






# Multidose transient transfection of human embryonic kidney 293 cells modulates recombinant adeno-associated virus2/5 Rep protein expression and influences the enrichment fraction of filled capsids

Prasanna Srinivasan<sup>1</sup>  | Christopher T. Canova<sup>2</sup> | Sha Sha<sup>3</sup> | Tam N. T. Nguyen<sup>4</sup> | John Joseph<sup>1</sup> | Jose Sangerman<sup>1</sup> | Andrew J. Maloney<sup>5</sup> | Georgios Katsikis<sup>6</sup> | Rui Wen Ou<sup>7</sup> | Moo Sun Hong<sup>8</sup>  | Jaclyn Ng<sup>9</sup> | Arella Yuan<sup>1</sup> | Daniel Antov<sup>1</sup> | Sally Song<sup>1</sup> | Wenyu Chen<sup>1</sup> | Caleb Neufeld<sup>1</sup> | Jacqueline M. Wolfrum<sup>1</sup> | Paul W. Barone<sup>1</sup>  | Anthony J. Sinskey<sup>7</sup>  | Stacy L. Springs<sup>1</sup> | Richard D. Braatz<sup>1,2</sup> 

<sup>1</sup>Center for Biomedical Innovation, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

<sup>2</sup>Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

<sup>3</sup>Ultragenyx Pharmaceutical Inc., Novato, Cambridge, USA

<sup>4</sup>BioNTech SE, Cambridge, Massachusetts, USA

<sup>5</sup>Amgen, Cambridge, Massachusetts, USA

<sup>6</sup>Anthology, Cambridge, Massachusetts, USA

<sup>7</sup>Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA

<sup>8</sup>School of Chemical and Biological Engineering, Seoul National University, Seoul, Republic of Korea

<sup>9</sup>Stanford University School of Medicine, Stanford, California, USA

## Correspondence

Anthony J. Sinskey, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA, USA.  
 Email: [asinskey@mit.edu](mailto:asinskey@mit.edu)

Stacy L. Springs and Richard D. Braatz,

## Abstract

Recombinant adeno-associated virus (rAAV) is a commonly used in vivo gene therapy vector because of its nonpathogenicity, long-term transgene expression, broad tropism, and ability to transduce both dividing and nondividing cells. However, rAAV vector production via transient transfection of mammalian cells typically yields a low fraction of filled-to-total capsids (~1%–30% of total capsids produced). Analysis of our previously developed mechanistic model for rAAV2/5 production attributed these low fill fractions to a poorly coordinated timeline between capsid synthesis and viral DNA replication and the repression of later phase capsid formation by Rep proteins. Here, we extend the model by quantifying the expression dynamics of total Rep proteins and their influence on the key steps of rAAV2/5 production using a multiple dosing transfection of human embryonic kidney 293 (HEK293) cells. We report that the availability of preformed empty capsids and viral DNA copies per cell are not limiting to the capsid-filling reaction. However, optimal expression of Rep proteins (<240 ± 13 ag per cell) enables enrichment of the filled capsid population (>12% of total capsids/cell) upstream. Our analysis suggests increased enrichment of filled capsids via regulating the expression of Rep proteins is possible but at the expense of per cell capsid titer in a triple plasmid transfection. Our study reveals an intrinsic limitation of scaling rAAV2/5 vector genome (vg) production and

Prasanna Srinivasan, Christopher T. Canova, Sha Sha, Tam N. T. Nguyen, and John Joseph contributed equally as first authors.

Arella Yuan, Daniel Antov, Sally Song, and Wenyu Chen contributed equally.

This is an open access article under the terms of the [Creative Commons Attribution-NonCommercial-NoDerivs](https://creativecommons.org/licenses/by-nc-nd/4.0/) License, which permits use and distribution in any medium, provided the original work is properly cited, the use is non-commercial and no modifications or adaptations are made.

© 2024 The Author(s). *Biotechnology and Bioengineering* published by Wiley Periodicals LLC.

Massachusetts Institute of Technology,  
Cambridge, MA, USA.  
Email: [ssprings@mit.edu](mailto:ssprings@mit.edu) and [braatz@mit.edu](mailto:braatz@mit.edu)

#### Funding information

Massachusetts Life Sciences Center;  
MathWorks Engineering Fellowship; U.S. Food  
and Drug Administration,  
Grant/Award Number: 1R01FD006584-02

underscores the need for approaches that allow for regulating the expression of Rep proteins to maximize vg titer per cell upstream.

#### KEYWORDS

Adeno-associated virus, gene therapy, mechanistic modeling, transfection, vector manufacturing

## 1 | INTRODUCTION

Recombinant adeno-associated virus (rAAV) is the most widely used *in vivo* gene therapy due to its nonpathogenicity, long-term expression, and broad tropism (Li & Samulski, 2020). Starting with the approval of LUXTURNA<sup>®</sup> in 2017, by 2023 five rAAV-based gene therapy products have been approved by the U.S. Food and Drug Administration (FDA, 2023). In addition, over 200 rAAV-based therapies are in clinical trials to treat various genetic disorders, some of which comprise a large patient population, such as with hemophilia, or require high vector genome (vg) dosage per patient, such as with Duchenne Muscular Dystrophy (Bulcha et al., 2021). The pace of clinical development and commercialization of rAAV-based gene therapy drives the demand for rAAV vector manufacturing, and therefore the need for manufacturing process development.

Several culture-based approaches have been developed for rAAV production (Sha et al., 2021), including transient transfection of HEK293 cells using multiple plasmids (Grimm et al., 1998; Matsushita et al., 1998; Xiao et al., 1998), stable integration of rAAV genes into HEK293 cell lines (Clark et al., 1995; Liu et al., 2000), infection of Sf9 cells with recombinant baculovirus (Kotin, 2011), and infection of BHK cells with Herpes Simplex Virus (Conway et al., 1997). Transient transfection of HEK293 cells remains the most flexible method to produce different serotypes of rAAV and is widely used to produce clinical material (Bulcha et al., 2021; Clément & Grieger, 2016). However, these cell cultures typically yield between 1% and 30% filled capsids in the crude harvest with the remaining lacking therapeutic transgene (Gimpel et al., 2021). The presence of empty particles in the final product increases the total capsid exposure, potentially triggers an adaptive immune response, and reduces the efficacy of gene transfer (Gao et al., 2014). Additionally, removal of empty capsids is challenging because of the structural similarities between filled and empty capsids, resulting in product loss and necessitating additional process development based on serotype (Gimpel et al., 2021; Wright, 2014).

The conventional triple transient transfection is performed in batch with all three plasmids delivered at the same time to cells once a target cell density has been achieved. Cotransfection of all three plasmids results in concomitant expression of four Rep isoforms (Rep78, Rep68, Rep52, and Rep40), three different capsid proteins (VP1, VP2, and VP3), and multiple helper proteins including assembly-activating protein and membrane-associated accessory proteins (Maurer & Weitzman, 2020). We previously developed a

mechanistic model to elucidate the dynamics of rAAV production in HEK293 cells via triple plasmid transfection. Our model predicted a poorly coordinated timeline between viral DNA (vDNA) replication and capsid production that we hypothesized contributes to poor capsid filling and that Rep proteins repressed capsid filling at later time points (Nguyen et al., 2021). Furthermore, several earlier publications reported that unregulated overexpression of Rep78/68 negatively impacts some key steps of rAAV production in cells (Chiorini et al., 1998; Jain et al., 2024; King et al., 2001; Li et al., 1997) in addition to promoting cytotoxicity (Schmidt et al., 2000).

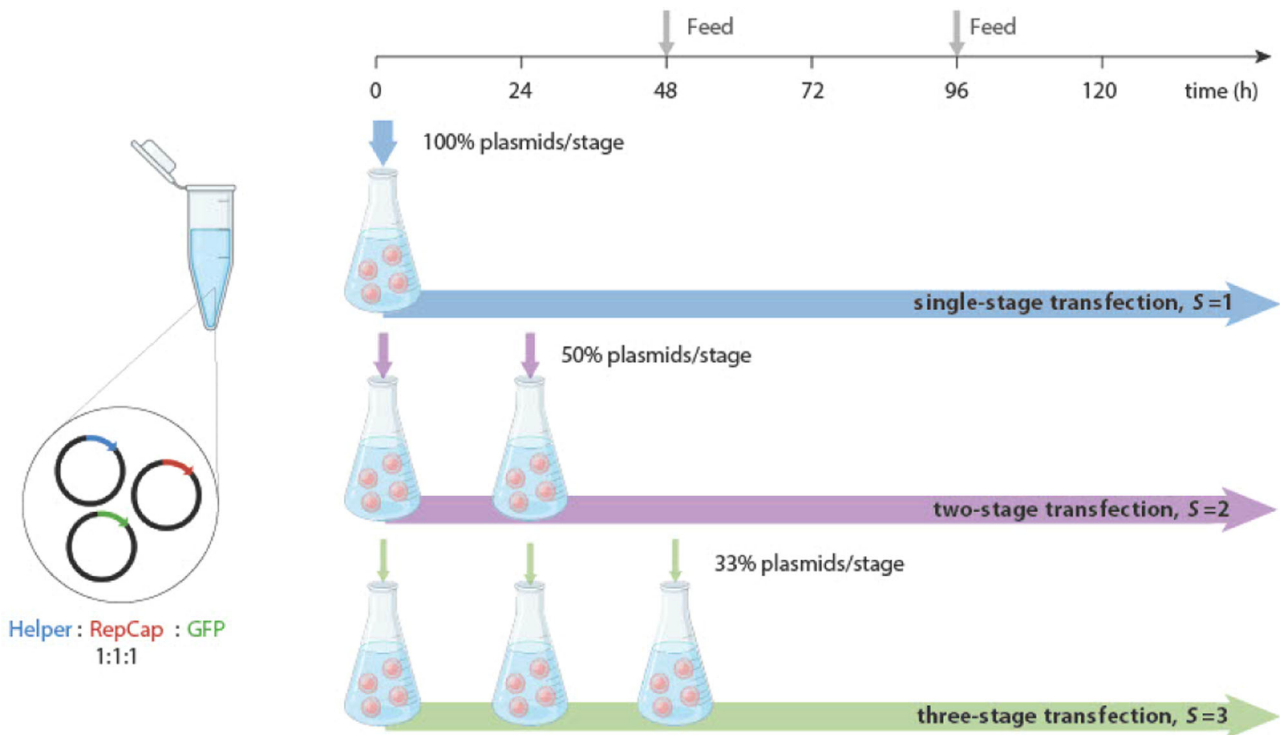
In this paper, we explore a multistage transfection strategy to understand the plasmid dosing effects on transient rAAV2/5 production. By differentially administering a set amount of plasmid DNA over multiple, time-separated doses we observe a modulation in Rep protein expression. Using the experimental data set, we improved our previously developed mechanistic model accounting for the effect of Rep protein expression on the key intermediate steps of rAAV2/5 capsid filling. Consistent with previous studies, the updated mechanistic model indicates that Rep proteins suppress both Rep and Cap protein synthesis while enhancing vDNA replication. This leads to a misalignment between capsid synthesis, vDNA replication, and encapsidation, the effects of which can be partially mitigated by reducing the intracellular concentration of Rep protein.

## 2 | RESULTS

### 2.1 | Experiments and mechanistic modeling of multistage triple transfection cell culture

We transfected HEK293 cells with plasmid (pAAV-GFP, pRC5, and pHelper at a 1:1:1 molar ratio) in three different experimental conditions to produce rAAV2/5 pseudotype (AAV vectors with an AAV5 capsid and a genome containing AAV2 inverted terminal repeats). In each condition, the total plasmid input remained constant but was evenly divided between stages in one, two, or three-stage transfections ( $S = 1, 2, 3$ , respectively), each 24 h apart (Figure 1). The 24 h interval between stages of plasmid additions was chosen to promote plasmid uptake into cells as it has been shown to peak 24 h after transfection (Carpentier et al., 2007).

In addition to the experiments, we developed a mechanistic model (Section 4) that describes the viral production process during triple transfection, extending our previous model (Nguyen



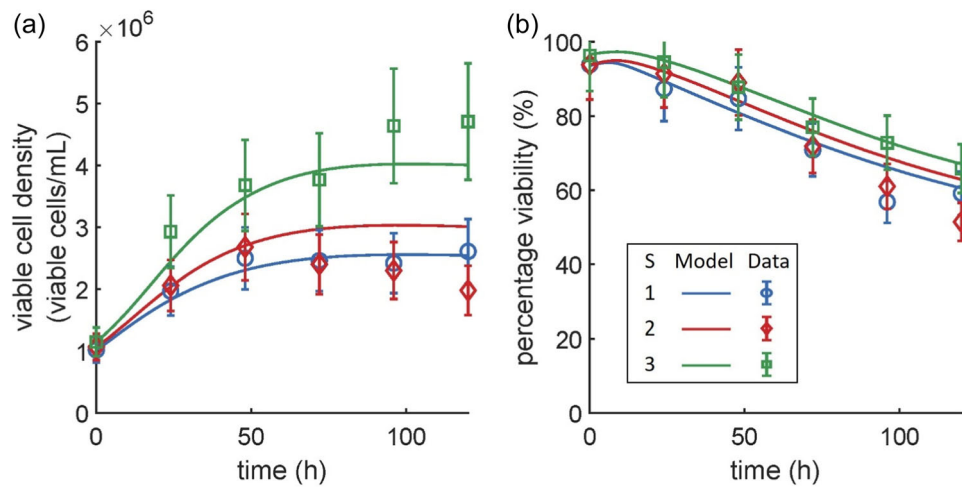
**FIGURE 1** Schematic of AAV vector production via multidose triple plasmid transfection. A transfection mix consisting of a 2:1 PEI:DNA mass ratio and 1:1:1 equimolar plasmid concentration was used. The cell density at  $t = 0$  h is  $\sim 1$  million cells/mL. Timeline of multidose transfection experimental design: one-stage transfection ( $S = 1$ ) represents the addition of  $2 \mu\text{g}$  of plasmids/mL of culture as a single dose at 0 h; two-stage transfection ( $S = 2$ ) represents the addition of  $1 \mu\text{g}$  plasmids/mL of culture at 0 and 24 h; three-stage transfection ( $S = 3$ ) represents the addition of  $0.7 \mu\text{g}$  plasmids/mL of culture at 0, 24, and 48 h. Each cell culture experiment was performed in triplicate ( $n = 3$ ) and measurements were taken at least twice. Glucose and glutamine were added at 48 and 96 h (indicated as “Feed”) to a target postfeed concentration of 5 g/L and 4 mM, respectively. AAV, adeno-associated virus; GFP, green fluorescent protein; PEI, polyethylenimine hydrochloride.

et al., 2021). We based our mechanistic model on a reaction network starting from transfection-mediated plasmid DNA delivery into the cell nucleus, leading to the synthesis of viral particles, and their release into media. The set of uncertain parameters were estimated to accurately fit the experimental data obtained at discrete time points to preserve the dynamic trends of multiple variables.

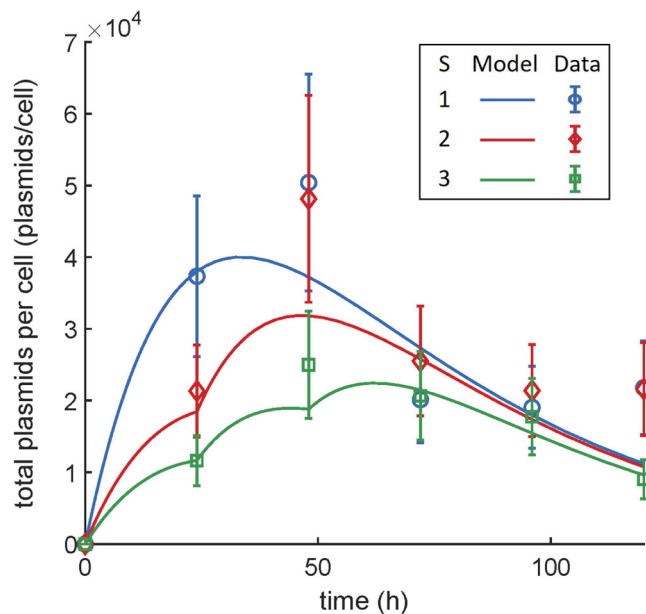
## 2.2 | Cell growth and viability improves, and plasmid uptake decreases with multiple doses

We first studied the effect of multi-stage transfection on cellular growth. Both cell density and viability improve with decreasing dose size. In particular,  $S = 3$  showed higher cell growth and viability than the  $S = 2$  and  $S = 1$  conditions (Figure 2a,b). This is inverse to the plasmid uptake values (Figure 3), in which  $S = 3$  had lower amounts of total plasmids per cell than both the  $S = 1$  and  $S = 2$  conditions for the first 72 h of culture after which the concentrations converge. We observed a statistically significant difference in the total plasmids per cell (Figure 3) between different dosing schemes ( $***p = 0.0004$  using two-way ANOVA analysis) with a reduced likelihood of measurement time influence for  $t > 72$  h ( $p > 0.9999$  using Tukey's multiple

comparison test for  $t > 72$  h). This suggests that differences in the plasmid uptake per cell via multi-dosing can possibly be attributed to events at the early time points ( $t < 72$  h). This result is complemented by the observed difference in the model-predicted total amount of DNA:PEI polyplexes taken up by the cells for different values of  $S$  when  $t < 72$  h (Supporting Information S1: Figure S1A,B). Notably the model predicts that the total plasmid DNA taken up by cells over the 120 h production timeline is roughly the same for all values of  $S$  (Supporting Information S1: Figure S1B). A possible explanation for this negligible predicted difference is that the higher growth rate of the  $S = 3$  condition balances the lower total plasmids per cell, resulting in a similar number of total plasmids taken up by the cells. Furthermore, we notice that the rate of plasmid DNA uptake depends on the state of the cells and the amount of plasmid DNA available. Although cells take up a larger total amount of the initial dose at 0 h for larger dose size (Supporting Information S1: Figure S1C), cells appear to take up a larger percentage of the initial dose with smaller dose size (Supporting Information S1: Figure S1D). In summary, multistage transfection can enable the modulation of the per cell availability of plasmid DNA (Figure 3), generating differential expression profiles that are useful for understanding rAAV2/5 production dynamics.



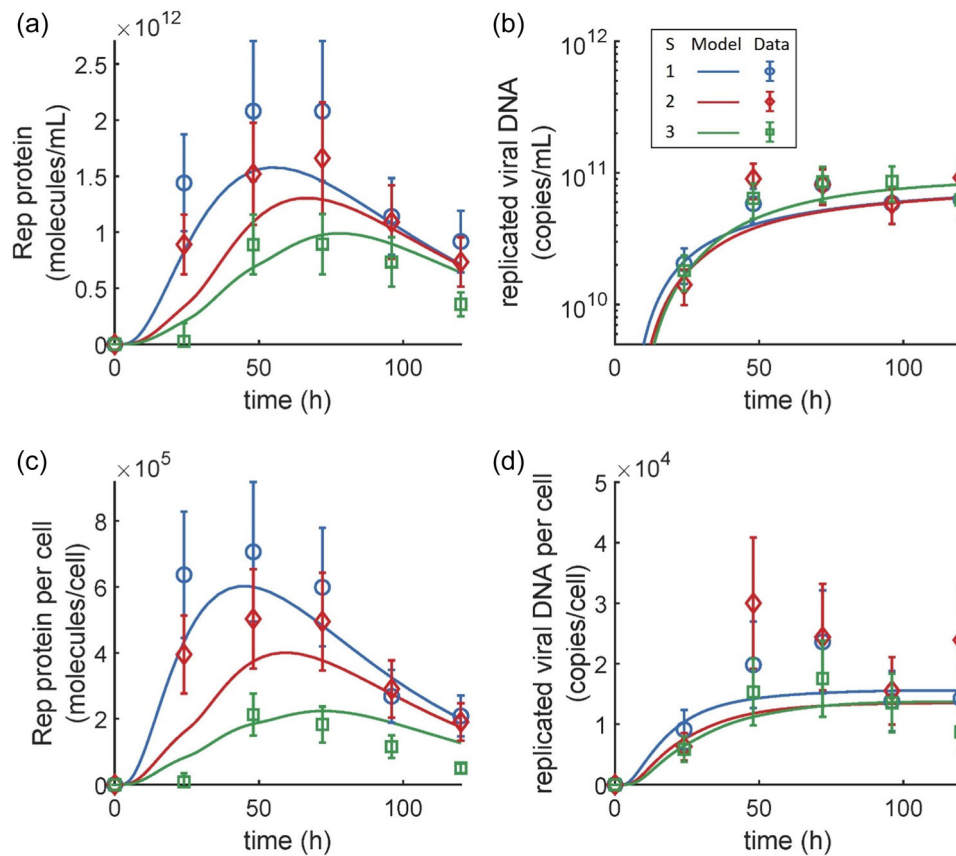
**FIGURE 2** Characterization of cell culture producing AAV via multidose triple plasmid transfection. (a) Viable cell density increased from  $t = 0$  h and reached a plateau for all dosing conditions. Model predictions agree closely with the experimental data. Two-way ANOVA analysis ( $\alpha = 0.05$ ) suggests a statistically significant difference in the viable cell density between the three different dosing schemes ( $***p < 0.0001$ ). Tukey's multiple comparison test indicates measurement time influences the statistical significance of the experimental data of different dosing schemes at the later stage of the production timeline ( $t > 96$ ) between  $S = 1$  and  $S = 3$  ( $**p = 0.0016$ ,  $**p = 0.003$  at 96 and 120 h, respectively) and  $S = 2$  and  $S = 3$  ( $***p = 0.0008$ ,  $***p = 0.0001$  at 96 and 120 h, respectively). (b) Cell viability decreases over time after transfection at  $t = 0$  h. Model predictions agree closely with the experimental data for all transfection conditions. Two-way ANOVA analysis ( $\alpha = 0.05$ ) suggests a statistically significant difference in the percent viability between the three different dosing schemes ( $***p = 0.0004$ ). Tukey's multiple comparison test indicates measurement time influences the statistical significance of the experimental data of  $S = 1$  and  $S = 3$  at the later stage of the production timeline ( $t > 96$  h;  $p = 0.0243$ ). AAV, adeno-associated virus.



**FIGURE 3** Dynamics of total plasmid uptake per cell after multidosing transient transfection. Two-way ANOVA analysis ( $\alpha = 0.05$ ) suggests significant difference in the total plasmids per cell between multidosing transfection stages ( $***p = 0.0004$ ). Tukey's multiple comparison test indicates that the differences between conditions are statistically insignificant  $t > 72$  h ( $p > 0.9999$ ).

### 2.3 | Rep proteins upregulate vDNA replication

Rep proteins influence several important key steps of rAAV particle production including vDNA replication (Li et al., 1997), vDNA targeting to capsids, vDNA packaging into capsids (Dubielzig et al., 1999; King et al., 2001; Wistuba et al., 1995), and Rep/Cap protein expression (Pereira et al., 1997; Trempe & Carter, 1988). We observe that the average total Rep protein concentration (Figure 4a) reaches a maximum between  $t = 48$  and  $72$  h irrespective of the number of transfection stages. Notably, we observe differential expression of Rep proteins per cell (Figure 4c) for the duration of the experiment in agreement with the total plasmids per cell concentration profiles (Figure 3). The difference in Rep protein concentration for  $S = 1, 2, 3$  correlates inversely with observed differences in cell growth and viability, which may be attributed to the cytotoxic effect of Rep proteins (Schmidt et al., 2000) and with other confounding impacts of transient transfection such as the presence of PEI. The upregulation of vDNA replication by Rep protein appears to reach a saturation limit in the dosing regimens considered (Figure 4b). For instance, we observe no significant difference in the total replicated vDNA even though Rep proteins were differentially expressed within this period. However, there is a statistically significant difference in the replicated viral DNA per cell between different dosing schemes ( $*p = 0.0203$  using two-way ANOVA analysis) with a post-hoc Tukey's multiple comparison test indicating statistically



**FIGURE 4** Dynamics of the Rep proteins and viral DNA synthesis. Values reported as volumetric concentrations were obtained directly from the experimental measurements and values reported on a per cell basis were calculated by dividing the volumetric concentrations with total cell density measurements. (a) Total Rep proteins produced per mL of culture. (b) Total replicated viral DNA synthesized per mL of culture. (c) Rep proteins per cell. Two-way ANOVA analysis ( $\alpha = 0.05$ ) indicates a statistically significant difference in the Rep proteins per cell between different dosing schemes ( $p < 0.0001$ ). Tukey's multiple comparison test indicates measurement time influence on the statistical significance of the experimental data from different dosing schemes for  $t \leq 72$  (\*\*\*\* $p < 0.0001$  at 24 and 48 h and \*\* $p = 0.001$  at 72 h between  $S = 1$  and  $S = 3$ ; \*\* $p = 0.0025$ , \* $p = 0.0293$ , \* $p = 0.0175$  at 24, 48, and 72 h, respectively, between  $S = 2$  and  $S = 3$ ). (d) Predicted replicated viral DNA per cell. Two-way ANOVA analysis ( $\alpha = 0.05$ ) indicates statistically significant difference in the replicated viral DNA per cell between different dosing schemes (\* $p = 0.0203$ ). Tukey's multiple comparison test indicates statistically insignificant differences at the measured time points for all conditions ( $p > 0.9$ ; [d]).

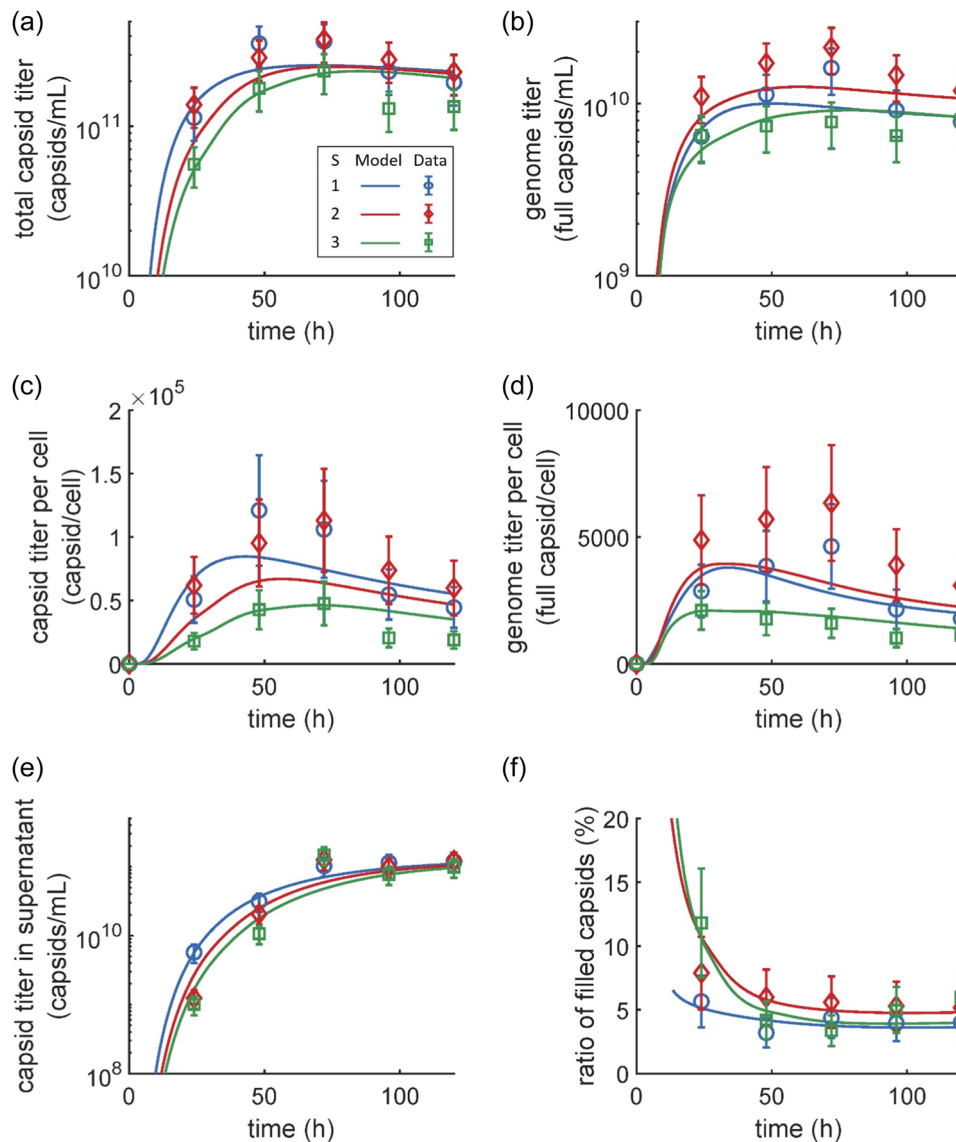
insignificant differences at the measured time points ( $p > 0.9$ ; Figure 4d).

## 2.4 | Rep protein expression impacts capsid synthesis and filling

The total capsid titer for  $S = 1$  and  $S = 2$  nearly overlap and exceed that of  $S = 3$  (Figure 5a). We observe a statistically significant difference in the total capsid titer per cell between different dosing schemes (\*\*\*\* $p < 0.0001$  using two-way ANOVA analysis) with a post-hoc Tukey's multiple comparison test indicating statistically significant differences at multiple timepoints when comparing  $S = 1$  and  $S = 2$  with  $S = 3$  (Figure 5c). Interestingly, both model and experimental data for  $S < 3$  indicate saturation of the capsid titer production per cell ( $p > 0.05$  between  $S = 1$  and  $S = 2$  using Tukey's multiple comparison test) suggesting that high expression of Rep

proteins limits capsid synthesis which is consistent with previous experimental observations (Li et al., 1997). Additionally, we observe a diminishing rate of capsid synthesis at higher plasmid concentrations that is represented in the mechanistic model using a saturation kinetic formulation for the pRC5 plasmid that encodes for the capsid proteins (Section 4, Equation (15)). This formulation suggests that a limitation in some unknown aspect of the cellular machinery also works to suppress capsid synthesis (Rajendra et al., 2015). The capsid titer in the supernatant plateaus roughly 72 h posttransfection (Figure 5e).

Similar to earlier studies, we observe that a majority of vDNA encapsidation occurs within the first 24 h of culture after which genome titer reaches steady state for all values of  $S$  (Figure 5b) (Xiao et al., 1998). Rep proteins govern several key steps of capsid filling including the preparation and targeting of single-stranded vDNA-Rep complexes and vDNA insertion into pre-formed capsids (King et al., 2001; Samulski & Muzyczka, 2014). That said, overexpression of Rep



**FIGURE 5** Dynamics of rAAV2/5 capsid and genome titers produced via multidosing transient transfection. (a) Total capsids per mL of cell culture. This measurement includes intracellular capsids and capsids released in the supernatant. Capsid production increases over time and begins to plateau after 48 h. Model predictions agree closely with the experimental data. (b) Vector genome produced per mL of culture. Interestingly, the  $S = 2$  condition has a larger vector genome concentration than the  $S = 1$  and  $S = 3$  conditions. The vector genome production follows similar trends to the overall capsid production. Model predictions agree closely with the experimental data. (c) Capsid titer per cell. Values reported on a per cell basis were calculated by dividing the volumetric concentrations with total cell density measurements. Two-way ANOVA analysis suggests a significant difference in the capsid titer per cell between different dosing schemes ( $***p < 0.0001$ ). Tukey's multiple comparison test indicates a measurement time influence on the statistical significance of the experimental data from different dosing schemes across the entire production timeline ( $**p = 0.0089$  at 24 h,  $****p < 0.0001$  at 48 and 72 h, and  $**p = 0.0064$  at 96 h between  $S = 1$  and  $S = 3$ ;  $***p = 0.0003$  at 24 h,  $****p < 0.0001$  at 48, 72, and 96 h and  $***p = 0.0009$  at 120 h between  $S = 2$  and  $S = 3$ ). (d) Genome titer per cell. Two-way ANOVA analysis ( $\alpha = 0.05$ ) suggests a significant difference in the vector genome titer per cell between different dosing schemes ( $***p < 0.0001$ ). Tukey's multiple comparison test indicates measurement time influence on the statistical significance of the experimental data from different dosing schemes across the entire production timeline ( $***p = 0.0002$  at 24 h,  $***p = 0.0005$  at 48 h,  $**p = 0.0015$  at 72 h,  $**p = 0.0011$  at 96 h and  $*p = 0.0175$  at 120 h between  $S = 1$  and  $S = 2$ ;  $****p < 0.0001$  at 24, 48, 72, and 96 h between  $S = 1$  and  $S = 3$ ;  $****p < 0.0001$  at 48 and 72 h, and  $*p = 0.0469$  at 96 h between  $S = 1$  and  $S = 3$ ). (e) Capsid titer concentration in the supernatant. (f) Ratio of filled-to-total capsids. This metric is the ratio of the genome titer concentration to the total capsid titer. The measured ratio of filled-to-total capsids is maximum at 24 h posttransfection for all dosing conditions. Fill fraction drops to  $< 10\%$  for  $t > 24$  h for all conditions. Two-way ANOVA analysis ( $\alpha = 0.05$ ) suggests a statistically significant difference in the fill fraction associated with different dosing schemes ( $p = 0.0246$ ). Tukey's multiple comparison test indicates a statistically significant difference at  $t = 24$  between  $S = 1$  and  $S = 3$  ( $*p = 0.0167$ ), suggesting the productive phase of filled capsid production is restricted within 24-h period.

protein has been shown to down-regulate capsid filling (King et al. 2001). The push-pull effect of the up- and downregulatory effects of Rep protein is demonstrated by the genome titer of the  $S = 2$  condition exceeding that of the  $S = 1$  condition (Figure 5d). A statistically significant difference in the vg titer per cell was observed between the calculated data set for different values of  $S$  ( $****p < 0.0001$  using two-way ANOVA analysis) with a post-hoc Tukey's multiple comparison test indicating statistically significant differences between all conditions across the entire production timeline (Figure 5d). The downregulation of filling by Rep also helps explain the trend in fill ratio, which peaks for all conditions early in the culture before decreasing to a baseline value (Figure 5f). These observations provided by the updated mechanistic model recapitulate previous conclusions (Nguyen et al., 2021) that there is a temporal misalignment between the capsid production and vDNA replication in triple plasmid transient transfection. This temporal misalignment can be partially corrected by dosing at lower plasmid concentrations, which limits the concentration of Rep proteins while maintaining a similar concentration of vDNA during the bulk of the capsid filling occurring early in the run (Figure 4d).

### 3 | DISCUSSION

We developed an integrated framework to advance our understanding of the molecular basis of rAAV2/5 production via transient transfection of HEK293 cells. Using experimental and modeling approaches, we analyzed the dynamics of critical molecular components such as intracellular plasmids (Figure 3), Rep proteins (Figure 4a), vDNA (Figure 4c), capsid titer (Figure 5a), and vg titer (Figure 5b). This analysis allows us to identify the potential underlying molecular mechanisms that govern plasmid-to-virion signaling at the single-cell level and better understand the intrinsic limitations of achieving filled capsid enrichment via triple plasmid transient transfection. Based on our findings, we propose a mechanism of Rep protein-mediated rAAV2/5 capsid filling via multidosing transfection schemes (Figure 6). Multidosing transfection schemes regulate Rep protein expression which influences the capsid titer and vDNA encapsidation rate during the initial productive phase ( $t < 48$  h).

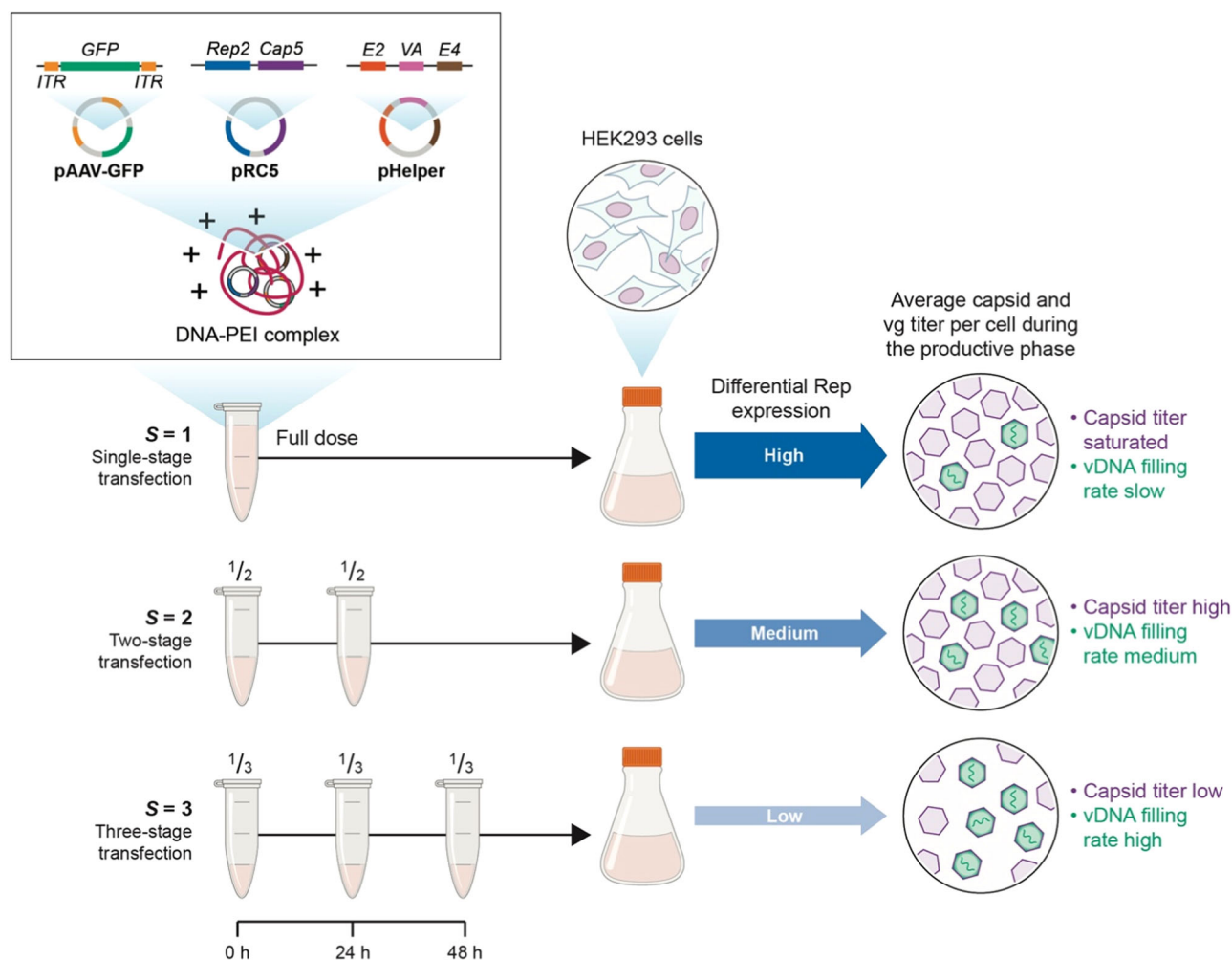
The plasmid uptake rate is modeled as a function of the viable cell density, extracellular plasmid concentration, and dose timing (Section 4, Equation 3). The number of binding sites of polyplexes increases with increasing viable cell density, which increases early in the culture before plateauing (Figure 2a). Despite the demonstrated decreasing ability of cells to uptake PEI:DNA complexes with multiple doses, multiple dosing transfection has been successfully demonstrated to improve the titers of virus-like particle production (Cervera et al., 2015; Fuenmayor et al., 2019). A key difference between these systems and AAV production, however, is the presence of Rep protein; when plasmids are dosed later in an AAV culture, they exhibit diminished benefits due to the repressive nature of the Rep protein that is already synthesized inside of the cells (Pereira et al., 1997; Trempe & Carter, 1988).

We used biophysical methods to orthogonally confirm the percentage of filled capsids at  $t = 24$  h. We first characterized the retention time of monomeric single particle AAVs using size exclusion chromatography (Fig. S2A). The percentage of filled capsids was then quantified based on the absorbance ratio (AR) at  $\lambda = 260$  nm and  $\lambda = 280$  nm at the single particle retention time (Sommer et al., 2003). The AR increases with increasing  $S$  (Supporting Information S1: Figure S2B), signifying an increase in the percentage of filled capsids with smaller initial doses of plasmid ( $S = 3 > S = 2 > S = 1$ ). These biophysical estimates of the percentage of filled capsids agree with the estimates made via biochemical approaches using ddPCR and ELISA (Supporting Information S1: Figure S2C). Furthermore, we observed that the stoichiometries of rAAV2/5 capsids produced via multidose transfection schemes were diverse, with the proportion of VP3 increasing with increasing values of  $S$  at  $t = 24$  (Supporting Information S1: Figure S3).

The production of rAAV via transient transfection of HEK293 cells is associated with the overexpression of Rep proteins that impacts both product yield and product quality. While multistage transfection and reducing the plasmid dose allow the modulation of Rep expression level to an extent, precise independent control of Rep protein expression is challenging because it is concomitantly expressed with viral and other accessory proteins. We achieved a titration of Rep protein concentrations through a multiple-transfection approach. For all conditions, the Rep protein concentration peaks between 48 and 72 h (Figure 4a) and drops at later time points ( $t > 72$  h), which is due not only to the reduction in the per cell availability of total plasmids but also to the degradation of Rep proteins and to inhibitory effects typical of late-run batch culture production, e.g., the accumulation of metabolic byproducts such as ammonia (Supporting Information S1: Figure S4) (Hansen & Emborg, 1994; Rajendra et al., 2011).

Rep proteins regulate multiple steps relevant to filled capsid production. The mechanistic modeling formulation for viral protein synthesis (Section 4, Equation 15) and Rep protein synthesis (Section 4, Equation 14) corroborate the demonstrated regulatory functions of Rep proteins on the cap p40 promoter (Trempe & Carter, 1988) and rep p5 promoter (Pereira et al., 1997). The mechanistic model formulation for vDNA replication (Section 4, Equation 16) likewise corroborates the enhancing function of Rep proteins in vDNA replication (Li et al., 1997). In all our experiments, plasmid dose levels are high enough that Rep proteins do not appear limiting.

Apparent turnover rates of Rep proteins for different values of  $S$  were compared to characterize the influence of Rep protein expression on the average encapsidation rate of pre-formed empty capsids. We defined the apparent turnover rate of the Rep proteins for the capsid filling reaction as the number of filled capsids produced per unit time-averaged mass of Rep proteins per hour. The estimated apparent turnover rate of Rep protein catalytic activity for producing filled rAAV2/5 capsids is orders of magnitude greater ( $\sim 730.0 \pm 60.8 \text{ fg}^{-1} \text{ h}^{-1}$ ) for  $S = 3$  when compared to that of  $S = 2$  ( $\sim 36.4 \pm 25.6 \text{ fg}^{-1} \text{ h}^{-1}$ ) or  $S = 1$  ( $\sim 13.3 \pm 2.6 \text{ fg}^{-1} \text{ h}^{-1}$ ) at 24 h, again indicating the positive impact of balanced Rep protein expression on rAAV2/5 capsid



**FIGURE 6** Mechanism of Rep proteins mediated capsid filling in the productive phase of cell culture via multi-dosing transfection schemes. Transcripts of the *Rep* (blue) and *Cap* (magenta) gene elements of the pRC5 plasmid produce their respective protein products represented with arrows (blue) and hexagons (magenta). The sizes of the blue arrows represent the concentration of Rep proteins. Filled capsids (green) containing single stranded vDNA cargo encoding the GFP transgene sequence were produced along with empty capsids (magenta). A high single dose of plasmid DNA:PEI complex ( $S = 1$ ) leads to a high expression of Rep proteins that likely causes saturation of the capsid titer and a slow encapsidation rate during the productive phase of rAAV2/5. For  $S = 2$ , a relatively decreased Rep proteins expression leads to an intermediate encapsidation rate without affecting the capsid titer. Lastly, for  $S = 3$ , low Rep proteins expression most likely leads to a high encapsidation rate but a low capsid titer during the productive phase of rAAV2/5. GFP, green fluorescent protein; PEI, polyethylenimine hydrochloride; rAAV2/5, recombinant adeno-associated virus2/5; vDNA, viral DNA.

filling (Supporting Information S1: Figure S5). It is interesting to note that despite the availability of sufficient empty capsids at  $t < 24$  h, the encapsidation rate seems to be affected by the availability of Rep protein complexes dictated by Rep protein expression. The apparent turnover rate early in the run (24 h) is higher than subsequent time points for all values of  $S$  and it is statistically significant between conditions ( $***p < 0.0001$  for two-way ANOVA analysis at  $t = 24$  h), again implicating the role of other factors that impede capsid filling reaction at the later stage of production ( $p > 0.99$  for two-way ANOVA analysis at all values of  $S$  for  $t > 24$  h).

The percentage of filled capsids increases with a decreasing amount of total plasmid DNA delivered to cells at 0 h. Similar observations were previously reported when applying design of

experiments (DOEs) (Fu et al., 2023) when considering plasmid DNA amount and DNA complexation variables. Performing DOEs and other process development activities can improve the filled-to-empty ratio output of the triple plasmid system, but only to an extent due to the coupled expression of the Rep and Cap proteins. For example, although we demonstrate tangible improvements to the percentage of filled capsids in our system, it remains within the threshold reported in the literature (~1%–30% of total capsids produced) (Gimpel et al., 2021). Because of this, methods that decouple the expression dynamics of Rep and Cap should be considered. For example, an earlier work (Ohba et al., 2023) reported high filled capsid fraction enrichment (about 50% of total capsids) using a chemically induced promoter to delay Cap protein expression for various rAAV



serotypes. An additional study used two sets of chemically induced promoters to independently control genome replication and Cap expression in a stable producer cell line and observed improvements in capsid filling with delayed Cap expression (Lu et al., 2024).

While the productive phase of rAAV2/5 vg is restricted to early time points posttransfection ( $t < 48$  h), we also analyzed the negative long-term effects of batch culture on culture productivity. Since the media is not exchanged in batch culture, ammonia and other toxic byproducts are able to accumulate (Supporting Information S1: Figure S4). Accumulation of these toxic byproducts contributes to the plateauing of many molecular species, such as capsid titer (Hansen & Emborg, 1994; Rajendra et al., 2011), which we analyzed in a representative pre-experiment test run ( $n = 2$ ) with the same feeding and transfection regimens as the experiment detailed in this study (Supporting Information S1: Figure S4). We captured this inhibitory effect in the mechanistic model using a monotonically decreasing inhibition factor (Materials and Methods,  $\theta_{inh}$ ) that restricts cellular growth, protein synthesis, and vDNA replication later in the culture. The use of a perfusion bioreactor system to continuously remove these toxic byproducts may extend some aspects of protein synthesis and vDNA replication, leading to increased productivity.

Although the kinetic model accurately captures the multistage triple transfection data, some model limitations should be considered. The model does not differentiate between the four Rep isoforms. In general, the pairs Rep78/68 and Rep52/40 have different functions; for example, Rep52/40 are responsible for packing the genome into the empty capsid while Rep78/68 support DNA replication (Sha et al., 2021). Future iterations should seek to incorporate the differential functions of the four Rep isoforms in an approach similar to a mechanistic model developed for the baculovirus expression system (Destro et al., 2023). The model is also limited by its lack of metabolic state information. Incorporating metabolite concentrations into the model formulation would allow for more direct comparison of different bioreactor operational models (e.g., perfusion vs. fed batch). In addition, we assume that the plasmid uptake rate and trafficking kinetics are the same for all three plasmids. However, the uptake rate of the plasmids may decrease with increasing plasmid size (5.4, 7.3, and 11.6 kbp for pGFP, pRC5, and pHelper, respectively, Supporting Information S1: Figure S6).

Our work provides additional mechanistic understanding of the key steps involved in rAAV2/5 capsid filling and exposes the intrinsic limitation of scaling rAAV2/5 production via triple plasmid transfection. Regulated expression of Rep proteins is necessary to optimize genome titer without compromising cellular growth and viability. Understanding the interplay between Rep protein expression, vDNA replication, and capsid synthesis can inform the engineering of new plasmids and future process development strategies.

## 4 | MATERIALS AND METHODS

No in vivo experiments were performed using model organisms as a part of this work. All cell culture experiments, and viral harvesting were performed following BL2 safety procedures documented in the

biological research registration approved by the MIT Environment, Health and Safety committee.

### 4.1 | Statistical analysis

We collected samples from biological experiments performed in triplicate ( $n = 3$ ). All analytical measurements reported were obtained from at least two technical repeats with appropriate positive and negative control measurements. Serial dilutions within linear working range of assays were performed wherever necessary to improve the accuracy and confidence of our measurements. We did not assume any statistical distribution for our data analysis. Since statistical power of limited number of datasets is reduced when measurement accuracy is compromised, only those data within the linear measurement range of the assays were considered for data analysis. Two-way ANOVA analysis (GraphPad PRISM 10 software) was applied to the experimental data sets to determine the statistical significance between the dose response profiles ( $\alpha = 0.05$ ). Tukey's multiple comparison test was performed to evaluate the statistical significance (based on adjusted  $p$ -value) of the main and interaction terms of ANOVA. The statistical significance of the experimental data set at the measured time points was assessed considering only the main terms ( $S = 1$  vs.  $S = 2$ ,  $S = 2$  vs.  $S = 3$ ,  $S = 3$  vs.  $S = 1$ ) at individual time points (24, 48, 72, 96, and 120 h). Interaction terms (24 h vs. 48 h, 24 vs. 72 h, etc.) were not considered because all the dosing experiments have a same production timeline starting at  $t = 0$  h without any time offset.

### 4.2 | Materials

FreeStyle™ 293-F cells (Cat. #R79007), culture medium FreeStyle™ F17 Expression Medium (Cat. #R12338018), and Gibco™ L-glutamine (200 mM, Cat. #25030081) were purchased from Thermo Fisher Scientific, Waltham, MA. Three plasmids—pAAV-GFP (Part #AAV-400), pAAV-RC5 (Part #VPK-425), and pHelper (Part #340202) were purchased from Cell Biolabs, San Diego, CA. These plasmids were used to produce rAAV2/5 (with genomes containing the inverted terminal repeat region from AAV2 and *rep* and *cap* genes from AAV2 and AAV5, respectively) particles carrying vDNA cargo. The vDNA encodes for enhanced green fluorescent protein under the control of CMV promoter. All three plasmids were amplified via bacterial transformation using One Shot™ Stbl3™ *E. coli* strain purchased from Thermo Fisher Scientific, Waltham, MA (Cat. #C737303). The plasmids were purified using Plasmid Plus Mega kit purchased from Qiagen (Cat. #12981) and filtered through 0.22- $\mu$ m polyvinylidene fluoride filter. The concentration of all plasmids was determined using NanoDrop™ ONE<sup>c</sup> spectrophotometers (Thermo Fisher Scientific, Waltham, MA) and stored at  $-20^{\circ}\text{C}$ . We used the transfection-grade linear polyethylenimine hydrochloride (PEI Max™; 40000 MW; Cat. #24765-1) purchased from Polysciences, Warrington, PA to prepare transfection reagent for transient transfection reactions. Briefly,

transfection reagent was prepared by dissolving lyophilized PEI Max™ in Milli-Q water at 1 g/L and adjusted to pH 7.0 by adding 1 N HCl (Sigma-Aldrich, Bioreagent, Cat. #H9892-100ML). The solution was filtered through a 0.22- $\mu$ m PES filter (Millipore Stericup Quick Release, Cat. #S2GPU05RE) and stored at 4°C. The AAV5 titration ELISA kits were purchased from Progen (Cat. #PRAAV5) and the kits for AAV Rep protein were bought from Cellbiolabs (Cat. #VPK-5118). The DNase I (Cat. #M0303S) and the Monarch® Genomic DNA purification kit (Cat. #T3010S) were purchased from New England Biolabs, Ipswich, MA. The ddPCR™ (droplet digital PCR) reagents and consumables namely, QX200 EvaGreen supermix tre (Cat. #1864035), droplet generation oil for EvaGreen (Cat. #1864112), ddPCR™ droplet reader oil (Cat. #1863004), DG32 automated droplet generator cartridges (Cat. # 1864108), and ddPCR 96-well semi-skirted plates (Cat. #12001925) were all purchased from Bio-Rad Laboratories Custom-made single-stranded DNA oligomers (synthesized using the services of Integrated DNA Technologies) dissolved in nuclease-free water as per manufacturer's instructions were used as primers for amplifying target sequences in ddPCR™ reactions. DMSO (Di-Methyl Sulfoxide, Cell Culture Reagent, Cat. #196055) purchased from MP Biomedicals and 10% SDS (Sodium Dodecyl Sulfate, Cat. #71736) purchased from Sigma Aldrich were used as additives for optimizing ddPCR™ reaction conditions.

### 4.3 | Cell culture

A vial of FreeStyle™ 293-F cells stored in liquid nitrogen was thawed in a 37°C water bath and transferred into pre-warmed 30-mL FreeStyle™ F17 expression medium supplemented with 4 mM glutamine in a 125 mL shake flask (Fisher brand sterile PC flasks, Cat. #PBV125). The culture was maintained at 37°C inside a humidified incubator (Thermo Scientific HERA cell VIOS 160i) with 5% CO<sub>2</sub> on a vibration resistant orbital shaker (ORBI SHAKER™ CO<sub>2</sub>, Benchmark) at 135 rpm. Subcultures were performed between 0.3 and 3 million cells/mL. At all times the seed culture cell viability was maintained above 90%. The seed culture was expanded before transfection according to the volume needed for the transfection experiments.

### 4.4 | Triple plasmid transient transfection and sample preparation for analytics

Before transfection, the seed cells were maintained in the exponential phase at a cell density between ~1 and 2 million cells/mL. Based on the total number of cells required for the experiment, the corresponding volume of the seed culture was transferred to 50-mL centrifuge tubes and centrifuged at 1000 rpm for 5 min. After discarding the supernatant, the pelleted cells were suspended in a fresh pre-warmed medium (FreeStyle™ F17 expression medium supplemented with 4-mM glutamine) in 125-mL shake flasks. In each shake flask, 20 mL of cells were added to target 1 million cells/mL cell

density. Three biological replicate cultures were performed for each experimental condition.

A transfection cocktail was freshly prepared with equimolar 1:1:1 ratio of pAAV-GFP:pAAV-RC5:pHelper for each transfection reaction. Plasmid dose as indicated in the transfection schemes (2, 1, or 0.7  $\mu$ g/mL) at PEI:DNA plasmid mass ratio of 2:1 was added. The transfection cocktail volume was 5% of the culture volume and it includes three components namely, plasmids, PEI, and FreeStyle™ F17 expression medium. The cocktail was prepared by adding the three components in the order of plasmids, culture medium, and PEI Max solution. The cocktail was vortexed vigorously for 10 s, incubated at room temperature for 10 min, and added directly to the cell culture. For the control culture, 20-mL culture was set up at 1 million cells/mL without adding the transfection cocktail. All the shake flasks were cultured in the same setting as seed cultures in the incubator for 120 h (5 days). At 48 and 96-h posttransfection, the metabolic profiles of the supernatant were measured to enable a feeding procedure. Glucose (300 g/L stock) and glutamine (200-mM stock) were added to the culture to bring the final glucose and glutamine concentration to 5 g/L and 4 mM, respectively.

A total of 2.2-mL sample representative of the culture volume was taken every 24 h from day 0 to day 5. Firstly, the culture solution in each shake flask was homogenously mixed by gently drawing the solution up and down thrice using a serological pipette. Next, 50  $\mu$ L of the sample drawn was immediately used for measuring cell count and viability (see Section 4.5). The rest of cell culture samples were aliquoted, 500  $\mu$ L in each microtube and centrifuged at 1000 rpm for 5 min. Supernatant was separated from the cell pellets collected at the bottom of the vial. The cell pellets were washed by gently suspending the cells in 500- $\mu$ L cold PBS followed by the removal of supernatant after centrifuging at 1000 rpm maintained at 4°C for 10 min. One aliquot reserved for intracellular plasmid and replicated DNA measurements was subjected to an additional wash step with 500  $\mu$ L cold PBS. This is done to improve the signal-to-noise ratio of the measurements by reducing the high background caused by abundant extracellular plasmids in DNA:PEI polyplexes. The supernatant and cell pellet samples prepared were stored at -20°C and -80°C, respectively, until analytical characterization experiments were performed.

### 4.5 | Cell density measurements and metabolite monitoring

To measure cell count and viability, we first mixed 50- $\mu$ L cell culture solution with 50- $\mu$ L 0.4% trypan blue stain (Gibco, Cat. #15250-061). Two technical measurements each by loading 10- $\mu$ L mixture into a Countess™ cell counting chamber slides (Thermo Fisher Scientific, Cat. #C10283) were done using Countess™ II Automated cell counter (Thermo Fisher Scientific). To analyze the metabolic profiles, the supernatant samples were thawed and screened using BioProfile™ FLEX2 (Nova Biomedical).

#### 4.6 | Average Rep proteins concentration per cell

One vial of cell pellets prepared from 500  $\mu$ L culture volume was thawed on ice. The lysis buffer (RIPA buffer and 10% protease inhibitor cocktail) was added using wide-bore tips to the cell pellets. The cell pellets were suspended in lysis buffer at a final cell density of 10 million cells/ml. This is followed by pipetting up and down a few times, and an incubation at 4°C for about 10 min. The crude lysate was then subjected to five alternative freeze-thaw cycles. Snap freezing was done in dry ice and thaw step was done in 37°C water bath. The solution was then centrifuged at 4°C for 20 min at 12,000 rpm. The supernatant collected was analyzed for estimating the Rep proteins concentration by ELISA. Briefly, samples were diluted by the diluent provided in the kit. AAV Rep2 standards provided were diluted in a range of 2–0.03 ng/mL. Two technical replicates of 100  $\mu$ L diluted standards or samples were loaded onto the antibody-coated plate and incubated at room temperature for 1 h on an orbital shaker at 120 rpm. The plate was then washed using 250  $\mu$ L 1 $\times$  wash buffer thrice, followed by incubation with 100  $\mu$ L of diluted biotinylated anti-AAV Rep primary antibody for 1 h at room temperature on an orbital shaker at 120 rpm. After incubation period, the solution was removed from the wells and the plate was washed by 250- $\mu$ L 1 $\times$  wash buffer thrice. This step was followed by another incubation with 100- $\mu$ L diluted streptavidin-enzyme conjugated secondary antibody for 1 h at room temperature on an orbital shaker at 120 rpm. After the last three washes using 250- $\mu$ L 1 $\times$  wash buffer, 100- $\mu$ L substrate solution was added, incubated at room temperature on an orbital shaker at 120 rpm. Finally, 100- $\mu$ L stop solution was added after the reaction color on the plate changed. The absorbance of the sample containing products of the reaction was read using a microplate reader at 450 nm (BioTek Instruments). According to the manufacturer, the standard Rep protein used in the kit is a conserved epitope of Rep across all four isoforms with a molecular weight of 17 kDa. The calibration curve obtained from the standards followed the trend of linear correlation between the concentration of the standards and their corresponding absorbance readouts. We estimated the average concentration of Rep proteins in the test samples analyzed from the calibration curve plotted using the standards.

#### 4.7 | Average capsid titer per cell

Capsid concentration was estimated from both cell pellets and the supernatant samples prepared previously. To quantify the capsids in the cells, one sample tube containing cells prepared from 500–13 $\mu$ L culture was thawed on ice. The pre-chilled lysis buffer (150-mM NaCl, 50-mM Tris-HCl, pH 8.5 containing 10% of protease inhibitor cocktail) was added to the cells. The ratio between the lysis buffer and the cell pellets was 100  $\mu$ L for 1 million total cells. The cells were suspended, briefly vortexed, and processed by a total of three cycles of freezing at ethanol bath (mixed by dry ice and 70% ethanol) for 10 min and thawing at 37°C water bath for 10 min. The sample was

clarified by centrifuging at 12,100 g, 4°C for 15 min. The supernatant was transferred to clean microtubes and followed by an AAV5 ELISA analysis. To quantify the capsids in the supernatant, an aliquot of the supernatant saved from cell culture was thawed and further clarified by centrifuging at 12,100 g, 4°C for 15 min. The clarified supernatant was then analyzed. Samples from either the cell lysate or the supernatant were diluted by the diluent buffer provided by the kit. AAV5 capsid standards provided by the ELISA kit were diluted to a range of  $1.20 \times 10^{10}$ – $1.88 \times 10^8$  capsids/mL. Two replicates of 100  $\mu$ L diluted standards or samples were loaded onto the antibody-coated plate and incubated at 37°C for 1 h. The plate was then washed by 200- $\mu$ L 1 $\times$  wash buffer three times, followed by incubation with 100  $\mu$ L of diluted biotinylated anti-capsid Rep antibody at 37°C for 1 h. The plate was washed by 200- $\mu$ L 1 $\times$  wash buffer three times and followed by further incubation with 100- $\mu$ L diluted streptozyme at 37°C for 1 h. After the last three washes using 200- $\mu$ L 1 $\times$  wash buffer, 100- $\mu$ L substrate solution was added, incubated at room temperature for 15 min. The 100- $\mu$ L stop solution was subsequently added. The final absorbance from the ELISA reactions was read by a microplate reader at 450 nm (BioTek Instruments). The standard curve was plotted using a four-parameter logistic (4PL) curve between the signals and the standard concentrations. The concentrations of capsids in samples (copy/mL) were subsequently calculated according to the standard curve.

#### 4.8 | Average plasmid copies and replicated vDNA per cell

We quantified the average copies/cell of all three plasmids and replicated viral DNA fragments using ddPCR assay. Copy number estimates using ddPCR is generally prone for a wide variability if the reaction conditions are not optimal. We therefore standardized the assay conditions to improve the accuracy of the measurements following MIQE guidelines (Bustin et al., 2009). We designed forward/reverse primers to specifically target and amplify a small region (between 75 and 200 bp) of GFP gene (*gfp\_for/rev*), plasmid backbone of pAAV-GFP (*gfp\_bb\_for/rev*), *rep* gene of pAAV-RC5 (*rc5\_for/rev*), and *E2A* gene of pHelper (*help\_for/rev*) plasmids without any cross-reactivity with other plasmids.

*gfp\_for*: 5'-GCAAAGACCCCAACGAGAAG-3'  
*gfp\_rev*: 5'-TCACGAACTCCAGCAGGACC-3'  
*gfp\_bb\_for*: 5'-GATGCCGCATAGTTAAGCC-3'  
*gfp\_bb\_rev*: 5'-GACGGTCACAGCTTGTCTG-3'  
*rc5\_for*: 5'-GAATCTGATTGAGCAGGCAC-3'  
*rc5\_rev*: 5'-GCATGTGGAAGTAGCTCTCTC-3'  
*help\_for*: 5'-CACGCCACGAGATTAGGTT-3'  
*help\_rev*: 5'-GAAACTCTTGCGGGCTTTG-3'

Our choice of target sequences amplified within chosen genetic elements was based on the criteria for ddPCR primer design recommended by the manufacturer. We further analyzed the composition of the primer sequence using IDT primer design tool to assess the feasibility of efficient amplification of the target sequences to

accurately quantify the gene copies. Differential free energy ( $G$ ) estimates of most hetero/homo dimers and hairpin structures of primers were in the acceptable range (between  $-3$  and  $-1$  kcal/mol) suitable for PCR amplification. However, the maximum  $G$  estimated for *gfp\_bb*, *rc5*, and helper primer sets were high ( $-6.7$ ,  $-6.3$ , and  $-9.82$  kcal/mol) and these correspond to hetrodimeric structures. Therefore, efficient disruption of primer dimer interaction is required to improve the amplification efficiency per cycle.

Before screening the actual test samples, we first optimized the reaction conditions to achieve linearity in ddPCR readouts within detectable range of the assay at high signal-to-noise ratio. We accomplished this by standardizing our assay conditions for each primer using the respective plasmids purified from *E. coli* culture as standard templates for calibration. We ensured the quality of all template plasmid DNA purified from *E. coli* preparation is high (AR (260 nm/280 nm)  $\sim 1.8$ – $2.0$ ). We noticed the ability of *rc5* and helper primer sets to amplify their target sequence dropped when primers were stored at  $4^{\circ}\text{C}$  for a long time (about a few months) or after a  $-20^{\circ}\text{C}$  freeze-thaw cycle. This was evident when the measured copy number estimates of pAAV-RC5 and pHelper plasmids agreed closely with theoretical estimates only when the respective primers were freshly reconstituted in Milli-Q water. However, when *rc5* and helper primers were subjected to a  $-20^{\circ}\text{C}$  freeze-thaw cycle, the plasmid copies estimate from the respective reactions dropped to  $\sim 80\%$  and  $\sim 50\%$  of their respective theoretical values. These results directly confirm the inhibitory effect of primer dimerization on target sequence amplification which in turn affects the gene copies estimation.

We overcome the primer-dimer interaction by adding different amounts of DMSO into the reaction mix and by using primers after heat shock. We found that 1% DMSO addition is sufficient to overcome the primer-dimer interaction after heat shock treatment of the primers (heating for 1 min at  $90^{\circ}\text{C}$  followed by cooling in ice). Using this approach, we consistently observed that the estimated copies of calibration standards agree very closely with theoretical values within 10%. We also noticed that linearization of dsDNA template (single site restriction digestion using HindIII-HF enzyme, Cat. #3104 L, New England Biolabs inc.) did not improve the estimates further (within  $\sim 3\%$ ) beyond the precision of ddPCR measurements. Next, we optimized the reaction mix formulation for the presence of PEI because test samples are prone for contamination with trace PEI that are carried after transient transfection reactions. PEI, when present in trace amounts, is known to compete with primer-template binding and can inhibit amplification of the target sequence (Zhang et al., 2020). This earlier work (Zhang et al., 2020) also reported several candidates that disrupt DNA:PEI interaction. We found that incubation of preformed DNA:PEI complex in 1:2 mass ratio prepared using standard templates in 20-mM NaOH and 2% SDS for 20 min overcame the inhibitory effects of PEI almost fully ( $>98\%$ ) if the primer-dimer interactions were disrupted by DMSO addition and heat shock treatment. Optimal reaction conditions for accurate quantification of the plasmid copies of a test sample are (i) template preparation by incubating nuclear DNA extract of HEK293 cells in

20-mM NaOH + 2%SDS for 20 min at a final concentration of  $1\text{ ng}/\mu\text{L}$ , (ii) heat shock treatment of the primers by heating for 1 min at  $90^{\circ}\text{C}$  followed by cooling in ice before addition into reaction mix, and (iii) addition of DMSO into the ddPCR reaction mix to a final concentration of 1%.

Test sample DNA was purified to estimate the quantity of pAAV-GFP, pAAV-RC5, and pHelper plasmids imported into cells at different time points of rAAV production. One sample tube containing cells from 500  $\mu\text{L}$  culture (with two PBS washes) was thawed on ice. The supernatant was discarded, and the cell pellets were collected at the bottom of the tubes. DNA was extracted from cell pellets by the Monarch™ genomic DNA purification kit following the instruction from the manufacturer. Briefly, cell pellets were suspended in 100- $\mu\text{L}$  cold PBS, mixed with 1- $\mu\text{L}$  proteinase K, 100- $\mu\text{L}$  lysis buffer, and incubated at  $56^{\circ}\text{C}$  for 5 min. Subsequently, the sample was mixed with 400- $\mu\text{L}$  binding buffer and transferred to a spin column and centrifuged at 1000  $g$ , room temperature for 3 min, followed by centrifugation at 13,000  $g$  for 1 min. The column was then placed on a clean tube and washed with 500- $\mu\text{L}$  wash buffer and a centrifugation at 13,000  $g$  for 1 min, followed by one more wash using 500- $\mu\text{L}$  wash buffer and a centrifugation at 18,000  $g$  for 1 min. The spin column was then transferred to a clean tube and mixed with 100- $\mu\text{L}$  elution buffer that was pre-heated at  $60^{\circ}\text{C}$ , and incubated for 1 min. The spin column was then centrifuged at 13,000  $g$  for 1 min. The eluent was collected was used for template preparation for ddPCR reaction.

DNA concentration of test samples (containing genomic DNA, plasmid DNA, and replicated vDNA) estimated using NanoDrop™ ONE<sup>c</sup> spectrophotometer varied within a range of two orders of magnitude (few  $\text{ng}/\mu\text{L}$  to hundreds of  $\text{ng}/\mu\text{L}$ ). Two confounding effects can be attributed to this variability. Firstly, the cell density varies at different time points. Next, DNA loss associated with purification also affects the estimated concentration. We therefore normalized the effect of cell density and nullified the effect of material loss in our assay by determining plasmid copies per pg of test sample DNA at an optimized reaction condition described earlier. We finally estimated the plasmid copies per cell by sixfold scaling of the estimated plasmid copies per pg of a test sample DNA because a single HEK293 cell contains about  $\sim 6$  pg of total genomic DNA.

Assay standardization experiments using plasmid standards consistently showed detector saturation at  $10^4$  copies/ $\mu\text{L}$ . This corresponds to a final concentration of  $\sim 0.05$ ,  $\sim 0.08$ , and  $\sim 0.12$  pg/ $\mu\text{L}$  of purified pAAV-GFP, pAAV-RC5, and pHelper standard plasmids respectively in the reaction mix. Assuming 1:1:1 stoichiometric uptake of all three plasmids, our assay design will not saturate the detector given the variability in the plasmid size and a typical per cell uptake ranging between  $\sim 10^4$  and  $10^5$  copies (Cohen et al., 2009). Hence, our assay design using pg of test sample DNA will be less likely to cause detector saturation at final concentration of reaction mix.

We prepared reaction mix for each sample by first incubating the test sample DNA for 20 min in 20-mM NaOH and 2% SDS at a final DNA concentration of  $1\text{ ng}/\mu\text{L}$ . Next, we diluted the incubation

reaction products 100-fold in DNase-free water to prepare 10-pg/ $\mu$ L stock of template for ddPCR reaction. The forward and reverse primers were mixed and diluted to a final stock concentration of 10  $\mu$ M. We used DNase-free water as a nontemplate negative control sample. In each well of a 96-well plate, reaction mixture consisting of the four components were added to 2.6- $\mu$ L water, namely, 11- $\mu$ L 2 $\times$  EvaGreen supermix, 0.22  $\mu$ L of 10- $\mu$ M primer solution after heat shock treatment, 2.2- $\mu$ L template, and 2.0- $\mu$ L DMSO. Aqueous droplets in oil prepared in QX200™ automated droplet generator using automated droplet generator cartridges were collected in ddPCR 96-well semi-skirted plates. The plates were sealed at the top using Aluminum foil before loading inside a thermal cycler (CFX96 Deep Well™ Real-Time System, Biorad). A thermal cycle program with multiple steps were used to amplify the target sequence: enzyme activation at 95°C for 5 min, 40 cycles comprising two steps (denaturation at 95°C for 30 s followed by primer annealing/extension at 60°C for 1 min), and lastly, a signal stabilization step comprising a cool down at 4°C for 5 min followed by heating at 90°C for 5 min. The amplified products were then read using QX200™ droplet reader controlled using QuantaSoft software for data acquisition.

We reduced the number of variables required to estimate the total replicated vDNA copies and plasmid copies so the error associated with the estimates can be minimized. The total replicated vDNA copies were estimated by subtracting the total GFP gene copies with the GFP gene copies associated with pAAV-GFP plasmid,

$$[\text{vDNA}] = [\text{GFP}] - [\text{GFP\_BB}]. \quad (1)$$

In this equation, [GFP] denotes the total GFP gene copies estimated using *gfp\_for/rev* primers and [GFP\_BB] represents the total GFP gene copies associated with pAAV\_GFP plasmid estimated using *gfp\_bb\_for/rev* primers.

The total plasmid copies taken up by the cells at any instant is given as

$$[\text{Tot}] = [\text{GFP\_BB}] + [\text{RC5}] + [\text{Helper}], \quad (2)$$

where [GFP\_BB], [RC5], and [Helper] represent copies of pAAV-GFP, pAAV-RC5, and pHelper plasmids estimated using their respective primer sets namely, *gfp\_bb\_for/rev*, *rc5\_for/rev*, and *help\_for/rev*.

#### 4.9 | vg titer per cell

To measure the copy number of AAV genome that was encapsidated in AAV particles remaining in cell pellets, one sample tube containing the cells from 500- $\mu$ L culture was thawed and lysed by three cycles of freeze and thaw, following the same procedures described in quantification of capsids in cells. To measure the copy number of AAV genome that was encapsidated in AAV particles that had been secreted to supernatant, the clarified supernatant could be directly used. For a 50- $\mu$ L reaction, 2- $\mu$ L of sample, 5  $\mu$ L of 10X DNase buffer, 5  $\mu$ L of DNase I (10 U), and 38- $\mu$ L DNase-free water were

mixed. The samples were incubated at 37°C for 2 h. A 50  $\mu$ L volume of 10-mM EDTA was added to the reaction to inhibit the activity of DNase I. The samples were diluted serially so that the readouts are less than the saturation level of equipment (10<sup>4</sup> copies/ $\mu$ L). All reactions for estimating genome copies were performed using *gfp\_for* and *gfp\_rev* primers.

#### 4.10 | Sample preparation for size exclusion chromatography

Triple plasmid transient transfection was performed in a 20 mL suspension HEK293 cell culture. Plasmid dose as indicated in the S = 1, 2, and 3 transfection schemes (2, 1, or 0.7  $\mu$ g/mL) at PEI:DNA plasmid mass ratio of 2:1 was added in equimolar concentration. Cells were harvested  $t = 24$  h posttransfection from all transfected flasks. The contents of the flasks were first centrifuged at 500 g for 3 min to pellet the cells and the supernatant was removed without disturbing the pelleted cells. AAVs from cell pellets were prepared using AAV-pro Purification Kkit Midi (all serotypes, TaKaRa Bio, Cat. #6675) following the manufacturer's protocol. Briefly, the cell pellet was suspended in 10 mL of AAV extraction Solution A plus in a 14-mL tube for 15 s by vortex agitation. The suspension was incubated at room temperature for 5 min and again vortexed for 15 s. This was followed by centrifugation at 4000 g for 10 min at 4°C. The supernatant was collected and transferred to a new 14-mL tube without disturbing the solid debris. About 1/10th of the volume in the vial of AAV Extraction solution B was added to the collected supernatant and mixed gently. To this mixture, 1/100th of the volume in the vial of Cryonase Cold-active nuclease was added and incubated for an hour at 37°C to digest the residual DNA fragments. The digested product was then treated with 1/10th of the volume in the vial of Precipitator A, vortexed for 10 s and incubated for 30 min at 37°C. The resultant mixture was vortexed for 10 s, treated with 1/20th of the volume in the vial of Precipitator B and vortexed again for 10 s followed by a centrifugation at 5000 g for 5 min at 4°C. The supernatant was collected and filtered through a Millex-HV 0.45-mm filter. The filtrate containing AAV particles was added to an Amicon-Ultra-15 100 kDa filter and spun at 2000g for 5 min at 15°C to ensure the sample volume was less than 1.5 mL. After discarding the filtrate, the sample collected inside the compartment was mixed with 5 mL of suspension buffer and centrifuged for 5 min at 15°C. This step was repeated four times to achieve a high degree of purity. The final wash step was concentrated to a volume of 500  $\mu$ L to achieve highly concentrated and pure AAV particles.

#### 4.11 | Biophysical characterization of AAV samples using size exclusion chromatography

Purified and concentrated AAV samples produced via transient transfection were used for biophysical characterization of AAV particles via size exclusion chromatography (SEC). 15  $\mu$ L of purified and

concentrated samples were injected into the SEC system (Agilent 1260 Infinity II). The SEC consists of an SRT SEC-100 column (Sepax, SN: 3A47249, 5 mm, 1000 Å, 4.6 × 300 mm) with 1x PBS as a mobile phase at 0.3 mL/min. UV absorbance traces at 260 nm, and 280 nm were recorded for 30 min of run time after injecting the samples into the column. The ratio of absorbances at 260 to 280 nm was used to quantify the percentage of filled capsids (~1.3 for filled capsids and ~0.6 for empty capsids). The AR computed at the retention time of AAV was used to quantify the percentage of filled capsids using the procedure described previously (Sommer et al., 2003).

#### 4.12 | Quantification of VP ratio using gel electrophoresis

About 40 µL of samples purified for size exclusion chromatography were used to estimate VP ratio using polyacrylamide gel electrophoresis (PAGE). We used 4%–15% Mini-PROTEAN TGX precast gels (Cat. #4561083, Bio-Rad) with 1x Tris/Glycine/SDS (Cat #1610732, Bio-Rad) as a running buffer. 30 µL of AAV sample was mixed with 10 µL of 4x Laemmli Buffer (Cat. # 1610747) and the mixture was boiled at 90°C for 15 min. 30 µL of the heated mixture was loaded into a well of TGX gel. BenchMark™ protein ladder (Cat. # 10747012, ThermoFisher Scientific) was used as a molecular weight standard. Electrophoretic separation was achieved by applying 100 V for 90 min. This was followed by SYPRO Ruby (Cat. # S12000, ThermoFisher Scientific) staining following the basic protocol of the manufacturer. Briefly, the basic protocol involves a fixing step using 100 mL of 50% methanol and 7% acetic acid for 30 min followed by an overnight staining step using 1x SYPRO Ruby stain. After overnight staining, the gel was washed using 100 mL of 10% methanol and 7% acetic acid for 30 min. Fluorescent signal from the washed gel was detected using Chemidoc imaging system (Bio-Rad) following SYPRO Ruby imaging protocol. Gray-scale intensity traces of the protein bands were obtained using ImageJ software (v1.53k) to quantify the VP ratio of capsids in the sample.

#### 4.13 | Mathematical modeling

We previously developed and validated a mechanistic model that describes the viral production process via triple transfection (Nguyen et al., 2021). Included in the reaction network were the nonviral gene delivery pathway and synthesis of viral particle components from plasmid DNA. With the availability of more data, both in terms of dynamics and types of measurements presented in this study, we extended the model to include another layer of information and more kinetic details to better reflect the viral production dynamics. Specific updates include extending the model from the cellular scale to the bioreactor scale (e.g., by including states such as viable cell density) and sampling over a longer duration (120 h) than the previous experiments (Nguyen et al., 2021). The complete set of state variables is outlined in Table 1.

Cell growth, cell death, and plasmid uptake are modeled by

$$r_{\text{uptake,PEI}} = f_{\text{uptake}} k_{\text{uptake,plasmid}} [\text{Plasmid}]_{\text{extracellular}} \frac{X_v}{X_v + K_{\text{uptake}}}, \quad (3)$$

$$\mu = \mu_{\text{max}} \theta_{\text{Inh}}, \quad (4)$$

$$\theta_{\text{Inh}} = \frac{K_{\text{Inh}}}{X_d + K_{\text{Inh}}}, \quad (5)$$

$$\frac{dX_v}{dt} = \mu X_v - k_{\text{death}} \frac{[\text{Rep}]}{[\text{Rep}] + K_d} X_v, \quad (6)$$

$$\frac{dX_d}{dt} = k_{\text{death}} \frac{[\text{Rep}]}{[\text{Rep}] + K_d} X_v, \quad (7)$$

$$\frac{d[\text{Plasmid}]_{\text{cell}}}{dt} = r_{\text{uptake,PEI}} - k_{\text{degrade,plasmid}} [\text{Plasmid}]_{\text{cell}}, \quad (8)$$

$$\frac{d[\text{Plasmid}]_{\text{extracellular}}}{dt} = -r_{\text{uptake,PEI}}, \quad (9)$$

$$\frac{df_{\text{uptake}}}{dt} = -k_f f_{\text{uptake}}. \quad (10)$$

The longer experiment time (120 h) leads to saturation kinetics for many of the cellular states that are modeled using Michaelis-Menten kinetic formations. Cellular growth is inhibited by the increasing concentration of toxic byproducts such as ammonia (Hansen & Emborg, 1994; Rajendra et al., 2011). The use of a monotonically decreasing metric ( $\theta_{\text{Inh}}$ ) effectively captures the combined inhibitory effects of a variety of species such as metabolites (e.g., ammonia) and PEI. The monotonic decrease in the effective plasmid uptake rate with time is modeled by the state variable  $f_{\text{uptake}}$ . This state variable captures the decreased ability of the cells to uptake plasmids that may be due to the agglomeration of already dosed plasmids or the decreased biological ability of the cells to uptake plasmids at later re-transfection events (Cervera et al., 2015; Hu et al., 2021). The model captures the cytotoxic effect of the Rep proteins in the formulation of the cellular death rate (Schmidt et al., 2000).

Plasmid trafficking is modeled by

$$\begin{aligned} \frac{d[\text{Plasmid}]_{\text{endosomal}}}{dt} &= \frac{1}{3} \frac{r_{\text{uptake,PEI}}}{X_v} \\ &\quad - (k_{\text{escape}} + k_{\text{degrade,plasmid}} + \mu) \\ &\quad [\text{Plasmid}]_{\text{endosomal}}, \end{aligned} \quad (11)$$

$$\begin{aligned} \frac{d[\text{Plasmid}]_{\text{cytosol}}}{dt} &= k_{\text{escape}} [\text{Plasmid}]_{\text{endosomal}} \\ &\quad + k_{\text{expel}} [\text{Plasmid}]_{\text{nucleus}} \\ &\quad - (k_{\text{entry,nuclear}} + k_{\text{degrade,plasmid}} + \mu) \\ &\quad [\text{Plasmid}]_{\text{cytosol}}, \end{aligned} \quad (12)$$

TABLE 1 Model parameters.

Parameter no.	Notation	Value	Confidence intervals	Unit	Method
1	$\mu_{\max}$	4.23E-02	-	$h^{-1}$	Fixed to control data
2	$K_{\text{uptake,plasmid}}$	6.65E-02	(6.31E-02, 7.01E-02)	$h^{-1}$	Fitted to data
3	$k_{\text{death}}$	7.42E-03	(6.90E-03, 7.98E-03)	$h^{-1}$	Fitted to data
4	$K_{\text{uptake}}$	4.03E+06	(3.75E+06, 4.33E+06)	live cell/mL	Fitted to data
5	$K_{\text{inh}}$	2.23E+05	(2.14E+05, 2.31E+05)	dead cell/mL	Fitted to data
6	$K_d$	7.55E+04	(6.78E+04, 8.41E+04)	Rep protein/cell	Fitted to data
7	$k_{\text{degrade,plasmid}}$	3.42E-02	(3.30E-02, 3.54E-02)	$h^{-1}$	Fitted to data
8	$k_{\text{escape}}$	6.00E-01	-	$h^{-1}$	Fixed to literature value (Varga et al., 2005)
9	$k_f$	1.00E-03	-	N/A	Fitted to literature data (Cervera et al., 2015)
10	$K_{\text{entry,nuclear}}$	4.30E-03	-	$h^{-1}$	Fixed to literature value (Nguyen et al., 2021)
11	$K_{\text{expel}}$	1.80E+01	-	$h^{-1}$	Fixed to literature value (Bishop et al., 2016)
12	$k_{\text{secrete,cyto}}$	1.02E-01	(8.99E-02, 1.15E-01)	$h^{-1}$	Fitted to data
13	$k_{\text{synthesis,Rep}}$	6.31E+03	(6.04E+03, 6.58E+03)	(Rep protein)/(plasmid)/h	Fitted to data
14	$K_{\text{RepSyn,Rep}}$	8.06E+06	(4.44E+06, 1.46E+07)	Rep protein/cell	Fitted to data
15	$k_{\text{synthesis,VP}}$	1.47E+06	(1.41E+06, 1.53E+06)	(VP protein/cell)/h	Fitted to data
16	$K_{\text{VPSyn,pRC}}$ , $K_{\text{RepSyn,pRC}}$	1.10E+01	-	plasmid/cell	Fixed to literature values (Wang et al., 2021)
17	$K_{\text{VPSyn,Rep}}$	1.08E+07	(6.31E+06, 1.84E+07)	Rep protein/cell	Fitted to data
18	$k_{\text{replication,DNA}}$	1.76E+03	(1.24E+03, 2.50E+03)	DNA/cell/h	Fitted to data
19	$K_{\text{DNARep,pHelper}}$	5.15E+00	(4.98E+00, 5.32E+00)	plasmid/cell	Fitted to data
20	$K_{\text{DNARep,Rep}}$	7.11E+03	(4.79E+03, 1.05E+04)	Rep protein/cell	Fitted to data
21	$K_{\text{DNApack,DNA}}$	2.80E+01	(2.57E+01, 3.05E+01)	DNA/cell	Fitted to data
22	$k_{\text{pack,DNA}}$	7.48E-02	(6.99E-02, 8.01E-02)	Full capsid/(Empty capsid DNA h)	Fitted to data
23	$k_{\text{degrade,Rep}}$	2.45E-02	-	$h^{-1}$	Fixed to literature value (Dahari et al., 2007; Nguyen et al., 2021; Redemann et al., 1989; Schwake et al., 2010)
24	$K_{\text{assembly}}$	7.50E-03	-	$h^{-1}$	Fitted to literature data (Yuan & Parrish, 2001)
25	$k_{\text{degrade,VP}}$	2.70E-01	-	$h^{-1}$	Fixed to literature value (Grosse et al., 2017)

(Continues)

TABLE 1 (Continued)

Parameter no.	Notation	Value	Confidence intervals	Unit	Method
26	$k_{\text{secrete,cell}}$	1.62E-02	(1.39E-02, 1.88E-02)	$\text{h}^{-1}$	Fitted to data
27	$k_{\text{degrade,cap}}$	5.99E-02	(5.63E-02, 6.37E-02)	$\text{h}^{-1}$	Fitted to data
28	$k_{\text{degrade,DNA}}$	4.87E-04	(4.24E-04, 5.59E-04)	$\text{h}^{-1}$	Fitted to data

$$\frac{d[\text{Plasmid}]_{\text{nucleus}}}{dt} = k_{\text{entry,nuclear}} [\text{Plasmid}]_{\text{cytosol}} - (k_{\text{expel}} + k_{\text{degrade,plasmid}} + \mu) [\text{Plasmid}]_{\text{nucleus}}. \quad (13)$$

Trafficking kinetics are assumed to be identical for all three plasmids (pRC5, pHelper, pAAV-GFP). It has been reported that plasmid in the nucleus can be expelled to the cytoplasm upon mitosis (Shimizu, 2005). Therefore, we added the kinetics of plasmid translocation from the nucleus to the cytoplasm described by the parameter  $k_{\text{expel}}$ .

The reaction network describing the synthesis of viral DNA, empty capsid, and genome-filled capsid from nuclear plasmids is structurally similar to the previous version of the model but with the inclusion of saturation kinetic formulations:

$$r_{\text{synthesis,Rep}} = k_{\text{synthesis,Rep}} [\text{Plasmid, Helper}]_{\text{nucleus}} \times \frac{[\text{Plasmid, RC}]_{\text{nucleus}}}{K_{\text{RepSyn,pRC}} + [\text{Plasmid, RC}]_{\text{nucleus}}} \frac{K_{\text{RepSyn,Rep}}}{K_{\text{RepSyn,Rep}} + [\text{Rep}]} \theta_{\text{Inh}}, \quad (14)$$

$$r_{\text{synthesis,VP}} = k_{\text{synthesis,VP}} \frac{[\text{Plasmid, RC}]_{\text{nucleus}}^2}{K_{\text{VPSyn,pRC}}^2 + [\text{Plasmid, RC}]_{\text{nucleus}}^2} \frac{K_{\text{VPSyn,Rep}}}{K_{\text{VPSyn,Rep}} + [\text{Rep}]} \theta_{\text{Inh}}, \quad (15)$$

$$r_{\text{replication,DNA}} = k_{\text{replication,DNA}} \frac{[\text{Rep}]}{K_{\text{DNArep,Rep}} + [\text{Rep}]} \times \frac{[\text{Plasmid, Helper}]_{\text{nucleus}}}{K_{\text{DNArep,pHelper}} + [\text{Plasmid, Helper}]_{\text{nucleus}}} \theta_{\text{Inh}}, \quad (16)$$

$$r_{\text{pack,DNA}} = k_{\text{pack,DNA}} [\text{EmptyCap}]_{\text{nucleus}} \frac{[\text{DNA}]_{\text{cell}}}{\left(1 + \frac{[\text{DNA}]_{\text{cell}}}{K_{\text{DNApack,DNA}}}\right)^2}, \quad (17)$$

$$\frac{d[\text{Rep}]}{dt} = r_{\text{synthesis,Rep}} - (k_{\text{degrade,Rep}} + \mu)[\text{Rep}], \quad (18)$$

$$\frac{d[\text{VP}]}{dt} = r_{\text{synthesis,VP}} - 60k_{\text{assembly}}[\text{VP}] - (k_{\text{degrade,VP}} + \mu)[\text{VP}], \quad (19)$$

$$\frac{d[\text{EmptyCap}]_{\text{nucleus}}}{dt} = k_{\text{assembly}}[\text{VP}] - r_{\text{pack,DNA}} - (k_{\text{secrete,cyto}} + \mu)[\text{EmptyCap}]_{\text{nucleus}}, \quad (20)$$

$$\frac{d[\text{EmptyCap}]_{\text{cyto}}}{dt} = k_{\text{secrete,cyto}}[\text{EmptyCap}]_{\text{nucleus}} - (k_{\text{secrete,cell}} + k_{\text{degrade,cap}} + \mu)[\text{EmptyCap}]_{\text{cyto}}, \quad (21)$$

$$\frac{d[\text{DNA}]_{\text{cell}}}{dt} = r_{\text{replication,DNA}} - r_{\text{pack,DNA}} - k_{\text{degrade,DNA}}[\text{DNA}]_{\text{cell}}, \quad (22)$$



$$\frac{d[\text{FullCap}]_{\text{nucleus}}}{dt} = r_{\text{pack,DNA}} - (k_{\text{secrete,cell}} + \mu)[\text{FullCap}]_{\text{nucleus}}, \quad (23)$$

$$\begin{aligned} \frac{d[\text{FullCap}]_{\text{cyto}}}{dt} &= k_{\text{secrete,cyto}}[\text{FullCap}]_{\text{nucleus}} \\ &- (k_{\text{secrete,cell}} + k_{\text{degrade,cap}} + \mu)[\text{FullCap}]_{\text{cyto}}. \end{aligned} \quad (24)$$

The model points to a multitude of Rep rate effects: Rep protein acts as an inhibitor of both Rep (Pereira et al., 1997) and viral protein synthesis (Trempe & Carter, 1988) while simultaneously boosting viral DNA replication. Interestingly, our system appears to demonstrate a combination inhibitory and enhancing effect of viral DNA on capsid filling; at low concentrations viral DNA positively impacts capsid filling while at higher concentrations viral DNA appears to have an inhibitory effect (Li et al., 1997; Prasad & Trempe, 1995). A potential mechanistic explanation for this effect is that early in the culture vDNA is limiting to viral capsid filling while later in the culture there are a greater proportion of dimerized replicated vDNA species that are energetically less favorable for insertion into preformed capsids (King et al., 2001; Prasad & Trempe, 1995). Viral protein synthesis is characterized by saturation kinetics for the RC plasmid, pointing to cellular machinery limitations in addition to the Rep inhibition effects (Rajendra et al., 2015). The synthesis and replication rates are negatively impacted by the inhibitory factor ( $\theta_{\text{inh}}$ ).

Since the reaction network in Equations (11)–(24) is in a single cell, the production rates were converted to the whole cell population per mL culture before being fit to the data. The concentration of Rep protein, replicated vDNA, total capsid, and genome capsid in the whole cell population are modeled in Equations (25)–(29) under the assumptions that (1) protein synthesis and degradation due to proteasome activities only occur in live cells, (2) vDNA copies encapsidated in capsids are not susceptible to degradation, (3) vDNA also degrades as genome-filled capsid degrades, and (4) capsid secretion to the extracellular media only occurs in live cells:

$$\frac{d[\text{Rep}]_{\text{culture}}}{dt} = r_{\text{synthesis,Rep}}X_v - k_{\text{degrade,Rep}}[\text{Rep}]_{\text{culture}}, \quad (25)$$

$$\begin{aligned} \frac{d[\text{DNA}]_{\text{culture}}}{dt} &= r_{\text{replication,DNA}}X_v - k_{\text{degrade,DNA}}[\text{DNA}]_{\text{cell}}X_v \\ &- (k_{\text{degrade,cap}} + k_{\text{secrete,cell}})[\text{FullCap}]_{\text{cyto}}X_v, \end{aligned} \quad (26)$$

$$\begin{aligned} \frac{d[\text{TotalCap}]_{\text{culture}}}{dt} &= k_{\text{assembly}}[\text{VP}]X_v - k_{\text{degrade,cap}} \\ &([\text{EmptyCap}]_{\text{cyto}} + [\text{FullCap}]_{\text{cyto}})X_v, \end{aligned} \quad (27)$$

$$\frac{d[\text{TotalCap}]_{\text{media}}}{dt} = k_{\text{secrete,cell}}([\text{EmptyCap}]_{\text{cyto}} + [\text{FullCap}]_{\text{cyto}})X_v, \quad (28)$$

$$\frac{d[\text{FullCap}]_{\text{culture}}}{dt} = r_{\text{pack,DNA}}X_v - k_{\text{degrade,cap}}[\text{FullCap}]_{\text{cyto}}X_v. \quad (29)$$

#### 4.14 | Parameter estimation

The cellular system is described by a system of ODEs,

$$\begin{cases} \frac{dx}{dt} = f(x(t), \theta) \\ \tilde{y}(t) = h(x(t), \theta) \end{cases} \quad (30)$$

where  $x(t)$  is the vector of model states,  $\theta$  is the vector of parameter values, and  $\tilde{y}$  is the vector of model outputs. These model outputs are the transformations that relate the internal states and parameters to the measured states.

Parameters were fit in three steps (see below for additional details). In the first step, parameters were estimated using maximum-likelihood estimation in the log space of the parameters (Beck & Arnold, 1977)

$$\hat{\theta} = \arg \min_{\theta} (\mathbf{y} - \tilde{\mathbf{y}}(\theta))^T \mathbf{V}_y^{-1} (\mathbf{y} - \tilde{\mathbf{y}}(\theta)), \quad (31)$$

where  $\hat{\theta}$  is a vector of the estimated parameters,  $\mathbf{y}$  is a vector containing the experimental observations,  $\tilde{\mathbf{y}}(\theta)$  is a vector containing model predictions with parameter set  $\theta$ , and  $\mathbf{V}_y$  is the measurement uncertainty covariance matrix. In the second and third steps, parameters were estimated using *maximum a posteriori* estimation

$$\hat{\theta} = \arg \min_{\theta} (\mathbf{y} - \tilde{\mathbf{y}}(\theta))^T \mathbf{V}_y^{-1} (\mathbf{y} - \tilde{\mathbf{y}}(\theta)) + (\theta - \mu)^T \mathbf{V}_{\mu}^{-1} (\theta - \mu), \quad (32)$$

where  $\mu$  is a vector of the parameter values predicted in the prior step and  $\mathbf{V}_{\mu}$  is a matrix of the covariance of the parameters fit in the prior step. This formulation instills a penalty for adjusting the parameters fit in previous steps. The values for  $\mu$  and  $\mathbf{V}_{\mu}$  are set such that the parameters being fit for the first time in the second or third steps do not instill any cost.

Optimization and simulation were performed using MathWorks (Waltham, MA) MATLAB™ software. The series of ordinary differential equations (ODEs) was solved using the stiff solver ode15s. Optimization was performed using a multistart approach with an interior point algorithm and 500 random initialization points. All parameters were scaled by  $\log(\theta)$  during optimization to improve the speed of searching the large parameter space and to constrain parameter values to positive values.

The sensitivities of the model states with respect to the parameters  $\left(\frac{\partial x}{\partial \theta}\right)$  were calculated using finite differences. The sensitivities can be subsequently transformed to the model output space

$$\frac{\partial \tilde{\mathbf{y}}}{\partial \theta} = \frac{\partial h(x, \theta)}{\partial x} \frac{\partial x}{\partial \theta} + \frac{\partial h(x, \theta)}{\partial \theta},$$

and to the log-transformed parameter space for each estimated parameter  $\theta_i$ ,

TABLE 2 Model species.

Species no.	Notation	Description	Unit
1	$X_v$	Viable cell density	cells/mL culture
2	$X_d$	Dead cell density	cells/mL culture
3	$[\text{Plasmid}]_{\text{cell}}$	Plasmid concentration per live cell	plasmid/cell
4	$[\text{Plasmid}]_{\text{extracellular}}$	Extracellular plasmid concentration	plasmid/mL culture
5	$f_{\text{uptake}}$	Plasmid uptake factor	N/A
6	$[\text{Plasmid, RC}]_{\text{nucleus}}$	RepCap plasmid in the nucleus	plasmid/cell
7	$[\text{Plasmid, Helper}]_{\text{nucleus}}$	Helper plasmid in the nucleus	plasmid/cell
8	$[\text{Plasmid}]_{\text{endosomal}}$	Endosomal plasmid - general	plasmid/cell
9	$[\text{Plasmid}]_{\text{cytosol}}$	Cytosolic plasmid - general	plasmid/cell
10	$[\text{Plasmid}]_{\text{nucleus}}$	Nuclear plasmid - general	plasmid/cell
11	$[\text{Rep}]$	Rep protein concentration in cells	Rep protein/cell
12	$[\text{VP}]$	Viral protein concentration in cells	Viral protein/cell
13	$[\text{EmptyCap}]_{\text{nucleus}}$	Empty capsid concentration in cell nucleus	capsid/cell
14	$[\text{EmptyCap}]_{\text{cyto}}$	Empty capsid concentration in cell cytosol	capsid/cell
15	$[\text{DNA}]_{\text{cell}}$	Viral DNA concentration in cell	DNA/cell
16	$[\text{FullCap}]_{\text{nucleus}}$	Full capsid concentration in cell nucleus	capsid/cell
17	$[\text{FullCap}]_{\text{cyto}}$	Full capsid concentration in cell cytosol	capsid/cell
18	$[\text{Rep}]_{\text{culture}}$	Total Rep protein concentration	Rep protein/mL culture
19	$[\text{DNA}]_{\text{culture}}$	Total viral DNA concentration	DNA/mL culture
20	$[\text{TotalCap}]_{\text{culture}}$	Total capsid concentration inside and outside the cells	capsid/mL culture
21	$[\text{TotalCap}]_{\text{media}}$	Total capsid concentration in the media	capsid/mL culture
22	$[\text{FullCap}]_{\text{culture}}$	Total full capsid concentration inside and outside the cells	capsid/mL culture

$$\frac{\partial \mathbf{x}}{\partial \log(\theta_i)} = \theta_i \frac{\partial \mathbf{x}}{\partial \theta_i} \quad (33)$$

$$CI_{i,100(1-\alpha)\%} = \hat{\theta}_i \pm t_{1-\alpha/2,df} \sqrt{V_{\hat{\theta}_i}} \quad (36)$$

Following each round of parameter fitting, the parameter covariance matrix ( $V_{\theta}$ ) is estimated as (Beck & Arnold, 1977)

$$V_{\theta} \approx \left( \frac{\partial \tilde{\mathbf{y}}^T}{\partial \theta} V_y^{-1} \frac{\partial \tilde{\mathbf{y}}}{\partial \theta} \right)^{-1} \quad (34)$$

and for the parameters fit in the previous round of fitting as

$$V_{\theta} \approx \left( \frac{\partial \tilde{\mathbf{y}}^T}{\partial \theta} V_y^{-1} \frac{\partial \tilde{\mathbf{y}}}{\partial \theta} + V_{\mu}^{-1} \right)^{-1} \quad (35)$$

where  $V_{\mu}$  is the parameter uncertainty prior determined in the previous round of fitting. Confidence intervals are subsequently calculated using the parameter variances

where  $\alpha$  is the significance level,  $t_{1-\alpha/2,df}$  is the associated two-tail Student's  $t$  value with  $df = N_y - N_{\theta}$  degrees of freedom where  $N_y$  is the number of experimental measurements and  $N_{\theta}$  is the number of estimated parameters, and  $V_{\hat{\theta}_i}$  is the diagonal value of the parameter covariance matrix associated with  $\hat{\theta}_i$ .

Parameters were estimated in three steps. In the first step, cell growth and plasmid uptake parameters ( $k_{\text{death}}$ ,  $K_{\text{Inh}}$ ,  $K_d$ ,  $k_{\text{uptake,plasmid}}$ ,  $K_{\text{uptake}}$ , and  $k_{\text{degrade,plasmid}}$ ) were estimated using viable cell density, dead cell density, and cellular plasmid content data. The maximum cellular growth rate ( $\mu_{\text{max}}$ ) was leveraged from negative control cell growth data where no plasmids were added to the culture (data not shown). Next, protein synthesis parameters ( $k_{\text{synthesis,Rep}}$ ,  $K_{\text{RepSyn,Rep}}$ ,  $K_{\text{VPSyn,Rep}}$ ,  $k_{\text{synthesis,VP}}$ ,  $k_{\text{degrade,cap}}$ ,  $k_{\text{secrete,media}}$ ) were estimated using Rep protein titer, total capsid titer, and extracellular capsid titer data. In the final fitting step, DNA replication and packing

parameters ( $k_{\text{replication,DNA}}$ ,  $K_{\text{DNArep,Rep}}$ ,  $K_{\text{DNApack,DNA}}$ ,  $k_{\text{pack,DNA}}$ ,  $k_{\text{secrete,media}}$ ,  $K_{\text{DNArep,pHelper}}$ , and  $k_{\text{degrade,DNA}}$ ) were estimated using total replicated viral DNA, genome titer data, and filled to empty capsid ratio data.

The complete set of parameter values and references are reported in Table 2 (Bishop et al., 2016; Dahari et al., 2007; Grosse et al., 2017; Redemann et al., 1989; Schwake et al., 2010; Varga et al., 2005; Wang et al., 2021; Werle et al., 2021; Yuan & Parrish, 2001).

## AUTHOR CONTRIBUTIONS

**Prasanna Srinivasan:** Conceptualization; methodology; validation; formal analysis; investigation; data curation; writing (original draft); writing (review and editing); visualization. **Christopher T. Canova:** Conceptualization; methodology; software; validation; formal analysis; investigation; data curation; writing (original draft); writing (review and editing); visualization. **Sha Sha:** Conceptualization; methodology; investigation; Data curation; writing (original draft); writing (review and editing); **Tam N. T. Nguyen:** Conceptualization; methodology; software; validation; formal analysis; investigation; data curation; writing (original draft); writing (review and editing); visualization. **John Joseph:** Validation; formal analysis; investigation; writing (review and editing). **Jose Sangerman:** Validation; investigation. **Andrew J. Maloney:** Conceptualization; methodology; writing (original draft); writing (review and editing). **Georgios Katsikis:** Formal analysis; writing (original draft); writing (review and editing). **Rui Wen Ou:** Investigation. **Moo Sun Hong:** Formal analysis; writing (original draft); writing (review and editing). **Jaclyn Ng:** Investigation. **Arella Yuan:** Investigation. **Daniel Antov:** Investigation. **Sally Song:** Investigation. **Wenyu Chen:** Investigation. **Caleb Neufeld:** Writing (original draft). **Jacqueline M. Wolfrum:** Formal analysis; resources; data curation; writing (original draft); writing (review and editing); funding acquisition. **Paul W. Barone:** Conceptualization; methodology; formal analysis; resources; data curation; writing (original draft); writing (review and editing); funding acquisition. **Anthony J. Sinskey:** Resources; data curation; writing (original draft); writing (review and editing); supervision; funding acquisition. **Stacy L. Springs:** formal analysis; resources; data curation; writing (original draft); writing (review and editing); funding acquisition. **Richard D. Braatz:** Conceptualization; methodology; resources; data curation; writing (original draft); writing (review and editing); funding acquisition.

## ACKNOWLEDGMENTS

This work was supported by the U.S. Food and Drug Administration (Grant ID: 1R01FD006584-02, Continuous Viral Vector Manufacturing based on Mechanistic Modeling and Novel Process Analytics). Tam N. T. Nguyen was partially supported by a MathWorks Engineering Fellowship. This research was also supported by a grant from the Massachusetts Life Sciences Center as part of the Building Breakthroughs Program. The authors thank Betsy Skrip of the MIT Center for Biomedical Innovation for creating the artwork for Figure 6.

## CONFLICT OF INTEREST STATEMENT

The authors declare no conflicts of interest.

## DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the corresponding author, Richard D. Braatz, upon reasonable request.

## ORCID

Prasanna Srinivasan  <http://orcid.org/0000-0002-4951-5799>

Moo Sun Hong  <http://orcid.org/0000-0003-2274-5030>

Paul W. Barone  <http://orcid.org/0000-0001-6802-6846>

Anthony J. Sinskey  <http://orcid.org/0000-0001-9433-4324>

Richard D. Braatz  <http://orcid.org/0000-0003-4304-3484>

## REFERENCES

- Beck, J. V., & Arnold, K. J. (1977). *Parameter estimation in engineering and science*. Wiley.
- Bishop, C. J., Majewski, R. L., Guiriba, T.-R. M., Wilson, D. R., Bhise, N. S., Quiñones-Hinojosa, A., & Green, J. J. (2016). Quantification of cellular and nuclear uptake rates of polymeric gene delivery nanoparticles and DNA plasmids via flow cytometry. *Acta Biomaterialia*, 37, 120–130. <https://doi.org/10.1016/j.actbio.2016.03.036>
- Bulcha, J. T., Wang, Y., Ma, H., Tai, P. W. L., & Gao, G. (2021). Viral vector platforms within the gene therapy landscape. *Signal Transduction and Targeted Therapy*, 6, 53. <https://doi.org/10.1038/s41392-021-00487-6>
- Bustin, S. A., Benes, V., Garson, J. A., Hellemans, J., Huggett, J., Kubista, M., Mueller, R., Nolan, T., Pfaffl, M. W., Shipley, G. L., Vandesompele, J., & Wittwer, C. T. (2009). The MIQE guidelines: minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry*, 55(4), 611–622. <https://doi.org/10.1373/clinchem.2008.112797>
- Carpentier, E., Paris, S., Kamen, A., & Durocher, Y. (2007). Limiting factors governing protein expression following polyethylenimine-mediated gene transfer in HEK293-EBNA1 cells. *Journal of Biotechnology*, 128(2), 268–280. <https://doi.org/10.1016/j.jbiotec.2006.10.014>
- Cervera, L., Gutiérrez-Granados, S., Berrow, N. S., Segura, M. M., & Gòdia, F. (2015). Extended gene expression by medium exchange and repeated transient transfection for recombinant protein production enhancement. *Biotechnology and Bioengineering*, 112(5), 934–946. <https://doi.org/10.1002/bit.25503>
- Chiorini, J. A., Wiener, S. M., Yang, L., Smith, R. H., Safer, B., Kilcoin, N. P., Liu, Y., Urcelay, E., & Kotin, R. M. (1998). The roles of AAV Rep proteins in gene expression and targeted integration. In K. I. Berns, & C. Giraud Eds, *Current topics in microbiology and immunology* (218th ed., pp. 25–33). Springer Berlin Heidelberg. [https://doi.org/10.1007/978-3-642-80207-2\\_2](https://doi.org/10.1007/978-3-642-80207-2_2)
- Clark, K. R., Voulgaropoulou, F., Fraley, D. M., & Johnson, P. R. (1995). Cell lines for the production of recombinant adeno-associated virus. *Human Gene Therapy*, 6(10), 1329–1341. <https://doi.org/10.1089/hum.1995.6.10-1329>
- Clément, N., & Grieger, J. C. (2016). Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Molecular Therapy - Methods & Clinical Development*, 3, 16002. <https://doi.org/10.1038/mtm.2016.2>
- Cohen, R. N., van der Aa, M. A. E. M., Macaraeg, N., Lee, A. P., & Szoka, F. C. (2009). Quantification of plasmid DNA copies in the nucleus after lipoplex and polyplex transfection. *Journal of Controlled Release*, 135(2), 166–174. <https://doi.org/10.1016/j.jconrel.2008.12.016>

- Conway, J. E., Zolotukhin, S., Muzyczka, N., Hayward, G. S., & Byrne, B. J. (1997). Recombinant adeno-associated virus type 2 replication and packaging is entirely supported by a herpes simplex virus type 1 amplicon expressing Rep and cap. *Journal of Virology*, 71(11), 8780–8789. <https://doi.org/10.1128/JVI.71.11.8780-8789.1997>
- Dahari, H., Ribeiro, R. M., Rice, C. M., & Perelson, A. S. (2007). Mathematical modeling of subgenomic hepatitis C virus replication in Huh-7 cells. *Journal of Virology*, 81(2), 750–760. <https://doi.org/10.1128/JVI.01304-06>
- Destro, F., Joseph, J., Srinivasan, P., Kanter, J. M., Neufeld, C., Wolfrum, J. M., Barone, P. W., Springs, S. L., Sinskey, A. J., Cecchini, S., Kotin, R. M., & Braatz, R. D. (2023). Mechanistic modeling explains the production dynamics of recombinant adeno-associated virus with the baculovirus expression vector system. *Molecular Therapy. Methods & Clinical Development*, 30, 122–146. <https://doi.org/10.1016/j.omtm.2023.05.019>
- Dubielzig, R., King, J. A., Weger, S., Kern, A., & Kleinschmidt, A. (1999). Adeno-associated virus type 2 protein interactions: Formation of pre-encapsidation complexes. *Journal of Virology*, 73(11), 8989–8998. <https://doi.org/10.1128/JVI.73.11.8989-8998.1999>
- FDA. (2023, June 30). *Approved cellular and gene therapy products*. <https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/approved-cellular-and-gene-therapy-products>
- Fu, Q., Lee, Y. S., Green, E. A., Wang, Y., Park, S. Y., Polanco, A., Lee, K. H., Betenbaugh, M., McNally, D., & Yoon, S. (2023). Design space determination to optimize DNA complexation and full capsid formation in transient rAAV manufacturing. *Biotechnology and Bioengineering*, 120(11), 3148–3162. <https://doi.org/10.1002/bit.28508>
- Fuenmayor, J., Cervera, L., Gòdia, F., & Kamen, A. (2019). Extended gene expression for Gag VLP production achieved at bioreactor scale. *Journal of Chemical Technology & Biotechnology*, 94(1), 302–308. <https://doi.org/10.1002/jctb.5777>
- Gao, K., Li, M., Zhong, L., Su, Q., Li, J., Li, S., He, R., Zhang, Y., Hendricks, G., Wang, J., & Gao, G. (2014). Empty virions in AAV8 vector preparations reduce transduction efficiency and may cause total viral particle dose-limiting side effects. *Molecular Therapy—Methods & Clinical Development*, 1, 9. <https://doi.org/10.1038/mtm.2013.9>
- Gimpel, A. L., Katsikis, G., Sha, S., Maloney, A. J., Hong, M. S., Nguyen, T. N. T., Wolfrum, J., Springs, S. L., Sinskey, A. J., Manalis, S. R., Barone, P. W., & Braatz, R. D. (2021). Analytical methods for process and product characterization of recombinant adeno-associated virus-based gene therapies. *Molecular Therapy—Methods & Clinical Development*, 20, 740–754. <https://doi.org/10.1016/j.omtm.2021.02.010>
- Grimm, D., Kern, A., Rittner, K., & Kleinschmidt, J. A. (1998). Novel tools for production and purification of recombinant adeno-associated virus vectors. *Human Gene Therapy*, 9(18), 2745–2760. <https://doi.org/10.1089/hum.1998.9.18-2745>
- Grosse, S., Penaud-Budloo, M., Herrmann, A.-K., Börner, K., Fakhiri, J., Laketa, V., Krämer, C., Wiedtke, E., Gunkel, M., Ménard, L., Ayuso, E., & Grimm, D. (2017). Relevance of assembly-activating protein for adeno-associated virus vector production and capsid protein stability in mammalian and insect cells. *Journal of Virology*, 91(20), e01198–17. <https://doi.org/10.1128/jvi.01198-17>
- Hansen, H. A., & Emborg, C. (1994). Influence of ammonium on growth, metabolism, and productivity of a continuous suspension Chinese hamster ovary cell culture. *Biotechnology Progress*, 10(1), 121–124. <https://doi.org/10.1021/bp00025a014>
- Hu, Y., Zhu, Y., Sutherland, N. D., Wilson, D. R., Pang, M., Liu, E., Staub, J. R., Berlinicke, C. A., Zack, D. J., Green, J. J., Reddy, S. K., & Mao, H.-Q. (2021). Size-controlled and shelf-stable DNA particles for production of lentiviral vectors. *Nano Letters*, 21(13), 5697–5705. <https://doi.org/10.1021/acs.nanolett.1c01421>
- Jain, N. K., Ogden, P. J., & Church, G. M. (2024). Comprehensive mutagenesis maps the effect of all single codon mutations in the AAV2 rep gene on AAV production. *eLife*, 12, RP87730. <https://doi.org/10.7554/eLife.87730.3>
- King, J. A., Dubielzig, R., Grimm, D., & Kleinschmidt, J. A. (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *The EMBO Journal*, 20(12), 3282–3291. <https://doi.org/10.1093/emboj/20.12.3282>
- Kotin, R. M. (2011). Large-scale recombinant adeno-associated virus production. *Human Molecular Genetics*, 20(R1), R2–R6. <https://doi.org/10.1093/hmg/ddr141>
- Li, C., & Samulski, R. J. (2020). Engineering adeno-associated virus vectors for gene therapy. *Nature Reviews Genetics*, 21(4), 255–272. <https://doi.org/10.1038/s41576-019-0205-4>
- Li, J., Samulski, R. J., & Xiao, X. (1997). Role for highly regulated rep gene expression in adeno-associated virus vector production. *Journal of Virology*, 71(7), 5236–5243. <https://doi.org/10.1128/jvi.71.7.5236-5243.1997>
- Liu, X., Voulgaropoulou, F., Chen, R., Johnson, P. R., & Clark, K. R. (2000). Selective Rep-Cap gene amplification as a mechanism for high-titer recombinant AAV production from stable cell lines. *Molecular Therapy*, 2(4), 394–403. <https://doi.org/10.1006/mthe.2000.0132>
- Lu, M., Lin, Y. C., Kuo, H. J., Cai, W., Ye, Q., Zhao, L., & Hu, W. S. (2024). Tuning capsid formation dynamics in recombinant adeno-associated virus producing synthetic cell lines to enhance full particle productivity. *Biotechnology Journal*, 19(3), 2400051. <https://doi.org/10.1002/biot.202400051>
- Matsushita, T., Elliger, S., Elliger, C., Podsakoff, G., Villarreal, L., Kurtzman, G., Iwaki, Y., & Colosi, P. (1998). Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Therapy*, 5(7), 938–945. <https://doi.org/10.1038/SJ.GT.3300680>
- Maurer, A. C., & Weitzman, M. D. (2020). Adeno-associated virus genome interactions important for vector production and transduction. *Human Gene Therapy*, 31(9–10), 499–511. <https://doi.org/10.1089/hum.2020.069>
- Nguyen, T. N. T., Sha, S., Hong, M. S., Maloney, A. J., Barone, P. W., Neufeld, C., Wolfrum, J., Springs, S. L., Sinskey, A. J., & Braatz, R. D. (2021). Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. *Molecular Therapy—Methods & Clinical Development*, 21, 642–655. <https://doi.org/10.1016/J.OMTM.2021.04.006/ATTACHMENT/OCA5EF7B-74E0-4F1C-979E-2F428A1FA5BA/MMC1.PDF>
- Ohba, K., Sehara, Y., Enoki, T., Mineno, J., Ozawa, K., & Mizukami, H. (2023). Adeno-associated virus vector system controlling capsid expression improves viral quantity and quality. *iScience*, 26(4), 106487. <https://doi.org/10.1016/j.isci.2023.106487>
- Pereira, D. J., McCarty, D. M., & Muzyczka, N. (1997). The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. *Journal of Virology*, 71(2), 1079–1088. <https://doi.org/10.1128/jvi.71.2.1079-1088.1997>
- Prasad, K.-M. R., & Trempe, J. P. (1995). The Adeno-Associated virus Rep78 protein is covalently linked to viral DNA in a preformed virion. *Virology*, 214(2), 360–370. <https://doi.org/10.1006/viro.1995.0045>
- Rajendra, Y., Kiseljak, D., Baldi, L., Hacker, D. L., & Wurm, F. M. (2011). Influence of glutamine on transient and stable recombinant protein production in CHO and HEK-293 cells. *BMC Proceedings*, 5(8), P35. <https://doi.org/10.1186/1753-6561-5-8-P35>
- Rajendra, Y., Kiseljak, D., Baldi, L., Wurm, F. M., & Hacker, D. L. (2015). Transcriptional and post-transcriptional limitations of high-yielding, PEI-mediated transient transfection with CHO and HEK-293E cells. *Biotechnology Progress*, 31(2), 541–549. <https://doi.org/10.1002/btpr.2064>

- Redemann, B. E., Mendelson, E., & Carter, B. J. (1989). Adeno-associated virus rep protein synthesis during productive infection. *Journal of Virology*, 63(2), 873–882. <https://doi.org/10.1128/jvi.63.2.873-882.1989>
- Samulski, R. J., & Muzyczka, N. (2014). AAV-mediated gene therapy for research and therapeutic purposes. *Annual Review of Virology*, 1(1), 427–451. <https://doi.org/10.1146/annurev-virology-031413-085355>
- Schmidt, M., Afione, S., & Kotin, R. M. (2000). Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53. *Journal of Virology*, 74(20), 9441–9450. <https://doi.org/10.1128/jvi.74.20.9441-9450.2000>
- Schwake, G., Youssef, S., Kuhr, J.-T., Gude, S., David, M. P., Mendoza, E., Frey, E., & Rädler, J. O. (2010). Predictive modeling of non-viral gene transfer. *Biotechnology and Bioengineering*, 105(4), 805–813. <https://doi.org/10.1002/bit.22604>
- Sha, S., Maloney, A. J., Katsikis, G., Nguyen, T. N. T., Neufeld, C., Wolfrum, J., Barone, P. W., Springs, S. L., Manalis, S. R., Sinskey, A. J., & Braatz, R. D. (2021). Cellular pathways of recombinant adeno-associated virus production for gene therapy. *Biotechnology Advances*, 49, 107764. <https://doi.org/10.1016/J.BIOTECHADV.2021.107764>
- Shimizu, N. (2005). Tracking of microinjected DNA in live cells reveals the intracellular behavior and elimination of extrachromosomal genetic material. *Nucleic Acids Research*, 33(19), 6296–6307. <https://doi.org/10.1093/nar/gki946>
- Sommer, J. üM., Smith, P. H., Parthasarathy, S., Isaacs, J., Vijay, S., Kieran, J., Powell, S. K., McClelland, A., & Wright, J. F. (2003). Quantification of adeno-associated virus particles and empty capsids by optical density measurement. *Molecular Therapy*, 7(1), 122–128. [https://doi.org/10.1016/S1525-0016\(02\)00019-9](https://doi.org/10.1016/S1525-0016(02)00019-9)
- Trempe, J. P., & Carter, B. J. (1988). Regulation of adeno-associated virus gene expression in 293 cells: Control of mRNA abundance and translation. *Journal of Virology*, 62(1), 68–74. <https://doi.org/10.1128/jvi.62.1.68-74.1988>
- Varga, C. M., Tedford, N. C., Thomas, M., Klibanov, A. M., Griffith, L. G., & Lauffenburger, D. A. (2005). Quantitative comparison of polyethylenimine formulations and adenoviral vectors in terms of intracellular gene delivery processes. *Gene Therapy*, 12(13), 1023–1032. <https://doi.org/10.1038/sj.gt.3302495>
- Wang, L., Chang, C. C., Sylvers, J., & Yuan, F. (2021). A statistical framework for determination of minimal plasmid copy number required for transgene expression in mammalian cells. *Bioelectrochemistry*, 138, 107731. <https://doi.org/10.1016/j.bioelechem.2020.107731>
- Werle, A. K., Powers, T. W., Zobel, J. F., Wappelhorst, C. N., Jarrold, M. F., Lyktey, N. A., Sloan, C., Wolf, A. J., Adams-Hall, S., Baldus, P., & Runnels, H. A. (2021). Comparison of analytical techniques to quantitate the capsid content of adeno-associated viral vectors. *Molecular Therapy. Methods & Clinical Development*, 23, 254–262. <https://doi.org/10.1016/j.omtm.2021.08.009>
- Wistuba, A., Weger, S., Kern, A., & Kleinschmidt, J. A. (1995). Intermediates of adeno-associated virus type 2 assembly: Identification of soluble complexes containing Rep and cap proteins. *Journal of Virology*, 69(9), 5311–5319. <https://doi.org/10.1128/JVI.69.9.5311-5319.1995>
- Wright, J. (2014). Product-related impurities in clinical-grade recombinant AAV vectors: Characterization and risk assessment. *Biomedicine*, 2(1), 80–97. <https://doi.org/10.3390/biomedicine2010080>
- Xiao, X., Li, J., & Samulski, R. J. (1998). Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *Journal of Virology*, 72(3), 2224–2232. <https://doi.org/10.1128/JVI.72.3.2224-2232.1998>
- Yuan, W., & Parrish, C. R. (2001). Canine parvovirus capsid assembly and differences in mammalian and insect cells. *Virology*, 279(2), 546–557. <https://doi.org/10.1006/viro.2000.0734>
- Zhang, S.-M., Roberts, M., Jones, M., Zeitler, J., Kilby, G., Liu, A., & White, J. R. (2020). A novel method for removing polyethyleneimine from biopharmaceutical samples: Improving assay sensitivity of residual DNA qPCR. *Biotechniques*, 68(6), 353–358. <https://doi.org/10.2144/btn-2020-0011>

## SUPPORTING INFORMATION

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

**How to cite this article:** Srinivasan, P., Canova, C. T., Sha, S., Nguyen, T. N. T., Joseph, J., Sangerman, J., Maloney, A. J., Katsikis, G., Ou, R. W., Hong, M. S., Ng, J., Yuan, A., Antov, D., Song, S., Chen, W., Neufeld, C., Wolfrum, J. M., Barone, P. W., Sinskey, A. J., ... Braatz, R. D. (2024). Multidose transient transfection of human embryonic kidney 293 cells modulates recombinant adeno-associated virus2/5 Rep protein expression and influences the enrichment fraction of filled capsids. *Biotechnology and Bioengineering*, 1–21. <https://doi.org/10.1002/bit.28828>