Molecular analysis of AAV5-hFVIII-SQ vector-genome-processing kinetics in transduced mouse and nonhuman primate livers

Choong-Ryoul Sihn,1 Britta Handside,1 Su Liu,1 Lening Zhang,1 Ryan Murphy,1 Bridget Yates,1 Lin Xie,1 Richard Torres,2 Chris B. Russell,2 Charles A. O’Neill,3 Erno Pungor,4 Stuart Bunting,1 and Sylvia Fong1

1Biology Research, BioMarin Pharmaceutical, Inc., Novato, CA, USA; 2Bioanalytical Sciences, BioMarin Pharmaceutical, Inc., Novato, CA, USA; 3Pharmacological Sciences, BioMarin Pharmaceutical, Inc., Novato, CA, USA; 4Analytical Chemistry, BioMarin Pharmaceutical, Inc., Novato, CA, USA

Valoctocogene roxaparvovec (AAV5-hFVIII-SQ) is an adenovector-associated virus serotype 5 (AAV5)-based gene therapy vector containing a B-domain-deleted human coagulation factor VIII (hFVIII) gene controlled by a liver-selective promoter. AAV5-hFVIII-SQ is currently under clinical investigation as a treatment for severe hemophilia A. The full-length AAV5-hFVIII-SQ is >4.9 kb, which is over the optimal packaging limit of AAV5. Following administration, the vector must undergo a number of genome-processing, assembly, and repair steps to form full-length circularized episomes that mediate long-term FVIII expression in target tissues. To understand the processing kinetics of the oversized AAV5-hFVIII-SQ vector genome into circular episomes, we characterized the various molecular forms of the AAV5-hFVIII-SQ genome at multiple time points up to 6 months postdose in the liver of murine and non-human primate models. Full-length circular episomes were detected in liver tissue beginning 1 week postdosing. Over 6 months, quantities of circular episomes (in a predominantly head-to-tail configuration) increased, while DNA species lacking inverted terminal repeats were preferentially degraded. Levels of duplex, circular, full-length genomes significantly correlated with levels of hFVIII-SQ RNA transcripts in mice and non-human primates dosed with AAV5-hFVIII-SQ. Altogether, we show that formation of full-length circular episomes in the liver following AAV5-hFVIII-SQ transduction was associated with long-term FVIII expression.

INTRODUCTION

Hemophilia A is an X-linked bleeding disorder caused by a deficiency in factor VIII (FVIII) coagulation protein activity.1,2 People with hemophilia A are susceptible to spontaneous and traumatic bleeding into soft tissues and joints that can result in painful, disabling arthropathy and impaired quality of life, including possible intracranial hemorrhage and early death.2,3 Hemophilia A is currently managed with chronic administration of exogenous FVIII, either prophylactically or in response to bleeding events.

Gene therapy with valoctocogene roxaparvovec (AAV5-hFVIII-SQ) is being developed for long-term management of severe hemophilia A.4-6 AAV5-hFVIII-SQ is a recombinant, replication-incompetent adenovector-associated virus serotype 5 (AAV5) gene therapy vector containing a genome comprising a single-stranded, codon-optimized, B-domain-deleted human FVIII gene (hFVIII-SQ) controlled by a hepatocyte-selective promoter, a synthetic polyadenylation (poly(A)) sequence, and AAV2-derived double-stranded inverted terminal repeats (ITRs) on the 5′ and 3′ ends.1-6 The total vector genome is more than 4.9 kb in length, exceeding the optimal packaging capacity of AAV vectors.7-9 As a result, incomplete genomes may be present in the vector preparation. In a murine model of hemophilia A treated with AAV5-hFVIII-SQ at doses from 2 × 1013 to 6 × 1013 vg/kg, normal to supraphysiological levels of circulating hFVIII-SQ protein were produced, and improvements in bleeding time and reductions in blood loss were achieved.4 In addition, comparable doses in primates produced similar levels of circulating hFVIII-SQ protein.2 In a phase 1/2 clinical trial, a single infusion of AAV5-hFVIII-SQ dosed at 6 × 1013 and 4 × 1013 vg/kg resulted in a clinically relevant reduction in bleeding through 5 and 4 years of follow-up postinfusion, respectively, for adult men with severe hemophilia A.5,10 The most common adverse events were transient, asymptomatic alanine aminotransferase increases that resolved without clinical sequelae.5,6,10 In a phase 3 trial, an infusion of 6 × 1013 vg/kg AAV5-hFVIII-SQ provided significantly increased FVIII activity and reduced bleeding at 52 weeks for adult men with severe hemophilia A.11 While these findings represent a substantial leap forward in the clinic, there are gaps in our biological understanding of the mechanisms behind long-term expressions of FVIII.

Multiple complex processes are involved in achieving successful and durable transgene expression using an AAV vector (Figure 1). Following dosing and target cell receptor-mediated uptake of the viral vector-genome-processing kinetics in transduced mouse and nonhuman primate livers
capsid, the single-stranded, linear vector DNA is transported into the nucleus. Vector-genome-processing can then occur, transforming the single-stranded vector DNA to double-stranded linear genomes via second-strand DNA synthesis. For vector genomes exceeding the packaging limit of the AAV capsid, the majority of vector genomes packaged are fragmented or clipped. To generate full-length vector genomes in cells, overlapping vector genomes of opposite strand polarities are first annealed, followed by second-strand synthesis and/or homologous recombination. This process is facilitated by Rad51C, a single-stranded DNA-binding protein that promotes strand annealing in response to double-strand breaks. The resultant double-stranded linear genomes are converted into stable monomeric and concatemeric circular episomes when the ITRs on the 5' and 3' ends of the linear vector genome structure undergo homologous recombination or non-homologous end-joining.

Because 97% of the overall genomes were degraded, we hypothesized that incomplete genomes lacking ITRs were being preferentially degraded over the first 8 weeks. To assess this, we used a variant of the ddPCR method, referred to as drop-phase ddPCR.

RESULTS

AAV5-hFVIII-SQ vector-genome-processing in mouse livers

To assess vector-genome-processing kinetics soon after vector administration, 8-week-old Rag2-/- FVIII-/- double knockout mice were used as an animal model of hemophilia A. Mice were dosed with 3.5 x 10^13 vg/kg AAV5-hFVIII-SQ (the size distribution of AAV5-hFVIII-SQ is shown in Figure S1), and livers were analyzed 24 h and 1, 3, 5, and 8 weeks thereafter. Quantitative real-time polymerase chain reactions (quantitative real-time PCR) using probes and primers specific to the hFVIII-SQ transgene were used to quantify hFVIII-SQ DNA abundance (SQ amplicon; Figure S2). Mice dosed with AAV5-hFVIII-SQ showed an initial sharp decline in overall hFVIII-SQ DNA in the first 3 weeks postadministration; by week 5 postdose, 97.3% of hFVIII-SQ vector genomes had been degraded compared with by 24 h postdose (Figure 2A). Comparison of levels of cytoplasmic and nuclear vector genomes using DNA in situ hybridization showed that an elimination of genomes within the cytoplasm was responsible for the majority of the decline in vector DNA observed in the first 3 weeks; vector genome signals in the nuclei also declined by 82% between 1 and 3 weeks (Figure S3). Consistent with droplet digital PCR (ddPCR) results, by 8 weeks, the vector DNA signals decreased 99% and 91% in the cytoplasm and in nuclei, respectively, compared with by 24 h postdose (Figure 2A). Comparison of levels of cytoplasmic and nuclear vector genomes using DNA in situ hybridization showed that an elimination of genomes within the cytoplasm was responsible for the majority of the decline in vector DNA observed in the first 3 weeks; vector genome signals in the nuclei also declined by 82% between 1 and 3 weeks (Figure S3). Consistent with droplet digital PCR (ddPCR) results, by 8 weeks, the vector DNA signals decreased 99% and 91% in the cytoplasm and in nuclei, respectively, compared with after 24 h. In the liver, levels of hFVIII-SQ RNA assessed by quantitative real-time PCR increased through 8 weeks postdose (Figure 2B) in contrast to the observed pattern of decreased DNA. Consistent with increasing liver hFVIII-SQ RNA levels, circulating plasma hFVIII-SQ protein, quantified via enzyme-linked immune-absorbent assay (ELISA), also increased over time (Figure 2C).

Because 97% of the overall genomes were degraded, we hypothesized that incomplete genomes lacking ITRs were being preferentially degraded over the first 8 weeks. To assess this, we used a variant of the ddPCR method, referred to as drop-phase ddPCR,
to determine the contiguity of vector genomes between the 5’ end, the 3’ end, and the middle (Figure S4). This method can detect and quantify the co-occurrence of separate target sequences individually labeled with different fluorescent probes that are present in a single droplet and therefore linked on a contiguous DNA strand. Various primers and probes were designed to generate short amplicons spanning approximately 1,000 base pairs internal to the 5’ ITR (eg, R1–R6 amplicons) or 3’ ITR (eg, R7–R11 amplicons) of AAV5-hFVIII-SQ. To confirm the validity of this method, we determined that full-length linear genomes and circular monomers and concatemers of known DNA forms could be detected (Figure S5). Then, each of these primers/probes was combined with a primer pair/probe that amplified the middle of the genome (SQ amplicon) in each duplex reaction (Figures 3A and S2). The relative proportions of vector genomes containing sequences spanning approximately 1 kb internal to either 5’ or 3’ ITR to vector genomes containing the SQ sequences increased significantly from 24 h to 3 weeks and increased only slightly from 3 weeks to 8 weeks post AAV5-hFVIII-SQ administration (Figures 3B–3F). Both vector genomes lacking either ITR but containing the SQ region (single positive in the drop-phase ddPCR reaction) and vector genomes containing the SQ region connected to an ITR (double positives) were rapidly lost during the first 3 weeks postdose, with the rate of loss decreasing substantially after week 3 (Figures 3G and 3H). Incomplete vector genomes (as defined by single positives), particularly those lacking either the 5’ or 3’ ITR, were preferentially degraded at a faster rate over 8 weeks than were vector genomes that contained either ITR (double positives).

To further investigate the mechanism and kinetics of degradation during this period, we examined vector structural connectivity with a set of 4 drop-phase amplicons spaced along the vector (amplicon locations are shown in Figure S2). We used a set of paired amplicon reactions to identify the fraction of each amplicon that was unconnected to any adjacent amplicons, representing potential intermediates on the degradation pathway, where sites are contained on small fragments disconnected from the rest of the vector genome. Partial vector fragments, containing one end or the other, are largely either degraded or assembled into more stable full-length forms after being released from the capsid. All 4 sites were predominantly connected in the first sample, from 24 h after dosing, when much of the DNA is likely still in the original vector form. If released DNA were rapidly degraded or assembled, we would only observe connectivity similar to the 24 h sample or increased connectivity. If released DNA were largely degraded by exonuclease activity from the exposed single- or double-stranded ends of the partial vector, we would observe a decrease in the relative amount of connected middle amplicons and an increase in disconnected ITR-proximal sites, while degradation from the ITR ends would create an increase in fragments that contain only the middle sites. Instead, all 4 amplicons had similar kinetics (Figure 3I), with each site reaching maximal fragmentation with about 60% unconnected to any other site, at 1-week postdose, when the majority of vector genomes were disappearing. As with the ITR proximal sites, the vector DNA remaining at 8 weeks was more highly connected. This suggests that vector genome degradation includes an active endonuclease fragmentation mechanism where released DNA is broken into smaller fragments that have sufficient persistence to be the majority form at the time point of maximal DNA degradation.

Circular full-length vector-genome-processing kinetics in mouse liver

We next examined the kinetics of the formation of circular full-length or ITR-fused AAV5-hFVIII-SQ vector genomes over 24 weeks in mice dosed with 2 × 10^{13}, 3.5 × 10^{13}, and 6 × 10^{13} vg/kg in 2 separate studies. DNA was extracted from mouse livers, and samples were digested with Plasmid Safe ATP-Dependent DNase (PS-DNase; Lucigen, Middleton, WI, USA) to hydrolyze linear DNA and isolate circular DNA. Full-length genomes were quantified using drop-phase ddPCR with primers/probes on the proximal end of the 5’ or 3’ ITRs (the D-segment) to generate

Figure 2. Assessment of AAV5-hFVIII-SQ vector genome levels and transgene expression in livers from mice dosed with 3.5 × 10^{13} vg/kg AAV5-hFVIII-SQ using quantitative real-time PCR

(A) Overall hFVIII-SQ vector DNA levels as measured by SQ amplicon in mouse liver over 8 weeks. (B) Liver hFVIII-SQ RNA levels over 8 weeks. (C) Plasma hFVIII protein levels over 8 weeks. Data are mean ± 1 standard deviation. At all timepoints, n = 10. (A) and (B) use a log scale y axis. hFVIII-SQ, human Factor VIII; hr, hour; wk, week.
amplicons R1 and R11 (Figure S2), thus spanning the entire vector genome unit (5’ D-segment, promoter, transgene, and poly(A) signal, 3’ D-segment). ITR-fused genome forms were identified using primers, and probe against a DNA sequence formed only when the 5’ and 3’ ITRs are joined by homologous recombination (Figure S6).

DNA samples were also digested with the restriction enzyme KpnI to enumerate each full-length head-to-tail vector genome units within concatemeric forms following PS-DNase treatment (Figures 4A and S6). No circular full-length vector genome DNA copies or ITR fusions were detected at 24 h postdosing; however, they were present at 1-week postdose and increased slowly through 8 weeks (Figure 4B).
Figure 4. Circular episome formation in mice treated with AAV5-hFVIII-SQ

(A) Schematic diagram showing enzymatic treatment of samples to reveal quantities of head-to-tail full-length genome units in circular episomes using drop-phase ddPCR. (B and C) Over 8 weeks postdosing with $3.5 \times 10^{13}$ vg/kg AAV5-hFVIII-SQ, levels of (B) full-length circular hFVIII-SQ vector DNA and (C) circular ITR-fused vector genomes in the liver. (D and E) Over 24 weeks, (D) full-length circular hFVIII-SQ vector DNA levels in livers from mice dosed with either $6 \times 10^{13}$ or $2 \times 10^{13}$ vg/kg and (E) circular ITR-fused vector genomes in livers of mice dosed with $6 \times 10^{13}$ vg/kg. (F) Schematic representation of circular episome structures using specific probes as detected by Southern blotting. (G) Southern blot showing circular episome vector concatemeric structure (left panel) and orientation (right panel). Full-length vector genomes were identified with linked R1-R11 amplicons and are defined as sequences including the D-segment of the 5' ITR, the promoter, the h-FVIII-SQ transgene, the polyadenylation signal, and the D-segment of the 3' ITR. The same primers and probes were used in both southern blots shown in (G). In (G), J1–J7 represent individual mice, hFVIII-SQ, human factor VIII; H, head (5' end); HH, head-to-head orientation; HT, head-to-tail orientation; TT, tail-to-tail orientation; M, monkey; PS, Plasmid Safe ATP-Dependent DNase; T, tail (3' end); vgDNA, vector genome DNA; wk, week; Veh, vehicle.
as did ITR fusion genomes (Figure 4C). Levels of both circular full-length episomes (Figure 4D) and ITR fusions (Figure 4E) remained stable through 24 weeks, and higher levels of circular full-length genomes resulted from the higher dose. Despite the reduction in overall vector genome DNA, repaired circular full-length and ITR-fused vector genome DNA species were formed and increased over time. Levels of circular full-length genomes were similar to levels of ITR-fused genomes, suggesting that most ITR-fused genomes are full-length genomes.

Southern blotting was performed to visualize the structure and orientation of genome forms using probes near the 5' (head) and 3' (tail) ends with monomeric/multimeric AAV5-hFVIII-SQ DNA used as a control (Figure S7). Consistent with ddPCR analysis, Southern blotting demonstrated that the circular episomes were present persistently at 4 through 24 weeks postdosing; bands migrated at sizes corresponding to full-length monomeric and concatemeric forms, mainly in the head-to-tail configuration (Figures 4Ga and S8).

Levels of hFVIII-SQ RNA transcripts were also assessed in mice. The quantity of RNA transcripts was significantly correlated with levels of overall hFVIII-SQ vector genomes as measured by the SQ amplicon (r = 0.46; p < 0.001; Figure 5A) but were more strongly correlated with full-length hFVIII-SQ vector genomes (r = 0.64; p < 0.0001; Figure 5B) and ITR-fused genomes (r = 0.60; p < 0.0001; Figure 5C).

DISCUSSION

In recent years, AAV-based gene transfer has rapidly advanced in the clinic, with multiple FDA-approved products on the market and even more in phase 3 trials. At the molecular level, however, much remains unknown. In this report, we studied the kinetics and structures of AAV5-hFVIII-SQ vector genomes in the livers of mice and non-human primates for up to 6 months following dosing. Consistent with previous reports, treatment of mice with AAV5-hFVIII-SQ was followed initially by a sharp decline in overall vector genome DNA: more than 97% of the hFVIII-SQ genome was degraded by 8 weeks postdose. The elimination of genomes in the cytoplasm and within the nuclei contributes to this decline. Incomplete genomes, particularly those lacking either 5' or 3' ITRs, were preferentially degraded at a faster rate. However, full-length circular episomes were observed in the liver 1-week postdose. While the total number of vector genomes present declined, the proportion of full-length duplex ITR fusions and circular episomes increased over time. Circular episomes were also present in non-human primates and persisted in the liver for up to 6 months, the longest time point assessed. These results are consistent with the proposed genome-processing model following AAV-mediated gene transduction, where most single-stranded vector genomes are destroyed, and a small portion are assembled into stable circular DNA.
episomes that resist degradation and mediate stable protein expression (Figure 1). Levels of FVIII RNA present in the mouse and non-human primate liver samples correlated with the levels of the full-length hFVIII-SQ genome capable of giving rise to transcription. Protein presence was first detected 3 weeks after dosing; the latency is expected due to the slow uncoating of the AAV5 capsid and the time required to process genomes into forms capable of giving rise to stable transcription and then the translation and secretion of sufficient protein quantities above the limit of detection. Head-to-tail circular genomes that form after AAV administration appear to be more stable than non-ITR, linear, and double-stranded genomes. Similar to previous studies using other AAV gene therapies, head-to-tail episomes are the primary forms of circular genomes in mice and non-human primates treated with AAV5-hFVIII-SQ; this configuration also appears to support stability of episomes and more efficient transcription of the transgene. Together, these data suggest that following AAV5-hFVIII-SQ transduction, the formation of circular episomes in the liver, particularly in the head-to-tail orientation, leads to efficient transcription and long-term FVIII expression in the liver.

In a canine model of severe hemophilia A, treatment with AAV-canine-FVIII-SQ resulted in persistent FVIII activity and correction of bleeding phenotypes for up to 10 years postdose. Analysis of terminal liver samples revealed more than 95% of remaining vector genomes were present in episomal forms, suggesting that episomes, not vector genomes integrated into the nuclear genome, may be the major DNA form responsible for long-term transgene expression; however, expression arising from integrated genomes cannot be ruled out. Nevertheless, levels of full-length episomes present in canine livers were positively associated with levels of plasma FVIII activity. Together with our results, these data demonstrate that circularized episomes can resist degradation and support long-term expression of transgenes, including oversized genomes such as AAV5-hFVIII-SQ.

In these preclinical studies, limitations include the relatively small sample sizes and the short maximum follow-up time of 6 months. Evidence of longer-term durability is pending additional follow-up. PS-DNase also degrades 25%–35% of circular plasmid DNA (data not shown), as the DNA extraction process likely linearizes and shears large concatemeric circular episomes, rendering them sensitive to PS-DNase. Therefore, the levels of circular vector DNA quantified following PS-DNase treatment represent the lowest possible value, but the proportion of linear versus circular genomes cannot be calculated accurately as it is unknown how many circular episomes are sheared during the extraction process. Additionally, expression arising from the integration of a small portion of full-length genomes into the host nuclear genome cannot be ruled out and may contribute to the results presented here.

Following treatment with AAV5-hFVIII-SQ, we observed dose-dependent increases in the quantities of circular full-length and ITR-fused vector genomes in the livers of mice and non-human primates that were associated with long-term expression of hFVIII-SQ. These results, obtained using a novel molecular technique, support the hypothesis that circular episomes formed in the weeks following AAV-vector-mediated gene transfer lead to long-term expression of episomes that resist degradation and mediate stable protein expression (Figure 1).
Molecular methods

DNA enzyme treatment procedures

DNA was extracted from frozen intact liver tissues using a MagMAX DNA Multi-Sample Ultra kit with the KingFisher Flex System (ThermoFisher Scientific, Waltham, MA, USA) following manufacturer’s instructions. Extracted DNA was diluted to 20 ng/μL. Samples were digested with PS-DNase (Lucigen, Middleton, WI, USA) to hydrolyze all linear forms of DNA and isolate circular DNA. The number of circular genomes is underestimated using PS-DNase because the DNA extraction process would likely shear/linearize some large concatemeric circular episomes, rendering them sensitive to PS-DNase (it is unclear the extent to which circular episomes are sheared during the extraction process), and PS-DNase has been shown to degrade 25%–35% of circular plasmid DNA (data not shown). Prior to droplet generation, 200 ng of total DNA was incubated for 16 h at 37°C with 50 units/μg of PS-DNase in 33 mM Tris-acetate (pH 7.5), 66 mM potassium acetate, 10 mM magnesium acetate, 0.5 mM dithiothreitol, and 1 mM adenosine triphosphate. PS-DNase was then inactivated with a 20 min incubation at 80°C, and samples were diluted to 2 ng/μL in 10 mM Tris-Cl (pH 8.5) and 0.05% pluronic F-68. For the ddPCR reactions, 5 μL of dilute sample was used.

ddPCR was also performed with another set of samples treated with both PS-DNase and KpnI restriction enzymes to quantify individual vector genome units within circular concatemers. DNA samples were treated with PS-DNase, heat inactivated, and diluted, as previously described. Prior to the droplet generation step, 4 units of KpnI-HF enzyme (New England Biolabs, Ipswich, MA, USA) were added to the ddPCR reaction mix and incubated at 37°C for 30 min.

ddPCR procedures

Quantities of vector genome forms in samples were measured with ddPCR, which captures individual DNA molecules in thousands of water-oil emulsion droplets prior to PCR amplification with fluorescent tags. Individual droplets are then counted as negative or positive using fluorescence, and Poisson statistics are applied to the fraction of positive droplets to estimate the copy number of target DNA molecules per sample. In this analysis, a variant of ddPCR called drop-phase ddPCR was performed to detect and quantify the levels of paired target sequences together on a single DNA molecule to measure the contiguity of the DNA molecule using 2 different fluorescent tags (FAM and HEX). The number of double-positive droplets is then calculated, and the total copy number of molecules with both target sequences is estimated using the software QuantaSoft (Bio-Rad, Hercules, CA, USA), which includes an algorithm that accounts for the probability that some double-positive droplets may occur due to chance. ddPCR of the endogenous gene TFRC (transferrin receptor protein 1) was also performed to provide a normalization reference for calculating vector copy numbers per diploid genome from the concentration of droplets containing HEX or FAM fluorophores.

The primer and probe sets used for ddPCR and drop-phase ddPCR are presented in Table S1. Fusions of the 5’ and 3’ ITRs (head-to-tail) were detected using forward and reverse primers located on...
the 5’ and 3’ ends of the linear genome that only produced an ampli-
cion when 5’ and 3’ ITRs were fused. Cross-ITR amplification was per-
formed in a 3-step ddPCR reaction with 35 s annealing at 59°C, 65 s
further extension at 72°C, and 30 s denaturation at 94°C. Amplifica-
tion was detected at high efficiency from circular constructs contain-
ing head-to-tail ITR fusions but not from linear constructs, head-to-
head constructs, or starting vector. Full-length vector genomes
able of giving rise to stable hFVIII-SQ transcription were detected
the ITR on the 5’ and 3’ ends of the genome, respectively (Figure S2).

Following treatment with KpnI and/or PS-DNase, 10 ng of DNA were
used in each ddPCR reaction. The reaction mixture contained 1x
dPCR Supermix for Probes without deoxyuridine triphosphate (Bio-Rad), 250 mM each of forward and reverse primers, 900 nM of
probes, and 5 μL of sample for a final reaction volume of 25 μL. A
Bio-Rad Auto Droplet Generator was used to generate droplets
from reaction mix and QX200 Droplet Generation Oil for Probes
(Bio-Rad), which were then transferred into a 96-well plate. PCR
was performed in a C1000 Touch Thermal Cycler (Bio-Rad) as fol-

drop-phase ddPCR reactions identifying the co-occurrence of
amplions of R1 and R11, which overlap with the D-segments of

One caveat of using ddPCR is that one droplet captures a single mole-
cule and counts it as one positive unit, regardless of whether it is a
monomer or multimeter containing multiple units of vector genomes.
This means the number of vg units could be undercounted if conca-
temers exist in the samples. To that end, we used the restriction
enzyme KpnI to make cuts within the circular genome and release in-
dividual vector genome DNA units from concatemers. Therefore,
an increase in the quantities of vector genomes upon KpnI digestion sug-
gests circular concatemers are present in the sample. In Figure S5B,
circular monomeric DNA was generated using the same linear
DNA fragment used previously and was subjected to drop-phase
dPCR using R1 and R11 primers and probes sets. Similar copy
numbers were observed with or without KpnI digestion in each com-
parison for the monomeric DNA. In contrast, if multimeric circular
controls as templates were used (Figure S5C), copy numbers increased upon KpnI digestion.

In situ hybridization to detect hFVIII-SQ DNA
Formalin-fixed paraffin-embedded mouse liver sections (5 μm) were
collected on Superfrost Plus slides using RNase-free conditions. An
RNAsec in situ hybridization protocol was performed using a Ven-
tana Discovery Ultra Autostainer (Tucson, AZ, USA). RNAsec
(Newark, CA, USA), Universal 2.5 Reagent Kit, and custom-generated
hFVIII-SQ probes. Slides were counterstained with 4’-6-diamidino-2-
phenylindole, and mounted with Fluoromount G. (catalogue number
[cat. no.] 17984-25, Electron Microscopy Sciences). Slides were
imaged on an Axio Scan.Z1 (Zeiss) slide scanner using a Plan-Apo-
chromat 20x/0.8 objective equipped with a Hamamatsu Orca Flash
camera. The intensity of hFVIII-SQ DNA in situ hybridization signal
within hepatocyte nuclei and/or hepatocyte cytoplasm was quantified
using Visiopharm image analysis software (Hoersholm, Denmark).

Southern blot procedures
Southern blotting was used to identify the configuration of circular
episomes. For each sample, 12 μg of DNA was digested with 50 units
each of EcoRI and HindIII restriction enzymes (New England Bio-
labs), which do not cut within vector genomes for 2 h, followed by
DS-DNase treatment at 37°C for 16 h. The reaction was halted by
heat inactivation at 80°C for 20 min; an additional set of samples
was also incubated at 37°C for 2 h with 10 units of KpnI. For all sam-

des, DNA was isolated with phenol-chloroform extraction, precipi-
tated in ethanol, and resuspended in 30 μL of nuclease-free water.

DNA samples were mixed with 6X gel loading dye (New En-
gland Biolabs), loaded into a 0.7% agarose gel containing 0.5X SYBR
Safe DNA Gel Stain dye (ThermoFisher Scientific), and electropho-
resis was performed at 30 to 35 V for 16 to 18 h at room temperature.
DNA was depurinated in 0.5 M NaOH and 1.5 M NaCl denaturing
buffer for 30 min at room temperature with gentle agitation, neutral-
ized in 1.5 M NaCl and 0.5 M Tris-HCl (pH 7.0) neutralizing buffer
for 30 min, and then soaked in 20X SSC transfer buffer (3M NaCl and
sodium citrate [pH 7.0]) for 30 min.

DNA was then transferred to a positively charged nylon mem-
brane using a Whatman Nytran SuPerCharge TurboBlotter system
Quantitative real-time PCR reactions of 20 µL were set up with 2X iQ Multiplex Powermix (Bio-Rad), 300 nM of forward and reverse primers, 150 nM of fluorescent probes targeting hFVIII-SQ cDNA, 100 ng of cDNA, and DNase/RNase-free water (Invitrogen, Carlsbad, CA, USA). Primer and probe sequences are listed in Table S3. Samples were loaded into clear 384-well PCR plates and run on a Roche Light Cycler 480 II (Roche, Basel, Switzerland) at 95°C for 2 min then for 50 cycles at 95°C for 3 s and 60°C for 20 s. Fluorescent signals at 483–533 nm were used to detect the hFVIII-SQ amplicon, and cross-point cycle values were calculated for each reaction by the Light Cycler 480 Software v. 1.5.1 using Abs quant/second derivative max analysis (Roche).

**hFVIII-SQ protein quantification**

Levels of hFVIII-SQ protein in sodium-citrate-anticoagulated plasma samples were measured using a sandwich ELISA with a human-specific anti-FVIII capture (GMA-9023, Green Mountain Antibodies, Burlington, VT, USA) and a detection (F8C-eIA, Affinity Biologicals, Ancaster, ON, Canada) antibody pair to specifically measure human FVIII. Nunc MaxiSorp high-binding black polypropylene plates (Thermo Fisher Scientific) were coated with 4 µg/mL of anti-FVIII (domain A2) antibodies. Samples were diluted 1:10 in Diluent Buffer (6% BSA in 1X TBS-T) and incubated for 1 h at room temperature. Sheep anti-FVIII antibodies conjugated to horseradish peroxidase were added and incubated for 1 h at room temperature. After a final wash, QuantaBlu substrate solution (Thermo Fisher Scientific) was used for detection. Relative fluorescent units were detected on a Molecular Devices FlexStation 3 instrument (Molecular Devices, San Jose, CA, USA), and concentrations were extrapolated from a standard curve prepared by spiking clinical grade recombinant B-domain-deleted hFVIII (Xyntha) in human FVIII-deficient plasma. Raw data was acquired using SoftMax Pro 5.4.1 (Molecular Devices).

**Statistical analysis**

Pearson’s correlation between levels of duplex linear and circular R1-R11 linked genomes per diploid genome or ITR-fused genomes and levels of FVIII-SQ transcript per µg of total RNA was performed using Prism software (GraphPad Software, San Diego, CA, USA).

**Distribution of materials and data**

Materials and protocols will be distributed to qualified scientific researchers for non-commercial academic purposes. The AAV5-hFVIII-SQ vector and the AAV5-hFVIII-SQ vector sequence are part of an ongoing development program, and they will not be shared.

**SUPPLEMENTAL INFORMATION**

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2021.12.004.

**ACKNOWLEDGMENTS**

Funding for this study was provided by BioMarin Pharmaceutical Inc. Medical writing support was provided by Sara Hawley, MS, and Micah Robinson, PhD, of BioMarin Pharmaceutical, and by Kathleen Pieper, PhD, and Atreju Lackey, PhD, of AlphaBioCom, LLC, and was
funded by BioMarin Pharmaceutical Inc. The research was performed in Reno, NV, and Novato, CA, in the USA.

AUTHOR CONTRIBUTIONS
S.F., E.P., C.A.O., C.B.R., and S.F. contributed to the study design. S.F. and S.S. oversaw the conduct of the study. L.Z., R.M., and L.X. participated in the conduct of the animal study. C.-R.S., B.H., R.T., S.L., and B.Y. performed the biochemical and molecular analyses. C.-R.S. and R.T. performed the statistical analysis. All authors contributed to the interpretation of the results, critically reviewed the manuscript during writing, and approved of the final draft for submission.

DECLARATION OF INTERESTS
C.-R.S., B.H., L.Z., R.M., B.Y., L.X., C.A.O., E.P., S.B., and S.F. are employees and stockholders of BioMarin Pharmaceutical. R.T. and S.L. are former employees of BioMarin Pharmaceutical and may hold stock.

REFERENCES
1. Iorio, A., Stonebraker, J.S., Chambost, H., Makris, M., Coffin, D., Herr, C., et al.; Data and Demographics Committee of the World Federation of Hemophilia (2019). Establishing the prevalence and prevalence at birth of hemophilia in males: a meta-analytic approach using national registries. Ann. Intern. Med. 171, 540–546.
2. Srivastava, A., Santagostino, E., Dougall, A., Kitchen, S., Sutherland, M., Pipe, S.W., Carcao, M., Mahlangu, J., Ragni, M.V., Windyga, J., et al. (2020). WFIN guidelines for the management of hemophilia, Third edition. Haemophilia. 26, 1–158.
3. Darby, S.C., Kan, S.W., Spooner, R.J., Giangrande, P.L., Hill, F.G., Hay, C.R., Lee, S.A., Ludlam, C.A., and Williams, M. (2007). Mortality rates, life expectancy, and aging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. J. Clin. Invest. 118, 815–825.
4. Bunting, S., Zhang, L., Xie, L., Bullens, S., Minihinkar, R., Fong, S., Sandza, K., Harmon, D., Yates, B., Handside, B., et al. (2018). Gene Therapy with BMN 270 results in therapeutic levels of FVIII in mice and primates and normalization of bleeding in hemophilia mice. Mol. Ther. 26, 496–509.
5. Rangarajan, S., Walsh, L., Lester, W., Perry, D., Madan, B., Laffan, M., Yu, H., Vettermann, C., Pierce, G.F., Wong, W.Y., et al. (2017). AAV5-factor VIII gene transfer in severe hemophilia A. N. Engl. J. Med. 377, 2519–2530.
6. Pasi, K.J., Rangarajan, S., Mitchell, N., Lester, W., Symington, E., Madan, B., Laffan, M., Russell, C.B., Li, M., Pierce, G.F., et al. (2020). Multiyear follow-up of AAV5-hFVIII-SQ gene therapy for hemophilia A. N. Engl. J. Med. 382, 29–40.
7. Albocca, M., Doria, M., Petrallo, M., Colella, P., Garcia-Hoyos, M., Gibbs, D., Kim, S.R., Maguire, A., Rex, T.S., Di Vicino, U., et al. (2008). Serotype-dependent packaging of large genes in adeno-associated viral vectors results in effective gene delivery in mice. J. Clin. Invest. 118, 1955–1964.
8. Lai, Y., Yue, Y., and Duan, D. (2010). Evidence for the failure of adeno-associated virus serotype 5 to package a viral genome > or = 8.2 kb. Mol. Ther. 18, 75–79.
9. Hirsch, M.L., Li, C., Bellon, I., Yin, C., Chavala, S., Pryadkina, M., Richard, I., and Samulski, R.J. (2013). Oversized AAV transduction is mediated via a DNA-PKcs-independent, Rad51C-dependent repair pathway. Mol. Ther. 21, 2205–2216.
10. Pasi, K.J., Laffan, M., Rangarajan, S., Robinson, T.M., Mitchell, N., Lester, W., Symington, E., Madan, B., Yang, X., Kim, B., et al. (2021). Persistence of haemostatic response following gene therapy with valoctocogene roxaparvovec in severe haemophilia A. Haemophilia. 27, 947–956.
11. Oezel, M.C., Mahlangu, J., Pasi, K., Gierrmasz, A., Leavitt, A.D., Laffan, M., et al. (2021). Efficacy and safety of valoctocogene roxaparvovec adeno-associated virus gene transfer for severe haemophilia A: results from the phase 3 GENEf8-1 trial [abstract]. Res. Pract. Thromb. Haemost. 5, 89–90.
12. 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.
33. Grimm, D., Pandey, K., Nakai, H., Storm, T.A., and Kay, M.A. (2006). Liver transduction with recombinant adeno-associated virus is primarily restricted by capsid serotype not vector genotype. J. Virol. 80, 426–439.

34. Schnepp, B.C., Chulay, J.D., Ye, G.J., Flotte, T.R., Trapnell, B.C., and Johnson, P.R. (2016). Recombinant adeno-associated virus vector genomes take the form of long-lived, transcriptionally competent episomes in human muscle. Hum. Gene Ther. 27, 32–42.

35. McIntosh, J., Lenting, P.J., Rosales, C., Lee, D., Rabbanian, S., Raj, D., Patel, N., Tuddenham, E.G., Christophe, O.D., McVey, J.H., et al. (2013). Therapeutic levels of FVIII following a single peripheral vein administration of rAAV vector encoding a novel human factor VIII variant. Blood 122, 3335–3344.

36. Nakai, H., Fuess, S., Storm, T.A., Meuse, L.A., and Kay, M.A. (2003). Free DNA ends are essential for concatemerization of synthetic double-stranded adeno-associated virus vector genomes transfected into mouse hepatocytes in vivo. Mol. Ther. 7, 112–121.

37. Yan, Z., Zhang, Y., Duan, D., and Engelhardt, J.F. (2000). Trans-splicing vectors expand the utility of adeno-associated virus for gene therapy. Proc. Natl. Acad. Sci. U S A 97, 6716–6721.

38. Yang, N.S., Wang, J.H., and Turner, J. (2004). Molecular strategies for improving cytokine transgene expression in normal and malignant tissues. Gene Ther. 11, 100–108.

39. Nakai, H., Storm, T.A., Fuess, S., and Kay, M.A. (2003). Pathways of removal of free DNA vector ends in normal and DNA-PKcs-deficient SCID mouse hepatocytes transduced with rAAV vectors. Hum. Gene Ther. 14, 871–881.

40. Snyder, R.O., Spratt, S.K., Lagarde, C., Bohd, D., Kaspar, B., Sloan, B., Cohen, L.K., and Danos, O. (1997). Efficient and stable adeno-associated virus-mediated transduction in the skeletal muscle of adult immunocompetent mice. Hum. Gene Ther. 8, 1891–1900.

41. Nakai, H., Thomas, C.E., Storm, T.A., Fuess, S., Powell, S., Wright, J.F., and Kay, M.A. (2002). A limited number of transducible hepatocytes restricts a wide-range linear vector dose response in recombinant adeno-associated virus-mediated liver transduction. J. Virol. 76, 11343–11349.

42. Song, S., Lu, Y., Choi, Y.K., Han, Y., Tang, Q., Zhao, G., Berns, K.I., and Flotte, T.R. (2004). DNA-dependent PK inhibits adeno-associated virus DNA integration. Proc. Natl. Acad. Sci. U S A 101, 2112–2116.

43. Bi, L., Lawler, A.M., Antonarakis, S.E., High, K.A., Gezhardt, J.D., and Kazazian, H.H., Jr. (1995). Targeted disruption of the mouse factor VIII gene produces a model of haemophilia A. Nat. Genet. 10, 119–121.

44. Shinkai, Y., Rathbun, G., Lam, K.P., Oltz, E.M., Stewart, V., Mendelsohn, M., Charon, J., Datta, M., Young, F., Stall, A.M., et al. (1992). RAG-2-deficient mice lack mature lymphocytes owing to inability to initiate V(DJ) rearrangement. Cell 68, 855–867.

45. Falese, L., Sandza, K., Yates, B., Triffault, S., Gangar, S., Long, B., Tsuruda, L., Carter, B., Vettermann, C., Zoog, S.J., et al. (2017). Strategy to detect pre-existing immunity to AAV gene therapy. Gene Ther. 24, 768–778.

46. Willson, T.A., Oakley, K.M., and Nagley, P. (1985). A simple method for constructing directly repeated multimeric DNA segments. Gene Anal. Techn. 2, 77–82.

47. Sambrook, J., and Russell, D.W. (2006). Purification of nucleic acids by extraction with phenol:chloroform. CSH Protoc. 2006. https://doi.org/10.1101/pdb.prot4455.