

ARTICLE

The impact of minimally oversized adeno-associated viral vectors encoding human factor VIII on vector potency *in vivo*

Sirkka Kyostio-Moore¹, Patricia Berthelette¹, Susan Piraino¹, Cathleen Sookdeo¹, Bindu Nambiar¹, Robert Jackson¹, Brenda Burnham¹, Catherine R O'Riordan¹, Seng H Cheng¹ and Donna Armentano¹

Recombinant adeno-associated viral (rAAV) vectors containing oversized genomes provide transgene expression despite low efficiency packaging of complete genomes. Here, we characterized the properties of oversized rAAV2/8 vectors (up to 5.4 kb) encoding human factor VIII (FVIII) under the transcriptional control of three liver promoters. All vectors provided sustained production of active FVIII in mice for 7 months and contained comparable levels of vector genomes and complete expression cassettes in liver. Therefore, for the 5.4 kb genome size range, a strong expression cassette was more important for FVIII production than the vector genome size. To evaluate the potency of slightly oversized vectors, a 5.1 kb AAVrh8R/FVIII vector was compared to a 4.6 kb (wild-type size) vector with an identical expression cassette (but containing a smaller C1-domain deleted FVIII) for 3 months in mice. The 5.1 kb vector had twofold to threefold lower levels of plasma FVIII protein and liver vector genomes than that obtained with the 4.6 kb vector. Vector genomes for both vectors persisted equally and existed primarily as high molecular weight concatemeric circular forms in liver. Taken together, these results indicate that the slightly oversized vectors containing heterogeneously packaged vector genomes generated a functional transgene product but exhibited a twofold to threefold lower *in vivo* potency.

Molecular Therapy — Methods & Clinical Development (2016) **3**, 16006; doi:10.1038/mtm.2016.6; published online 24 February 2016

INTRODUCTION

Recombinant AAV (rAAV) vectors are gaining interest as gene delivery vehicles for treating genetic diseases. Their small packaging capacity is a limitation that has hindered their consideration for delivering large transgenes such as factor VIII (FVIII), dystrophin, dysferlin, and cystic fibrosis transmembrane conductance regulator. Early studies have defined their packaging limit at 4.7 to 4.8 kb although this result was later challenged by the packaging of 6 kb genomes.^{1,2} However, more recent studies have indicated that the packaged vector genomes for AAV2, AAV5, or AAV8 serotype vectors do not exceed 5.2 kb and that genomes from oversized vectors are typically deleted at the 5' ends of both strand polarities.^{3–6}

Proof of therapeutic concept for hemophilia A has been shown using oversized rAAV ("fragmented") vectors in murine and canine models of the disease. Both the production of therapeutic levels of active FVIII protein and subsequent correction of the bleeding disorder was demonstrated.^{3,7–11} A strategy employing two wild-type sized vectors to separately encode the FVIII heavy and light chains eliminated the packaging of deleted vector genomes but showed an imbalance of heavy and light chain production.¹² To overcome this issue, an alternative system that generated full-length expression cassettes *in vivo* following concatemer formation and intron splicing ("split" or "trans-splicing vectors") was developed.¹³ Comparison of the split vector system to that of an oversized 6.2 kb rAAV6 vector expressing luciferase in mouse muscle, eye, and liver revealed

that the latter generated higher levels of expression.¹⁴ Moreover, as the production and delivery of two vectors is more complex, the oversized vectors currently represent a more practical approach for gene transfer of large transgenes. Hence, a better understanding of the merits and limitations of oversized vectors would be helpful for their further development.

The goals of our study were to characterize the attributes and challenges associated with generating moderately oversized rAAV vectors. We examined the relative importance of the size of the genome and the selection of the expression cassette on the packaging efficiency and gene transfer activity. To accomplish this objective, we generated varying sizes of rAAV vectors encoding FVIII and tested their activity in the context of either a pre-viral plasmid or a rAAV vector. We observed delayed formation of double-stranded genomes from oversized vectors and their overall potency for gene transfer and expression was reduced twofold to threefold.

RESULTS

Activity of FVIII-encoding plasmid expression cassettes *in vitro* and *in vivo*

We constructed a number of FVIII expression cassettes using B-domain deleted human FVIII and minimally sized liver-specific promoters, introns, and polyadenylation (polyA) signals (Figure 1a and Supplementary Table S1). The selection of promoters tested included a liver promoter (LP1),¹⁵ mouse or human transthyretin

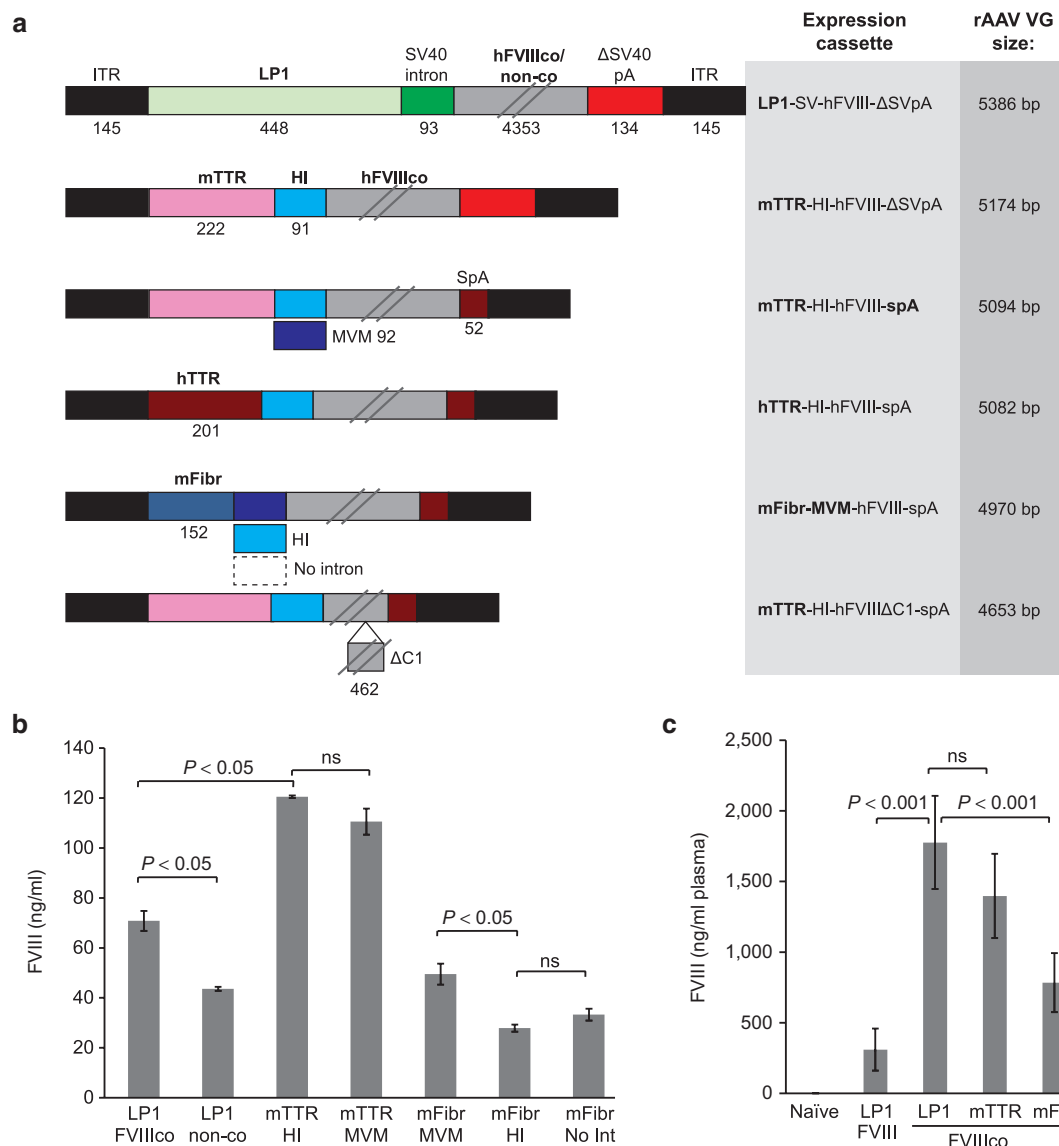


Figure 1 Generation and testing of human FVIII expression cassettes in previral plasmids. **(a)** Description of FVIII expression cassettes used in previral plasmids. Vector genome elements flanked by AAV2 ITRs and their sizes (bp) in each construct are shown (to-scale except FVIII cDNA). **(b)** FVIII expression *in vitro*. Human Huh7 cells were transfected with previral plasmids followed by quantitation of FVIII protein levels by ELISA in culture media. The experiment was performed four times using $n = 3$ for each construct with comparable results. **(c)** FVIII expression *in vivo*. Previral plasmids were administered by high-volume injection into female C57BL/6 mice, and FVIII plasma levels were measured by ELISA on day 1. The experiment was performed three times using $n = 7-10$ mice for each construct and the result from one representative study is shown. Values in each graph represent the average \pm SD, and significance was calculated using Student's *t*-test with Bonferroni correction. hFVIIIco, codon-optimized human FVIII Δ B cDNA; non-co, native FVIII DNA sequence; ITR, inverted terminal repeat; LP1, liver promoter; mTTR, mouse transthyretin promoter; hTTR, human transthyretin promoter; HI, hybrid intron; MVM, minute virus of mouse intron; Δ SV40 pA, truncated simian virus 40 polyA; spA, synthetic polyA; mFibr, modified human fibrinogen promoter; Δ C1, deletion of the hFVIII C1 domain; VG, vector genome.

promoter (mTTR, hTTR),^{7,16} and a novel modified human fibrinogen (mFibr) promoter with sizes of 448, 222, and 152 bp, respectively. All of the promoters contained an intron sequence (from minute virus of mice, simian virus (SV) 40, or human IgG- β globin hybrid intron (HI)) and a short synthetic polyA signal (spA) or SV40 polyA.^{17,18} The expression cassettes were flanked by inverted terminal repeats (ITRs) from AAV2 resulting in vector sizes ranging from 5.0 to 5.4 kb. The plasmid vectors were tested for FVIII production by transient transfection into Huh7 cells (Figure 1b). The data showed that the highest levels of FVIII were obtained when the mTTR promoter (comparable levels with the HI or minute virus of mice intron) was used followed by the LP1 and mFibr promoters. The hTTR and mFibr

promoters exhibited a similarly low activity when tested *in vitro* (Supplementary Figure S1). Since the small spA signal provided comparable activity as the larger SV40 pA signal sequence, we elected to use spA in the subsequent constructs (Supplementary Figure S1). For the mFibr promoter, the incorporation of an intron or its elimination had little effect on FVIII levels (Figure 1b). Lastly, when we tested the effect of codon-optimization of human FVIII in the context of the LP1 promoter, we observed a modest 1.6-fold increase in FVIII production *in vitro*.

The expression cassettes in the plasmid vectors were next tested by hydrodynamic tail vein injections into normal C57BL/6 mice (Figure 1c). In contrast to the modest effect seen *in vitro*, the use

of the codon-optimized hFVIII cassette generated sixfold higher plasma FVIII levels than the nonoptimized sequence at 24 hours post administration. Comparable plasma FVIII levels were obtained with the mTTR and LP1 promoters while the lowest levels were obtained with the mFibr promoter. Moreover, the inclusion of an intron in the mFibr construct also enhanced FVIII expression (not shown).

Generation and characterization of oversized rAAV8/FVIII vectors

The rAAV8 vectors with different genome sizes noted in Figure 1 (ranging between 5.0 and 5.4 kb) were generated using the triple transfection method and purified over a CsCl gradient. Analysis of the viral preparations showed comparable levels of VP1, VP2, and VP3 capsid proteins for each construct except for the LP1-FVIII construct which showed a slightly higher intensity (Figure 2a and Supplementary Table S2). All vector lots also exhibited comparable heat stability and exposure of the VP1 N-terminal epitope (recognized by an A1 antibody) to that noted for a control AAV8 vector encoding a 4.7 kb-sized (wild-type sized) factor IX (FIX) expression cassette (Figure 2b,c). These observations suggested that no conformational changes were generated in the capsid structure despite the attempt to package larger genomes.

The quantitation of vector yields revealed a trend toward lower vector yields with greater genome sizes (Figure 3a). The yields for the largest vector, AAV8/LP1-FVIII (5.4 kb), were approximately fourfold lower than those for the 4.7 kb AAV8/LP1-FIX vector. Increasing vector genome size also resulted in an increase in the packaging of the plasmid backbone, as determined by the presence of increasing amounts of the ampicillin resistance gene in the preparations (Figure 3a). Southern blot analysis of the viral preparations using a FVIII-specific probe showed that the majority of genomes were ~4.6 kb in size, indicating that genomes greater than 5.0 kb were not effectively packaged (not shown). There was also a correlation between the size of the expression cassette and an increase in the signal for the plasmid backbone (ampicillin gene) that is indicative of aberrant packaging (not shown). The quality of the viral vectors was also assessed using analytical ultracentrifugation to determine the homogeneity of the preparation.¹⁹ All oversized vectors (greater than 4.7 kb) exhibited two major peaks with sedimentation values of ~84S and 102S (Figure 3b). The 102-103S species represent capsids harboring a wild-type size vector genome ~4.6–4.7 kb in agreement with the Southern blot analysis (data not shown), while the 84-85S species represent capsids containing smaller, aberrant genomes, predominantly ~2.2 kb in size.¹⁹ Interestingly, we observed increasing proportions of viral vectors in the 102S peak as the size of the vector genome was decreased. These results are in agreement with the size-dependent differences observed in the qPCR and Southern blot analyses noted above.

Human FVIII production from oversized rAAV8 vectors *in vivo*

All oversized AAV8/FVIII vectors (between 5.0 and 5.4 kb) administered at 3×10^{11} vector genomes (vg)/mouse produced high plasma levels of FVIII protein that were sustained over 7 months following a single tail vein injection (Figure 4a). Analysis of the FVIII activity at day 141 postinjection showed that it was functional, as assessed by a Coatstest FVIII assay (Figure 4b). The activity results also showed the same ratio and rank order for the constructs as was obtained by the FVIII ELISA method suggesting similar specific activities. Analysis of the livers at 7 months posttreatment revealed that all livers harbored similar numbers of vector genomes irrespective of

the vector used (Figure 4c). Analysis of FVIII mRNA and protein levels at this time point indicated a similar rank order for expression, as noted earlier for their corresponding previral plasmids (Figures 1c and 4d,e). Therefore, as the number of vector genomes in the livers of mice transduced by the three vectors was comparable, the differential FVIII mRNA and protein levels measured were likely due to differences in the strength of the expression cassettes. Taken together, these data indicate that each viral vector generated full-length vector genomes and that the amount of FVIII production was primarily dependent on the design of the expression cassette as opposed to the vector genome size.

Potency of rAAVrh8R/FVIII vectors *in vivo*

As the expression of FVIII levels from the mTTR promoter was comparable to that from the LP1 promoter, the mTTR promoter with the 5.1 kb expression cassette was selected for further evaluation of the potency of an oversized vector. Its potency was compared to a wild-type size 4.6 kb vector that contained an identical transcriptional cassette except for a deletion of the FVIII C1 domain. This modification was introduced to minimize differences in DNA sequences from the 5.1 kb vector, especially close to the terminal ends (Figure 1a). To address the potential for reduced expression and/or faster clearance of the C1-domain deleted FVIII, previral plasmids harboring the 5.1 and 4.6 kb vector genomes were evaluated *in vitro* and *in vivo* (via high volume injections into C57BL/6 mice) for FVIII expression. The FVIII Δ C1 protein was detected both *in vitro* and *in vivo* at 0.2- to 0.3-fold lower levels than that of the intact FVIII construct (Figure 5 a,b). Hence, lower production of FVIII rather than an increased clearance was likely responsible for the observed differences between these two vector configurations. The *in vivo* stability of the FVIII Δ C1 protein was evaluated by determining the size of the expressed FVIII protein in plasma using an antibody specific to the 90 kDa heavy chain of FVIII (Figure 5c). No evidence of degradation of FVIII Δ C1 protein was observed; however, this variant was mainly seen as a single-chain protein (expected size ~150 kDa), while the intact FVIII construct comprised a mixture of single-chain and cleaved heavy/light-chain FVIII. The production of primarily the single-chain form has been observed previously for human FVIII Δ B from rAAV vectors *in vivo* and for recombinant canine FVIII Δ B.^{11,20}

We next constructed the 5.1 and 4.6 kb mTTR/FVIII vector with and without the C1 domain, respectively, using an alternate serotype vector (AAVrh8R). Southern blot analysis of the genomes isolated from purified 5.1 and 4.6 kb vectors showed that the genome sizes were more homogenous for the 4.6 kb vector compared to the 5.1 kb vector (Figure 5d). Based on the dot blot analysis, the 5.1 kb vector contained completely packaged 3' ends of vector genomes of both polarities which is consistent with packaging initiating at the 3' ends.^{2,21,22} However, the 5' ends in the 5.1 kb vector genomes were present at lower levels than in the 4.6 kb vector, indicating reduced encapsidation of the 5' termini of the genomes in the 5.1 kb vector (Figure 5e).

The two vectors were administered intravenously at 3×10^{11} vg/mouse. Mice treated with the 4.6 kb vector generated plasma FVIII levels that were approximately twofold higher than those treated with the oversized 5.1 kb vector (statistically significant at all time points measured; Figure 6a). In addition, the expression kinetics appeared to be slower for the 5.1 kb vector. While the ratio of FVIII protein per vector genome was comparable for the 4.6 kb vector on days 3 and 90, there was a threefold increase in this ratio for the 5.1 kb vector over the same time interval. At day 90 postinjection, FVIII mRNA and vector genome levels in the liver were quantitated

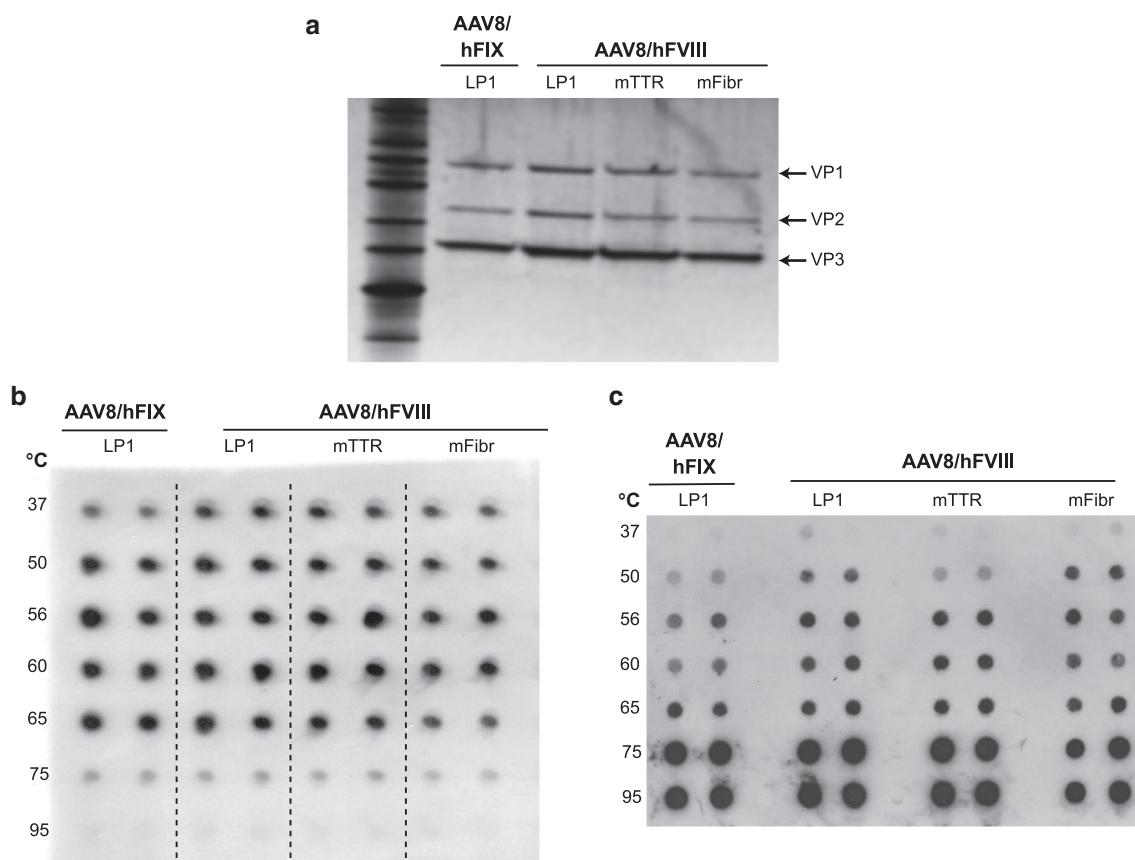


Figure 2 Characterization of rAAV8 vectors containing oversized FVIII expression cassettes. All vectors were generated by triple transfection and purified using a CsCl_2 gradient. **(a)** Capsid analysis by SDS-PAGE gel. A total of 1×10^{10} vg virus/lane (based on LP1 or FVIII-A1 qPCR) was resolved by 4–12% Tris-Bis SDS-PAGE gel and followed by staining with Sypro protein stain. The locations for capsid proteins VP1, VP2, and VP3 are shown on the right. Quantitation of staining intensities is shown in Supplementary Table S2. **(b)** Heat stability of AAV8 vectors. Each vector was heated at the indicated temperature and loaded on a membrane at 1×10^9 vg/dot, and the amount of intact vector was determined by detection with a conformation-specific AAV8 capsid antibody. **(c)** VP1 detection after heat treatment. Vectors were processed as described above, and the exposure of the N-terminus of the VP1 protein was analyzed using an A1 antibody. LP1, 5.4 kb LP1-hFVIIIco or 4.6 kb LP1-hFIXco vector; mTTR, 5.1 kb mTTR-hFVIIIco vector; mFibr, 5.0 kb mFibr-hFVIIIco vector; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis.

(Figure 6b). Similar to the FVIII protein levels, there was a twofold higher level of vector genomes in the 4.6 kb vector compared to the oversized vector. However, the mRNA levels in the livers of mice transduced by both vectors were not significantly different. Administration of a lower dose (4×10^{10} vg/mouse) of the 5.1 kb vector resulted in a 4.4-fold lower level of mRNA, indicating a dose-dependent effect. This dose response was also demonstrated for plasma FVIII activity between the high- and low-dose treatment groups with the lower dose cohort exhibiting approximately threefold lower levels of FVIII activity (Figure 6c). Mice given the 4.6 kb vector (encoding the C1-domain deleted hFVIII) generated FVIII activity that was not different from that of naive mice, confirming that this FVIII protein variant was not active (not shown).

Quantity and persistence of genomes in the liver nuclei

The quantitation of vector genome copies in the isolated liver nuclei of AAVrh8R/5.1 kb-transduced mice showed that they contained twofold to threefold lower levels than those administered the smaller AAVrh8R/4.6 kb vector on both days 3 and 90 posttreatment (Figure 7a). This difference in copy number was similar to those obtained quantitating vector genomes in liver sections indicating that vector genomes detected resided in the nucleus (Figure 6b). Both vectors had comparable persistence over the 3-month study

period as no significant difference in vector genome levels in liver were noted between days 3 and 90 posttreatment.

Vector genome integrity was determined by digesting liver DNA using restriction enzymes predicted to cleave sequences located within the promoter, FVIII cDNA, and polyadenylation sites (Figure 7b). If the incomplete plus or minus strand genomes were converted to double-stranded (ds) forms, the resulting genomes would be missing the 5' or 3' termini of the expression cassette (*i.e.*, the promoter region and polyA site) and be transcriptionally inactive. However, the Southern blot results showed the presence of complete expression cassette as the correct size fragments were detected for the oversized (1.8, 2.8 kb) vector in all of the animals tested ($n = 4\text{--}5/\text{group}$) both on days 3 and 90 (fragment sizes for the control 4.6 kb vector were 1.8, 2.3 kb; Figure 7c,d). This observation indicated that the extreme 5' ends of the genome had been regenerated *in vivo*. However, the pattern at day 3 was more heterogeneous, suggesting that the process leading to the generation of full-length ds DNA forms had not been completed at this time point (Figure 7c). By day 90, the vector genomes had all been processed into uniform ds genomes with no apparent truncations or deletions within the expression cassette (Figure 7d). As expected, no human FVIII genomes were detected in the DNA from naive mice. Fainter signals were observed for the 5.1 kb vector, which was consistent

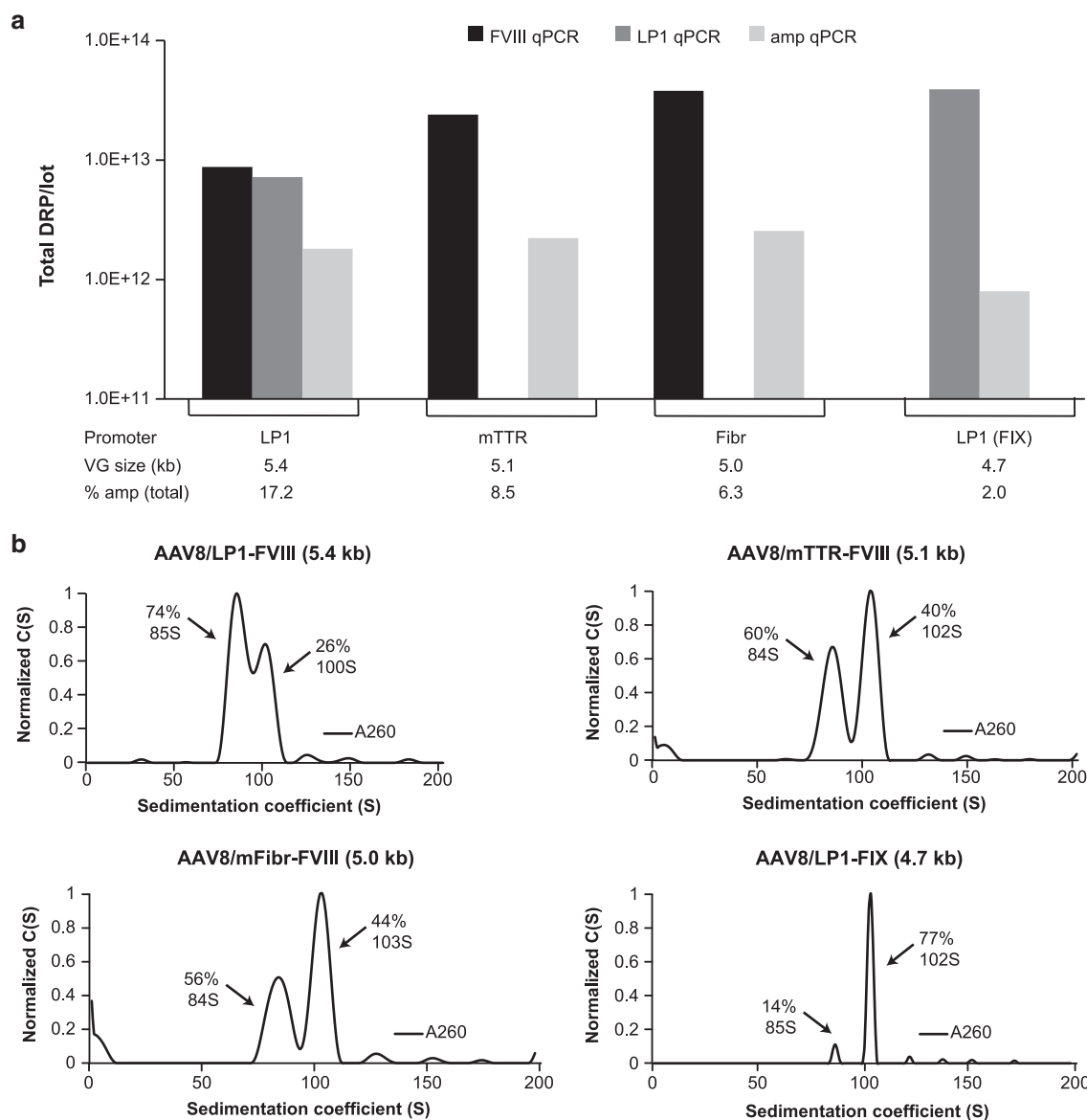


Figure 3 Characterization of packaged genomes in AAV8/hFVIIIco vectors. **(a)** Quantitation of total vector yields and the amount of aberrant packaging. Vectors were quantitated for packaged FVIII (FVIII-A1 or LP1) and ampicillin sequences (amp) using qPCR analysis. Promoters, vector genome (vg) sizes (kb) and the percentages of aberrant (amp) packaging are shown on the bottom. **(b)** Analytical ultracentrifugation of purified vectors. Vector mass (S) and percentage (%) for each distinct peak is shown. Abbreviations are the same as in Figure 2.

with the twofold lower vector genome copies (detected by qPCR) compared with the 4.6 kb vector (Figure 7a).

The structures of the vector genomes in the liver generated by the 5.1 and 4.6 kb viral vectors at days 3 and 90 postadministration were also analyzed (Figure 8a). On day 3, vector DNA was heterogeneous in size and primarily linear, as indicated by its sensitivity to digestion with plasmid-safe DNase (Figure 8b). Only small amounts of ds DNA were evident, as indicated by the inability to generate distinct fragments following XbaI digestion (cuts double-stranded vector genomes once). On day 90, vector-derived DNA existed mainly as high molecular weight (HMW) circular species, as it was resistant to plasmid-safe DNase treatment (Figure 7c). Digestion by XbaI reduced the HMW concatemers to a single unit-length band of either 4.6 kb or 5.1 kb for the wild-type and oversized vectors, respectively, indicating the presence of double-stranded, multimeric forms of the vector episomes. Taken together, these data

indicate that the oversized 5.1 kb vector generated persistent double-stranded and circular HMW genomes similar to the wild-type size vector.

DISCUSSION

Recombinant AAV vectors are gaining favor as gene transfer vectors of choice for select diseases due to their demonstrated safety profile and capacity to confer long-term transgene expression.^{6-12,14} Despite the limitation in packaging expression cassettes that exceed 4.7 kb in size, they have nevertheless been successfully used as oversized or fragmented rAAV vectors to express large transgenes *in vivo*.^{6-12,14} However, these vectors are challenging both for production and for the quality of the vector preparations. In this study, we generated a number of oversized rAAV vectors encoding a human B-domain deleted FVIII and characterized their quality and hepatic transduction ability.

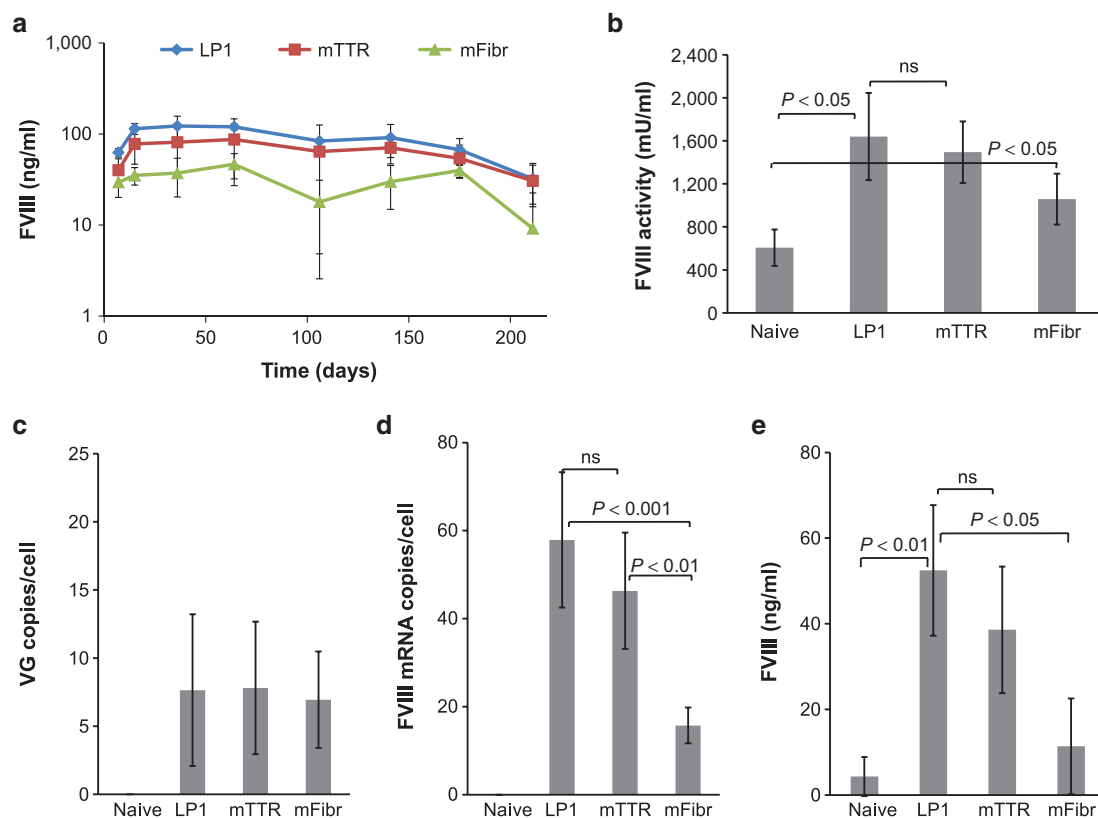


Figure 4 Human FVIII expression from oversized rAAV8 vectors *in vivo*. Vectors were administered at 3×10^{11} vg/mouse into female C57BL/6 mice ($n = 6-8$ mice/treatment group). **(a)** Human FVIII levels in plasma over 7 months. Plasma FVIII levels were quantitated by FVIII ELISA, and endogenous mouse FVIII protein levels were subtracted. **(b)** Plasma FVIII activity levels 141 days post-rAAV8 vector administration. Activity was measured by Coatest assay. **(c)** Vector genome copies in liver on day 211. Vector genomes were quantitated by qPCR using primers specific to human FVIIIco. **(d)** FVIII mRNA levels in liver on day 211. Vector derived hFVIII transcript levels were quantitated by qPCR using primers specific to human FVIIIco. **(e)** Plasma FVIII protein levels on day 211. FVIII levels were measured by ELISA. Values in all graphs represent the average \pm SD, and significance was calculated using Student's *t*-test with Bonferroni correction.

Our data showed that the generation of these oversized rAAVs by a triple transfection method resulted in viral lots of varying yield and quality. All the oversized rAAV vectors with 5.0 to 5.4 kb size genomes packaged mostly 4.6 kb (wild-type) size genomes. The yields attained for the largest vector tested (5.4 kb) were typically fourfold lower than those for the control 4.7 kb rAAV-FIX vector. Moreover, the fidelity of viral packaging decreased with increasing genome size, as indicated by the presence of encapsidated DNA impurities. As packaging has been shown to occur from the 3' end of the single-stranded genome,^{2,21,22} the encapsidated oversized genomes are expected to lack the 5' ITR which may contribute to the infidelity. In addition to aberrant packaging, the oversized vectors also incorporated large portions of subgenomic-length FVIII-specific DNA. This incorporation may be due to a higher incidence of premature packaging stops in the oversized vectors, which suggests a more complex process for encapsidation than a simple "head-full" mechanism. Covalent attachment of Rep78 to the 5' end of the single-stranded genome to mediate terminal resolution has been proposed as a packaging signal for newly synthesized DNA and a bridge from DNA to the preassembled viral particles.^{23,24} The spacing between the 5' bound Rep and the 3' ends used for encapsidation may be critical for packaging correct unit-length genomes. In the absence of this, packaging may stall and result in cleavage of single-stranded genomes by an unknown mechanism generating distinct fragment sizes that are often seen with oversized

rAAV vector lots.^{2,3,5,6,11} A better understanding of this process may help to decrease the heterogeneity of oversized viral vector preparations.

The prevailing strategy to develop rAAV vectors with large transgenes is to minimize the genome size to improve packaging fidelity. However, our data demonstrated that smaller genome sizes are not necessarily a solution particularly if the reduction in size compromises gene expression. We noted that the *in vivo* potency of oversized vectors, as measured by plasma FVIII levels, was more dependent on the cassette design as opposed to the genome size. Comparable levels of vector genomes were detected in the liver at seven months posttreatment, indicating that the efficiency of the process for generating these stable vector genomes was independent of their size. Thus, our data suggest that improvements to the expression cassette (promoter, intron, and polyadenylation signals) can be beneficial even if they increase the vector genome size.

To evaluate the potency of a slightly oversized vector, we generated an expression cassette containing the same elements as that used in the 5.1 kb vector but encoding a C1-domain deleted FVIII. A concern with the use of FVIII Δ C1 was the potential for more rapid clearance of the FVIII protein *in vivo*. The half-life of FVIII is enhanced by binding to von Willebrand factor which limits its interactions with phospholipid surfaces and protects it from inactivation by plasma proteases.²⁵ The interaction of von Willebrand factor to FVIII is mediated by the $\alpha 3$ acidic region and C2 domains of FVIII though the FVIII C1 domain is known to enhance this interaction.²⁵⁻²⁷ Although our

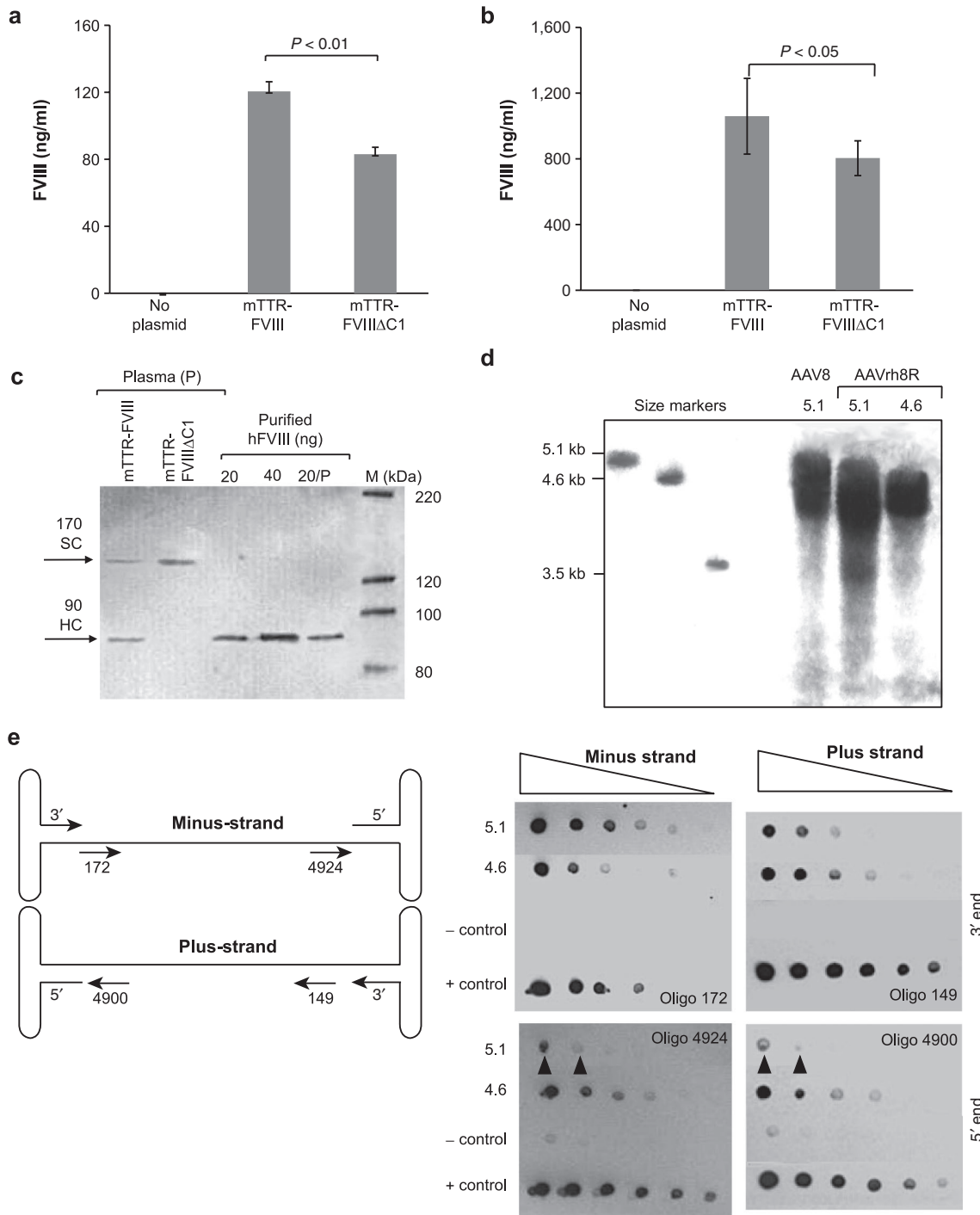


Figure 5 Characterization of 4.6 and 5.1 kb vectors in previral plasmids and viral vectors. FVIIIΔC1 and intact FVIII protein production from the 4.6 and 5.1 kb previral plasmids, respectively, *in vitro* (a) and *in vivo* (b). For comparison *in vitro*, plasmids were transfected into Huh7 cells. For analysis *in vivo*, plasmids were administered by high-volume injection into male C57BL/6 mice. FVIII levels were measured by ELISA in the conditioned media 48 hours later or in plasma on day 1, respectively. (c) Analysis of FVIII proteins in plasma after delivery of previral plasmid vectors. Plasma samples were analyzed by western blot using antibody specific to the heavy-chain of human FVIII as described in Materials and Methods. Purified hFVIIIΔB was run as control. Molecular weight markers and location for single-chain and heavy-chain FVIII are shown. (d) Analysis of packaged FVIII vector genome sizes in AAVrh8R vectors with 4.6 or 5.1 kb genomes. Genomes were extracted from purified rAAVrh8R vectors, and the DNA was run on an alkaline agarose gel and detected by Southern blot analysis using 2.7 kb FVIII-specific probe. Size markers (3.5, 4.6, and 5.1 kb) are shown on the left. A rAAV8 vector with a 5.1 kb genome is shown for comparison. (e) Characterization of packaged vector genome terminal ends by DNA dot blot analysis. Serial dilutions of vectors were applied onto membranes (starting at 2.4×10^9 vg) and hybridized with 3' end-labeled oligonucleotide probes specific to the 3' or 5' terminal ends of the vector genomes of both polarities (location shown in the diagram, left). The signal intensity of the 5.1 kb vector was compared to that of the 4.6 kb vector (completely packaged), and the differences observed are indicated (triangles). pUC57-based plasmids with ssLP1-hFVIIIco and mTTR-FVIII were used as negative and positive controls, respectively.

studies showed that the stability of FVIIIΔC1 and intact FVIII protein was similar, it is possible that the observed twofold difference in plasma FVIII levels as measured by protein levels is an underestimation. However, a similar twofold to threefold difference between the 5.1 and 4.6 kb vectors was detected when the vector genome levels in liver was measured which is consistent with the noted differences in plasma FVIII levels. Overall, this potency difference for the 5.1 kb oversized vector was lower than that noted in other reports with much larger genomes (6 to 8 kb) in rAAV vectors. For example, Hirsch *et al.*¹⁴ reported a 30-fold lower luciferase activity from a 6.2 kb AAV8/luciferase vector compared to a 3.9 kb size vector 12 to 15 days post vector administration into mice. *In vitro*, Wu *et al.*⁵ demonstrated a 77- to 1,600-fold lower transduction of 293 cells with AAV2/LacZ vectors ranging in size from 5.2 to 8.7 kb and containing incompletely packaged expression cassettes. Interestingly, only 4- to 15-fold losses in potency were obtained with vectors containing a LacZ expression cassette with 5 kb from either ITR independent of vector genome size (5.2 to 8.7 kb).

They are multiple explanations for lower levels of cellular vector genomes by oversized vectors. This could be due to reduced capsid uncoating although we did not observe any physical stability differences *in vitro*. Oversized vectors may also be more sensitive to proteasomal degradation based on the findings that the use of proteasome inhibitors increased the transduction activity of oversized vectors.^{2,8} Alternatively, proteasome inhibitors increased vector transfer into the nucleus, which in turn enhanced the genome correction process *in vivo*. A genome correction process is required because the amount of 5.1 kb full-length genomes encapsidated represented only a small portion of the total packaged DNA. Depending on the efficiency of this process, it is likely that a certain threshold level of vector genomes in the nucleus is required for the regeneration of full-length expression cassettes. Therefore, a higher vector dose may be needed for oversized vectors. The effect of vector dose on the genome correction process has been previously demonstrated *in vitro*.⁵ Finally, despite a reduction in the number of vector genomes in the liver compared to wild-type size rAAV genome, the 5.1 kb oversized vector was not compromised for persistence of these genomes in liver. Our data on vector genomes on day 3 and day 90 after vector administration demonstrated that the oversized vectors conferred comparable maintenance of genome copies in the liver to that observed for a wild-type sized vector.

Both the annealing of complementary single-stranded plus and minus strands and homologous recombination between double-stranded fragments have been suggested as potential mechanisms for regenerating ds transcriptionally active genomes for oversized vectors.⁵ The annealing of complementary strands is the major process active in the mouse liver.²⁸ Whether this process or homologous recombination was responsible for correcting the partial genomes in our studies is not clear. All rAAV vectors require the formation of circular genomes by intermolecular recombination at the ITRs for long-term vector persistence which is facilitated by host DNA recombination and repair factors.²⁹⁻³³ In particular, the presence of HMW concatemers correlates with long-term gene expression in both muscle and liver.^{15,33-35} As the majority of the vector genomes in liver in our studies after 3 months were HMW circular forms, this finding indicates that oversized vectors were able to generate similar genomic structures as was observed for the wild-type sized vectors.

Currently, no oversized rAAV vectors have been used in the clinic. For wild-type size rAAV vectors that have been evaluated in clinical trials thus far, vector homogeneity and minimal packaging of

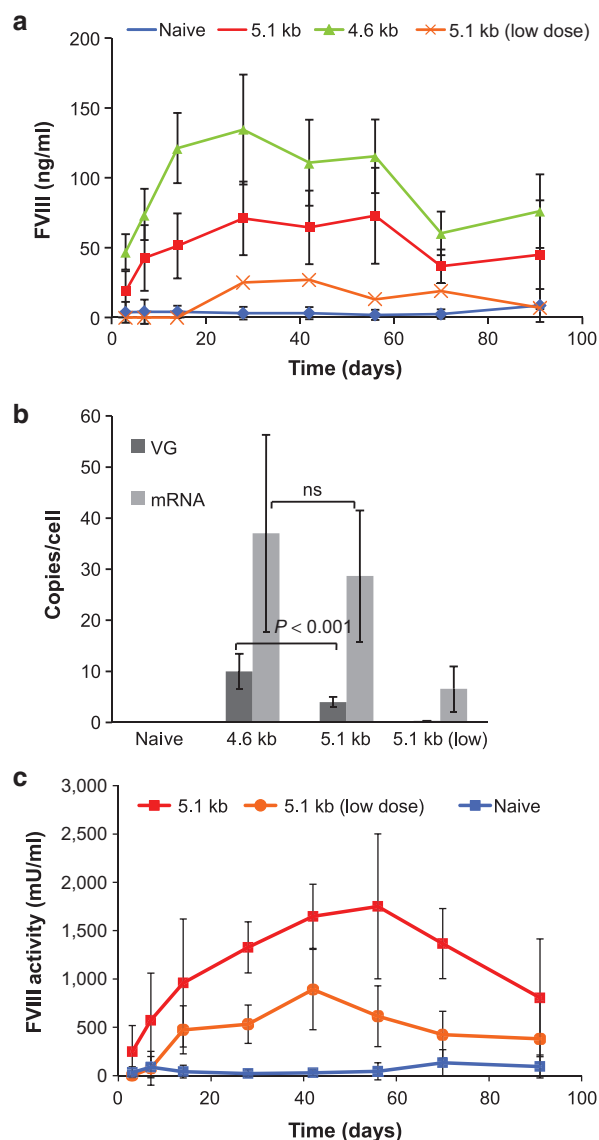


Figure 6 Potency analysis of the oversized 5.1 kb rAAVrh8R vector *in vivo*. (a) Human FVIII protein levels in plasma after AAVrh8R/FVIIIco virus administration. Vectors were administered at 3×10^{11} vg/mouse (4.6 and 5.1 kb) and 4×10^{10} vg mouse (5.1 kb; low dose) into male C57BL/6 mice ($n = 7-10$ mice/group) and plasma FVIII levels were determined by ELISA. (b) Vector genome and hFVIII mRNA levels in liver on day 90. Levels were quantitated by qPCR using primers specific to human FVIIIco. (c) Plasma FVIII activity levels. FVIII activity levels were measured by Coatest assay. Values in all graphs represent the average \pm SD, and significance was calculated using Student's *t*-test with Bonferroni correction.

nonvector DNA have been critical quality end points.^{36,37} Published data as well as those reported here have suggested that increasing vector genome size can adversely affect vector quality in ways that may be inconsistent with established standards specifically with regard to heterogeneity in encapsidated genome size and packaging of the plasmid backbone. Packaging of nonvector DNA sequences poses potential safety risks with regard to transfer of antibiotic resistance, plasmid- and host-derived oncogenes, and immunogenic polypeptides.³⁶ While these concerns may be mitigated by demonstrating a lack of expression of packaged sequences, implications of these quality attributes on safety can only be fully assessed via further preclinical testing. No stringent *in vivo* safety

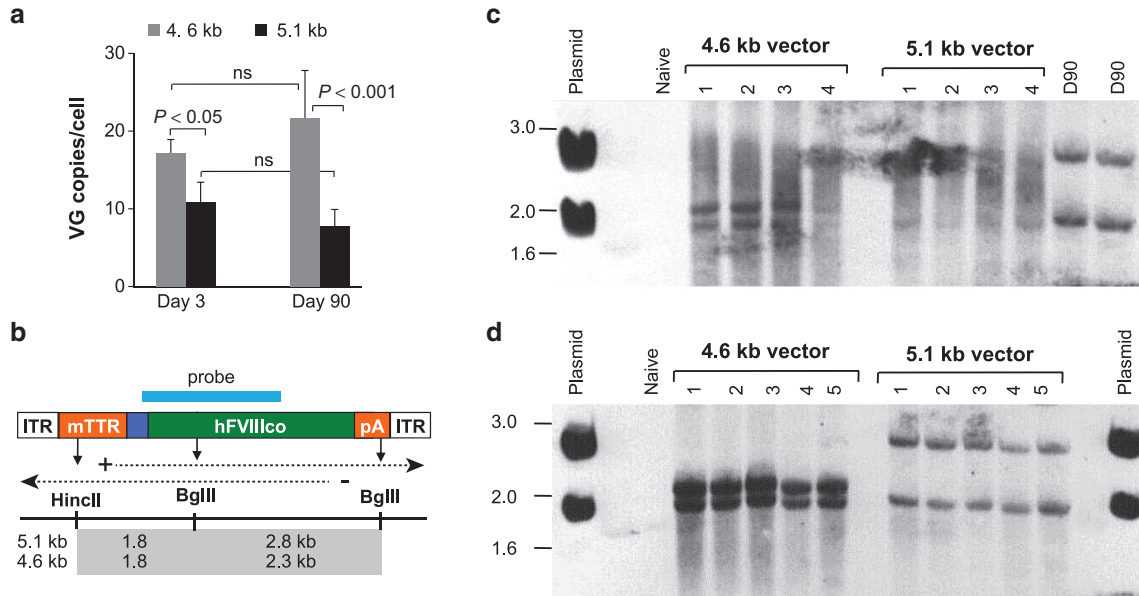


Figure 7 Characterization of AAVrh8R/FVIII vector genome levels and their integrity in liver. Liver nuclei from 4.6 and 5.1 kb vector-treated mice ($n = 4-5$ mice/group) were harvested, and the DNA was isolated. **(a)** Vector genome levels in livers on days 3 and 90. Vector genomes were quantitated by qPCR using primers specific to human FVIIIco ($n = 7-10$ /group). **(b)** Diagram for the analysis of vector integrity. Locations of HincII and BglIII restriction enzyme sites and the Southern probe in the full-length mTTR-FVIII genome are shown. **(c)** Analysis of vector DNA integrity in liver on day 3. Liver DNA was digested with HincII and BglIII to test for the presence of the complete expression cassette. A plasmid with the 5.1 kb mTTR/hFVIIIco genome was used as a positive control for the correct fragment sizes (1.8 and 2.8 kb). Two mice from day 90 were included in the day 3 analysis for size comparison. **(d)** Analysis of vector DNA integrity in liver on day 90. The analysis was similar to that of day 3.

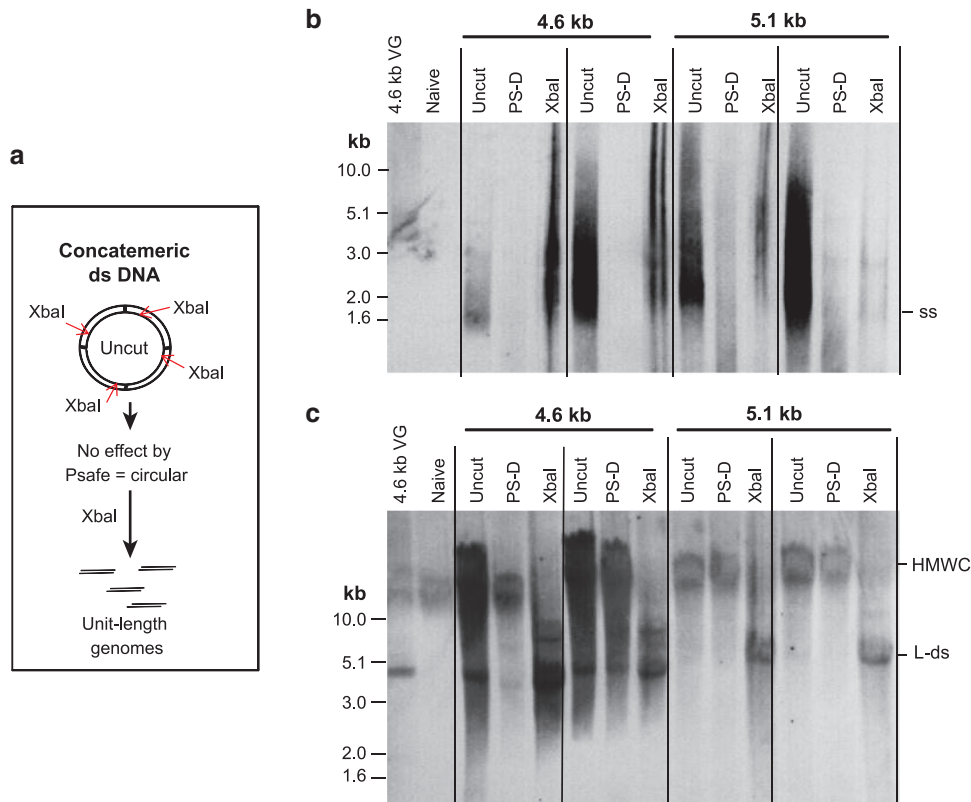


Figure 8 Characterization of persistent AAVrh8R/hFVIIIco vector-derived molecular forms in liver. Description of analyses for **(a)** liver nuclear DNA and results from **(b)** day 3 and **(c)** day 90 samples. Two animals were tested from the 4.6 kb and 5.1 kb vector treatment groups. Southern blots were hybridized with a 2.7 kb FVIII fragment shown in Figure 7b. The 4.6 kb purified vector DNA was included as a size marker. Uncut, untreated DNA; PS-D, digestion with plasmid-safe DNase; Xbal, digestion with a single cutter restriction enzyme for ds vector genome; HMWC, high-molecular-weight concatemers; L-ds, linear double-stranded form; ss, single-stranded form.

studies of oversized vectors have been reported though they have been successfully used in large animals and in some instances monitored cases over a 10-year period.^{7,11} It is possible that modifications to upstream processes such as the use of alternate vector production platforms may improve the quality of oversized vectors. To that end, we are currently investigating the production of oversized vectors via the producer cell line platform (manuscript in preparation). However, as features of AAV biology such as capsid structure and mechanism of genome encapsidation has set a natural limit for packagable genome size, "oversized" vector packaging might be more easily achieved using related viruses that can package larger genomes.³⁸ Alternatively, a strategy relying on intracellular processing of multiple wild-type size genomes to re-generate the full-length expression cassette is an option. The latter will provide the quality needed but would require both more efficient *in vivo* processes and vastly enhanced vector production capacities.

In summary, despite heterogeneously packaged vector genomes, all the minimally oversized vectors we tested were able to generate functional protein and engender the presence of genomes in the liver that persisted over a period of several months irrespective of their size. However, their potency measured by these endpoints was reduced twofold to threefold *in vivo*. More research is needed to understand the *in vivo* correction processes shortly after cellular uptake for these vectors to look for ways to enhance their biological activity. Additionally, significant improvements in vector quality as well as increased production yields to satisfy the anticipated high vector doses would be required to facilitate their use in the clinic.

MATERIALS AND METHODS

Construction of FIXco and FVIIIco expression plasmids

A self-complementary (sc) LP1 (liver promoter)-hFIXco genome was synthesized (Genscript, Piscataway, NJ) according to the published sequence¹⁵ and was cloned into the pUC57 plasmid. A single-stranded (ss) version of LP1-hFIXco was generated by cloning a "stuffer" sequence (2,150 bp) downstream of the expression cassette followed by KpnI and a wild-type ITR. For the generation of ssLP1-FVIIIco, a codon-optimized human FVIIIΔB-SQ with a truncated SV40 polyA signal was synthesized (GeneArt/Invitrogen, Grand Island, NY). This sequence was cloned as an Xba and KpnI fragment into the ssLP1-hFIXco vector to replace hFIXco with hFVIIIco. Plasmid with mouse TTR¹⁶ promoter was generated by replacing the ITR-LP1 promoter region with the mTTR-HI⁷ as a SacI and XbaI fragment. The SV40 polyA was replaced with a Spel-KpnI fragment containing a 52 bp spA.¹⁷ The mTTR-HI was generated such that the promoter could be changed as a MluI-NheI fragment and introns as NheI-XbaI fragments. Other promoters cloned were the human TTR¹⁶ and a modified human fibrinogen promoter (mFibr; Figure 1 and Supplementary Table S3). A minute virus of mice intron was tested as an alternative intron.¹⁸ A 4.6 kb mTTR-FVIII expression cassette was generated by deleting the coding region of the C1 domain of FVIII.

In vitro cell cultures

All tissue culture reagents were obtained from Irvine Scientific (Santa Ana, CA) or Invitrogen. Human liver carcinoma cells (Huh7) (3×10^5 cells/well) were plated on six-well dishes in Dulbecco's modified Eagle's medium with high glucose, 10% fetal bovine serum, and 10 ml/l Pen Strep (10 units/ml penicillin and 10 μg/ml streptomycin). Plasmids (2 μg) were transfected with Lipofectamine (Invitrogen), and conditioned media was harvested 72 hours later and assayed for FVIII levels by enzyme-linked immunosorbent assay (Human FVIII:C ELISA; Enzyme Research Laboratories, South Bend, IN).

Construction of rAAV vectors

All rAAV vectors were generated by the triple transfection method, and vector genomes were packaged as AAV8 or AAVrh8R capsid serotypes.³⁹ Vectors were purified using single CsCl gradient or column chromatography. Vectors were quantified by a real-time (RT) qPCR assay (7500 Real-Time PCR System; Applied Biosystems, Foster City, CA) specific to LP1 promoter or FVIIIco

sequences using a standard curve of serially diluted linearized hFVIIIco plasmid DNA. The level of plasmid backbone packaging was measured by quantitating presence of DNA for ampicillin resistance gene (*amp^r*) located 1,182 bp from the 3' ITR of the previral plasmid). All primers and probes used in qPCR are shown in Supplementary Table S3 (FVIII primers/probe are located in the A1-A2 junction in a region of 1,123 to 1,246 bp from the start codon of FVIII). The values are expressed as vector vg per ml.

rAAV vector characterization *in vitro*

For the capsid analysis, vectors were mixed with loading dye and a reducing agent and were heated at 70 °C for 10 minutes. A total of 1×10^{10} vg virus/lane (based on qPCR titer) was resolved by 4–12% Tris-Bis sodium dodecyl sulfate polyacrylamide gel electrophoresis (1.5 mm) in 4-morpholinepropanesulfonic acid buffer. The gel was stained overnight with Sypro protein gel stain (Invitrogen). The density of the signals was quantitated using ImageJ (<http://rsb.info.nih.gov/ij>).

To evaluate capsid stability, purified viruses were diluted in phosphate-buffered saline (PBS) and heated at various temperatures for 10 minutes. The samples were cooled on ice and 1×10^9 vg/dot was applied onto nitrocellulose membranes. The membranes were rinsed with PBS with Tween (PBST) and blocked in 0.15% I-Block (ThermoFisher Scientific/Life Technologies, Carlsbad, CA) overnight. The membranes were then incubated with a conformational antibody to the AAV8 capsid (4.5.3)⁴⁰ or with A1 antibody (American Research Products, Waltham, MA) at 1 μg/ml at room temperature for 1 hour. A goat antimouse IgG-horseradish peroxidase conjugate (1:10,000 dilution; Jackson ImmunoResearch, West Grove, PA) was used as a secondary antibody and incubated at room temperature for 1 hour. The membranes were then rinsed in PBST, and signals were developed using an enhanced chemiluminescence detection kit (ThermoFisher Scientific/Life Technologies).

To analyze the packaged vector DNA, genomes were extracted from purified viruses. Briefly, vectors were digested with DNase followed by proteinase K digestion, phenol-chloroform extraction, and precipitation in the presence of glycogen. For Southern blot analysis, the vector genomes were run on 1% alkaline (NaOH) agarose gels and then transferred and cross-linked onto Hybond membranes (Amersham/GE Healthcare, Westborough, MA). Membranes were probed with a 2.7 kb alkaline phosphatase-labeled hFVIII cDNA fragment (A1–A3 domains) at 55 °C overnight, detected with CDP-Star chemiluminescent reagent (GE Healthcare), and exposed to X-ray film. For characterization of vector genomes by DNA dot blot analysis, the purified genomes were applied by twofold serial dilutions onto membranes (starting at 2.4×10^9 vg; six dilutions). Strand-specific oligonucleotide probes were 3'-end labeled with digoxigenin (DIG)-ddUTP, and the membranes were processed according to manufacturer's protocol (DIG Oligo3'-End Labeling Kit, Roche Diagnostics, Indianapolis, IN). Probe sequences specific to the 3' and 5' termini of vector genomes used are shown in Supplementary Table S3.

Analytical ultracentrifugation analysis was performed to reveal the different capsid species within the vector preparations as previously described.¹⁹ Briefly, 2×10^{12} vg of purified vector was buffer exchanged into PBS, pH 7.2, followed by adjustment of vector concentration to OD₂₆₀ of between 0.2 and 0.8. Vectors were then analyzed using a Proteome Lab XL-1 (Beckman Coulter, Indianapolis, IN) and percentage of genome containing virions was determined by absorbance detection method and SEDFIT using the continuous-size C(S) distribution method. The results were plotted as the normalized differential coefficient distribution value, C(S), versus sedimentation coefficient (S).

Animal studies and plasma analyses

All protocols were approved by the Genzyme Institutional Animal Care and Use Committee (IACUC) and were conducted in accordance with humane guidelines for animal care and use. Female or male C57BL/6 mice (Taconic, Hudson, NY) were housed in groups and fed standard irradiated rodent chow and water *ad libitum*. Plasmids (30 μg/mouse) were administered by hydrodynamic high-volume injections into the tail vein using sterile saline. Recombinant AAV vectors were injected at 4×10^{10} or 3×10^{11} vg/mouse by tail vein injections. Animals were bled by retro-orbital sinus bleeding into sodium citrate or EDTA, and plasma was stored frozen until analysis. For liver collections, mice were opened with a disposable scalpel, and a small liver section from the left medial lobe was removed with disposable forceps and scissors (all instruments were changed between individual animals). For some animals, the remainder of the liver was removed for liver nuclei isolation.

Plasma FVIII levels were measured using FVIII ELISA (Enzyme Research Laboratories) or FVIII activity assay (Chromogenix Coatest SP4 FVIII; DiaPharma, West Chester Township, OH) according to manufacturer's protocol and adjusted to 96-well format. Activity values were measured as % FVIII activity compared to normal plasma and converted into mU/ml. Pooled normal human plasma (Innovative Research, Novi, MI) was used as standard.

Plasma FVIII size was analyzed by western blot analysis. Briefly, plasma samples were resolved on 4–12% Tris-Bis sodium dodecyl sulfate polyacrylamide gel electrophoresis using 4-morpholinepropanesulfonic acid buffer. The gel was stained overnight with Sypro protein gel stain (Invitrogen). Proteins were transferred to nitrocellulose membrane by electroblotting, and the membrane was then blocked in PBS/0.2% I-Block/0.1% Tween20 at 4 °C overnight. The membranes were incubated with a mouse monoclonal antibody to human FVIII (AbD Serotec MCA467; Raleigh, NC) at 1 µg/ml at room temperature for 1 hour. A goat antimouse IgG-horseradish peroxidase conjugate (1:50,000 dilution; Jackson ImmunoResearch) was used as a secondary antibody and incubated at room temperature for 1 hour. The membranes were then rinsed in PBST, and signals were developed using an enhanced chemiluminescence detection kit (Thermo Scientific/Life Technologies). Purified hFVIIIΔB (RayBiotech, Norcross, GA) was run as positive control.

Liver DNA isolation and characterization

Mouse liver DNA was isolated from a section of the left medial lobe or from liver nuclei. For the former, frozen sections (0.116–0.216 g) were homogenized in 1 ml of RLT+plus β-mercaptoethanol and 1 mm zirconia beads for 1 minute. A total of 250 µl of the homogenate was transferred to 750 µl of TRIzol reagent (for RNA, ThermoFisher Scientific) or DNA STAT-60 (for DNA; Tel-Test, Friendswood, TX). DNA was isolated according to the DNA STAT-60 protocol up to the aqueous phase, after which the samples were treated with RNase and the DNA was purified using PureLink columns (Invitrogen) according to the manufacturer's protocol.

Liver nuclei were isolated using a Nuclei PURE Prep isolation kit (Sigma-Aldrich, St Louis, MO). In brief, fresh tissue (1 g/sample) was rinsed with PBS and homogenized in lysis solution containing dithiothreitol and Triton X-100 for 30 seconds at 20,000 rpm using a hand-held homogenizer (IKA Works, Wilmington, NC). Homogenates were diluted 1:2 with 1.8 mol/l sucrose, layered over 1.8 mol/l sucrose cushions in Beckman SW28 ultracentrifuge tubes, and spun at 13,000 rpm for 45 minutes. Nuclei were re-suspended in 400 µl of PBS followed by DNA extraction using DNA STAT-60. Vector copies in liver DNA were quantitated using qPCR (TaqMan assay) and hFVIII-specific primers/probe as described for rAAV vector quantitation and expressed as vg copies per cell (using 5 pg dsDNA per diploid cellular genome).

For liver genome characterization by Southern blot, 20 to 40 µg of genomic DNA (1.6 × 10⁸ vg or 10 to 20 vg/cell based on the qPCR) was analyzed with or without restriction digestion with various enzymes (digests were precipitated in the presence of glycogen). Plasmids digested with the same enzymes were included as dsDNA size markers and were run in the presence of 40 µg of genomic DNA from naive mice. Southern blotting and detection were performed as described above.

Liver RNA isolation and analysis

From the liver homogenates, 250 µl was transferred to 750 µl of TRIzol containing beads and then homogenized for 1 minute. RNA was isolated using the TRIzol protocol with RNA binding columns (Promega, Madison, WI). cDNA was generated using a High-Capacity cDNA Reverse Transcription kit (Applied Biosystems, Carlsbad, CA). Transcript levels were determined using quantitative RT-PCR with primers (Applied Biosystems) specific for human FVIII (Supplementary Table S3).

Statistical analysis

In vitro data consisted of groups of three independent observations and were performed a minimum of two times. *In vivo* data consisted of groups of 4–10 animals. The data were analyzed for significance by unpaired Student's *t*-test. For multiple comparisons, the Bonferroni correction was applied. *P* < 0.05 was considered statistically significant.

CONFLICT OF INTEREST

All authors were employees of Sanofi Genzyme at the time of the study. The authors declare no conflict of interest.

ACKNOWLEDGMENTS

We thank Alison Schroerer for help with the illustrations. We also appreciate the support from Samuel Wadsworth and Karen Vincent. We thank members of Virus Production and Comparative Medicine for their technical assistance with some of the studies.

REFERENCES

- 1 Dong, JY, Fan, PD and Frizzell, RA (1996). Quantitative analysis of the packaging capacity of recombinant adeno-associated virus. *Hum Gene Ther* **7**: 2101–2112.
- 2 Grieger, JC and Samulski, RJ (2005). Packaging capacity of adeno-associated virus serotypes: impact of larger genomes on infectivity and postentry steps. *J Virol* **79**: 9933–9944.
- 3 Lu, H, Chen, L, Wang, J, Huack, B, Sarkar, R, Zhou, S et al. (2008). Complete correction of hemophilia A with adeno-associated viral vectors containing a full-size expression cassette. *Hum Gene Ther* **19**: 648–654.
- 4 Dong, B, Nakai, H and Xiao, W (2010). Characterization of genome integrity for oversized recombinant AAV vector. *Mol Ther* **18**: 87–92.
- 5 Wu, Z, Yang, H and Colosi, P (2010). Effect of genome size on AAV vector packaging. *Mol Ther* **18**: 80–86.
- 6 Grose, WE, Clark, KR, Griffin, D, Malik, V, Shontz, KM, Montgomery, CL et al. (2012). Homologous recombination mediates functional recovery of dysferlin deficiency following AAV5 gene transfer. *PLoS One* **7**: e39233.
- 7 Jiang, H, Lillcrap, D, Patarroyo-White, S, Liu, T, Qian, X, Scallan, CD et al. (2006). Multiyear therapeutic benefit of AAV serotypes 2, 6, and 8 delivering factor VIII to hemophilia A mice and dogs. *Blood* **108**: 107–115.
- 8 Monihan, PE, Lotrop, CD, Sun, J, Hirsch, ML, Kafri, T, Kantor, B, et al. (2010). Proteasome inhibitors enhance gene delivery by AAV virus vectors expressing large genomes in hemophilia mouse and dog models: a strategy for broad clinical application. *Mol Ther* **18**: 1907–1916.
- 9 Ishiwata, A, Mimuro, J, Mizukami, H, Kashiwakura, Y, Takano, K, Ohmori, T et al. (2009). Liver-restricted expression of the canine factor VIII gene facilitates prevention of inhibitor formation in factor VIII-deficient mice. *J Gene Med* **11**: 1020–1029.
- 10 Sabatino, DE, Lange, AM, Altynova, ES, Sarkar, R, Zhou, S, Merricks, EP et al. (2011). Efficacy and safety of long-term prophylaxis in severe hemophilia A dogs following liver gene therapy using AAV vectors. *Mol Ther* **19**: 442–449.
- 11 McIntosh, J, Lenting, PJ, Rosales, C, Lee, D, Rabbanian, S, Raj, D et al. (2013). Therapeutic levels of FVIII following a single peripheral vein administration of rAAV vector encoding a novel human factor VIII variant. *Blood* **121**: 3335–3344.
- 12 Scallan, CD, Liu, T, Parker, AE, Patarroyo-White, SL, Chen, H, Jiang, H et al. (2003). Phenotypic correction of a mouse model of hemophilia A using AAV2 vectors encoding the heavy and light chains of FVIII. *Blood* **102**: 3919–3926.
- 13 Chao, H, Sun, L, Bruce, A, Xiao, X and Walsh, C. (2003). Expression of human factor VIII by splicing between dimerized AAV vectors. *Mol Ther* **6**: 716–722.
- 14 Hirsch, ML, Li, C, Bellon, I, Yin, C, Chavala, S, Pryadkina, M et al. (2013). Oversized AAV transduction is mediated via a DNA-PKcs-independent, Rad51C-dependent repair pathway. *Mol Ther* **21**: 2205–2216.
- 15 Nathwani, AC, Gray, JT, Ng, CY, Zhou, J, Spence, Y, Waddington, SN et al. (2006). Self-complementary adeno-associated virus vectors containing a novel liver-specific human factor IX expression cassette enable highly efficient transduction of murine and nonhuman primate liver. *Blood* **107**: 2653–2661.
- 16 Costa, EL and Darnell, JE Jr. (1986). Transcriptional control of the mouse prealbumin (transferrin) gene: both promoter sequences and a distinct enhancer are cell-specific. *Mol Cell Biol* **6**: 4697–4708.
- 17 Levitt, N, Briggs, D, Gil, A and Proudfoot, NJ (1989). Definition of an efficient synthetic poly(A) site. *Genes Dev* **3**: 1019–1025.
- 18 Wu, Z, Zhang, T, Yin, C, Yin, F, Van Dyke, T et al. (2008). Optimization of self-complementary AAV vectors for liver-directed expression results in sustained correction of hemophilia B at low vector dose. *Mol Ther* **16**: 280–289.
- 19 Burnham, B, Nass, S, Kong, E, Mattingly, M, Woodcock, D, Song, A et al. (2015). Analytical ultracentrifugation as an approach to characterize recombinant adeno-associated viral vectors. *Hum Gene Ther Methods* **26**: 228–242.
- 20 Siner, JI, Iacobelli, NP, Sabatino, DE, Ivanciu, L, Zhou, S, Poncz, M et al. (2013). Minimal modification in the factor VIII B-domain sequence ameliorates the murine hemophilia A phenotype. *Blood* **121**: 4396–4403.
- 21 King, JA, Dubielzig, R, Grimm, D and Kleinschmidt, JA (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J* **20**: 3282–3291.
- 22 Kapranov, P, Chen, L, Dederich, D, Dong, B, He, J, Steinmann, KE et al. (2012). Native molecular state of adeno-associated viral vectors revealed by single-molecule sequencing. *Hum Gene Ther* **23**: 46–55.
- 23 Snyder, RO, Im, DS and Muzyczka, N (1990). Evidence for covalent attachment of the adeno-associated virus (AAV) rep protein to the ends of the AAV genome. *J Virol* **64**: 6204–6213.

- 24 Prasad, KM and Trempe, JP (1995). The adeno-associated virus Rep78 protein is covalently linked to viral DNA in a preformed virion. *Virology* **214**: 360–370.
- 25 Lenting, PJ, Christophe, OD and Guéguen, P (2010). The disappearing act of factor VIII. *Haemophilia* **16**: 6–15.
- 26 Saenko, EL and Scandella, D (1997). The acidic region of the factor VIII light chain and the C2 domain together form the high affinity binding site for von willebrand factor. *J Biol Chem* **272**: 18007–18014.
- 27 Chiu, PL, Bou-Assaf, GM, Chhabra, ES, Chambers, MG, Peters, RT, Kulman, JD *et al.* (2015). Mapping the interaction between factor VIII and von Willebrand factor by electron microscopy and mass spectrometry. *Blood* **126**: 935–938.
- 28 Nakai, H, Storm, TA and Kay, MA (2000). Recruitment of single-stranded recombinant adeno-associated virus vector genomes and intermolecular recombination are responsible for stable transduction of liver *in vivo*. *J Virol* **74**: 9451–9463.
- 29 Nakai, H, Yant, SR, Storm, TA, Fuess, S, Meuse, L and Kay, MA (2001). Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction *in vivo*. *J Virol* **75**: 6969–6976.
- 30 Yue, Y and Duan, D (2003). Double strand interaction is the predominant pathway for intermolecular recombination of adeno-associated viral genomes. *Virology* **313**: 1–7.
- 31 Yan, Z, Zak, R, Zhang, Y and Engelhardt, JF (2005). Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes. *J Virol* **79**: 364–379.
- 32 Choi, VW, McCarty, DM and Samulski, RJ (2006). Host cell DNA repair pathways in adeno-associated viral genome processing. *J Virol* **80**: 10346–10356.
- 33 Duan, D, Sharma, P, Yang, J, Yue, Y, Dudus, L, Zhang, Y *et al.* (1998). Circular intermediates of recombinant adeno-associated virus have defined structural characteristics responsible for long-term episomal persistence in muscle tissue. *J Virol* **72**: 8568–8577.
- 34 Schnepf, BC, Clark, KR, Klemanski, DL, Pacak, CA and Johnson, PR (2003). Genetic fate of recombinant adeno-associated virus vector genomes in muscle. *J Virol* **77**: 3495–3504.
- 35 Thomas, CE, Storm, TA, Huang, Z and Kay, MA (2004). Rapid uncoating of vector genomes is the key to efficient liver transduction with pseudotyped adeno-associated virus vectors. *J Virol* **78**: 3110–3122.
- 36 Wright, JF (2008). Manufacturing and characterizing AAV-based vectors for use in clinical studies. *Gene Ther* **15**: 840–848.
- 37 Allay, JA, Sleep, S, Long, S, Tillman, DM, Clark, R, Carney, G *et al.* (2011). Good manufacturing practice production of self-complementary serotype 8 adeno-associated viral vector for a hemophilia B clinical trial. *Hum Gene Ther* **22**: 595–604.
- 38 Yan, Z, Keiser, NW, Song, Y, Deng, X, Cheng, F, Qiu, J *et al.* (2013). A novel chimeric adeno-associated virus 2/human bocavirus 1 parvovirus vector efficiently transduces human airway epithelia. *Mol Ther* **21**: 2181–2194.
- 39 Rabinowitz, JE, Rolling, F, Li, C, Conrath, H, Xiao, W, Xiao, X *et al.* (2002). Cross-packaging of a single adeno-associated virus (AAV) type 2 vector genome into multiple AAV serotypes enables transduction with broad specificity. *J Virol* **76**: 791–801.
- 40 Murphy, SL, Bhagwat, A, Edmonson, S, Zhou, S and High, KA (2008). High-throughput screening and biophysical interrogation of hepatotropic AAV. *Mol Ther* **16**: 1960–1967.



This work is licensed under a Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>

Supplementary Information accompanies this paper on the *Molecular Therapy—Methods & Clinical Development* website (<http://www.nature.com/mtm>)