

Andrei Hutanu^{1,2} 
 David Boelsterli³
 Claudio Schmidli³
 Cristina Montealegre³
 Mike H. N. Dang Thai¹
 Balazs Bobaly³
 Marius Koch³
 Maria A. Schwarz^{2,3}

¹Analytical Development and Quality Control, Pharma Technical Development Biologics Europe, University of Basel, Basel 4056, Switzerland

²Analytical Development and Quality Control, Pharma Technical Development Biologics Europe, University of Basel, Basel 4056, Switzerland

³Solvias AG, Kaiseraugst, Switzerland

Received September 18, 2021

Revised October 25, 2021

Accepted November 13, 2021

Research Article

Stronger together: Analytical techniques for recombinant adeno associated virus

With recent FDA approval of two recombinant adeno-associated virus (rAAV)-based gene therapies, these vectors have proven that they are suitable to address monogenic diseases. However, rAAVs are relatively new modalities, and their production and therapy costs significantly exceed those of conventional biologics. Thus, significant efforts are made to improve the processes, methods, and techniques used in manufacturing and quality control (QC). Here, we evaluate transmission electron microscopy (TEM), analytical ultracentrifugation (AUC), and two modes of capillary electrophoresis (CE) for their ability to analyze the DNA encapsidated by rAAVs. While TEM and AUC are well-established methods for rAAV, capillary gel electrophoresis (CGE) has been just recently proposed for viral genome sizing. The data presented reflect that samples are very complex, with various DNA species incorporated in the virus, including small fragments as well as DNA that is larger than the targeted transgene. CGE provides a good insight in the filling of rAAVs, but the workflow is tedious and the method is not applicable for the determination of DNA titer, since a procedure for the absolute quantification (e.g., calibration) is not yet established. For estimating the genome titer, we propose a simplified capillary zone electrophoresis approach with minimal sample preparation and short separation times (<5 min/run). Our data show the benefits of using the four techniques combined, since each of them alone is prone to delivering ambiguous results. For this reason, a clear view of the rAAV interior can only be provided by using several analytical methods simultaneously.

Keywords:

Analytical ultra-centrifugation / Capillary gel electrophoresis / Capillary zone electrophoresis / Characterization / Transmission electron microscopy
 DOI 10.1002/elps.202100302



Additional supporting information may be found online in the Supporting Information section at the end of the article.

1 Introduction

The adeno-associated virus (AAV) is a small, non-enveloped ssDNA virus of 25 nm in diameter that belongs to the family of parvoviridae [1]. The 4.7 kb ssDNA wild-type genome is

surrounded by a capsid with icosahedral symmetry which is formed by three viral proteins (VP1–3) [2]. The capsid proteins are all coded by one gene (Cap) and are formed through alternative splicing and are thus only differing in their N-terminus. Additionally, the genome contains four different regulatory proteins (Rep) and is flanked by 145 nucleotide inverted terminal repeats (ITR), which form a T-shaped hairpin through base pairing [3]. The genome can either be a sense (+) or anti-sense (-) DNA strand [4]. Genome packaging into an empty capsid just needs ITRs, when all helper functions, Rep and Cap are provided by plasmids. In order to produce recombinant AAV (rAAV), 95% of the AAV genome can thus be removed and replaced by a transgene [5]. The two main production methods use transfection of mammalian Human Embryonic Kidney 293 (HEK293) cells or the baculovirus expression vector system (BEVS) for infecting *Spodoptera frugiperda* (Sf9) insect cells [6]. While the Sf9 system is easier to handle and shows better scalability, it is reported that rAAVs from this source tend to have a different expression of viral proteins, different

Correspondence: Andrei Hutanu, Analytical Development and Quality Control, Pharma Technical Development Biologics Europe, F. Hoffmann-La Roche AG, Grenzacherstrasse 124, 4070 Basel, Switzerland
 E-mail: andrei.hutanu@roche.com

Abbreviations: AEX, anion exchange chromatography; AUC, analytical ultra-centrifugation; CPA, corrected peak area; CQA, critical quality attribute; DBA, dye based assay; GC/mL, genome copies per mL; GMP, good manufacturing practice; HMW, high molecular weight; ITR, inverted terminal repeat; MWCO, molecular weight cutoff; QC, quality control; rAAV, recombinant adeno associated virus; S1 or S2, supplier 1 or 2; TEM, transmission electron microscopy; VP/mL, virus particles per mL; VP1–3, virus proteins 1–3

posttranslational modifications, and different methylation of genomes, which may lead to a decreased infectivity [7,8].

The ability of site-specific integration of DNA into the host cell genome [9–11] and the missing pathogenicity [12] are attributes, that predestine AAVs to the use as gene delivery vehicles for gene therapy. With the FDA approval of two rAAV based gene therapies in 2017 and 2019 [13], these viruses have proven that they are suitable vectors to address monogenic diseases. In particular, rAAVs have shown remarkable improvements for patients with spinal muscular atrophy [14] and retinal diseases [15], but considerable research is also made for genetic remedies that affect the central nervous system, muscles, and liver [16]. Since rAAVs are relatively new modalities and production costs remain high, significant efforts are currently being made to improve the processes, methods, and techniques used in manufacturing and QC. The goal of analytical development for QC of biopharmaceuticals is to identify physicochemical properties that influence safety and efficacy and set specific ranges or limits that safeguard the high product quality that is required for therapeutic use. Most of these critical quality attributes (CQAs) have been identified for rAAVs as well and are discussed in recent reviews [17,18] (<https://alliancerm.org/manufacturing/a-gene-2021>). In this work we will focus on the CQAs discussed in the following section.

For appropriate dosing, the virus titer or concentration of the AAV sample is a product-related CQA. There are three general approaches to determine the titer. The genome titer (genome copies per mL: GC/mL) describes the capsids filled with the correct genome and is thus used for clinical dosing [19]. When the capsid concentration (viral particles per mL: VP/mL) is measured regardless of the content, the capsid titer is estimated. To get a measure of the biological activity, the infectivity titer is evaluated by different *in vitro* cell-based assays [20,21].

Production yields remain the main challenge for manufacturing rAAV because the majority of capsids tend to not contain the transgene [22]. For this reason, fast and automated methods for the estimation of the full to an empty ratio (or content ratio: full to all) are critical from a QC perspective. Capsids may also be filled with a part of the transduction cassette, the plasmid backbone, or wild-type sequences. This is expected to lead to immunotoxicity or genotoxicity [23].

Even though physicochemical methods are only rarely used in the characterization of viral vector systems during production, they are considered to play an important role in the future. Hereby, they will cover CQAs as the content ratio, the virus (protein) and genome titer, purity, or aggregation, to name just a few. TEM [24,25] and AUC [24,26] are two orthogonal techniques, that can detect different viral populations. Both have already been widely applied with their specific advantages and drawbacks. TEM is a very flexible method as it offers visual information for content ratio, protein impurities, and aggregation. However, inconsistent staining can lead to ambiguous results of capsid fillings and turnaround times are high. AUC, on the other hand, easily provides quantitative results for different viral species with a high

resolution. Its limitations include high sample consumption and low throughput. Additionally, both methods seem to be challenging to implement in a good manufacturing practice (GMP) environment. Capillary gel electrophoresis (CGE) was recently proposed as a method that can separate different size variants of DNA in rAAV samples and could even have the prospect for genome titer estimation [27]. Related CGE methods are extensively used for the assessment of protein impurities of biopharmaceuticals including rAAVs [28–30]. In this publication, we discuss the advantage of the synergistic application of four methods, TEM, AUC, and two submodes of CE (see Table 1 [24,31–37]), focusing on the quantity, quality, and state of the genome. HPLC methods such as RP-HPLC, SEC or AEX, or other CE submethods as cIEF are not considered here. The applicability of these approaches and their place in the rAAV analytical toolbox is under investigation, although some reports consider these techniques already available [17,18]. The juxtaposition of these four methods revealed an in-depth understanding of the filling of complex rAAV samples and shed new light on the strengths and weaknesses of the CGE approach.

2 Materials and methods

General: HPLC grade water was prepared using a Q-POD® Ultrapure Water Remote Dispenser (Cat. no. ZMQSP0D01) by Merck Millipore/Merck KGaA (Darmstadt; Germany). Purified AAV2 and AAV3 (approx. 2×10^{13} GC/mL by qPCR) and the ssDNA genome were obtained from different commercial manufacturers. The incorporated genomes should express an enhanced green fluorescent protein (EGFP) with a Cytomegalovirus (CMV) promoter. Genome sizes have a theoretical size of 2501 bp for Supplier 1 and 2565 bp for Supplier 2. The samples from different sources are abbreviated with S1 (Supplier 1) and S2 (Supplier 2, e.g., AAV2 S2). Samples were formulated in 1x PBS with 0.001% Pluronic F-68.

TEM: Samples were stained with 1.5% (w/v) uranyl acetate in H₂O (Merck 8473; discontinued). TEM Grids: 200 mesh, copper, formvar/carbon-coated (Cat. no. S162) were from Plano GmbH (Wetzlar; Germany). Sample preparation: 4 μ L of each sample was incubated for 30 s on a glow-discharged carbon-coated TEM grid. After a blotting step, grids were washed three times with Milli-Q water and negatively stained two times with 4 μ L of 1.5% uranyl acetate. After every staining and washing step, excess liquid was removed using a blotting paper. Image acquisition was performed on a Philips CM100 TEM operated at 80 kV.

AUC: Analysis was performed on an Optima analytical ultracentrifuge from Beckman Coulter (Indianapolis, USA). The samples were loaded into AUC cells equipped with quartz windows and 12 mm double-sector charcoal-filled EPON centerpieces. A sample volume of 370 μ L was loaded in the sample sector, whereas 390 μ L of formulation buffer was loaded in the reference sector. The AUC cells were equilibrated at 15°C for 1 h before starting the rotation at the indicated rotor speed. The sedimentation was monitored

Table 1. Overview of Methods used in this work and the CQAs they address

Technique	CQA	Specific results	Comments	Requirements test sample	Quality standard	Possible Alternatives
TEM	Content ratio	Evaluation of empty, filled and disrupted AAV Evaluation of impurities	Uneven staining may lead to ambiguous results Simple visual interpretation	$> 10^{12}$ VP/mL	non-GMP	AUC AEX [24,31] ELISA/qPCR [32] OD [32]
AUC	Content ratio	Estimation of filled and empty species Estimation of overfilled and partial filled species Estimation of degraded AAV	Empty material could be critical Re-homogenization and reutilization may be possible	$> 10^{12}$ VP/mL ca. 400 μ L Highly purified material No aggregates	non-GMP	TEM AEX [24,31] ELISA/qPCR [32] OD [32]
CGE _{DNA}	Quality of DNA Genome titer	Identity of genome (based on size/migration time) Size variants Quantification of target variant	Non-covalent labeling needed (BGE) CE separation profile depends on test sample preparation conditions	$> 10^{11}$ GC/mL > 0.2 ng/ μ L	GMP	AGE [19,33] AUC
CZE _{DNA}	Quantity of free DNA Genome titer	Specific DNA quantification based on q/r	Evaluation of encapsidated / free DNA Evaluation of DNA release under stress conditions	$> 10^{11}$ GC/mL > 0.2 ng/ μ L	GMP	qPCR [19,34] DBA [35–37]

AEX, anion exchange chromatography; qPCR, quantitative polymerase chain reaction; DBA, dye based assay

by UV absorbance at 230, 260, and 280 nm. Scans were collected every 60 s with a radial position data spacing of 10 μ m. The sedimentation coefficient and relative amounts of each species were determined by processing the data with SEDFIT (v16.1c) using c(s) model (source1). The relative amount of each subpopulation of the AAV samples was acquired by integrating the respective peak area in the sedimentation plot.

CE: Amicon Ultra-0.5 mL Centrifugal Filters with a 100 kDa cutoff (Cat. no. UFC5100BK), Benzonase Nuclease (Cat. no. E1014), 0.5 M EDTA (Cat. no. 15 575 020), polyvinylpyrrolidone (PVP; Cat. no. 437 190), 10 \times Tris borate EDTA (TBE) buffer (Cat. no. 574 795), Tris Base (Cat. no. T6066) and Urea (Cat. no. U6504) were supplied by Sigma–Aldrich/Merck KGaA (Darmstadt; Germany). DNase I 10 \times Reaction Buffer (Cat. no. AM8170G), PBS tablets (Cat. no. 18 912 014), Pluronic F-68 (Cat. no. 24 040 032), SYBRTM Green II RNA Gel Stain 10 000 \times concentrate in DMSO (Cat. no. S7564), UltraPureTM DNase/RNase-Free Distilled Water (Cat. no. 10 977 015) were purchased from Thermo Fisher Scientific (Waltham, USA). The QIAquick PCR purification Kit used for purification was from Qiagen (Germantown, USA). Acetic acid 99.7% (Cat. no. LC101001) was sourced from Fisher Chemical (Hampton; USA). Capillaries with an inner diameter of 50 μ m (Cat. no. TSP-050375) and 100 μ m (Cat. no. TSP-100375) were from Polymicro Technologies /Molex LLC (Phoenix, USA).

Analysis was carried out using a SCIEX PA800 Plus system (Brea; USA) equipped with a solid-state laser with an excitation wavelength of 488 nm and a 520 nm bandpass emission filter (Cat. no. 65–699) from Edmund Optics (Barrington; USA), a 30 kV power supply and a temperature-controlled autosampler ($\pm 2^\circ$ C). Data were acquired and an-

alyzed using 32 Karat software 10.3. **CGE:** AAV samples received from the different manufacturers were purified following the instructions from the QIAquick PCR purification kit but with two washing steps of the QIAquick column. Before injection in CE, the sample was heated at 70 $^\circ$ C for 2 min followed by 5 min in ice. The DNA digestion with benzonase and later centrifugation to remove benzonase and degraded material was performed as in [27]. After centrifugation, the sample was directly purified with the QIAquick PCR purification kit and heated at 70 $^\circ$ C for 2 min followed by 5 min in ice. The separation gel buffer consisted of 1% PVP, 4 M Urea in 1 \times TBE solution with 1:25000 diluted SYBR Green II [27,38]. A bare fused silica capillary with a 100 or 50 μ m internal diameter and 30 cm effective length was used for the separation. The separation voltage was 6 kV using reverse polarity. The samples were injected by applying -4 kV during 2–6 s. Capillary temperature was set to 25 $^\circ$ C and 10 $^\circ$ C was used for the autosampler. **CZE:** The separation buffer was prepared by dissolving Tris base to a final concentration of 25 mM and adjusting the pH to 8.0 ± 0.05 with acetic acid. SYBR-Green II was added to a final dilution of 1:10000 shortly before the analysis run. Purified linear construct which was diluted to the desired concentration with Nuclease-free distilled water served as a positive control, while formulation buffer served as blank. The same procedure as for the AAV sample was followed. Normal heat-shock: for CZE analysis 10 μ L AAV sample was heated to 70 $^\circ$ C for 2 min and placed on ice for 5 min immediately after. Variations in heating temperature and duration are described in a later section. Benzonase digest: to determine DNA content inside and outside capsids, the benzonase digest was performed as in the CGE with subsequent removal of the DNase by centrifugal filters. The residue was collected

and heated to 70°C for 2 min and afterward placed on ice for 5 min. Analysis: a neutral capillary (Cat. no. 477 441) from SCIEX (Brea; USA) with an I.D. of 50 µm and cut to a total length of 30 cm and 20 cm length to window was employed. The capillary cartridge was kept at 25°C while samples were stored at 10°C. Prior to each injection, the capillary was rinsed with water and equilibrated with separation buffer at 50.0 psi for 1 min each. Samples were injected at 0.5 psi for 10 s and separated for 4 min with the separation voltage set to –30 kV.

After each sequence, the capillary was rinsed with water for 10 min at 50 psi. For long-term storage, the capillary was kept in the refrigerator with the capillary tips placed in water. Each capillary was used for several hundred injections without loss of performance.

3 Results and discussion

In the following sections, we evaluate TEM, AUC, and two modes of CE for their ability to analyze the DNA encapsidated by rAAVs. The methods discussed below are based on generally different separation and/or detection principles but also the “target molecule” is different. For TEM and AUC analysis, the intact viral systems are investigated whereas for CE methods the genome is set on focus. For this purpose, the virus has to be destroyed with suitable methods or the conditions have to be chosen in such a way that DNA can escape from the virus. Starting from the TEM and bridging to the AUC results, we take a deeper look into the viral filling, which to a significant extent, changes the physicochemical properties of the virus itself. Not only the weight (sedimentation coefficient), but also the charge inside and outside, and the zeta potential to name a few properties.

3.1 Transmission electron microscopy

Transmission electron microscopy (TEM) is a well-established method for providing visual information on AAVs. This allows an image-based morphological assessment of the capsid content [39,40]. The structural characterization relies on good contrast, which is typically achieved with heavy metal staining salts. The salt diffuses into the vacancies of the empty AAV capsids, and due to the electron-dense staining material, the core of the capsid appears dark in the TEM images. The higher the filling level (amount of encapsidated DNA), the less space within the capsid is available for the staining material, therefore the brighter the core regions of the AAVs appear (Fig. 1A). It should be noted, that in spite of the simple principle, automatic full/empty identification of AAV particles is still challenging, especially for partially filled particles. Novel software solutions use deep learning algorithms for quantitative assessments (e.g., from Vironova: <https://cellculturedish.com/automated-integrity-analysis-of-aav-and-adenovirus-particles-using-minitem/>). These approaches provide quite accurate results. Furthermore, TEM images reveal potential aggregates, impurities such

as cell debris (Fig. 1C) broken particles (Fig. 1D), residual DNA (Fig. 1E). As one of the most abundant proteins with a size >500 kDa the 20S proteasome core is a common contamination that can easily be seen with TEM (Fig. 1F).

3.2 Analytical ultracentrifugation

As in a common centrifuge, species are separated in an AUC experiment by exposure to a centrifugal force. Depending on their size and weight, species sediment at a specific velocity through the AUC cell. Empty, full, partially filled and overfilled rAAV capsids can in this way be separated from each other. The architecture of the AUC allows furthermore the monitoring of the separation in real-time using, for example, absorbance detection [41]. Analysis of the data yields the intensity of the individual species as a function of their sedimentation coefficient, providing a sedimentation plot (Fig. 2A). The intensity is related to the extinction coefficients of the AAV capsid and the DNA, which both have a profound absorbance in the UV range (DNA maximum at 260 nm and capsid maximum at 280 nm). AUC is therefore mainly used to shed light on the content ratio as a CQA by analyzing at these wavelengths. As indicated before, the manufacturing of AAVs is still challenging and solutions often contain mixtures of species.

As a typical example, the sedimentation plot of the full AAV3 S2 shows five distinct signals of different intensity (Fig. 2A, line 1). The most intensive signal is detected with a sedimentation coefficient of 70 Svedberg. The relative standard deviation from six AUC measurements of this Svedberg value accounts to 0.1%. It is assigned to the full AAV3 species, in line with the TEM image that mostly shows this species (Fig. 1A). Furthermore, the intensity at 260 nm is larger than at 280 nm (Ratio 260/280 is approx. 1.24) indicating that the capsid contains DNA and is not empty. This argument is based on the fact that the capsid alone has its maximum absorbance at 280 nm (Fig. S1) and lower absorbance at 260 nm (Fig. 2A). By definition, the 260/280 ratio is therefore below 1 when looking at protein samples alone such as empty AAV capsids. On the contrary, the ratio for DNA is above 1 since its absorbance at 260 nm is larger than at 280 nm. A ratio of 1.24 for an AAV species in AUC can thus only stem from DNA-loaded capsids. The second most intense signal is detected at around 80 Svedberg, which means that it contains even more DNA than the full species. It is therefore assigned to an over-filled capsid as well as the signal between 90 and 100 Svedberg. Signals of smaller intensity can be detected below the full species of 70 Svedberg, indicating that they contain less or no DNA. A comparison to an empty AAV3 S2 batch shows a pronounced intensity at 50 Svedberg that is consequently associated with the empty capsid (Fig. 2A, line 2). Furthermore, the intensity ratio between 260 nm and 280 nm is 0.58, that is, a range expected for an empty capsid [41]. This also shows that 260/280 ratios above approximately 0.6 (and not only above 1) indicate that the corresponding capsids contain traces of DNA, that is, they are partially filled. This is the case

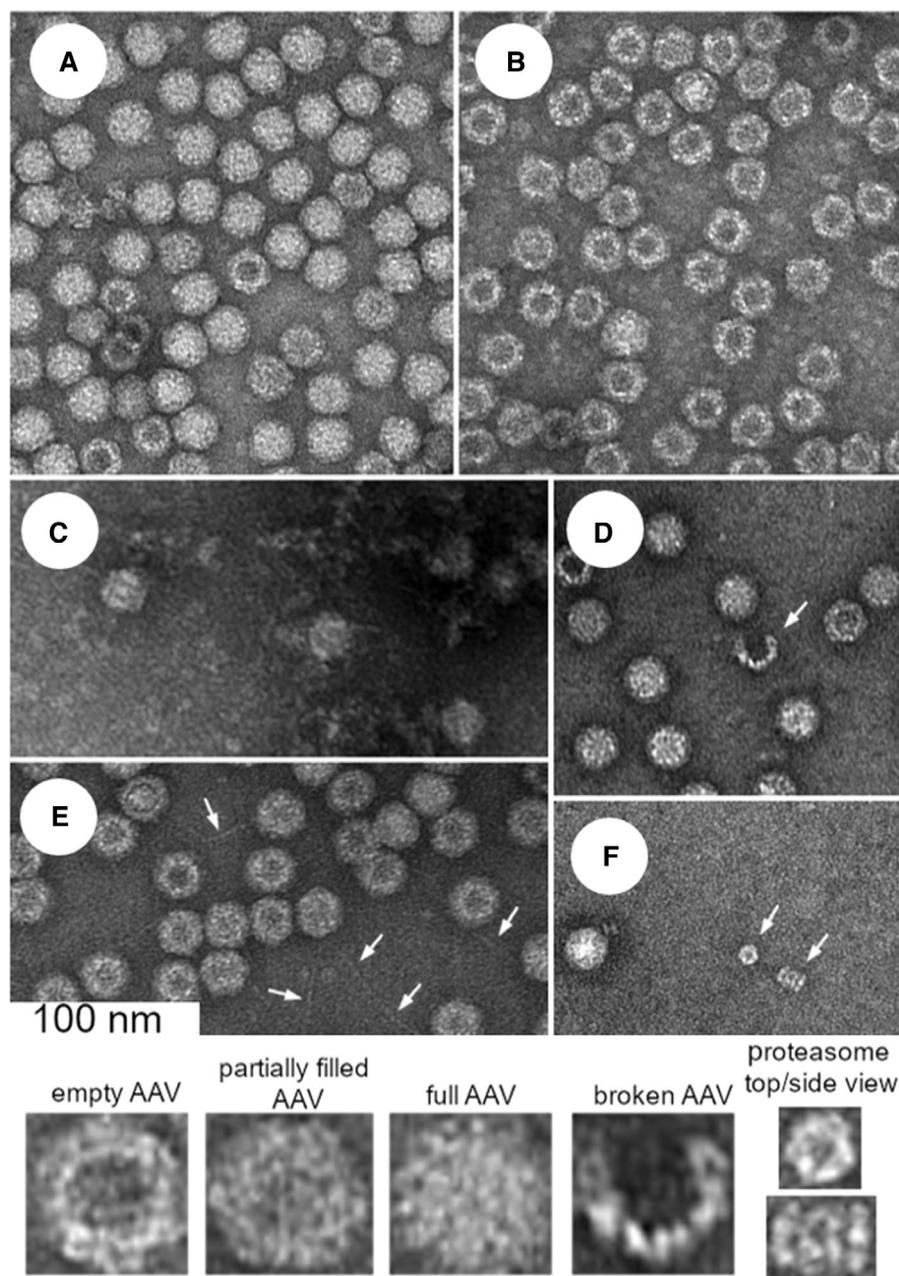


Figure 1. TEM images of AAV samples (upper part) with enlarged excerpts (lower part). (A) AAV3 S2 full, showing capsids with and without filling. (B) AAV3 S2 empty, showing capsids with and without filling. (C) AAV2 S1 full, showing impurities and aggregates in the background. (D) AAV2 S1 showing a broken AAV, (E) AAV3 S2 full, stressed, showing DNA in the background, (F) AAV2 S2 empty, after AUC showing residual proteasome impurities in the background.

for the signal at around 60 Svedberg, which falls in between empty and full capsids. The 260/280 ratio accounts to 1.22 for this signal.

Similar to the full AAV3 S2 batch, the empty AAV3 S2 batch showed five signals at 50, 60, 70, 80, and 90–100 Svedberg, but with different relative intensities. Here, the most intensive signal has already been assigned to the empty AAV3 and the remaining signal assignment follows the argumentation as for the full AAV3 S2 batch. It should be noted, that the AUC measurements of the AAV2 serotype were not repeatable and the samples degraded during the AUC run. Reuse of the solution was thus not possible and neither was an interpretation of the collected data.

To further ensure these results, the full and empty AAV3 S2 samples were mixed in a ratio of 1:1 and also analyzed. A separation of the empty and full capsids was visible (Fig 2A). The absorbance signal of the raw data consisted of an OD of 0.4 for both wavelengths at 280 and 260 nm. The theoretical relative content of empty and full capsids should be therefore approximately 50%.

AUC experiments were also conducted with a thermally stressed full AAV3 S2 batch (stressed at 70°C for 5 min). Slower and faster sedimenting species as in unstressed samples were present at 34 Svedberg and over 200 Svedberg respectively (Fig. 2B). We propose that the signal for the newly appearing but slower sedimenting species is caused by

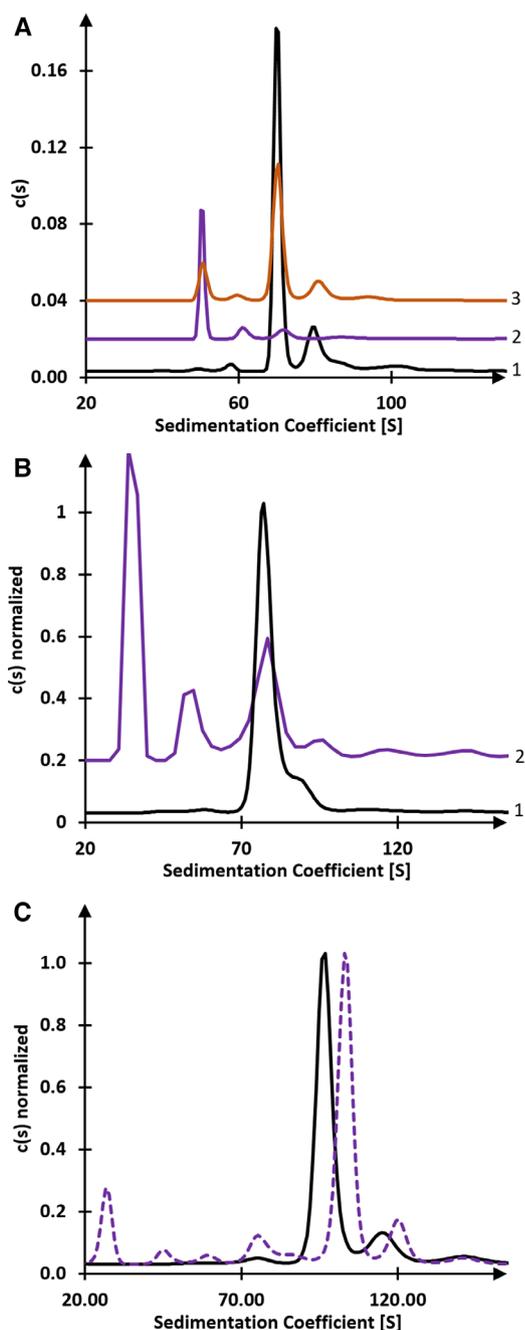


Figure 2. (A) Sedimentation distribution plot of AAV3 S2 showing the full batch (line 1), the empty batch (line 2), and the 1:1 mixture (line 3). The rotor speed was set to 20 000 rpm at 15°C and detection was conducted at 260 nm. Full and empty batches were diluted to a concentration of 7.80×10^{12} GC/mL based on the declared titer. (B) Normalized sedimentation distribution plot of AAV3 S2 showing the unstressed full batch (line 1) and the stressed full batch (line 2). Thermal stressing was carried out at 70°C for 5 min, the rotor speed was set to 16 000 rpm at 20°C. Detection was conducted at 260 nm with a sample concentration of 2.80×10^{12} GC/mL. (C) Normalized and corrected to standard conditions (water at 20°C) sedimentation distribution plot comparing full AAV3 batches from supplier 1 (dashed purple) and 2 (solid black). The rotor speed was set to 16 000 rpm at 20°C. The detection was conducted at 260 nm with a sample concentration of 2.80×10^{12} GC/mL.

capsid-fragments (34 Svedberg) and empty capsids (50 Svedberg) as detected in the stressed TEM samples (Fig. 1E). The faster sedimenting species are most probably aggregates. Furthermore, the intensity of the full capsid signal decreased whereas the signal for the empty capsid increased (Fig. 2B). The change indicates that the DNA diffuses/flows out of the capsid during thermal stressing (DNA release), which is also shown by TEM (Fig. 1E) and this can be exploited for genome quantification via CZE (see Fig. 4 later).

The collected data of the AAV3 serotype were compared between suppliers one and two. There is a slight shift in the sedimentation rate visible of the fully packed virus (Fig. 2C). This could be caused either by different packed DNA size or a general difference in the shape/weight of the vector. However, no significant differences of the packed DNA between the two distributors would be expected. By analyzing the content of the virus proteins VP1, VP2, and VP3 via CE-SDS, it was observed that the distributor with the faster sedimenting species showed a higher ratio of VP1 (not shown). This could explain the difference in the sedimentation behavior due to the fact that the VP1 protein has a higher molecular weight, compared to the VP2 and VP3 proteins.

3.3 Capillary gel electrophoresis

In the following, CGE data are used to support and complement the TEM and AUC results. With a CGE analysis, it is possible to separate all DNA size species, depending on their sequence length and conformation. In addition, it is generally possible to quantify by signal intensity using an appropriate reference standard. CGE data are complementary to qPCR and can be considered a useful extension in order to identify and characterize the genome, as CGE can distinguish between different product-related size variants, but also process-related variants as an oversized transgene or host cell DNA impurities.

The gel used in this analysis consisted of 1% PVP, 4 M Urea in $1 \times$ TBE buffer with 1:25000 diluted SYBR Green II [27,38]. While the electric field strength is an important parameter to consider in DNA analysis due to its high impact on separation efficiency and resolution, the optimum strength is mostly dependent on DNA size [42]. A low electric field of 150 V/cm together with 25°C capillary temperature were applied to have good resolution for a higher size range in around 30 min separation time. In addition, intermolecular conformation and intermolecular interaction can have a considerable impact on the separation profile. Thus, denaturing conditions during separation (4 M Urea in BGE) are indispensable [43]. Under these conditions, the separation of nine RNA transcripts from the 0.2 to 10 kb Transcript RNA Marker was possible with a resolution of 200 bases in the range of AAV genome (2 to 5 kb), although non-linear resolution is expected in this range [44].

The full AAV3 S1 and AAV3 S2 samples showed a similar profile with the main peak as the 2.5 kb ssDNA (Fig. 3 lines 2–4), which was also confirmed by spiking experiments

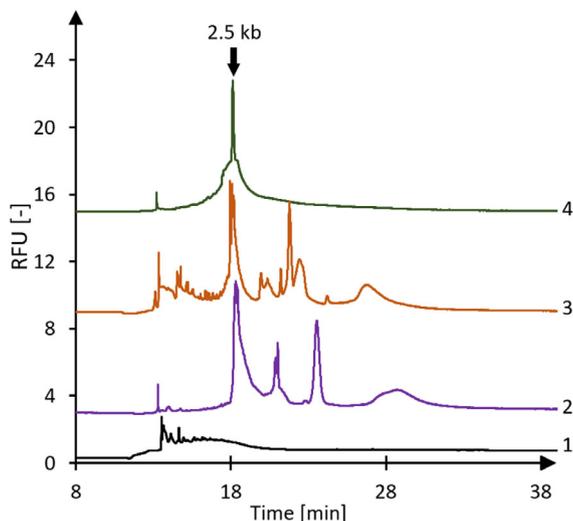


Figure 3. Comparison of CGE-LIF separation of (1) empty AAV3 S2, (2) full AAV3 S2, (3) full AAV2 S1 and (4) ssDNA genome from S1. Conditions for all lines: BGE: 1% PVP, 4 M Urea in 1×TBE with 1:25 000 SYBR Green II, Detection at 520 nm, Separation: 30/40 cm fs capillary; –6 kV at 25°C capillary temperature. Samples were purified with the QIAquick PCR purification kit. For further conditions, see chapter 2.

and migration time comparison. Larger impurities and a low amount of smaller DNA fragments can also be found in both samples. As Fig. 3 shows, empty AAV3 samples did not show any clear genome material in comparison with the respective full samples, confirming that these viral vectors do not contain the target transgene.

Benzonase was used to determine if the nucleic acid impurities observed in Fig. 3 are encapsidated or present outside the capsid, probably originated from host-cell or plasmid DNA. To initially check the performance of the benzonase treatment and subsequent filtration, the ssDNA material was submitted to this process with the later purification step with the QIAquick kit. No peak was observed in the sample treated with benzonase, which confirms the applicability of the conditions used (data not shown). The same samples from Fig. 3 were also tested by using these two preparations (with/without benzonase treatment) but no significant changes were observed. This allowed confirming that the observed peaks are only caused by encapsidated DNA.

Although comparable profiles were always obtained for all samples, variations in peak shape and relative intensity of the larger variants were observed between different sample preparations. In order to prove, that the observed peak profile is not caused by dsDNA as a method artifact, strategies were tested to reduce or control the variants, as the addition of 10% (v/v) formamide, 10% (v/v) DMSO, or 5 min sample sonication prior to the heating step at 70°C. Only formamide introduced a reduction of high molecular weight (HMW) peaks (Fig. S2). Thus, we assess that the observed peaks are related to different encapsulated ssDNA species.

When comparing CGE data with AUC and TEM some interesting observations can be made. The AAV3 S2 empty sample showed some full capsids in TEM (Fig. 1B) and two filled species in AUC (Fig. 2B). CGE also shows that there is DNA inside these samples, which contains smaller fragments than the target genome. These small-sized DNA fragments can also be seen in the full samples from both suppliers, although to a different extent. Most likely they are assessed as correctly filled in TEM and can be determined as partially filled by AUC, but may also add up as correctly filled in this technique. We assume that these capsids contain ssDNA that incorporates some hairpin, mimicking the ITR structure. Interestingly, samples from supplier two contained more of the low molecular weight DNA forms (Fig. 3 line 3) and also showed many partially filled species in AUC (Fig. 2C). Samples from both manufacturers also contain ssDNA that is larger than 2.5 kb, which is generally possible since the AAV capsid can incorporate up to 5 kb [45]. A common wrong filling takes place when the plasmid backbone including the ITRs is packaged inside the virus [46]. We would expect a plasmid backbone size of approx. 2.8 kb for Supplier 1 and 4.5 kb for Supplier 2. These differences can also be seen in different migration times of the HMW DNA in Fig. 3 (lines 2 and 3), which seems to fit this theory. An approximate look at the CPAs from Fig. 3 suggests that more than 30% of the viral capsids are filled with DNA larger than intended. Although one additional species can be expected due to the packaging of the plasmid backbone, both suppliers show at least two DNA species that are inside the virus. In AUC heavy species are detected too (80 and 90 and 100 Svedberg Fig. 2A) but the fraction seems to be smaller than CGE suggests. It is thus very likely that these AUC species are caused by the incorporation of a wrong DNA. Besides the capsid backbone, it remains unclear what these DNA species may be.

3.4 Capillary zone electrophoresis

The charge density in oligonucleotides is uniform and constant for nucleic acid molecules of varying lengths. Therefore, they typically migrate based on the charge to hydrodynamic radius ratio at the same velocity within an electrical field in a free solution. Thus, the separation of nucleic acid species, according to their number of nucleotides or nucleotide sequence, requires either the addition of a sieving matrix (see CGE) or the addition of nucleic acid species to the BGE that can separate the ssDNA in a sequence-selective manner (affinity capillary electrophoresis). Otherwise, the signal intensity can be evaluated in order to get information on the total genome concentration. Here, we start with the discussion on how a simple CZE could be used in order to quantify nucleic acid species getting a result that correlates to the genome titer. However, it should be noted that this type of analytics is in its infancy and still needs many investigations, both CZE with and without additives in the BGE.

As already mentioned, the workflow for CGE showed a poor reproducibility in signal intensity, which most likely is

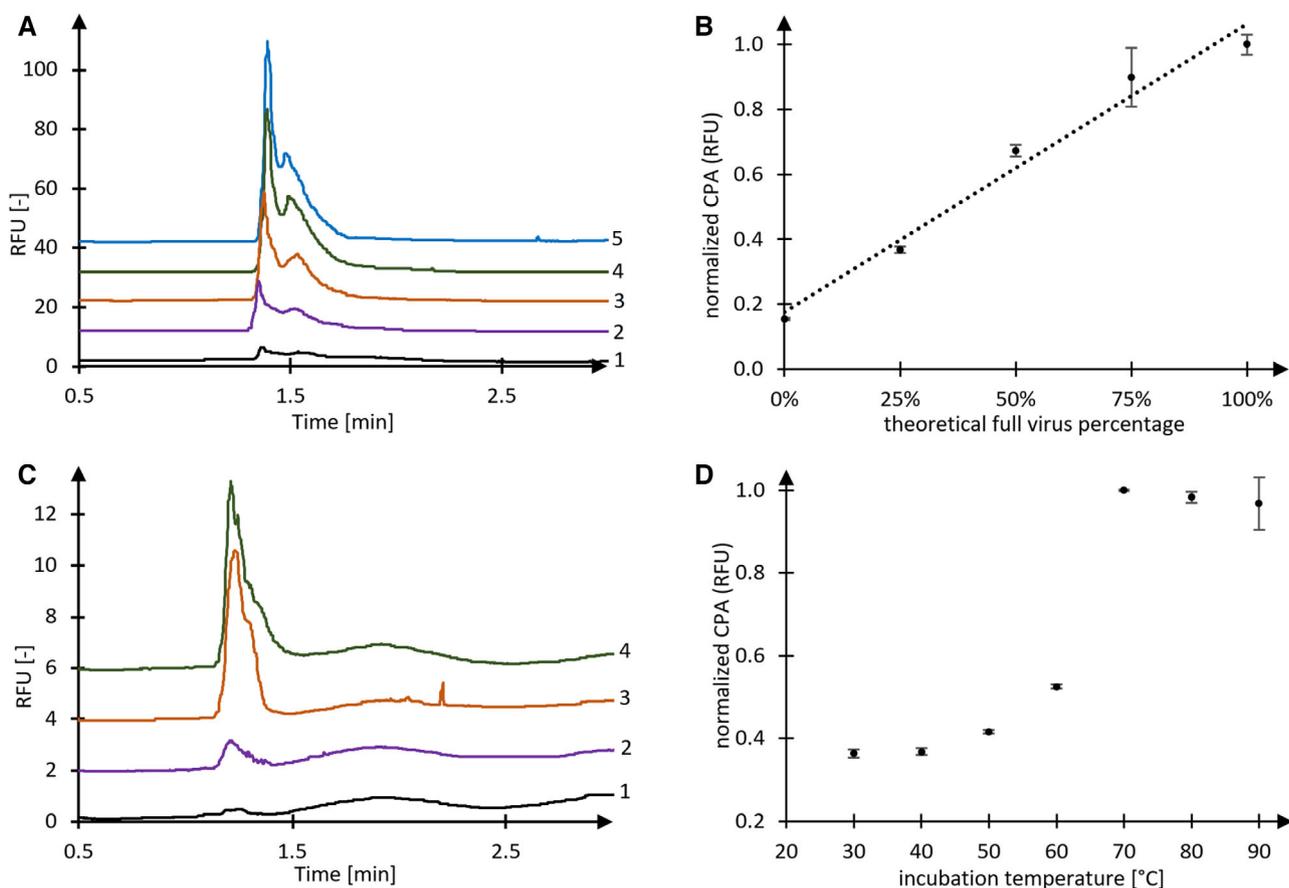


Figure 4. (A) Electropherogram of different mixtures of full and empty AAV2, S1 samples. Total concentration of each sample: 10^{13} VP/mL. Line 1: 0% full; Line 2: 25% full; Line 3: 50% full; Line 4: 75% full; Line 5: 100% full. (B) Linearity of determined CPAs from A) Error bars represent standard deviation from four measurements. CPAs were normalized by setting the highest CPA (line 5 in A) as reference. (C) Electropherograms of differently treated full AAV2, S1 samples. Line 1: benzonase; no heat stress; no DNA Line 2: no benzonase; no heat stress; DNA impurities; Line 3: benzonase; heat stress; released encapsidated DNA Line 4: no benzonase; heat stress; released encapsidated DNA + impurities. (D) CPAs of differently stressed full AAV2, S1 samples. Stress conditions: 2 min of indicated temperature. Error bars represent standard deviation from two measurements. CPAs were normalized by setting the highest CPA (70°C measurement) as reference. Conditions for all measurements: BGE: Tris 25 mM; SYBR Green II 1:10 000; Detection at 520 nm, Separation: 20/30 cm neutral coated capillary; -30 kV at 25°C capillary temperature.

caused by the two purification steps. CZE could be a good alternative to determine the DNA titer, whereby the test sample preparation is clearly simpler and no purification steps are needed. The analysis is based on the release of the DNA after rupture of full capsid forced by short thermal stress (70°C, 2 min). Furthermore, the analysis of free DNA which is already outside of the virus system, or DNA leaked after specific stress conditions could be an interesting tool for the characterization and evaluation of the status of the virus shell.

As a proof of concept, we have diluted an AAV full sample with a formulation buffer (not shown) or mixed with an empty probe, and determined the linearity of the total peak area (Fig. 4A). For the analysis of the signal intensity and thus the concentration of DNA, the entire peak area was evaluated. Using a very simple BGE consisting of 25 mM Tris and 1:10000 SYBR Green II, a good correlation of virus concentration and the fluorescence signal was observed in the range

of 10^{11} and 10^{13} GC/mL ($R^2 = 0.98$). The mixture with empty virus has a stable viral particle count but different DNA concentration and indicates that the method could also be used for estimation of the content ratio when combined with a capsid protein/particle measurement technique. We have found that the CZE profile, showing a sharp signal followed by a broader one, strongly depends on the salt concentration inside the sample. Experiments conducted with pure DNA diluted in different concentrations of PBS have shown, that at low salt concentrations only one peak is obtained, while a rising concentration leads to a more complex profile due to binding salt ions (not shown). To ensure that the detected peaks were indeed DNA, we have performed a benzonase digest after temperature stress which was able to remove all observed peaks.

For further assessment of applications of the CZE method, we have performed a similar workflow as for CGE,

with the aim to detect the DNA fraction located outside of the virus and/or the DNA inside the virus. For this purpose, the sample was either treated with benzonase with subsequent removal of the DNase via centrifugal filters (100 kDa MWCO) or just filtered for comparison. Both preparations were split and one half was stressed for 2 min at 70°C and the other half remained untreated. Figure 4C shows that the method is indeed able to detect DNA outside of the capsid that is cut by benzonase (Lines 1 and 2) and is thus suitable to estimate DNA impurities that are not inside the capsids. Using suitable DNA reference standards, quantification could also be carried out, but it must first be thoroughly clarified which conditions allow stable DNA analysis.

Another interesting application for this assay would be all kinds of stability studies or studies related to storage conditions (e.g., freeze/thaw, impact of UV irradiation) or formulation development. This setup was used to ensure that unstressed viral samples are not leaking DNA when a strong electric field (1 kV/cm) is applied, by comparing the peak size of a DNA standard with rAAV sample peak size from 83 to 1000 V/cm. Our results showed no DNA release at any field strength, indicating that the method is indeed suitable for rAAV. Additionally, we have estimated the release of viral DNA through thermal stress for a temperature ranging from 30–90°C and stress duration of 2–20 min (electropherograms not shown). Fig. 4D shows that the critical temperature for DNA release is between 60 and 70°C, while little DNA is released below 60°C. This is in line with the results obtained in TEM (Fig. 1E) which showed free DNA and more empty particles (also demonstrated by AUC Fig. 2B) when the virus was stressed at 70°C. There was no crucial difference detected between different durations meaning that the process of DNA release is very fast and is completed within 2 min. Our data shows how this approach can be used to detect different conditions which lead to DNA release, possibly helping to understand the viral unpacking process.

4 Concluding remarks

Focusing on the quantity, quality, and state of the genome this work has shown the effectiveness of a synergistic approach of orthogonal physicochemical methods. While TEM and AUC set the focus on the state of the complete viral system, with respect to genome filling, the presented CE methods can be used for the analysis of the genome itself. TEM and AUC delivered comparable results, without major discrepancies between the viral entities observed. All species identified by TEM were detected by AUC as well. AUC is capable of providing an overall picture on the heterogeneity of the samples, which can be used for the determination of relative amounts of empty, partially filled, filled, and overfilled capsids. TEM generates impressive images and can be beneficially used for the visualization of impurities, like proteasomes or broken virus capsids. CE techniques can help to clarify results obtained with TEM and AUC, which alone would be difficult to interpret.

For example, partially filled capsids might be challenging to distinguish from correctly filled ones by TEM and AUC, as was demonstrated here for the samples that were supposed to be containing empty particles. In case the determination of the filling consistency by AUC or TEM remains challenging, CGE can quickly reveal an approximate size and relative amount of the encapsidated DNA strands. Our results showed that if fragments were identified in CGE, partially filled species always appeared in the AUC. Capsids containing larger DNA strands than intended cannot be detected by TEM at all, and are only seen by CGE and AUC. Although we would expect only one oversized DNA form, that would represent the plasmid backbone, AUC and CGE have revealed at least two additional species in samples from both suppliers. CGE has thus shown that it provides comparable and complementary results to AUC. In our opinion, it is conceivable that CGE will play an important role for the assessment of rAAV DNA purity in the future.

We found that CGE provides a good insight in the filling of rAAVs, but for estimation of the genome titer, the workflow is tedious and the process of calibration is not yet established. For this reason, we propose a simplified CZE workflow that can be used under native conditions with short separation times (< 5 min/run). The detection principle relies on DNA release under heat stress (2 min at 70°C). The method was shown to be linear in a broad range and provides reproducible results. Although dye based assays are using the same principle with 96 well-plate in a fluorescence plate reader have already been reported [35–37], the current CZE approach offers the benefit of selectively detecting the pure DNA, since it is very likely that DNA and viral proteins are separated under these conditions. This should lead to less artifacts and better reproducibility. DNA concentration could be determined more accurately and precisely, when using an appropriate reference standard, and this way support the understanding of DNA release under different conditions.

The complexity of rAAVs as next-level biopharmaceuticals is much higher compared to conventional protein biologics, e.g., monoclonal antibodies. Thus, relying on a single analytical technique increases the risk of observing ambiguous data or even misleading results in quality control. The four methods presented here offer a comprehensive analytical package to assess potential CQAs related to the capsid filling, and their use should therefore be considered in combination. Further developments and improvements, especially for CE-based methods are expected, since these approaches can be easily automated, offer fast analysis, and can be operated with low sample amounts.

The authors wish to thank for their contribution to this project Dr. Bernd Moritz for his useful and constructive suggestions during the planning and development of this research work, Dr. Adelheid Rohde, Dr. Manuel Gregoritz and Dr. Jan Stracke for their valuable support, and Gianni Morson for setting up our transmission electron microscope and establishing TEM as a technique at Solvias AG. Resources provided for this research by Solvias AG are gratefully acknowledged.

The authors have declared no conflict of interest.

Data availability statement

The data that support the findings of this study are available on request from the corresponding author. The data are not publicly available due to privacy or ethical restrictions.

5 References

- [1] Naso, M. F., Tomkowicz, B., Perry, W. L., Strohl, W. R., *BioDrugs* 2017, *31*, 317–334.
- [2] Berns, K. I., Giraud, C., In: Berns, K. I., Giraud, C. (Eds.), *Adeno-Associated Virus (AAV) Vectors in Gene Therapy*, Springer, Berlin 1996 pp. 1–23.
- [3] Lusby, E., Fife, K., Berns, K. I., *J. Virology* 1980, *34*, 402–409.
- [4] Mayor, H. D., Torikai, K., Melnick, J. L., Mandel, M., *Science* 1969, *166*, 1280–1282.
- [5] Samulski, R. J., Chang, L.-S., Shenk, T., *J. Virology* 1989, *63*, 3822–3828.
- [6] Ayuso, E., Mingozzi, F., Bosch, F., *Curr. Gene Ther.* 2010, *10*, 423–436.
- [7] Rumachik, N. G., Malaker, S. A., Poweleit, N., Maynard, L. H., Adams, C. M., Leib, R. D., Cirolia, G., Thomas, D., Stammes, S., Holt, K., *Mol. Ther. Methods Clin. Dev.* 2020, *18*, 98–118.
- [8] Kondratov, O., Marsic, D., Crosson, S. M., Mendez-Gomez, H. R., Moskalenko, O., Mietzsch, M., Heilbronn, R., Allison, J. R., Green, K. B., Agbandje-McKenna, M., *Mol. Ther.* 2017, *25*, 2661–2675.
- [9] Kotin, R. M., Siniscalco, M., Samulski, R. J., Zhu, X. D., Hunter, L., Laughlin, C. A., McLaughlin, S., Muzyczka, N., Rocchi, M., Berns, K. I., *Proc. Natl. Acad. Sci. USA* 1990, *87*, 2211–2215.
- [10] Daya, S., Cortez, N., Berns, K. I., *J. Virol.* 2009, *83*, 11655–11664.
- [11] Wang, D., Tai, P. W., Gao, G., *Nat. Rev. Drug Discovery* 2019, *18*, 358–378.
- [12] Boutin, S., Monteilhet, V., Veron, P., Leborgne, C., Benveniste, O., Montus, M. F., Masurier, C., *Hum. Gene Ther.* 2010, *21*, 704–712.
- [13] Kuzmin, D. A., Shutova, M. V., Johnston, N. R., Smith, O. P., Fedorin, V. V., Kukushkin, Y. S., van der Loo, J. C., Johnstone, E. C., *Nat. Rev. Drug Discovery* 2021, *20*, 173–174.
- [14] Al-Zaidy, S., Pickard, A. S., Kotha, K., Alfano, L. N., Lowes, L., Paul, G., Church, K., Lehman, K., Sproule, D. M., Dabbous, O., *Pediatr. Pulmonol.* 2019, *54*, 179–185.
- [15] Miraldi Utz, V., Coussa, R. G., Antaki, F., Traboulsi, E. I., *Ophthalmic Genet.* 2018, *39*, 671–677.
- [16] Keeler, A. M., Flotte, T. R., *Annu. Rev. Virology* 2019, *6*, 601–621.
- [17] Wright, J. F., *Biotechnol. J.* 2021, *16*, 2000022.
- [18] Gimpel, A. L., Katsikis, G., Sha, S., Maloney, A. J., Hong, M. S., Nguyen, T. T., Wolfrum, J., Springs, S. L., Sinskey, A. J., Manalis, S. R., *Mol. Ther. Methods Clin. Dev.* 2021, *20*, 740–754.
- [19] d’Costa, S., Blouin, V., Broucque, F., Penaud-Budloo, M., Fçranois, A., Perez, I. C., Le Bec, C., Moullier, P., Snyder, R. O., Ayuso, E., *Mol. Ther. Methods Clin. Dev.* 2016, *3*, 16019.
- [20] Zhen, Z., Espinoza, Y., Bleu, T., Sommer, J. M., Wright, J. F., *Hum. Gene Ther.* 2004, *15*, 709–715.
- [21] François, A., Bouzelha, M., Lecomte, E., Broucque, F., Penaud-Budloo, M., Adjali, O., Moullier, P., Blouin, V., Ayuso, E., *Mol. Ther. Methods Clin. Dev.* 2018, *10*, 223–236.
- [22] Penaud-Budloo, M., François, A., Clément, N., Ayuso, E., *Mol. Ther. Methods Clin. Dev.* 2018, *8*, 166–180.
- [23] Davé, U. P., Cornetta, K., *Mol. Ther.* 2021, *29*, 418–419.
- [24] Fu, X., Chen, W.-C., Argento, C., Clarner, P., Bhatt, V., Dickerson, R., Bou-Assaf, G., Bakhshayeshi, M., Lu, X., Bergelson, S., *Hum. Gene Ther. Methods* 2019, *30*, 144–152.
- [25] Horowitz, E. D., Rahman, K. S., Bower, B. D., Dismuke, D. J., Falvo, M. R., Griffith, J. D., Harvey, S. C., Asokan, A., *J. Virol.* 2013, *87*, 2994–3002.
- [26] Burnham, B., Nass, S., Kong, E., Mattingly, M., Woodcock, D., Song, A., Wadsworth, S., Cheng, S. H., Scaria, A., O’Riordan, C. R., *Hum. Gene Ther. Methods* 2015, *26*, 228–242.
- [27] Luo, J., Guttman, A., *Adv. Biopharm. Anal.* 2020, *33*, 33–38.
- [28] Zhang, C. X., Meagher, M. M., *Anal. Chem.* 2017, *89*, 3285–3292.
- [29] Zhang, Z., Park, J., Barrett, H., Dooley, S., Davies, C., Verhagen, M. F., *Hum. Gene Ther.* 2021, *32*, 628–637.
- [30] Kurasawa, J. H., Park, A., Sowers, C. R., Halpin, R. A., Tovchigrechko, A., Dobson, C. L., Schmelzer, A. E., Gao, C., Wilson, S. D., Ikeda, Y., *Mol. Ther. Methods Clin. Dev.* 2020, *19*, 330–340.
- [31] Qu, G., Bahr-Davidson, J., Prado, J., Tai, A., Cataniag, F., McDonnell, J., Zhou, J., Hauck, B., Luna, J., Sommer, J. M., *J. Virol. Methods* 2007, *140*, 183–192.
- [32] Smith, P. H., Parthasarathy, S., Isaacs, J., Vijay, S., Kieran, J., Powell, S. K., McClelland, A., Wright, J. F., *Mol. Ther.* 2003, *7*, 122–128.
- [33] Fagone, P., Wright, J. F., Nathwani, A. C., Nienhuis, A. W., Davidoff, A. M., Gray, J. T., *Hum. Gene Ther. Methods* 2012, *23*, 1–7.
- [34] Wang, Y., Menon, N., Shen, S., Feschenko, M., Bergelson, S., *Mol. Ther. Methods Clin. Dev.* 2020, *19*, 341–346.
- [35] Piedra, J., Ontiveros, M., Miravet, S., Penalva, C., Monfar, M., Chillon, M., *Hum. Gene Ther. Methods* 2015, *26*, 35–42.
- [36] Xu, J., DeVries, S. H., Zhu, Y., *Hum. Gene Ther.* 2020, *31*, 1086–1099.
- [37] Bee, J. S., O’Berry, K., Zhang, Y. Z., Phillippi, M. K., Kaushal, A., DePaz, R. A., Marshall, T., *J. Pharm. Sci.* 2021, *110*, 3183–3187.
- [38] Khandurina, J., Chang, H.-S., Wanders, B., Guttman, A., *BioTechniques* 2002, *32*, 1226–1230.
- [39] Chen, H., *Microsc. Microanal.* 2007, *13*, 384–389.
- [40] Zhou, Z. H., *Structure* 2012, *20*, 1286–1288.

- [41] Schuck, P., Zhao, H., Brautigam, C. A., Ghirlando, R., *Basic principles of analytical ultracentrifugation*, CRC Press, Boca Raton, FL 2016.
- [42] Heller, C., *Electrophoresis* 2000, 21, 593–602.
- [43] Song, J. M., Yeung, E. S., *Electrophoresis* 2001, 22, 748–754.
- [44] Durney, B. C., Crieffield, C. L., Holland, L. A., *Anal. Bioanal. Chem.* 2015, 407, 6923–6938.
- [45] Wu, Z., Yang, H., Colosi, P., *Mol. Ther.* 2010, 18, 80–86.
- [46] Chadeuf, G., Ciron, C., Moullier, P., Salvetti, A., *Mol. Ther.* 2005, 12, 744–753.