, after 5 days, cell supernatants were assayed for  $\alpha$ -Gal A activity. The nmol/h/mL  $\alpha$ -Gal A activity was calculated based on fluorescent activity using a standard curve. Each virus sample was assessed for potency in triplicate, with each of the replicates being assayed in duplicate, and the standard deviation of the six data points was used to plot error bars.

chromatography (mid-scale process). The yield of rAAV6 using the SGMO Helper was 2.5-fold higher compared with the yield of rAAV6 using Bac-RepCap (Table 1). The increased rAAV6 yield using the SGMO Helper compared with Bac-RepCap using mid-scale process is consistent with increased yields observed with the small-scale process, demonstrating the increase is not an artifact of the production scale or purification process (Table 1).

Analysis of purified capsids by SDS-PAGE and Coomassie staining indicated similar purity with column chromatography purification as previously observed with CsCl density gradient purification (Fig-

ure 4A). Densitometry analysis of the Coomassie-stained SDS-PAGE gel for capsid ratios showed that, similar to CsCl-purified rAAV6, Bac-RepCap incorporated less VP1 (1:1.7:82) compared with the SGMO Helper (1:1.2:24) (Table 1). We also analyzed the purified capsids using capillary electrophoresis (CE-SDS) and determined that the capsid ratios also showed increased VP1 incorporation with the SGMO Helper (1:2.4:30) compared with Bac-RepCap (1:1.5:52) (Figure S7; Table S5). The capsid ratios from the mid-scale process are similar to the capsid ratios determined using the small-scale process (Table 1), suggesting that capsid assembly is consistent in shake flask and bioreactor environments.

The purification process could impact the number of empty capsids present in an AAV preparation and, indeed, several groups have observed enrichment of full capsids using density gradient purification.<sup>27–29</sup> In contrast, current column purification schema purify AAV based on capsid epitope recognition and do not have the ability to distinguish between empty and full capsids. The amount of empty capsids in column purified AAV preparations is likely more representative of the actual empty:full capsid ratio, since, in contrast to density gradients, the method does not typically enrich for full capsids. We assayed the capsid titer of rAAV6 produced using the mid-scale process by ELISA and found that, Bac-RepCap-derived rAAV6 was 100% full, whereas the SGMO Helper-derived rAAV6 was 98% full (Table S7). The % full capsids observed with column purified rAAV6 produced using Bac-RepCap and SGMO Helper is significantly higher than the number of full capsids observed with 293 AAV production, which can contain up to 90% empty capsids.<sup>30</sup>

The potency of rAAV6 material from the scaled-up production process was tested by transducing HepG2 cells using an MOI of 300K, 600K, and 900K. Five days post-transduction, tissue culture supernatants were assayed for the presence of  $\alpha$ -Gal A activity, as described in the materials and methods section. The 293-derived and SGMO Helper-derived rAAV6 had similar levels of  $\alpha$ -Gal A activity, whereas Bac-RepCap-derived rAAV6 had reduced  $\alpha$ -Gal A activity (Figure 4B; Table S6). These results demonstrate that the increased potency of SGMO Helper-derived rAAV6 observed with material from the small-scale process can be replicated using a mid-scale process and suggests that the SGMO Helper will generate higher yields, improved capsid ratios, and greater potency in a larger scale production environment.

#### Long-read sequence analysis of rAAV6 from scaled-up manufacturing

Long-read sequencing has recently emerged as an additional tool for studying packaged vector genomes and identification of non-vector genome sequences.<sup>31–33</sup> Prior to sequence analysis, duplicate samples from the scaled-up manufacturing process were analyzed by SDS-PAGE to determine their respective capsid ratios (Figure S8A). Consistent with the previously described preparations, densitometry analysis of the Coomassie-stained SDS-PAGE gels showed AAVs produced using Bac-RepCap incorporated less VP1 (1:1.8:99 and 1:1.4:88) compared with AAVs produced using the SGMO Helper

**Table 3. Long-read NGS analysis of encapsidated DNA**

Target aligned against	Filter step	Bac-RepCap	SGMO Helper	Bac-RepCap	SGMO Helper
Vector genome	0	97.65	98.90	97.66	97.16
Rep/Cap	1	0.66	0.00	0.60	0.00
Bacmid	2	1.49	0.64	1.39	2.03
Sf9 genome	3	0.04	0.04	0.05	0.10
Total mapped reads:		99.84	99.58	99.70	99.29
Percentage of mapped reads that are vector genome:		97.81	99.32	97.95	97.85
Percentage of unmapped reads:		0.16	0.42	0.30	0.71

Capsid-associated DNA was sequenced and mapped back to the vector genome, Rep/Cap, bacmid, or Sf9 genome. Values are presented as percentage of total sequences that mapped to a particular target. Percentage of unmapped reads are reported. Unmapped reads were used to BLAST search against the NCBI genome collection and greater than 75% of unmapped reads mapped to the *Shigella* sp. PAMC 28750, a *Shigella* isolate from Antarctic lichens.

(1:0.7:34 to 1:1:24) (Table 1). We assayed the capsid titer of the rAAV6 by ELISA and found that both the Bac-RepCap-derived and SGMO Helper-derived rAAV6 preparations contained 100% full capsids (Table S7). The potency of rAAV6 material was tested by transducing HepG2 cells using an MOI of 300K, 600K, and 900K. Five days post-transduction, tissue culture supernatants were assayed for the presence of  $\alpha$ -Gal A activity, as described in the [materials and methods](#) section. Consistent with the previous scaled-up preparations, 293-derived and SGMO Helper-derived rAAV6 had similar levels of  $\alpha$ -Gal A activity, whereas Bac-RepCap-derived rAAV6 had reduced  $\alpha$ -Gal A activity (Figure S8B). These results demonstrate that the scaled-up preparations used for long-read sequencing analysis are similar to the previously described preparations.

Encapsidated DNA from duplicate scaled-up Bac-RepCap and SGMO Helper-derived rAAV6 preps was interrogated using the SMRT bell library preparation and PacBio Sequel instrument. Prior to sequence analysis, AAV preps are treated with Benzonase during the purification process and DNase I prior to DNA extraction, so sequence reads should only originate from encapsidated DNA. Consistent with the short-read sequencing results using small-scale preps (Table 2), long-read sequencing mapped 97.81% and 97.95% of reads to the vector genome in rAAV6 preps derived from Bac-RepCap and 97.85% to 99.32% of reads to the vector genome in rAAV6 preps derived from the SGMO Helper (Table 3). To identify contaminating DNA sequences, we compared the sequence reads that did not map to the vector genome with a panel of expected contaminating DNAs, using a workflow where individual sequence reads are sequentially compared with reference sequences and binned according to whether the sequence matches the reference sequence. The first filter contained the AAV2 *Rep* and AAV6 *Cap* genes used to produce the rAAV6 and mapped 0.60% to 0.66% of reads in samples derived from Bac-RepCap compared with 0% of reads in samples derived from the SGMO Helper (Table 3). The second filter contained the bacmid sequence and mapped 1.39% to 1.49% of reads in samples derived from Bac-RepCap compared with 0.64% to 2.03% of reads in samples derived from the SGMO Helper (Table 3). The last filter was the Sf9 genome and mapped 0.04% to 0.05% of reads in samples

derived from the Bac-RepCap compared with 0.04% to 0.10% of reads in samples derived from the SGMO Helper (Table 3). The final analysis performed a BLAST search using the remaining unmapped reads, which represented 0.16% to 0.30% of all reads in samples from the Bac-RepCap and 0.42% to 0.71% of all reads in samples from the SGMO Helper (Table 3). Results from the BLAST search indicated that greater than 75% of unmapped reads mapped to the *Shigella* sp. PAMC 28750, a *Shigella* isolate from Antarctica, suggesting a possible environmental contamination during rAAV production, purification, or sample preparation for NGS analysis. These results demonstrate that the SGMO Helper packages AAV vector genomes with the same fidelity as a previously described Sf9-baculovirus AAV production system.

In addition to encapsidated DNA, long-read sequencing is purported to be useful for analyzing vector genomes for the presence or absence of truncations.<sup>31,32</sup> We analyzed sequences that mapped to the vector genome for truncations and compared the frequency and location of truncations from AAV produced using the Bac-RepCap or SGMO Helper. Vector genome reads that contained both ITRs comprised 35% to 37% of the total reads with both the Bac-RepCap and SGMO Helper-derived AAVs (Table 4). Vector genomes containing only the left ITR or right ITR comprised 24% to 28% and 30% to 36% of the total reads with both the Bac-RepCap and SGMO Helper-derived AAVs, respectively (Table 4). Vector genomes containing no ITR or ITR only comprised 2% to 3% or 0.2% to 0.3% of the total reads with both the Bac-RepCap and SGMO Helper-derived AAVs, respectively (Table 4). We performed a t test with multiple comparisons of each category of truncation to compare AAVs produced using the Bac-RepCap or SGMO Helper and found there was no significant difference in ITR content (Table 4). We employed an alternative method to examine the homogeneity of the packaged vector genomes by performing alkaline gel electrophoresis on the samples used for long-read sequencing and compared them qualitatively with rAAV6 produced with the same vector genome using the 293 cells. We found there was a prominent band corresponding to the expected vector genome size for all samples, with no additional prominent bands consistent with the truncations determined by long-read

**Table 4. Quantitation of ITR content in sequencing reads that map to the vector genome**

Sample	Total mapped reads	Both ITR reads (% total reads)	Left ITR-only reads (% total reads)	Right ITR-only reads (% total reads)	No ITR reads (% total reads)	ITR-only reads (% total reads)
Bac-RepCap	146,721	52,634 (35.9%)	41,721 (28.4%)	45,138 (30.8%)	5,024 (3.4%)	453 (0.3%)
Bac-RepCap	187,604	69,519 (37.1%)	52,108 (27.8%)	57,551 (30.7%)	6,073 (3.2%)	462 (0.2%)
SGMO Helper	205,347	64,478 (31.4%)	55,963 (27.3%)	75,853 (36.9%)	4,185 (2.0%)	508 (0.2%)
SGMO Helper	152,342	53,898 (35.4%)	37,727 (24.8%)	51,672 (33.9%)	5,838 (3.8%)	368 (0.2%)
	Significant? (t test, alpha = 0.05)	No (p = 0.867)	No (p = 0.995)	No (p = 0.457)	No (p = 0.638)	No (p = 0.807)

Values are presented as read count or percentage of the total reads used in this analysis; t tests were used to compare each category of ITR content for each sample.

sequencing (Figure S9). While the sequencing results demonstrate there is no significant difference in the number of truncations observed by long-read sequencing analysis of AAVs produced using the Bac-RepCap or SGMO Helper, the alkaline gel electrophoresis analysis suggests that the number of truncations determined by long-read sequencing may be overestimated.

Truncations mapped using the long-read sequencing data were further interrogated to determine whether there were specific sites where truncations may occur at a higher frequency. Indeed, six sites were associated with truncations at a higher frequency than elsewhere in the vector genome, all of which corresponded to the ITR sequence. Truncations were observed at nucleotides 1 and 2,772, which are the terminal nucleotides of the vector genome (Table 5; Figure S10). Surprisingly, the truncations at these positions occurred in 100% of the reads, which corresponded to a significant dropoff in read depth in the ITR termini (Table 5; Figure S10). The remaining four truncations occurred within the ITR sequence and map to the transition from the A domain to the B or C domain (nucleotide 27), the transition from the C' to B domain or B' to C domain (nucleotide 46), to a site in the A' domain adjacent to the Rep Binding Element (RBE) (nucleotide 74), or to the base of an alternative ITR secondary structure formed by the D and A' domains, adjacent to the RBE (nucleotide 2,677) (Table 5; Figure S10).<sup>34</sup> The frequency of truncation at these positions ranged from 20% to 36% of the total reads at each position (Table 5). We performed a t test with multiple comparisons of each truncation to compare AAVs produced using the Bac-RepCap or SGMO Helper and found there was no significant difference in the frequency of truncations at each location (Table 5). The above results suggest that there is no significant difference in capsid content with AAVs produced using the SGMO Helper compared with the previously described Bac-RepCap from the Kotin lab.

#### **In vivo testing of rAAV6 derived from the SGMO Helper**

We have extensively tested rAAV6 derived from the SGMO Helper for yield, capsid ratio, and potency using multiple cell lines, vector genomes, production volumes, and purification methods (Figures 2 and 4). In order to confirm that our observations translate to improvements in *in vivo* potency, we designed a study to test the potency of rAAV6 produced using the SGMO Helper in C57/Bl6 mice. For the *in vivo* study, we utilized a vector genome that carries the GLA

cDNA from cynomolgus macaques (cynoGLA) under the control of a liver-specific promoter. Using the small-scale process, we generated rAAV6 using Bac-RepCap and the SGMO as well as the HEK293 production system. Using the cynoGLA transgene, the yield of rAAV6 from the SGMO Helper was 4-fold higher than the yield of rAAV6 from Bac-RepCap (Table 1). Analysis of rAAV6 by SDS-PAGE and Coomassie stain showed similar purity and densitometry analysis indicated that the capsid ratios of rAAV6 from Bac-RepCap (1:1.5:59) and the SGMO Helper (1:0.9:27) were consistent with previous observations and demonstrated increased VP1 incorporation with rAAV6 from the SGMO Helper (Table 1; Figure 5A).

The potency of rAAV6 generated in 293 and Sf9 cells was tested in a mouse hepatocyte cell line (Hepa1-6) using an MOI of 100K, 300K, and 600K. Five days post-transduction, tissue culture supernatants were assayed for GLA enzyme activity. Consistent with our observations with human hepatocellular carcinoma cell lines and vector genomes containing either huGLA or huFIX cDNAs, we observed an increase in  $\alpha$ -Gal A activity with rAAV6 produced using the SGMO Helper compared with rAAV6 produced using Bac-RepCap at an MOI of 300K and 600K (Figure 5B; Table S6).  $\alpha$ -Gal A activity was similar, with rAAV6 produced using the SGMO Helper and rAAV6 produced in 293 cells at an MOI of 300K and 600K. In contrast to our previous observations, rAAV6 from 293 and Sf9 cells had similar potency at an MOI of 100K, suggesting that the 100K MOI may represent the lower limit for quantitation of  $\alpha$ -Gal A activity using Hepa1-6 cells.

The same materials from the above analyses were formulated to deliver a 5E12 vg/kg (low dose) or 2E13 vg/kg (high dose) rAAV6 in mice. Formulated material from the low-dose or high-dose group was administered to six male C57/Bl6 mice by tail vein injection and observed immediately post-dose for adverse effects. After 28 days, mice were euthanized and exsanguinated. Blood samples were aliquoted, frozen, and transferred to SGMO Therapeutics for analysis. Blood samples were diluted using 1X PBS and assayed for  $\alpha$ -Gal A activity as described in the materials and methods section. The 293, Bac-RepCap, and the SGMO Helper low-dose groups had similar levels of  $\alpha$ -Gal A activity, while the 293 and SGMO Helper high-dose groups had higher levels of  $\alpha$ -Gal A activity compared with the Bac-RepCap high-dose group (Figure 5C). Statistical analysis using a two-way

**Table 5. Mapping of long-read sequencing truncations to the vector genome**

Nucleotide position (1–2,772)	Bac-RepCap	Bac-RepCap	SGMO Helper	SGMO Helper	Significant? (t test)
2,772	107,421 (100%)	137,442 (100%)	173,535 (100%)	124,655 (100%)	No (p = 0.451)
2,677	76,927 (35%)	106,331 (36%)	121,382 (35%)	94,742 (36%)	No (p = 0.495)
1	68,183 (100%)	88,652 (100%)	91,370 (100%)	68,655 (100%)	No (p = 0.926)
74	59,316 (29%)	79,540 (30%)	81,865 (30%)	62,415 (30%)	No (p = 0.865)
46	29,130 (20%)	38,902 (21%)	38,299 (21%)	31,084 (21%)	No (p = 0.922)
27	28,626 (29%)	34,801 (28%)	34,849 (27%)	28,627 (29%)	No (p = 0.996)

The top six truncations identified by long-read sequencing are included and the frequency of each truncation is reported. The percentage of truncations vs. total reads at each position is reported in parentheses; t tests were used to compare the frequency of truncations in each sample.

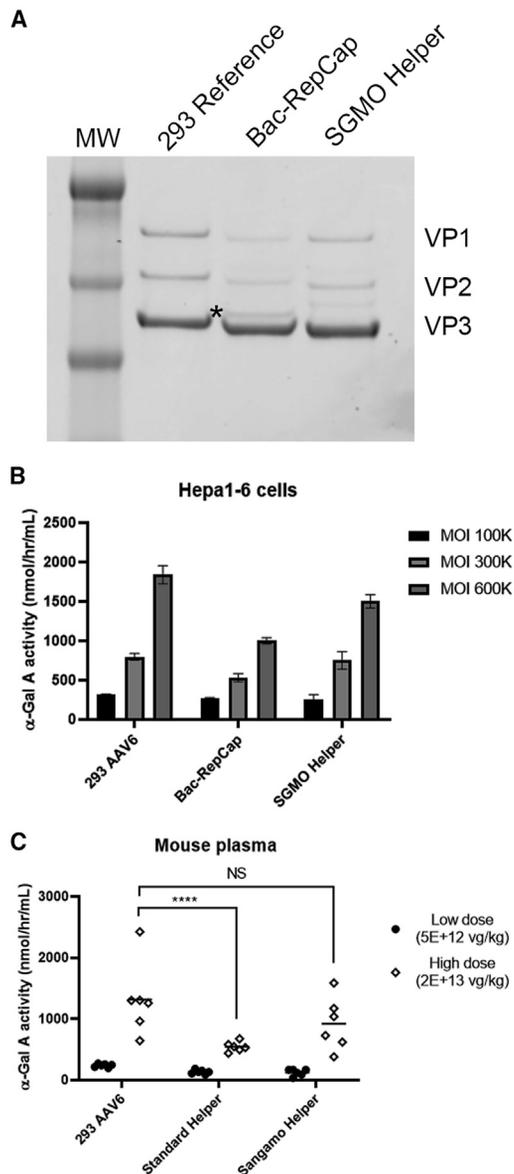
ANOVA and Tukey's multiple comparisons demonstrated that there was no significant difference in  $\alpha$ -Gal A activity in the 293 and SGMO Helper high-dose groups, whereas there was a significant difference in  $\alpha$ -Gal A activity between 293 and Bac-RepCap high-dose groups ( $p = 0.0018$ ). Although  $\alpha$ -Gal A activity was higher with the SGMO Helper high-dose group compared with the Bac-RepCap high-dose group, the results were not statistically significant ( $p = 0.298$ ). The data demonstrate that the improvements in potency observed with *in vitro* testing is consistent with observations *in vivo* and confirm that the SGMO Helper produces rAAV6 with comparable potency to 293-derived rAAV6.

## DISCUSSION

We have described the development and testing of the SGMO Helper, a novel helper system for AAV6 production in Sf9 cells capable of producing higher yields and improved potency compared with the Bac-RepCap developed in the Kotin lab. In addition to separating the ORFs for the Rep and Cap proteins, the SGMO Helper incorporates several improvements: an ACG start codon and strong Kozak context for VP1 translation, incorporation of the AAV2 PLA2 domain into the AAV6 capsid protein, and the removal of a baculovirus-specific proteolytic cleavage site in the VP1 and VP2 unique regions. We found these modifications improved yield using small-scale and mid-scale production processes, demonstrating that the SGMO Helper is a scalable system for AAV production. The elimination of the baculovirus-specific proteolytic cleavage leads to production of more consistent capsids that preserve the complete VP1 protein, including the PLA2 domain required for endosomal escape. Consistent with other published studies, the increased incorporation of intact VP1 into capsids observed with the SGMO Helper correlates with an increase in potency.<sup>6–11,20,21</sup> We observed increases in both *in vitro* and *in vivo* potency, suggesting that our improvements have the potential to translate to clinical improvements. While we did not examine the relative biodistribution of Bac-RepCap and SGMO Helper rAAV6, no changes were made to the VP3 sequence in the SGMO Helper, so the PLA2 domain and AAV9 transplant are unlikely to impact the tissue tropism. Indeed, multiple structural studies have shown that the outward facing amino acid residues of the capsid are from the common VP3 sequence shared by all three VP proteins.<sup>15,35,36</sup> Furthermore, genetic studies where sequences from different serotypes are swapped also demonstrate that *in vivo* tissue

tropism is primarily driven by VP3 sequences.<sup>37,38</sup> Last, we demonstrated that the SGMO Helper system is genetically stable over serial passaging, with no differences in potency observed with rAAV6 produced using early, mid, or late passage BIICs, although decreases in AAV productivity and Rep and Cap protein expression were observed. We have performed head-to-head comparisons using an AAV9 capsid in the Bac-RepCap and SGMO Helper gene architectures and observed similar improvements to yield, capsid ratio, and potency (manuscript in preparation).

Sequence analysis of the encapsidated DNA demonstrated that the Bac-RepCap and SGMO Helpers produce capsids with similar profiles of non-vector DNA and vector genomes. While sequencing of encapsidated DNA is useful for characterization of capsid content, sequencing data for vector genome integrity must be interpreted with caution.<sup>39</sup> The short-read sequencing method used for this publication relies on PCR amplification during the library preparation, which could introduce bias into the library, and does not provide single molecule resolution.<sup>40,41</sup> Long-read sequencing provides single molecule resolution, but also relies on complex library preparation that requires annealing of negative and positive strand copies of the vector genomes, introducing potential template bias from either strand. The problem with introduced template bias is of particular concern since the predominant sequence remains constant (promoter, transgene, and polyA signal), but the ITRs vary in their flip and flop orientation. A mismatch of the flip and flop orientations would lead to imperfect sequence complementarity in the ITRs that could be impacted by the library preparation or sequencing process. The library preparation requires enzymatic treatments to repair DNA damage, resection, or fill-in of DNA overhangs, A-tailing, and ligation, all of which can lead to the introduction of non-templated nucleotides during DNA repair, removal, or addition of extraneous nucleotides during end repair, or ligation of multiple vector genomes. Indeed, a recent publication examining the accuracy of long-read sequencing of mRNA transcript termini demonstrated inconsistencies with long-read sequencing data of well-annotated transcripts, which could also apply to single-stranded DNA species containing complex secondary structures at their termini.<sup>42</sup> Last, it is critical to demonstrate the relevance of these sequence analyses to AAV biology using orthogonal techniques that analyze vector genome integrity and potency. The complex secondary structures in



**Figure 5. Testing of AAV6 produced using Bac-RepCap and the SGMO Helper for *in vivo* potency**

(A) SDS-PAGE and Coomassie blue stained gel of purified AAV6. The individual Cap (VP1, VP2, and VP3) proteins are indicated on the panel. The proteolytic cleavage fragment observed with AAV6 produced using Bac-RepCap is indicated by an asterisk. AAV8 produced using the 293 production system (293 Reference) was also included for comparison. The capsid ratio for each production run is reported in Table 1.

(B) *In vitro* potency of AAV6 produced using 293 cells or Sf9 cells (Bac-RepCap and SGMO Helper) on Hepa1-6 cells. Hepa1-6 cells were transduced using the indicated MOI and, after 5 days, cell supernatants were assayed for  $\alpha$ -Gal A activity. The nmol/h/mL  $\alpha$ -Gal A activity was calculated based on fluorescent activity using a standard curve. Each virus sample was assessed for potency in triplicate, with each of the replicates being assayed in duplicate, and the standard deviation of the six data points was used to plot error bars. (C) Mouse plasma levels of  $\alpha$ -Gal A activity from mice dosed with 5E+12 vg/kg or 2E+13

the ITRs could cause deletion of those sequences during library preparation or sequencing and be the root cause of the left and right ITR-only reads detected by long-read sequencing (Table 4). Indeed, while the long-read sequencing data suggest the presence of vector genomes containing ITR deletions in both the Bac-RepCap and SGMO Helper AAV preparations (Tables 4 and 5), the potency of the AAV from the SGMO Helper is significantly higher than the AAV from Bac-RepCap and similar to AAV produced using 293 cells (Figure S8). Since the only measurable difference between the preparations is the capsid ratio and potency (Table 1; Figure S8), but the ITR content is similar (Tables 4 and 5), the capsid ratio is more likely to influence AAV potency in this case than the presence/absence of ITRs or sequence truncations. To date, studies applying long-read sequencing techniques to AAV vector genomes have not utilized orthogonal techniques to examine AAV vector quality other than DNA gel electrophoresis, thus lack the data to correlate vector sequence and potency.<sup>31,32</sup> These results highlight the need to interpret AAV sequencing results with caution, as well as the need for purpose-built technology for sequencing individual vector genomes with high fidelity.

The utility of the Sf9-baculovirus system for AAV production is now well-established, but differences in the gene architecture of each system can impact yield and potency. Early efforts to produce AAV in Sf9 cells utilized up to five baculoviruses to deliver the *Rep* and *Cap* genes in addition to the vector genome (reviewed in Lubelski et al.<sup>43</sup>). One version suffered from genetic instability over time due to the presence of tandem *Rep* sequences that led to recombination and deletion of the *Rep* genes.<sup>5</sup> We have avoided this issue by placing the *Rep* genes on separate baculoviruses, preventing recombination events from occurring during cloning or passaging of the baculovirus. In contrast to the Urabe et al. baculovirus system, the improved baculovirus production system (here referred to as the Bac-RepCap) was shown to be genetically stable but relied on a leaky scanning mechanism to express the *Rep* and *Cap* proteins.<sup>5</sup> As a result, the amount of VP1 translation directly impacts the amount of VP3 translation, limiting the ability to increase VP1 incorporation without decreasing VP3 incorporation and reducing total AAV yield.<sup>9</sup> The SGMO Helper improves the ability to fine-tune the level of VP1 expression by maintaining the *VP1* gene on a separate baculovirus from the *VP2/3* gene while maintaining AAV yield.

Manufacturing of commercialized AAV products has so far been limited to natural AAV serotypes (AAV1, AAV2, and AAV9), but advances in AAV engineering demand greater flexibility for manufacturing of non-natural capsids. Mammalian cell AAV production relies on RepCap plasmids that utilize a combination of alternative promoters, RNA splicing, and leaky ribosomal scanning to

vg/kg of 293 AAV6 or Sf9 AAV6 produced using Bac-RepCap or the SGMO Helper. Twenty-eight days post-administration, mice were euthanized and plasma samples were collected to assay for  $\alpha$ -Gal A activity. The nmol/h/mL  $\alpha$ -Gal A activity was calculated based on fluorescent activity using a standard curve. p values were calculated using a two-way ANOVA with Tukey's multiple comparisons test (\*\*\*\* =  $p < 0.0001$ ).

express the full complement of Rep or Cap proteins from overlapping ORFs, whereas the Bac-RepCap developed in the Kotin lab for AAV production in Sf9 cells utilizes leaky ribosomal scanning from overlapping ORFs to express Rep and Cap proteins.<sup>1,5</sup> In both cases, the overlapping ORFs limit the ability to make modifications to an individual Rep or Cap protein. Separating the ORFs for the VP1 and VP2/3 genes increases the flexibility of the Sf9-baculovirus production system, providing a platform for the generation of mosaic capsids, where Cap proteins from different serotypes are used to assemble AAV capsids, deletion of VP2 to assemble capsids containing only VP1 and VP3, incorporation of VP2 proteins with N-terminal fusions, or domain insertion into VP1 only.<sup>38,44–49</sup> The separation of the Rep78 and Rep52 genes provides additional options for modification. Indeed, a recent study suggests that mosaic AAV2 Rep proteins containing Rep78 sequences from serotypes 1, 6, and 8 reduce the percentage of empty capsids in rAAV preparations.<sup>50</sup> In addition, the timing and level of Rep expression can be modulated to optimize rAAV production, either by modifying the translation start context or by using alternative baculovirus or insect promoters.

In summary, the SGM0 Helper offers distinct advantages over existing Sf9-baculovirus systems for AAV production by providing improved yields and potency. Indeed, the utility of the system for production of diverse AAV serotypes is an active area of investigation. Furthermore, the gene architecture of the SGM0 Helper lends itself to production of engineered AAVs by increasing options for modification of the Rep and Cap expression cassettes.

## MATERIALS AND METHODS

### Cell lines

Sf9 cells were obtained from ThermoFisher Scientific (12659017) and maintained in Sf-900 III SFM media according to the manufacturer's recommendation. HepG2 (HB-8065) and Hepa1-6 (CRL-1830) cells were obtained from ATCC and HuH-7 cells were obtained from the JCRB Cell Bank (JCRB0403). Mammalian cell lines were maintained in DMEM (Corning) supplemented with 10% fetal bovine serum (Seradigm).

### Transfer vector and bacmid generation

The pFastBac Dual Expression Vector was purchased from ThermoFisher. Site-directed mutagenesis primers (For: CTCGTTCCGCTGCAGGACAGAA, Rev: TTCTGTCTGCAGGCGAACGAG) were used to create an SbfI restriction site downstream of the HSV tk poly A site. The newly formed plasmid was then digested with BamHI for 60 min at 37 °C and column purified with Qiagen PCR Purification Kit. IDT synthesized Rep52 and Rep78 were then cloned into the digested fragment using Takara's Infusion Assembly kit, transformed into competent DH5 alpha cells, and plated on LB-Amp. Colonies were selected and plasmid DNA isolated using Qiagen Miniprep kit and subsequently sequenced with Illumina Nextera XT Library. Positive clones were then digested with XmaI and purified using Qiagen PCR Purification Kit. VP1 was then cloned into the Rep52 construct and VP2/3 was cloned into the Rep78 construct using Takara's Infusion Assembly kit. Colonies were selected and

plasmid DNA was obtained using Qiagen Mini-prep kit and sequenced with Illumina XT Library. Positive sequenced clones were selected and banked. The Rep52/VP1 plasmid was then digested with SbfI and the vector genome was cloned using T4 DNA Ligase (New England Biolabs). Colonies were once again selected, grown in LB-Amp, plasmid purified, and sequenced. The resulting transfer plasmids were used to generate bacmids by transformation of DH10Bac cells (ThermoFisher) according to the manufacturer's protocol. Positive clones identified by blue/white screening were grown in LB-Kan media and prepared according to the manufacturer's protocol. Prepared bacmids were sequence confirmed and used to generate baculovirus as described below.

### Baculovirus and BIIC production

Recombinant baculovirus was generated by transfection of Sf9 cells using the TransIT-Insect transfection reagent (Mirus Bio) according to the manufacturer's instructions. Seventy-two hours post-transfection, the baculovirus infectious titer was determined using the BacPAK Baculovirus Rapid Titer Kit (Takara Bio). BIIC production was performed as described in Wasilko and Lee.<sup>23</sup> Briefly, naive Sf9 cells were infected with baculovirus using an MOI of 0.1 and, after 72 h, assayed for greater than 2- $\mu$ m increase in cell diameter using a Countess Cell Counter (ThermoFisher). BIIC cultures that showed a greater than 2- $\mu$ m increase in cell diameter were harvested and frozen in medium containing fresh Sf-900 III SFM media (ThermoFisher), 10% DMSO (Sigma), and 10 mg/mL recombinant human serum albumin (Sigma) using a Mr. Frosty freezing container (ThermoFisher).

### Serial passaging of baculovirus

Serial passaging of baculovirus was performed by thawing of BIIC, diluting 1:100 in fresh media, and inoculating a naive culture of Sf9 cells with a 1:100 volume of diluted BIIC. After 72 h, a 1-mL sample of the culture was harvested for sequencing and western blot analysis. In addition, a 1:100 volume was used to inoculate a fresh naive culture of Sf9 cells. This process was repeated for six passages and sequencing and western blot analysis was performed for each passage. Supernatant from passages three and six was used to generate BIICs, as described above. The BIICs from passages three and six were used to initiate AAV production, as described below.

### AAV production

To initiate AAV production, naive Sf9 cells were inoculated with BIICs derived from the Rep78-VP2/3 Helper and the Rep52-VP1 Helper containing an AAV vector genome. After 6 days, cells were collected by centrifugation, resuspended in TMS (50 mM Tris, 150 mM NaCl, 2 mM MgCl<sub>2</sub>) and freeze/thawed three times by incubation in a dry ice-ethanol bath and 37 °C waterbath. Crude viral lysate was treated with Denarase nuclease (c-LEcta), clarified by centrifugation and precipitated by the addition of PEG/NaCl (31% PEG-8000, 2 M NaCl) and incubation on ice. Precipitate was collected by centrifugation, the pellet was resuspended in HSS (50 mM HEPES, 1.15 M NaCl, 20 mM EDTA), and the suspension was applied to a CsCl density gradient composed of 1.3 g/cm<sup>2</sup> and 1.5 g/cm<sup>2</sup> CsCl

in 1XDPBS. The gradient was centrifuged for 18 h at 28,000 rpm and the AAV band was collected and dialyzed into formulation buffer (1XDPBS, 1% sucrose, 0.05% Kolliphor, 35 mM NaCl, pH 7.1).

#### AAV titering

Five microliters of purified virus was added to 45  $\mu$ L of DNase I mix (ThermoFisher) and incubated at 37 °C for 60 min. The samples were then diluted in 1x GeneAmp PCR Buffer (ThermoFisher) at three different dilutions. Twenty microliters of the diluted samples was added to the TaqMan Universal Master Mix (ThermoFisher) in a 96-well plate with a gene-specific primer probe alongside a quantitative standard curve. Plates were sealed and samples were heated to 50 °C for 2 min, denatured at 95 °C for 10 min, and followed by 40 cycles (95 °C for 15 s, 60 °C for 1 min) in the QuantStudio 3 Real-Time PCR System (ThermoFisher). Analysis was performed using QuantStudio Design and Analysis software, back-calculating for the dilution factor.

#### SDS-PAGE and Coomassie staining

AAV samples were analyzed by SDS-PAGE and Coomassie staining by combining approximately 2E11 vg with NuPAGE LDS Sample Buffer (ThermoFisher) and Sample Reducing Agent (ThermoFisher) and incubating at 95 °C. Samples were loaded on NuPAGE 4%–12%, Bis-Tris gels (ThermoFisher) along with the SeeBlue Plus 2 Pre-stained protein standard (ThermoFisher) and electrophoresed for 100 min at 120 V. Gels were removed and stained using the SimplyBlue SafeStain (ThermoFisher) according to the rapid staining protocol. Stained gels were imaged using a Bio-Rad ChemiDoc Imaging System and analyzed using the Image Lab Software package (Bio-Rad Laboratories).

#### CE-SDS analysis

AAV samples (1.25E12 vg per sample) were run on the Maurice instrument (ProteinSimple) according to the manufacturer's instructions. The area under the curve for the peaks corresponding to VP1, VP2, and VP3 was determined and used to calculate the capsid ratio.

#### Capsid ELISA

The capsid titer was determined using an AAV6-specific capture ELISA (ProGen Biotechnik) according to the manufacturer's instructions. The vg:capsid ratio was calculated by dividing the vg titer (vg/mL) by the capsid titer (capsid/mL).

#### Western blot analysis

Cell pellets were collected, lysed using RIPA buffer (Sigma), and clarified by centrifugation. Samples were combined with NuPAGE LDS Sample Buffer (ThermoFisher) and Sample Reducing Agent (ThermoFisher) and incubated at 95 °C. Samples were loaded on NuPAGE 4%–12%, Bis-Tris gels (ThermoFisher) along with the SeeBlue Plus 2 Pre-stained protein standard (ThermoFisher) and electrophoresed for 100 min at 120 V. Gels were transferred to PVDF membranes using an iBlot 2 Gel Transfer Device (ThermoFisher) according to the manufacturer's protocol. Membranes were probed using either an antibody to AAV Rep protein (clone 259.5, American Research Products, Inc.) or AAV Cap protein (B1, Progen Biotechnik) and detected using a DyLight800-tagged goat

anti-mouse secondary antibody (ThermoFisher). Membranes were imaged using a Bio-Rad ChemiDoc Imaging System (Bio-Rad Laboratories).

#### Sequencing of transfer vectors, bacmids, baculoviruses, and AAV

Transfer vectors and bacmids were prepared for sequencing using the Illumina Nextera XT kit according to the manufacturer's instructions. For baculoviruses and DNA from AAV preparations, samples were extracted using the Qiagen Blood and Tissue Kit Spin-Column protocol for Animal Blood or Cells, then prepared for sequencing using the Illumina Nextera XT kit. Nextera sample libraries were quantitated using a Qubit according to the manufacturer's instructions, then diluted and loaded onto an Illumina Miseq according to manufacturer's instructions.

#### Mapping of DNA contaminants in AAV preparations

Briefly, demultiplexed sample reads were adapter trimmed using trim\_galore (0.6.6) to remove Nextera adapter sequence, filtered for sequence length  $\geq 50$  base pairs (bp), and quality filtered such that all quality scores are  $\geq 20$  using a custom python script. R1 reads were then sequentially aligned using bowtie2 in end-to-end mode vs. a variety of possible nucleic acid contaminant templates.<sup>51</sup> At each step, hits (defined as reads aligning end-to-end with at least 92% of the read bases aligning, and with an edit distance of less than 8% of length) are removed, and all other reads proceed to alignment against the next template, for all templates in the alignment list. The reason for not using a single merged target database is to avoid multimapped alignments. Reads that might align perfectly to multiple templates during the process will be counted as originating from the first template aligned against (e.g., a transgene segment that has homology to a human gene will be counted as aligning to the transgene if it is first in alignment order). Chimeric reads (e.g., 50-bp transgene, 50-bp foreign DNA) will not be counted as hits and will remain in the unmapped portion of reads after all alignments are complete, as will reads arising from a contaminant not in the alignment list.

#### Long-read sequencing of encapsidated DNA

Viral DNA extraction, library preparations, sequencing reactions, and bioinformatics analysis were conducted at Azenta Life Sciences. (South Plainfield, NJ, USA).

#### Alkaline gel electrophoresis

Samples (2E11 vg total) were combined with 6X Purple Loading Buffer (New England Biolabs) and 10X alkaline buffer (500 mM NaOH, 10 mM EDTA), heated to 95 °C and cooled to room temperature. Samples were run on a 1% agarose gel in 1X alkaline gel buffer overnight at 20 V. The gel was neutralized in 1X TAE (ThermoFisher) for 30 min, stained with ethidium bromide (Sigma) and visualized using a Bio-Rad ChemiDoc Imaging System (Bio-Rad Laboratories).

#### In vitro potency assay

For the  $\alpha$ -galactosidase A (GLA) potency assay, HepG2, HuH-7, or Hepa1-6 cells were plated at a sub-confluent density in 96-well plates

and incubated overnight. Cells were transduced using the indicated MOI and incubated for 6 days. Cell supernatants were collected and the GLA activity was detected by incubating with 5 mM 4-Methylumbelliferyl- $\alpha$ -D-galactopyranoside (Sigma) for 1 h at 37 °C, the reaction was stopped by the addition of 0.5 M Glycine, pH 10.7, and the fluorescent signal was detected using a fluorescent plate reader. The fluorescent signal was converted to nmol/h/mL using a standard curve derived from an 8-point 2X dilution of 1 mM 4-Methylumbelliferyl (Sigma).

For the FIX potency assay, HepG2 cells were plated at a sub-confluent density in 48-well plates and incubated overnight. Cells were transduced at an MOI of 1E06 vg/cell and incubated for 6 days. Cell supernatants were collected and the FIX level was determined using the VisuLize Factor IX Antigen Kit (Affinity Biologicals, Inc.) according to the manufacturer's instructions.

### Mouse study

The study complied with all applicable sections of the current version of the Final Rules of the Animal Welfare Act regulations (CFR, Title 9), and the Guide for the Care and Use of Laboratory Animals, Institute for Laboratory Animal Research, National Research Council, Eighth Edition. The study did not unnecessarily duplicate previous experiments and the sponsor did consider alternatives to the use of live animals. Procedures used in the study were designed with the consideration of the well-being of the animals. The protocol and any amendments or procedures involving the care or use of animals in the study were reviewed and approved by the Testing Facilities Institutional Animal Care and Use Committee (IACUC) before the initiation of such procedures. Test articles formulated at the indicated concentration were administered to male wild-type C57BL/6 mice by tail vein injection. All mice were euthanized at 29-days post-administration and plasma samples were collected at necropsy. Plasma samples were diluted in 1XDPBS and assayed for GLA activity, as described above.

### DATA AND CODE AVAILABILITY

Data generated can be found within the published article and its [supplemental information](#).

### SUPPLEMENTAL INFORMATION

Supplemental information can be found online at <https://doi.org/10.1016/j.omtm.2024.101228>.

### ACKNOWLEDGMENTS

These studies were supported by Sangamo Therapeutics. The authors would like to acknowledge the Technical Development, Non-clinical, and Research teams for their insightful comments and critical data and manuscript review. The authors would also like to thank Anna Pavlova and Austin Lim for their assistance with the CE-SDS analysis and Anastasia Friedberger for her assistance with long-read sequence analysis.

### AUTHOR CONTRIBUTIONS

H.T. performed cloning of the SGMO Helper; A.M.W., Y.M., H.-O.N., D.S., and T.W. performed AAV production and testing; S.B. performed huFIX assays; G.E. performed ELISAs; K.M. and G.P. managed the mouse study; D.S. and G.L. performed bioinformatic analysis; M.T. and J.P. performed mid-scale AAV production; and R.S. provided technical guidance and detailed manuscript review.

### DECLARATION OF INTERESTS

A.M.W., Y.M., H.-O.N., D.S., S.B., G.E., K.M., G.L., M.T., H.T., and J.P. are current employees and stockholders of Sangamo Therapeutics. G.P., D.S., R.S., and T.W. are former employees of Sangamo Therapeutics and may hold stock.

### REFERENCES

- Berns, K.I., and Parrish, C.R. (2013). Parvoviridae. In *Field's Virology*, D.M. Knipe and P. Howley, eds. (Lippincott Williams and Wilkins), pp. 1768–1791.
- Wang, D., Tai, P.W.L., and Gao, G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. *Nat. Rev. Drug Discov.* *18*, 358–378.
- Meier, A.F., Fraefel, C., and Seyffert, M. (2020). The Interplay between Adeno-Associated Virus and its Helper Viruses. *Viruses* *12*.
- Joshi, P.R.H., Venereo-Sanchez, A., Chahal, P.S., and Kamen, A.A. (2021). Advancements in molecular design and bioprocessing of recombinant adeno-associated virus gene delivery vectors using the insect-cell baculovirus expression platform. *Biotechnol. J.* *16*, e2000021.
- Smith, R.H., Levy, J.R., and Kotin, R.M. (2009). A simplified baculovirus-AAV expression vector system coupled with one-step affinity purification yields high-titer rAAV stocks from insect cells. *Mol. Ther.* *17*, 1888–1896.
- Mietzsch, M., Hering, H., Hammer, E.-M., Agbandje-McKenna, M., Zolotukhin, S., and Heilbronn, R. (2017). OneBac 2.0: Sf9 Cell Lines for Production of AAV1, AAV2, and AAV8 Vectors with Minimal Encapsulation of Foreign DNA. *Hum. Gene Ther. Methods* *28*, 15–22.
- Kondratov, O., Marsic, D., Crosson, S.M., Mendez-Gomez, H.R., Moskalenko, O., Mietzsch, M., Heilbronn, R., Allison, J.R., Green, K.B., Agbandje-McKenna, M., and Zolotukhin, S. (2017). Direct Head-to-Head Evaluation of Recombinant Adeno-associated Viral Vectors Manufactured in Human versus Insect Cells. *Mol. Ther.* *25*, 2661–2675.
- Aslanidi, G., Lamb, K., and Zolotukhin, S. (2009). An inducible system for highly efficient production of recombinant adeno-associated virus (rAAV) vectors in insect Sf9 cells. *Proc. Natl. Acad. Sci. USA* *106*, 5059–5064.
- Kohlbrenner, E., Aslanidi, G., Nash, K., Shklyae, S., Campbell-Thompson, M., Byrne, B.J., Snyder, R.O., Muzyczka, N., Warrington, K.H., Jr., and Zolotukhin, S. (2005). Successful production of pseudotyped rAAV vectors using a modified baculovirus expression system. *Mol. Ther.* *12*, 1217–1225.
- Urabe, M., Nakakura, T., Xin, K.-Q., Obara, Y., Mizukami, H., Kume, A., Kotin, R.M., and Ozawa, K. (2006). Scalable generation of high-titer recombinant adeno-associated virus type 5 in insect cells. *J. Virol.* *80*, 1874–1885.
- Urabe, M., Ding, C., and Kotin, R.M. (2002). Insect cells as a factory to produce adeno-associated virus type 2 vectors. *Hum. Gene Ther.* *13*, 1935–1943.
- Galibert, L., Savy, A., Dickx, Y., Bonnin, D., Bertin, B., Mushimiyimana, I., van Oers, M.M., and Merten, O.-W. (2018). Origins of truncated supplementary capsid proteins in rAAV8 vectors produced with the baculovirus system. *PLoS One* *13*, e0207414.
- Grieger, J.C., Snowdy, S., and Samulski, R.J. (2006). Separate basic region motifs within the adeno-associated virus capsid proteins are essential for infectivity and assembly. *J. Virol.* *80*, 5199–5210.
- Sonntag, F., Bleker, S., Leuchs, B., Fischer, R., and Kleinschmidt, J.A. (2006). Adeno-associated virus type 2 capsids with externalized VP1/VP2 trafficking domains are

- generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus. *J. Virol.* *80*, 11040–11054.
15. Kronenberg, S., Böttcher, B., von der Lieth, C.W., Bleker, S., and Kleinschmidt, J.A. (2005). A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini. *J. Virol.* *79*, 5296–5303.
  16. Girod, A., Wobus, C.E., Zádori, Z., Ried, M., Leike, K., Tijssen, P., Kleinschmidt, J.A., and Hallek, M. (2002). The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. *J. Gen. Virol.* *83*, 973–978.
  17. Bleker, S., Sonntag, F., and Kleinschmidt, J.A. (2005). Mutational analysis of narrow pores at the fivefold symmetry axes of adeno-associated virus type 2 capsids reveals a dual role in genome packaging and activation of phospholipase A2 activity. *J. Virol.* *79*, 2528–2540.
  18. Galibert, L., Jacob, A., Savy, A., Dickx, Y., Bonnin, D., Lecomte, C., Rivollet, L., Sanatine, P., Boutin Fontaine, M., Le Bec, C., and Merten, O.W. (2021). Monobac System-A Single Baculovirus for the Production of rAAV. *Microorganisms* *9*, 1799.
  19. Kearse, M.G., and Wilusz, J.E. (2017). Non-AUG translation: a new start for protein synthesis in eukaryotes. *Genes Dev.* *31*, 1717–1731.
  20. Bosma, B., du Plessis, F., Ehlert, E., Nijmeijer, B., de Haan, M., Petry, H., and Lubelski, J. (2018). Optimization of viral protein ratios for production of rAAV serotype 5 in the baculovirus system. *Gene Ther.* *25*, 415–424.
  21. Warrington, K.H., Jr., Gorbatyuk, O.S., Harrison, J.K., Opie, S.R., Zolotukhin, S., and Muzyczka, N. (2004). Adeno-associated virus type 2 VP2 capsid protein is nonessential and can tolerate large peptide insertions at its N terminus. *J. Virol.* *78*, 6595–6609.
  22. Pijlman, G.P., van den Born, E., Martens, D.E., and Vlak, J.M. (2001). Autographa californica baculoviruses with large genomic deletions are rapidly generated in infected insect cells. *Virology* *283*, 132–138.
  23. Wasilko, D., and Lee, S.E. (2006). TIPS: Titerless Infected-Cells Preservation and Scale-Up. *Bioprocess J.* *5*, 29–32.
  24. Oyama, H., Ishii, K., Maruno, T., Torisu, T., and Uchiyama, S. (2021). Characterization of Adeno-Associated Virus Capsid Proteins with Two Types of VP3-Related Components by Capillary Gel Electrophoresis and Mass Spectrometry. *Hum. Gene Ther.* *32*, 1403–1416.
  25. Onishi, T., Nonaka, M., Maruno, T., Yamaguchi, Y., Fukuhara, M., Torisu, T., Maeda, M., Abbatello, S., Haris, A., Richardson, K., et al. (2023). Enhancement of recombinant adeno-associated virus activity by improved stoichiometry and homogeneity of capsid protein assembly. *Mol. Ther. Methods Clin. Dev.* *31*, 101142.
  26. Walker, B.J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., Cuomo, C.A., Zeng, Q., Wortman, J., Young, S.K., and Earl, A.M. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* *9*, e112963.
  27. Strobel, B., Miller, F.D., Rist, W., and Lamla, T. (2015). Comparative Analysis of Cesium Chloride- and Iodixanol-Based Purification of Recombinant Adeno-Associated Viral Vectors for Preclinical Applications. *Hum. Gene Ther. Methods* *26*, 147–157.
  28. Ayuso, E., Mingozzi, F., and Bosch, F. (2010). Production, purification and characterization of adeno-associated vectors. *Curr. Gene Ther.* *10*, 423–436.
  29. Lock, M., Alvira, M., Vandenberghe, L.H., Samanta, A., Toelen, J., Debyser, Z., and Wilson, J.M. (2010). Rapid, simple, and versatile manufacturing of recombinant adeno-associated viral vectors at scale. *Hum. Gene Ther.* *21*, 1259–1271.
  30. Urabe, M., Xin, K.-Q., Obara, Y., Nakakura, T., Mizukami, H., Kume, A., Okuda, K., and Ozawa, K. (2006). Removal of Empty Capsids from Type 1 Adeno-Associated Virus Vector Stocks by Anion-Exchange Chromatography Potentiates Transgene Expression. *Mol. Ther.* *13*, 823–828.
  31. Zhang, J., Guo, P., Yu, X., Frabutt, D.A., Lam, A.K., Mulcrone, P.L., Chrzanowski, M., Firman, J., Pouchnik, D., Sang, N., et al. (2022). Subgenomic particles in rAAV vectors result from DNA lesion/break and non-homologous end joining of vector genomes. *Mol. Ther. Nucleic Acids* *29*, 852–861.
  32. Tran, N.T., Lecomte, E., Saleun, S., Namkung, S., Robin, C., Weber, K., Devine, E., Blouin, V., Adjali, O., Ayuso, E., et al. (2022). Human and Insect Cell-Produced Recombinant Adeno-Associated Viruses Show Differences in Genome Heterogeneity. *Hum. Gene Ther.* *33*, 371–388.
  33. Brimble, M.A., Winston, S.M., and Davidoff, A.M. (2023). Stowaways in the cargo: Contaminating nucleic acids in rAAV preparations for gene therapy. *Mol. Ther.* *31*, 2826–2838.
  34. Brister, J.R., and Muzyczka, N. (1999). Rep-mediated nicking of the adeno-associated virus origin requires two biochemical activities, DNA helicase activity and transesterification. *J. Virol.* *73*, 9325–9336.
  35. Xie, Q., Bu, W., Bhatia, S., Hare, J., Somasundaram, T., Azzi, A., and Chapman, M.S. (2002). The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc. Natl. Acad. Sci. USA* *99*, 10405–10410.
  36. Govindasamy, L., DiMattia, M.A., Gurda, B.L., Halder, S., McKenna, R., Chiorini, J.A., Muzyczka, N., Zolotukhin, S., and Agbandje-McKenna, M. (2013). Structural insights into adeno-associated virus serotype 5. *J. Virol.* *87*, 11187–11199.
  37. Büning, H., and Srivastava, A. (2019). Capsid Modifications for Targeting and Improving the Efficacy of AAV Vectors. *Mol. Ther. Methods Clin. Dev.* *12*, 248–265.
  38. Chai, Z., Zhang, X., Dobbins, A.L., Frost, E.A., Samulski, R.J., and Li, C. (2019). Chimeric Capsid Proteins Impact Transduction Efficiency of Haploid Adeno-Associated Virus Vectors. *Viruses* *11*.
  39. Wilmott, P., Lisowski, L., Lisowski, L., Alexander, I.E., Logan, G.J., Lisowski, L., Logan, G.J., and Logan, G.J. (2019). A User's Guide to the Inverted Terminal Repeats of Adeno-Associated Virus. *Hum. Gene Ther. Methods* *30*, 206–213.
  40. Benjamini, Y., and Speed, T.P. (2012). Summarizing and correcting the GC content bias in high-throughput sequencing. *Nucleic Acids Res.* *40*, e72.
  41. Penaud-Budloo, M., Lecomte, E., Guy-Duché, A., Guy-Duché, A., Roulet, A., Lopez-Roques, C., Tournaire, B., Cogné, B., Léger, A., Léger, A., et al. (2017). Accurate Identification and Quantification of DNA Species by Next-Generation Sequencing in Adeno-Associated Viral Vectors Produced in Insect Cells. *Hum. Gene Ther. Methods* *28*, 148–162.
  42. Calvo-Roitberg, E., Daniels, R.F., and Pai, A.A. (2023). Challenges in identifying mRNA transcript starts and ends from long-read sequencing data. Preprint at bioRxiv. <https://doi.org/10.1101/2023.07.26.550536>.
  43. Lubelski, J., Hermens, W., and Petry, H. (2014). Insect cell-based recombinant adeno-associated virus production: molecular process optimization. *Bioprocess J.* *13*, 6–11.
  44. Kawabata, H., Konno, A., Matsuzaki, Y., and Hirai, H. (2023). A blood-brain barrier-penetrating AAV2 mutant created by a brain microvasculature endothelial cell-targeted AAV2 variant. *Mol. Ther. Methods Clin. Dev.* *29*, 81–92.
  45. Lux, K., Goerlitz, N., Schlemminger, S., Perabo, L., Goldnau, D., Endell, J., Leike, K., Kofler, D.M., Finke, S., Hallek, M., and Büning, H. (2005). Green Fluorescent Protein-Tagged Adeno-Associated Virus Particles Allow the Study of Cytosolic and Nuclear Trafficking. *J. Virol.* *79*, 11776–11787.
  46. Krooss, S.A., Dai, Z., Schmidt, F., Rovai, A., Fakhiri, J., Dhingra, A., Yuan, Q., Yang, T., Balakrishnan, A., Steinbrück, L., et al. (2020). Ex Vivo/In vivo Gene Editing in Hepatocytes Using “All-in-One” CRISPR-Adeno-Associated Virus Vectors with a Self-Linearizing Repair Template. *iScience* *23*, 100764.
  47. (2023). A mosaic adeno-associated virus vector as a versatile tool that exhibits high levels of transgene expression and neuron specificity in primate brain. *Nat. Commun.* *14*, 1–17.
  48. Hauck, B., Xu, R.R., Xie, J., Wu, W., Ding, Q., Sipler, M., Wang, H., Chen, L., Wright, J.F., and Xiao, W. (2006). Efficient AAV1-AAV2 hybrid vector for gene therapy of hemophilia. *Hum. Gene Ther.* *17*, 46–54.
  49. Hoffmann, M.D., Zdechlik, A.C., He, Y., Nedrud, D., Aslanidi, G., Gordon, W., and Schmidt, D. (2023). Multiparametric domain insertional profiling of adeno-associated virus VP1. *Mol. Ther. Methods Clin. Dev.* *31*, 101143.
  50. Mietzsch, M., Eddington, C., Jose, A., Hsi, J., Chipman, P., Henley, T., Choudhry, M., McKenna, R., and Agbandje-McKenna, M. (2021). Improved Genome Packaging Efficiency of Adeno-associated Virus Vectors Using Rep Hybrids. *J. Virol.* *95*, e0077321.
  51. Langmead, B., Trapnell, C., Pop, M., and Salzberg, S.L. (2009). Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* *10*, R25.