Molecular Therapy
Methods & Clinical Development
Original Article

Generation of high-quality single-stranded DNA for full-length and truncated genome standards of recombinant adeno-associated viruses

Julia Manz,[1] Raphael Ruppert,[1] Markus Haindl,[1] Jürgen Hubbuch,[2] Tobias Graf,[3] and Johannes Pschirer[1]

[1]Gene Therapy Technical Development, Roche Diagnostics GmbH, 82377 Penzberg, Germany; [2]Institute of Process Engineering in Life Sciences, Section IV: Biomolecular Separation Engineering, Karlsruhe Institute of Technology, 76131 Karlsruhe, Germany; [3]Pharma Technical Development Analytics, Roche Diagnostics GmbH, 82377 Penzberg, Germany

The integrity of the single-stranded DNA (ssDNA) genome of recombinant adeno-associated viral vectors (rAAVs) for gene therapy applications is a critical quality attribute. To monitor this attribute, robust analytical methods, such as long-read sequencing or multiplex PCR, are needed. Due to the heterogeneity of the packaged rAAV payloads, the availability of appropriate ssDNA standards could clearly facilitate the development of these methods and may be used for system suitability testing. However, due to a genome length of up to approximately 4.7 kilobases and the secondary structure of the inverted terminal repeats, such standards are not trivial to produce. Here, we introduce a versatile method that starts with plasmid DNA containing the rAAV genome of interest. After its linearization, the sticky ends are filled in with biotinylated deoxynucleotide triphosphates, and the DNA is bound to streptavidin-coupled beads. With two further restriction enzymes, the plasmid backbone is removed, and finally, the ssDNA is obtained by alkaline strand separation. The quality and purity of the received ssDNA were verified using double-strand synthesis followed by size determination and long-read sequencing. With the presented method, it is possible to generate both full-length and truncated ssDNA standards within one day, contributing to improved quality control of rAAVs.

INTRODUCTION

Adeno-associated viruses (AAVs) are non-enveloped parvoviruses consisting of an icosahedral capsid and a single-stranded DNA (ssDNA) genome of about 4.7 kilobases (kb).[1] The ssDNA can either be in plus or in minus orientation, with both polarities packed in capsids at equal frequency. Palindromic sequences called inverted terminal repeats (ITRs) flank the coding region of the genome and form T-shaped hairpin structures[2] that are crucial for viral replication and packaging.[3]

Due to numerous advantages such as their reduced pathogenicity and the ability to also transduce post-mitotic and therefore non-dividing cells, recombinant AAVs (rAAVs) are promising vectors for gene therapy applications.[4] Over the past years, an increasing number of rAAV-based gene therapies have gained marketing authorization.[5,6] The transient-transfection manufacturing process of rAAVs typically consists of the following steps: production of the necessary plasmids that provide the genetic information for the rAAV generation, co-transfection of the plasmids in producer cells, cell cultivation, purification of the vectors, and formulation.[7,8] To confirm the drug's quality, critical quality attributes (CQAs) are established and consistently monitored by appropriate analytical methods. For instance, the integrity of the ssDNA payload from ITR to ITR, representing the active pharmaceutical ingredient (API) of rAAVs, needs to be assessed thoroughly with regard to the occurrence of mutations, truncations, and rearrangements. Such events could for instance adversely affect the drug's potency.[9,10] Suitable methods to monitor this CQA are long-read next-generation sequencing (NGS) and multiplex PCR.[11] Development of these analytical methods currently relies largely on available rAAV material, whose packaged DNA payload is heterogeneous due to its highly complex manufacturing process. Hence, a promising alternative to obtain high-quality ssDNA representative of the desired payload would be its generation decoupled from the actual rAAV production. Multiple approaches for producing ssDNA have been described,[12] and an overview of existing strategies and their limitations with regard to rAAV DNA is provided below.

ssDNA synthesis via phosphoramidite chemistry is limited to a length of approximately 200 bases due to the significantly lower incorporation efficiency compared to DNA replication and the risk of depurination as a side reaction.[13–16] Using *in vitro* transcription followed by reverse transcription and RNase H treatment (ivTRT), ssDNA up to ~1.5 kb can be obtained.[17,18] Still, this length is below the genome size of common rAAVs, and, moreover, important parts such as the ITRs and the promoter would be missing in the transcript and thus also in the final ssDNA. Rolling circle amplification (RCA) overcomes the length limitation by using a circular template and a primer in

Received 12 June 2025; accepted 29 September 2025;
https://doi.org/10.1016/j.omtm.2025.101603.

Correspondence: Johannes Pschirer, Roche Diagnostics GmbH, Nonnenwald 2, 82377 Penzberg, Bavaria, Germany.
E-mail: johannes.pschirer@roche.com

an isothermal polymerase reaction, yielding long concatemeric ssDNA.[19,20] However, since rAAV genomes are not circular, they cannot directly serve as templates for RCA. Using plasmid DNA containing the vector genome (pITR) would amplify not only the desired sequence but also the backbone. Besides, the product would not be multimeric ssDNA but concatemeric ones. Another strategy described in literature is based on bacterial production of ssDNA with the help of phagemids and helper bacteriophages.[21] Still this approach bears a similar disadvantage like RCA using pITRs as templates: the phagemid and thus the ssDNA would also contain sequences necessary for amplification. Moreover, the requirement of an ssDNA origin-of-replication incorporation complicates the technique.[12] Asymmetric PCR offers an efficient solution to generate ssDNA of predefined lengths with sizes greater than 15 kb.[22] However, secondary structures and high GC content, as both encountered in the ITRs, are problematic for an accurate PCR,[23] rendering this technique ill-suited for the preparation of rAAV ssDNA standards. Another interesting approach described in literature starts with plasmid DNA and makes use of Golden Gate intramolecular ligation to obtain plasmid DNA that only contains the region of interest. Next, a nickase and an exonuclease convert the circular double-stranded DNA (dsDNA) to ssDNA, and Cas9 is used for linearization. However, the linearization efficiency was only about 50%,[24] demonstrating that separation of DNA strands depending on enzymatic activity is not very efficient.

Despite a multitude of workflows reported to generate ssDNA, there is still a lack of suitable methods to obtain long ssDNA with complex structures as encountered in rAAVs. In this study, we describe the development of a new approach relying on plasmid DNA as a readily available starting material. It involves the incorporation of biotinylated deoxynucleotide triphosphates (dNTPs), which can then be used for efficient strand separation by exploiting the strong interaction between streptavidin and biotin. For the latter step, a similar approach was described by Wakimoto et al.[25] The suitability of the presented versatile protocol is demonstrated by confirming the correct length, identity, and purity of the generated rAAV ssDNA construct by capillary electrophoresis and long-read NGS.

RESULTS

Development of a method to produce rAAV ssDNA based on biotin-streptavidin interaction

In an earlier attempt to generate rAAV ssDNA, we used pITRs, linearized them near one end of the vector genome, and dephosphorylated them. A second cleavage near the other end of the rAAV sequence resulted in two dsDNA fragments, with one phosphorylated and one dephosphorylated 5′ end. Next, lambda exonuclease, which digests DNA strands with 5′ phosphate, was utilized to generate ssDNA.[26] However, analogous to the work from Strawn et al., the separation efficiency of DNA strands by means of enzymes was unsatisfactory.[24] Thus, other mechanisms for strand separation were evaluated, including the exploitation of the biotin streptavidin interaction, one of the strongest occurring in nature.[27] A similar approach was conducted in a study using PCR with one biotinylated primer followed by separation of the resulting PCR products with streptavidin-coupled magnetic beads.[25] Yet, PCR cannot be used for the generation of rAAV ssDNA. Thus, we assessed different alternative biotinylation approaches. In this regard, the envisioned strategy was to linearize a pITR by restriction digest near one end of the rAAV genome sequence and to use the cleavage site to introduce biotin modifications. After binding of the biotinylated DNA to streptavidin-coupled magnetic beads, two additional restriction digests are performed: one near the other end of the vector genome and one near the biotinylated backbone end, resulting in the detachment of the plasmid backbone from the beads. Eventually, the rAAV ssDNA of interest is released via alkaline strand separation by treating the streptavidin beads with sodium hydroxide (NaOH).

In a first attempt, this approach was realized by ligating biotinylated hairpin adapters to the sticky ends of the linearized plasmids. These adapters are self-complementary DNA sequences with a biotin modification and sticky ends fitting to those of the linearized pITRs. While we were able to obtain the desired ssDNA with this method, a significant amount of the rAAV dsDNA was also present (data not shown). Various adaptations of the protocol were tested, such as increasing the ligation time for the hairpin adapters from 1 to 8 h, elongating the wash steps to get rid of unwanted DNA fragments, and a heat denaturation step of the linearized plasmid to exclude a potential interfering effect of the ITRs' secondary structure on ligation efficiency. In addition, the NaOH step was substituted by incubating the dsDNA bound to streptavidin beads with helicase to prevent alkaline conditions during strand separation that could impair the biotin-streptavidin interaction leading to co-elution of dsDNA. Moreover, a second incubation of the obtained mix of ssDNA and dsDNA with streptavidin-coupled magnetic beads to get rid of potentially remaining biotinylated strands was tested. However, none of these adaptations resulted in any significant improvement (data not shown).

Consequently, an alternative approach for introducing the biotin tag into the linearized DNA was tested. For this, the 5′ overhangs were filled in with a mix of both biotinylated and non-biotinylated dNTPs using Klenow fragment (3′→5′ exo-), thereby circumventing the ligation step and enabling the incorporation of more than one biotin moiety on each end of the linearized plasmid. To remove the remainder of free biotinylated dNTPs from the solution that would compete with the biotinylated DNA for streptavidin binding, a cleanup with ProNex magnetic beads was performed. In the presence of polyethylene glycol and salts, DNA binds non-specifically to these beads, with longer nucleic acids having a higher affinity.[28] By contrast, dNTPs have very low affinity to these beads and can be easily washed away. The rest of the workflow remained as described above: binding of the biotinylated DNA to streptavidin-coupled beads, removal of the backbone by double-digest, and alkaline strand separation (see Figure 1 for the entire process).

Concentration and size evaluation of the produced rAAV ssDNA

The concentration of the ssDNA obtained by this adapted workflow was measured with the Qubit ssDNA assay kit, and the

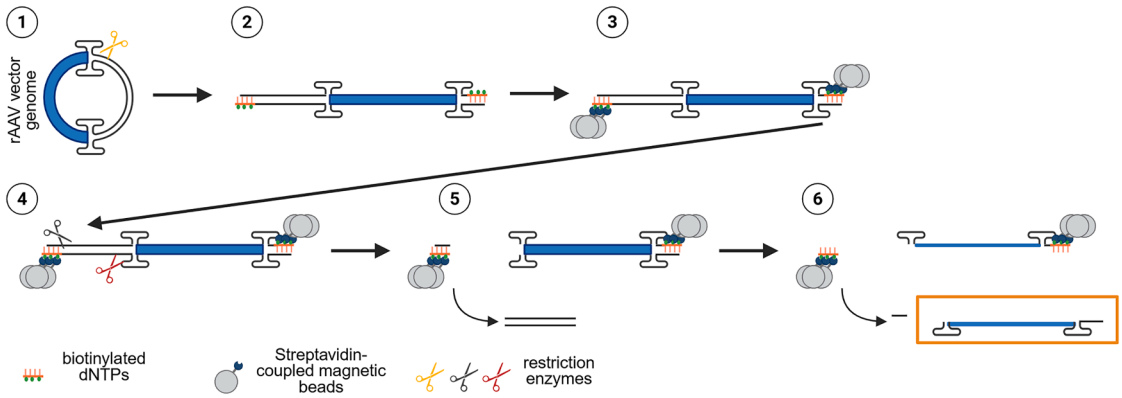

Figure 1. Schematic overview of the rAAV ssDNA generation process

A pITR is linearized near one end of the rAAV genome (1). After filling in the sticky ends with biotinylated dNTPs (2), the biotin-tagged molecules are bound to streptavidin-coupled beads (3). To detach the plasmid backbone from the beads, two further restriction digests are performed (4), and the plasmid backbone is washed away (5). Finally, after alkaline strand separation, the desired rAAV ssDNA (inside the orange box) can be obtained from the supernatant (6).

recovery relative to the input plasmid DNA was calculated (Figure 2A). In total, nine preparations were produced throughout this study with a median value of 34%. Besides evaluating the quality of the generated ssDNA (lane A3), the DNA size of the pITR1 after linearization (lane A1) and the backbone fragment in the supernatant after the double-digest of the bound biotinylated DNA (lane A2) were analyzed using a 5300 Fragment Analyzer (Figure 2B). The expected lengths are summarized in Table 1, and the cutting positions of the restriction enzymes are highlighted in Figure S1. Lane A1 shows two distinct bands, one intense band at ∼6,000 base pairs (bp) corresponding to the pITR1 length of 6,634 bp and one band with lower intensity at ∼75 bp, which can be attributed to the activator oligonucleotide that is needed by the linearization restriction enzyme. In lane A2, the major band is located at ∼2,000 bp and thus at the expected backbone length of 2,069 bp. However, also a weak signal at the length of the double-stranded rAAV (dsrAAV) genome can be observed. For the obtained ssDNA, one distinct band is visible in lane A3. Since the method and the ladder are designed for dsDNA, an estimation concerning the length of the observed band is not possible.

The method described here can easily be adapted to other plasmids with no changes required besides the selection of the restriction enzymes. To demonstrate this, we replicated the protocol with a second pITR (pITR2), resulting in a similar recovery value of 31% and good quality and purity (Figure S2, lanes B1–B3), which emphasizes the versatility of our method.

Verification of the generated ssDNA by double-strand synthesis and NGS

To verify the genome length of the obtained rAAV ssDNA, the elution fraction from the streptavidin beads was subjected to double-strand synthesis using the ITR with the free 3′-OH group as a primer. Since the 5′ ITR should be displaced for this instead of being digested, a polymerase with strand displacement function but lacking 5′→3′ exonuclease activity was employed for this task. As part of a previous study, various such polymerases were tested: Bst 3.0, Bsm, Bsu, Deep Vent, EquiPhi29, and phi29. Since Bst 3.0 performed best in converting rAAV ssDNA to dsDNA without producing artificial truncations, this polymerase was chosen (unpublished data). The product length of the synthesis was analyzed on the 5300 Fragment Analyzer (Figure 2B, lane A4). Besides the dsDNA, displayed as a marked band at the expected length of 4,565 bp, a faint smear and a weak band at about the length of the linearized plasmid (6,634 bp) could be detected. The verification experiment with pITR2 resulted in one band at the expected length (Figure S2, lane B4).

To confirm the results from the Fragment Analyzer regarding the genome length and, furthermore, assess the sequence and purity, single-molecule, real-time (SMRT) sequencing was performed. For sequencing, dsDNA is required, which is obtained by thermal annealing of AAV plus and minus genomes according to the manufacturer's protocol. However, our method results in ssDNA of only one polarity and is hence unsuitable for this approach. Therefore, we used the obtained dsDNA from the double-strand synthesis as input material for the library preparation. Analogous to the protocol from the manufacturer, hairpin adapters were ligated to the DNA to produce circular SMRTbell templates for the sequencing run. Since the 3′ ITR has been used as primer during dsDNA synthesis, the employed dsDNA only featured one open end, and thus, just one adapter could be ligated. As an SMRT read is defined as the sequence between two adapters, the special structure of our library molecules was expected to result in read lengths twice as long as our rAAV genome, i.e., approximately 9 kb as shown in Figure 3A. The output file of the sequencer contained 110.103 reads with 98.1% of mapped bases fitting to the rAAV and 0.83% to the pITR1 backbone reference sequence (Table S1).

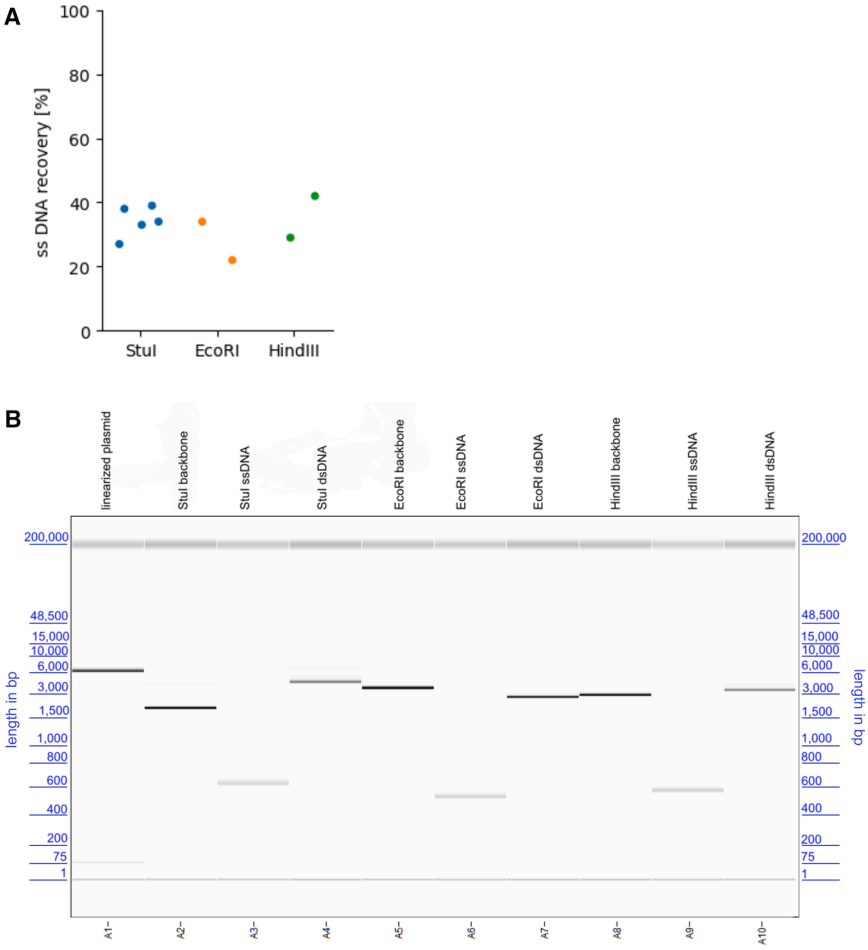

Figure 2. ssDNA recovery and gel image of the produced rAAV ssDNA, dsDNA, and intermediate products using pITR1 as starting material

(A) The concentration of the obtained rAAV ssDNA was measured using the Qubit ssDNA Assay Kit, and the recovery was calculated. The results for the full-length (StuI) and truncated (EcoRI and HindIII) genomes are depicted in different colors. (B) The gel image shows the sizing of the generated rAAV ssDNA, dsDNA, and intermediate products on the 5300 Fragment Analyzer using the Agilent DNF-464 HS Large Fragment 50 kb Kit. The upper and lower bands correspond to the upper and lower markers of the kit. All dsDNA molecules migrated to the expected positions on the agarose gel. **Lane A1**: dsDNA plasmid after linearization. **Lane A2**: dsDNA backbone after double-digest with StuI and PacI. **Lane A3**: ssDNA obtained by the digest with StuI and PacI and alkaline strand separation. **Lane A4**: dsDNA resulting from double-strand synthesis using the ssDNA from lane A3. **Lane A5**: dsDNA backbone after double-digest with EcoRI and PacI. **Lane A6**: ssDNA obtained by the digest with EcoRI and PacI and alkaline strand separation. **Lane A7**: dsDNA resulting from double-strand synthesis using the ssDNA from lane A6. **Lane A8**: dsDNA backbone after double-digest with HindIII and PacI. **Lane A9**: ssDNA obtained by the digest with HindIII and PacI and alkaline strand separation. **Lane A10**: dsDNA resulting from double-strand synthesis using the ssDNA from lane A9.

In addition, an Integrative Genomics Viewer (IGV) display was compiled to visualize the mapping of some exemplary reads to the reference pITR1 sequence (Figure 3B). In the rAAV genome region, ranging from position 2083 up to the end of the reference sequence, an almost even coverage is visible with no mutations in the consensus sequence (see the gray coverage plot). The colorful parts on the left of the shown alignments are again due to the special structure of our library with only one adapter that results in reads twice as long as the rAAV genome. Consequently, only half of the read can be primarily mapped against the reference sequence, and the other half is soft-clipped.

Production of truncated rAAV genomes

Next, we assessed whether our method is also suitable for the generation of truncated rAAV ssDNA. For the production of these ssDNA constructs, the enzyme that cuts near the second ITR was replaced by a restriction endonuclease that cuts specifically inside the rAAV genome, while all other steps of the procedure remained unchanged. Again, the backbone fragment in the supernatant after the double-digest of the bound biotinylated DNA and the obtained ssDNA was analyzed using the 5300 Fragment Analyzer. Figure 2B lanes A5 and A6 show the obtained bands for the backbone and the truncated ssDNA rAAV genome, respectively. The band in lane A5 matches with the expected length of 3,920 bp (see Table 1). Lane A6 contains one band for the ssDNA. To verify the length of this product, the second strand was synthesized using a primer which binds directly at the end of the restriction site. The resulting product has the expected length of about 2,714 bp (see Table 1; Figure 2B, lane A7). The same steps were also performed using another enzyme in the double-digest, theoretically resulting in a 3,191 bp backbone fragment and truncated rAAV ssDNA with a length of 3,343 bases (see Table 1). Again, the obtained bands in Figure 2B lanes A8–A10 fit these expectations.

DISCUSSION

The generation of ssDNA serving as potential rAAV genome standard is challenging due to its length of up to 4.7 kb and the propensity of the GC-rich ITRs (~70%) to form strong secondary structures.[29] To overcome this current limitation, a new approach is described in this study, involving the linearization of an rAAV genome containing plasmid, followed by introducing a biotin tag and binding to magnetic streptavidin beads. After two further restriction digests, the desired rAAV ssDNA can be released by alkaline strand separation.

Table 1. Expected lengths resulting from the different combinations of restriction enzymes for pITR1

Restriction enzyme combination	Expected "backbone" length (bp)	Expected ssDNA length (bases)
StuI, PacI	2,069	4,565
EcoRI, PacI	3,920	2,714
HindIII, PacI	3,191	3,443

bp, base pairs; dsDNA, double-stranded DNA; pITR, plasmid DNA containing the vector genome; ssDNA, single-stranded DNA.

For the functionalization with biotin, we first ligated biotinylated hairpin adapters to the linearized plasmid. However, this approach resulted in not only the desired ssDNA but also a significant amount of dsDNA, even after multiple attempts to optimize the method. The problems arising from the use of biotinylated hairpin adapters may be attributed to the inherent properties of the ITRs. To obtain rAAV ssDNA, the plasmids are linearized in a way that one of the ITRs becomes exposed. These sequences form strong secondary structures,[3] which in turn can compromise the performance of certain enzymes. For instance, T4 RNA ligase efficiency is impaired by secondary structures.[30,31] In general, secondary structures might interfere with ligation efficiency by impeding the hybridization of the sticky ends and by sterical hindrance of the enzyme. To resolve this constraint, the ligation of biotinylated adapters was replaced by filling in the sticky ends with biotinylated dNTPs catalyzed by a DNA polymerase. This approach bears also the advantage that more than one biotin tag can be incorporated, and thus the individual biotin-streptavidin binding affinity accumulates, resulting in high avidity.[32,33] This is particularly relevant for the NaOH treatment step to dissociate and hence elute the non-biotinylated from the biotinylated strand, as strong alkaline conditions can also destabilize the biotin-streptavidin interaction[34] and thus result in a mixture of rAAV ssDNA and dsDNA. Indeed, with the fill-in approach, only one band was visible for the ssDNA product on the fragment analyzer (Figure 2B, lane A3).

This strategy works well and produces ssDNA of high quality with a median recovery rate of 34% with no significant difference between full-length and truncated genomes. Nevertheless, increasing this value could be the subject of future work. One reason for the relatively low yield could be that not all molecules were biotinylated. If they were not biotinylated at all, they would not bind to streptavidin beads and hence be lost right at the beginning of the protocol. Alternatively, if not all linearized plasmids were filled in at the ITR site, these specific molecules would only bind to the beads with the backbone side. Consequently, not only the plasmid backbone but also the rAAV genome would be released after the double-digest. This hypothesis was backed up by the observation that a band at the length of the dsrAAV genome was visible in the supernatant sample (Figure 2B lane A2). While this might lower the overall ssDNA yield, the purity of the obtained ssDNA remained unaffected (only one band in Figure 2B, lane A3). Thus, it seems that the sticky ends that were filled in contained a sufficient number of biotinylated dNTPs to withstand the NaOH treatment. Potential mitigation strategies to reduce the loss of rAAV genomes could involve further optimization of the filling-in step, by fine-tuning either the incorporation time or the employed ratio of biotinylated and non-biotinylated dNTPs.[35] This applies particularly to the linearization site next to the ITRs, whose secondary structure might have an impact on incorporation efficiency, especially when substituting dNTPs by their modified derivatives, possibly through steric hindrance.

Regardless of these suggested measures to increase the ssDNA yield, it is important to note that the primary focus of this work was to develop a tool capable of generating rAAV ssDNA with high purity that can be used for analytical purposes. To achieve this, several washing steps were included in the protocol that might have a detrimental effect on the overall recovery. In light of these requirements, a median recovery of 34% seems appropriate and represents a significant improvement over our previous attempts (unpublished data).

Since there is currently no readily available analytical method to reliably determine the length of ssDNA, double-strand synthesis and SMRT sequencing were applied to confirm the length, identity, and purity of the ssDNA product. With 98.1% of mapped bases fitting to the rAAV genome (Table S1), high purity could be proven. Although the sequence identity could be confirmed (Figure 3B) and we mainly detected dsDNA of the expected length, a relatively small amount of additional constructs with varying genome lengths were apparent in both the gel image (Figure 2B, lane A4) and the read length plot (Figure 3A). This could be due to the use of Bst 3.0 DNA polymerase for double-strand synthesis. Some of its properties such as a high-strand displacement activity and a lacking 5′->3′ exonuclease activity are advantageous for employing this polymerase for the double-strand synthesis using the ITR as a primer. However, its usage was reported to also be associated with a number of side products. In this regard, one application of Bst 3.0 is loop-mediated isothermal amplification (LAMP), leading to long concatemers of the target sequence connected by single-stranded loops.[36,37] In addition, non-specific products have also been observed and discussed for Bst polymerase.[38] Due to the structural similarity of the loops in the LAMP protocol and the T-shaped structure of the ITRs in the rAAV genome, it is conceivable that such unspecific products could also occur during our dsDNA synthesis. Consequently, we presume that the smear is an artifact of the double-strand synthesis and not due to variations in the length of the original ssDNA product. This hypothesis is corroborated by the observation that only one band was visible for the ssDNA product (Figure 2B lane A3) and that the other double-strand DNA syntheses using template

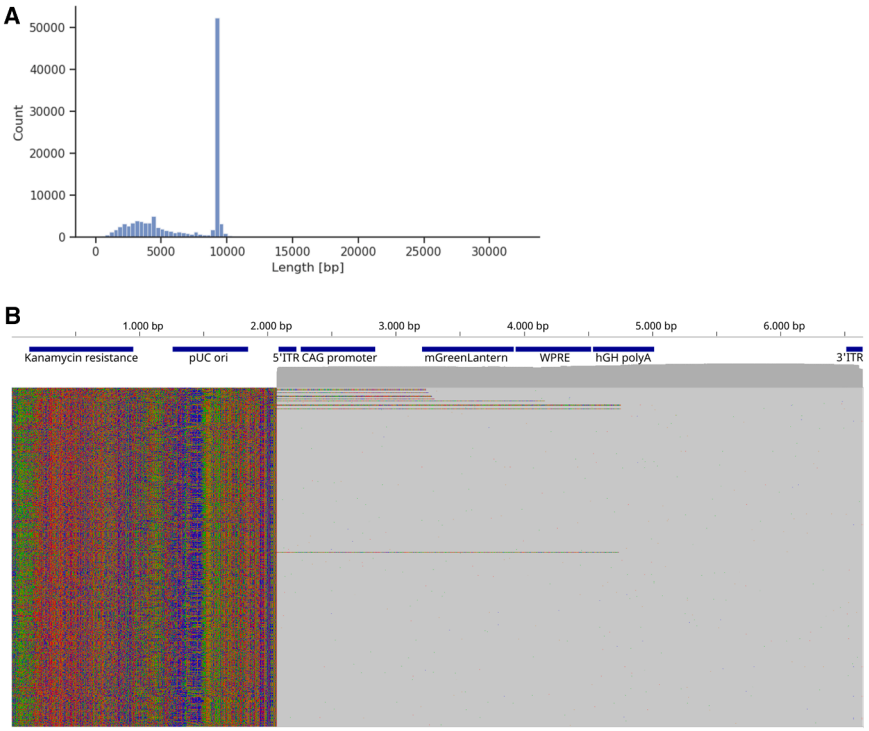

Figure 3. Results from sequencing the produced full-length rAAV ssDNA genome

The generated ssDNA genome from pITR1 was converted to dsDNA using the 3′ ITR as a primer. The obtained product was used as input for SMRTbell library preparation and sequenced on the Sequel IIe. (A) Distribution of the lengths of the raw reads. Since the ITR was used as a primer, the read lengths are twice as long as the length of the original ssDNA. (B) IGV display of some exemplary reads of the generated rAAV genomes aligned to the reference sequence. Matches are depicted in gray and mismatches in different colors. Due to the usage of the 3′ ITR as a primer, the reads are twice as long as the rAAV genome reference, resulting in soft-clipping on the left-hand side. The overall coverage of the reference genome is shown above the reads.

specific primers instead of the T-shaped ITRs resulted in one clear dsDNA band (Figure 2B, lanes A7 and A10).

For rAAVs, there are not only full and empty capsids but also those containing only a partial ssDNA payload. These truncated genomes could have a negative impact on the efficacy and furthermore raise safety concerns as they might lead to the expression of aberrant or truncated proteins.[11,39] In addition, truncated genomes can lead to erroneous estimations of vector genome titers by quantitative PCR (qPCR) or droplet digital PCR (ddPCR) and thus to incorrect dosing.[40] Using our proposed approach, it is possible to generate not only full rAAV genomes but also partial rAAV ssDNA. This is achieved by replacing the second restriction endonuclease with an enzyme that specifically cuts inside the rAAV genome instead of next to the ITR region, as demonstrated in two examples (see Figure 2B, lanes A6, A7, A9, and A10). Such truncated ssDNA preparations could, for example, be used for evaluating their effect on the results of different analytical methods or for method development including length bias determination, validation, and calibration. In addition, it is conceivable that readily available truncated ssDNA genomes can enhance the overall understanding of rAAVs and their efficacy, for example, by assessing the impact of these alterations on potency.

Beyond these applications, the produced ssDNA could potentially also be used for *in vitro* packaging in pre-assembled capsids. It is hypothesized that these systems might allow to package ssDNA above the current limitation of about 5 kb. Furthermore, they could circumvent the risk of contamination with adventitious viruses or other substances (e.g., proteins, metabolites, or reagents used for lysis) that might occur in cell culture.[41–43]

Taken together, our method is capable of reproducibly generating both full-length and truncated rAAV ssDNA genomes with high purity. In this regard, the produced ssDNA constructs can serve as analytical standards for method development and benchmarking different analytical techniques.

MATERIALS AND METHODS

All enzymes were purchased from NEB (Ipswich, Massachusetts, US).

Plasmid information

The used pITR1 was produced in-house. It has a length of 6,634 bp and contains an rAAV genome with 4,565 bp in length coding for mGreenLantern as a transgene. A plasmid map containing the restriction sites used for the ssDNA production can be found in Figure S1. The pITR2 used to prove the versatility of our method was also produced in-house. It has a length of 5,249 bp, including an rAAV genome of 3,083 bp.

Linearization and biotinylation of the plasmid

Ten units (u) of PaqCI/µg DNA in 1× rCutSmart was used to cut the pITRs near one ITR in the plasmid backbone. PaqCI activator was added in a 1:1 ratio to the enzyme. The digest was performed for 1 h at 37°C, followed by heat inactivation for 20 min at 65°C. To fill in the sticky ends, 30 u Klenow fragment (3′ →5′ exo-)/µg plasmid DNA in 1× NEBuffer 2.0 was added. The mixture was supplemented with dTTP, dGTP (NEB, Ipswich, Massachusetts, US), biotinylated dATP, and biotinylated dCTP (Thermo Fisher Scientific, Waltham, Massachusetts, US) to a final concentration of 0.05 mM each. The reaction took place for 1 h at 37°C. To remove free biotinylated dNTPs, 2× volume/volume room

temperature adjusted ProNex Beads (Promega, Madison, Wisconsin, US) were added. After 15-min incubation, the tube was placed in a magnetic separation rack, the supernatant was removed, and the beads were washed twice with 80% ethanol. Next, the tube was removed from the magnetic separation rack, and the beads were resuspended in 100 µL nuclease-free water. After 12 min incubation, the supernatant containing the purified DNA was saved in a new tube and immediately used for the next step or stored at 4°C until further use.

Binding of the biotinylated DNA to streptavidin-coupled beads

Fifty microliters of Dynabeads MyOne Streptavidin C1 (Thermo Fisher Scientific, Waltham, Massachusetts, US)/µg plasmid DNA was transferred to a new tube and washed with 0.7 mL 1× Binding & Washing Buffer (5 mM Tris-HCl [pH 8], 0.5 mM EDTA, 1 M NaCl). The tube was placed in a magnetic separation rack, the supernatant was removed, and the beads were resuspended in 50 µL 1× Binding & Washing Buffer. This washing step was repeated twice for a total of three washes. At the end, the beads were resuspended in 100 µL 2× Binding & Washing Buffer.

One hundred microliters of biotinylated DNA was added to 100 µL of washed streptavidin beads and incubated for 30 min on a rotator. Afterward, the tube was placed in a magnetic separation rack, and the supernatant was removed. After removing the tube from the magnetic separation rack, the beads were resuspended in 250 µL 1× Binding & Washing Buffer and incubated for 10 min on a rotator. The tube was placed back in the magnetic separation rack, and the supernatant was removed. The washing step was repeated twice for a total of three washes.

Generation of the rAAV ssDNA

The beads were resuspended in 49 µL nuclease-free water/µg plasmid DNA. After addition of 28 u of StuI, PacI, and rCutSmart (to 1×), the mixture was incubated at 37°C and 400 rpm for 1 h. For the production of truncated rAAV ssDNA, StuI was replaced by either 28 u of EcoRI HF or HindIII HF. For the verification of the method with pITR2, AccI was used instead of PacI. After the digestion, the tube was placed in a magnetic separation rack, and the supernatant was removed. The beads were resuspended in 100 µL 1× SSC Buffer (Roche Diagnostics GmbH, Mannheim, Germany)/µg plasmid DNA and incubated for 10 min on a rotator. Next, the tube was placed back in the magnetic separation rack, the supernatant was removed, and the beads were resuspended in 40 µL 0.15 M NaOH/µg plasmid DNA. After 10-min incubation on a rotator, the tube was placed in the magnetic separation rack, and the supernatant containing the non-biotinylated ssDNA strand (in this case, the minus strand) was transferred in a new tube. The mixture was neutralized by adding 4.4 µL 10× TE buffer (pH 8) and 4.6 µL 1.25 M acetic acid.

Double-strand synthesis

Double-strand synthesis of the full-length rAAV ssDNA using the 3′ ITR as a primer was carried out based on the DNA extension protocol of Zhang et al.[44] Deviating from the described method, Bst 3.0 was used instead of Bst 2.0. Moreover, no BSA was added, and the mixture was incubated for 1 h at 65°C followed by 5 min at 80°C. Double-stranding of the truncated genomes followed the same protocol with the following exceptions: primers complementary to the 5′ end of the ssDNAs were ordered from IDT (Coralville, Iowa, US) and added to the ssDNA in 10-fold excess. Just before the extension at 65°C, an annealing step for 1 min at the appropriate temperature for the respective primer was added.

Determination of ssDNA recovery

The concentration of the obtained rAAV ssDNA was measured with the Qubit ssDNA Assay Kit (Thermo Fisher Scientific, Waltham, Massachusetts, US) according to the manufacturer's protocol. The ssDNA recovery was calculated with Equation 1:

$$Recovery\ [\%] = \frac{Obtained\ ssDNA\ amount\ [ng]}{Used\ dsDNA\ amount\ [ng] \times 0.5 \times Portion\ rAAV\ genome} \times 100 \quad \text{(Equation 1)}$$

Length determination using the 5300 Fragment Analyzer

For the length determination on the 5300 Fragment Analyzer (Agilent Technologies, Santa Clara, US), the Agilent DNF-464 HS Large Fragment 50 kb Kit was used. All steps were carried out according to the manufacturer's protocol.

Next-generation sequencing

The double-stranded rAAV genomes (see double-strand synthesis) were used as input for SMRTbell library preparation. Starting from step 5, the instructions of PacBio's protocol "Preparing multiplexed AAV SMRTbell libraries using SMRTbell prep kit 3.0" were followed. However, to prevent loss of short DNA fragments, the bead ratio was increased from 1.3× volume/volume to 2× for each cleanup. Moreover, the binding and elution times for the cleanups were prolonged to 15 min. The SMRTbell library was sequenced on the Sequel IIe instrument (Pacific Biosciences, Menlo Park, US) with the standard settings for the application "adeno-associated virus."

Bioinformatics analysis

The reference sequence for the pITR1 was internally available as FASTA file. Sequences of possible contaminations were retrieved from the NCBI database (human: GenBank: GCA_000001405.15; *Escherichia coli*: GenBank: GCA_904425475.1). Since only reads with a Phred score of at least 20 were contained in the raw sequencing data,

no subsequent filtering was necessary. For the alignment, the minimap2 SMRT wrapper for PacBio data (pbmm2, v.1.16.0) was used with the option—unmapped in order to have all reads in the resulting alignment file. Using pysam (v.0.19.1), the aligned bam file was read in a pandas (v.1.4.1) DataFrame. The length plot of the primary reads was generated with seaborn (v.0.11.2), a Python visualization library based on matplotlib. IGV (v.2.16.0) was used for the visualization of the reads.

DATA AND CODE AVAILABILITY

PacBio sequencing data have been deposited to the Sequence Read Archive (SRA) under the accession number PRJNA1275844.

ACKNOWLEDGMENTS

We are grateful for valuable discussions and constant support from various colleagues of the laboratories at Roche Diagnostics GmbH in Penzberg, Germany. Special thanks goes to the Nucleic Acid Science team, which provided us with the necessary plasmid material. Graphical abstract and Figure 1 were created with BioRender.com.

AUTHOR CONTRIBUTIONS

J.M. and J.P., conceptualization; J.M., data curation, formal analysis, investigation, methodology, validation, and visualization; J.P. and R.R., project administration; M.H. and R.R., resources; J.M. and J.P., software; J.P. and J.H., supervision; J.M, J.P., and T.G., writing – original draft; all authors, writing – review & editing.

DECLARATION OF INTERESTS

J.M. and J.P. declare the following conflict of interest: a patent regarding the described method to produce ssDNA was filed on 01/20/2025.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2025.101603.

REFERENCES

1. Balakrishnan, B., and Jayandharan, G.R. (2014). Basic Biology of Adeno-Associated Virus (AAV) Vectors Used in Gene Therapy. Curr. Gene Ther. *14*, 86–100.
2. Srivastava, A., Lusby, E.W., and Berns, K.I. (1983). Nucleotide sequence and organization of the adeno-associated virus 2 genome. J. Virol. *45*, 555–564.
3. Wilmott, P., Lisowski, L., Alexander, I.E., and Logan, G.J. (2019). A User's Guide to the Inverted Terminal Repeats of Adeno-Associated Virus. Hum. Gene Ther. Methods *30*, 206–213. https://doi.org/10.1089/hgtb.2019.276.
4. Pattali, R., Mou, Y., and Li, X.-J. (2019). AAV9 Vector: a Novel modality in gene therapy for spinal muscular atrophy. Gene Ther. *26*, 287–295. https://doi.org/10.1038/s41434-019-0085-4.
5. Qie, B., Tuo, J., Chen, F., Ding, H., and Lyu, L. (2025). Gene therapy for genetic diseases: challenges and future directions. MedComm (2020) *6*, e70091. https://doi.org/10.1002/mco2.70091.
6. Zhao, Q., Peng, H., Ma, Y., Yuan, H., and Jiang, H. (2025). In vivo applications and toxicities of AAV-based gene therapies in rare diseases. Orphanet J. Rare Dis. *20*, 368. https://doi.org/10.1186/s13023-025-03893-z.
7. Samulski, R.J., and Muzyczka, N. (2014). AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. Annu. Rev. Virol. *1*, 427–451. https://doi.org/10.1146/annurev-virology-031413-085355.
8. Wang, J.-H., Gessler, D.J., Zhan, W., Gallagher, T.L., and Gao, G. (2024). Adeno-associated virus as a delivery vector for gene therapy of human diseases. Signal Transduct. Target. Ther. *9*, 78. https://doi.org/10.1038/s41392-024-01780-w.
9. Maynard, L.H., Smith, O., Tilmans, N.P., Tham, E., Hosseinzadeh, S., Tan, W., Leenay, R., May, A.P., and Paulk, N.K. (2019). Fast-Seq: A Simple Method for Rapid and Inexpensive Validation of Packaged Single-Stranded Adeno-Associated Viral Genomes in Academic Settings. Hum. Gene Ther. Methods *30*, 195–205. https://doi.org/10.1089/hgtb.2019.110.
10. Chen, Y., Hu, S., Lee, W., Walsh, N., Iozza, K., Huang, N., Preston, G., Drouin, L.M., Jia, N., Deng, J., et al. (2024). A Comprehensive Study of the Effects by Sequence Truncation within Inverted Terminal Repeats (ITRs) on the Productivity, Genome Packaging, and Potency of AAV Vectors. Microorganisms *12*, 310. https://doi.org/10.3390/microorganisms12020310.
11. Kontogiannis, T., Braybrook, J., McElroy, C., Foy, C., Whale, A.S., Quaglia, M., and Smales, C.M. (2024). Characterization of AAV vectors: A review of analytical techniques and critical quality attributes. Mol. Ther. Methods Clin. Dev. *32*, 101309. https://doi.org/10.1016/j.omtm.2024.101309.
12. Hao, M., Qiao, J., and Qi, H. (2020). Current and Emerging Methods for the Synthesis of Single-Stranded DNA. Genes *11*, 116. https://doi.org/10.3390/genes11020116.
13. Kosuri, S., and Church, G.M. (2014). Large-scale *de novo* DNA synthesis: technologies and applications. Nat. Methods *11*, 499–507. https://doi.org/10.1038/nmeth.2918.
14. Efcavitch, J.W., and Heiner, C. (1985). Depurination As a Yield Decreasing Mechanism in Oligodeoxynucleotide Synthesis. Nucleosides Nucleotides *4*, 267. https://doi.org/10.1080/07328318508077883.
15. LeProust, E.M., Peck, B.J., Spirin, K., McCuen, H.B., Moore, B., Namsaraev, E., and Caruthers, M.H. (2010). Synthesis of high-quality libraries of long (150mer) oligonucleotides by a novel depurination controlled process. Nucleic Acids Res. *38*, 2522–2540. https://doi.org/10.1093/nar/gkq163.
16. Li, S., Tan, W., Jia, X., Miao, Q., Liu, Y., and Yang, D. (2024). Recent advances in the synthesis of single-stranded DNA *in vitro*. Biotechnol. J. *19*, 2400026. https://doi.org/10.1002/biot.202400026.
17. Miura, H., Gurumurthy, C.B., Sato, T., Sato, M., and Ohtsuka, M. (2015). CRISPR/Cas9-based generation of knockdown mice by intronic insertion of artificial microRNA using longer single-stranded DNA. Sci. Rep. *5*, 12799. https://doi.org/10.1038/srep12799.
18. Miura, H., Quadros, R.M., Gurumurthy, C.B., and Ohtsuka, M. (2018). Easi-CRISPR protocol for creating knock-in and conditional knockout mouse models using long ssDNA donors. Nat. Protoc. *13*, 195–215. https://doi.org/10.1038/nprot.2017.153.
19. Johne, R., Müller, H., Rector, A., van Ranst, M., and Stevens, H. (2009). Rolling-circle amplification of viral DNA genomes using phi29 polymerase. Trends Microbiol. *17*, 205–211. https://doi.org/10.1016/j.tim.2009.02.004.
20. Ali, M.M., Li, F., Zhang, Z., Zhang, K., Kang, D.-K., Ankrum, J.A., Le, X.C., and Zhao, W. (2014). Rolling circle amplification: a versatile tool for chemical biology, materials science and medicine. Chem. Soc. Rev. *43*, 3324–3341. https://doi.org/10.1039/c3cs60439j.
21. Zhou, B., Dong, Q., Ma, R., Chen, Y., Yang, J., Sun, L.-Z., and Huang, C. (2009). Rapid isolation of highly pure ssDNA from phagemids. Anal. Biochem. *389*, 177–179. https://doi.org/10.1016/j.ab.2009.03.044.
22. Veneziano, R., Shepherd, T.R., Ratanalert, S., Bellou, L., Tao, C., and Bathe, M. (2018). In vitro synthesis of gene-length single-stranded DNA. Sci. Rep. *8*, 6548. https://doi.org/10.1038/s41598-018-24677-5.
23. Green, M.R., and Sambrook, J. (2019). Polymerase Chain Reaction (PCR) Amplification of GC-Rich Templates. Cold Spring Harb. Protoc. *2019*, pdb.prot095141. https://doi.org/10.1101/pdb.prot095141.
24. Strawn, I.K., Steiner, P.J., Newton, M.S., and Whitehead, T.A. (2022). A method for generating user-defined circular single-stranded DNA from plasmid DNA using Golden Gate intramolecular ligation. Preprint at bioRxiv. https://doi.org/10.1101/2022.11.21.517425.
25. Wakimoto, Y., Jiang, J., and Wakimoto, H. (2014). Isolation of Single-Stranded DNA. Curr. Protoc. Mol. Biol. *107*, 2.15.1–2.15.9. https://doi.org/10.1002/0471142727.mb0215s107.
26. Manz, J., Pschirer, J., Ruppert, R., and Kappaun, F. Method for the generation of single stranded DNA, application No. PCT/EP2024/073796, filing date 08/26/2024.
27. Liu, F., Zhang, J.Z.H., and Mei, Y. (2016). The origin of the cooperativity in the streptavidin-biotin system: A computational investigation through molecular dynamics simulations. Sci. Rep. *6*, 27190. https://doi.org/10.1038/srep27190.

28. Hawkins, T. (1998). DNA purification and isolation using magnetic particles. US patent US5705628A, filed September 20, 1994 and granted January 6, 1998.

29. Earley, L.F., Conatser, L.M., Lue, V.M., Dobbins, A.L., Li, C., Hirsch, M.L., and Samulski, R.J. (2020). Adeno-Associated Virus Serotype-Specific Inverted Terminal Repeat Sequence Role in Vector Transgene Expression. Hum. Gene Ther. *31*, 151–162. https://doi.org/10.1089/hum.2019.274.

30. Zhuang, F., Fuchs, R.T., Sun, Z., Zheng, Y., and Robb, G.B. (2012). Structural bias in T4 RNA ligase-mediated 3′-adapter ligation. Nucleic Acids Res. *40*, e54. https://doi.org/10.1093/nar/gkr1263.

31. Fuchs, R.T., Sun, Z., Zhuang, F., and Robb, G.B. (2015). Bias in Ligation-Based Small RNA Sequencing Library Construction Is Determined by Adaptor and RNA Structure. PLoS One *10*, e0126049. https://doi.org/10.1371/journal.pone.0126049.

32. Oostindie, S.C., Lazar, G.A., Schuurman, J., and Parren, P.W.H.I. (2022). Avidity in antibody effector functions and biotherapeutic drug design. Nat. Rev. Drug Discov. *21*, 715–735. https://doi.org/10.1038/s41573-022-00501-8.

33. Karush, F. (1976). Multivalent binding and functional affinity. Contemp. Top. Mol. Immunol. *5*, 217–228. https://doi.org/10.1007/978-1-4684-8142-6_8.

34. Kilili, G.K., Tilton, L., and Karbiwnyk, C.M. (2016). [Letter to the Editor] NaOH Concentration and Streptavidin Bead Type are Key Factors for Optimal DNA Aptamer Strand Separation and Isolation. Biotechniques *61*, 114–116. https://doi.org/10.2144/000114449.

35. Paul, N., and Yee, J. (2010). PCR Incorporation of Modified dNTPs: The Substrate Properties of Biotinylated dNTPs. Biotechniques *48*, 333–334. https://doi.org/10.2144/000113405.

36. Notomi, T., Okayama, H., Masubuchi, H., Yonekawa, T., Watanabe, K., Amino, N., and Hase, T. (2000). Loop-mediated isothermal amplification of DNA. Nucleic Acids Res. *28*, e63. https://doi.org/10.1093/nar/28.12.e63.

37. Kashir, J., and Yaqinuddin, A. (2020). Loop mediated isothermal amplification (LAMP) assays as a rapid diagnostic for COVID-19. Med. Hypotheses *141*, 109786. https://doi.org/10.1016/j.mehy.2020.109786.

38. Rolando, J.C., Jue, E., Barlow, J.T., and Ismagilov, R.F. (2020). Real-time kinetics and high-resolution melt curves in single-molecule digital LAMP to differentiate and study specific and non-specific amplification. Nucleic Acids Res. *48*, e42. https://doi.org/10.1093/nar/gkaa099.

39. Troxell, B., Jaslow, S.L., Tsai, I.-W., Sullivan, C., Draper, B.E., Jarrold, M.F., Lindsey, K., and Blue, L. (2023). Partial genome content within rAAVs impacts performance in a cell assay-dependent manner. Mol. Ther. Methods Clin. Dev. *30*, 288–302. https://doi.org/10.1016/j.omtm.2023.07.007.

40. Prantner, A., and Maar, D. (2023). Genome concentration, characterization, and integrity analysis of recombinant adeno-associated viral vectors using droplet digital PCR. PLoS One *18*, e0280242. https://doi.org/10.1371/journal.pone.0280242.

41. Zhou, X., and Muzyczka, N. (1998). In Vitro Packaging of Adeno-Associated Virus DNA. J. Virol. *72*, 3241–3247.

42. Le, D.T., Radukic, M.T., and Müller, K.M. (2019). Adeno-associated virus capsid protein expression in Escherichia coli and chemically defined capsid assembly. Sci. Rep. *9*, 18631. https://doi.org/10.1038/s41598-019-54928-y.

43. King, J.A., Dubielzig, R., Grimm, D., and Kleinschmidt, J.A. (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. EMBO J. *20*, 3282–3291. https://doi.org/10.1093/emboj/20.12.3282.

44. Zhang, J., Yu, X., Chrzanowski, M., Tian, J., Pouchnik, D., Guo, P., Herzog, R.W., and Xiao, W. (2024). Thorough molecular configuration analysis of noncanonical AAV genomes in AAV vector preparations. Mol. Ther. Methods Clin. Dev. *32*, 101215. https://doi.org/10.1016/j.omtm.2024.101215.