Molecular Therapy
Methods & Clinical Development
Original Article

Mechanistic modeling explains the production dynamics of recombinant adeno-associated virus with the baculovirus expression vector system

Francesco Destro,[1] John Joseph,[2] Prasanna Srinivasan,[2] Joshua M. Kanter,[3] Caleb Neufeld,[2] Jacqueline M. Wolfrum,[2] Paul W. Barone,[2] Stacy L. Springs,[2] Anthony J. Sinskey,[2,4] Sylvain Cecchini,[3] Robert M. Kotin,[3,5] and Richard D. Braatz[1,2]

[1]Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; [2]Center for Biomedical Innovation, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; [3]Gene Therapy Center, University of Massachusetts Chan Medical School, Worcester, MA 01655, USA; [4]Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139, USA; [5]Carbon Biosciences, Waltham, MA 02451, USA

Current manufacturing processes for recombinant adeno-associated viruses (rAAVs) have less-than-desired yields and produce significant amounts of empty capsids. The increasing demand and the high cost of goods for rAAV-based gene therapies motivate development of more efficient manufacturing processes. Recently, the US Food and Drug Administration (FDA) approved the first rAAV-based gene therapy product manufactured in the baculovirus expression vector system (BEVS), a technology that demonstrated production of high titers of full capsids. This work presents a first mechanistic model describing the key extracellular and intracellular phenomena occurring during baculovirus infection and rAAV maturation in the BEVS. The model predictions are successfully validated for in-house and literature experimental measurements of the vector genome and of structural and non-structural proteins collected during rAAV manufacturing in the BEVS with the TwoBac and ThreeBac constructs. A model-based analysis of the process is carried out to identify the bottlenecks that limit full capsid formation. Vector genome amplification is found to be the limiting step for rAAV production in Sf9 cells using either the TwoBac or ThreeBac system. In turn, vector genome amplification is hindered by limiting Rep78 levels. Transgene and non-essential baculovirus protein expression in the insect cell during rAAV manufacturing also negatively influences the rAAV production yields.

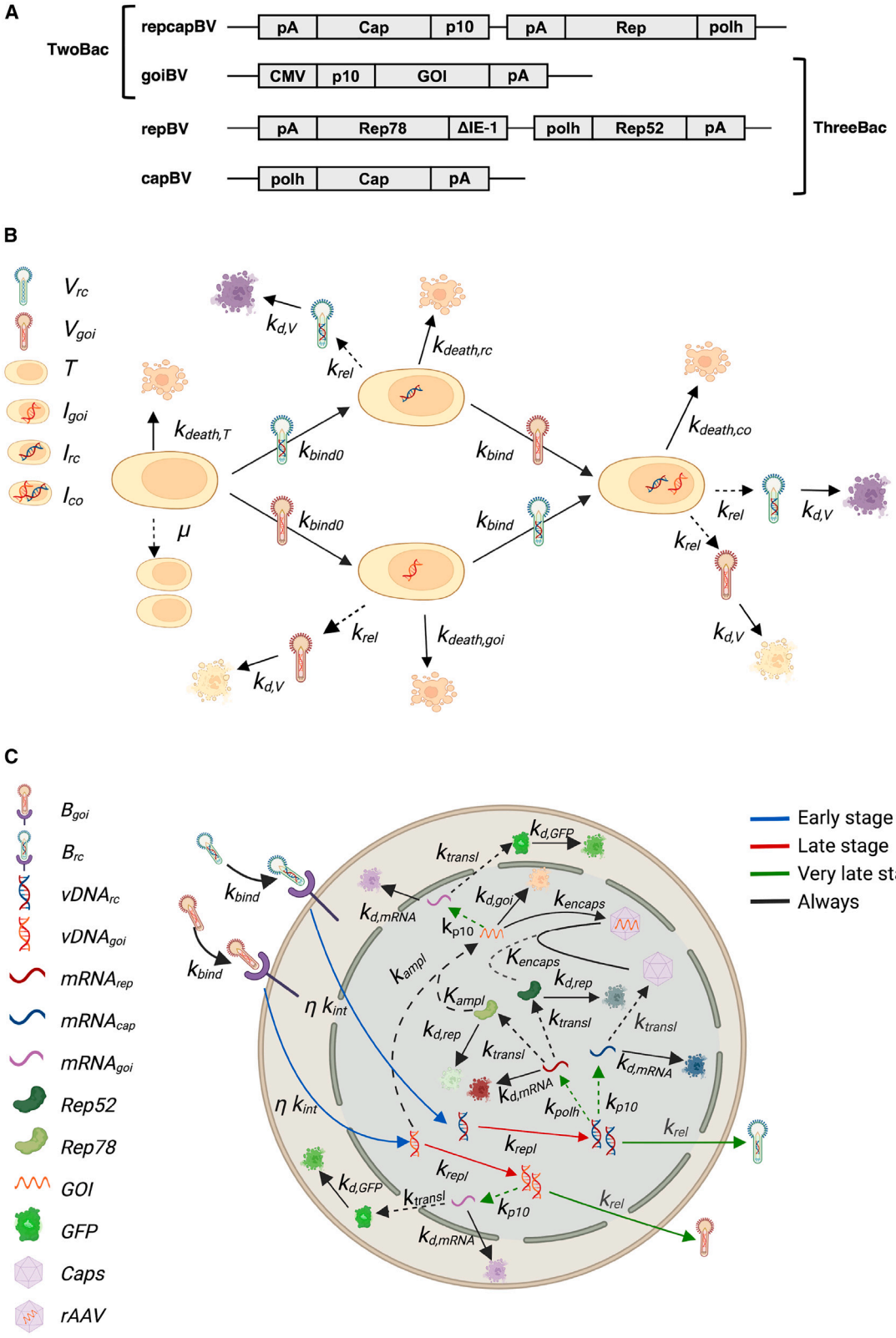
INTRODUCTION

Recombinant adeno-associated virus (rAAV) is an important vector for *in vivo* and *ex vivo* gene therapy. The first gene therapy treatments approved in the US[1] and in the European Union[2] rely on rAAV vectors, and currently more than 200 clinical trials worldwide involve rAAV-based therapies.[3] Considering that up to 30 new rAAV therapies are expected to be launched by 2025,[4] the demand for rAAVs is estimated to soon exceed the production capacity, potentially limiting access to these therapeutics for pre-clinical and clinical trials and for treating patients.[5] At the same time, there is a substantial need to reduce the cost of goods for manufacturing rAAV-based gene therapies, often exceeding $1 million per dose.[6] These issues motivate the development of more efficient processes for rAAV manufacturing.

State-of-the-art technology for rAAV manufacturing is based on either mammalian or insect cell lines, into which the genes for AAV production are introduced through transfection or infection processes. The most productive processes for mammalian cells are based on suspension cultures, with AAV genes introduced using either transient transfection with plasmids or infection with recombinant human herpes simplex virus type 1.[7,8] These systems produce large amounts of empty rAAV capsids (up to 80%–95% of the total),[9,10] which are associated with a low rAAV volumetric titer and extra cost and processes for enriching the filled particles before administering the vector to the patient. In contrast, rAAV manufacturing through recombinant baculovirus (BV) infection of insect cell lines can reliably achieve filled-to-empty capsid ratios ranging from 50%–80%.[11,12] The cell line derived from *Spodoptera frugiperda* (Sf9) is conventionally used with the BV expression vector system (BEVS).[13,14] Compared with mammalian cell-based processes, insect cells have multiple advantages. Production of recombinant proteins with the BEVS in Sf9 suspension cultures is well established, including for vaccine manufacturing.[15] In addition, insect cells are intrinsically more resistant to contamination from human pathogens than mammalian cell lines,[16] and, in insect cells, many mammalian promoters are inactive or attenuated; hence, detrimental effects caused by transgene expression during rAAV production are averted. BV infection propagates through virus budding in the same cells that produce rAAVs, resulting in near-quantitative infection of the cell culture. In contrast, transient transfection does not result in cell-to-cell spread of plasmid; hence, transfection efficiency is a major

Received 4 February 2023; accepted 30 May 2023;
https://doi.org/10.1016/j.omtm.2023.05.019.

Correspondence: Richard D. Braatz, Department of Chemical Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA.
E-mail: braatz@mit.edu



(legend on next page)

determinant of productivity. In addition, transient transfections require 1 mg per liter of plasmid DNA (pDNA) or more, which represents a substantial financial component of the cost of goods for large-scale, clinical-grade production. Multiple clinical rAAV vectors for gene therapies are already produced in Sf9 cells with the BEVS,[2] and the US Food and Drug Administration (FDA) recently approved, for the first time, a gene therapy product manufactured with the BEVS (HEMGENIX).[17,18] Different versions of the BEVS have been developed to produce rAAV with Sf9 cells through infection from recombinant *Autographa californica* multiple nuclear polyhedrosis viruses (AcMNPV): ThreeBac,[19] TwoBac,[20] OneBac,[21] and MonoBac.[22] In each version of the BEVS, the genetic elements for producing rAAVs (*Rep*, *Cap*, and the vector genome template)[8] are split among different recombinant BVs or integrated into the host genome. The first study demonstrating production of rAAVs in insect cells introduced the ThreeBac system,[19] involving coinfection of Sf9 cells with three viruses (Figure S1). The three recombinant BVs (repBV, capBV, and goiBV) carry the genetic information necessary for rAAV production[8] in four cassettes (Figure 1A): *Rep52* and *Rep78* in repBV, *Cap* in capBV, and the inverted terminal repeat/gene of interest (ITR/GOI) cassette in goiBV.[10] The ThreeBac system, as originally described, demonstrated poor passage stability because of inverted repeats of homologous regions of the *Rep52* and *Rep78* genes in repBV. Since then, different improved BV constructs for rAAV manufacturing have been proposed.[23,24] The TwoBac system[20] was the first set of constructs to overcome the passage instability limitation of ThreeBac. In TwoBac, Sf9 cells are coinfected by only two recombinant BVs: repcapBV and goiBV (Figure 1). While goiBV in TwoBac is equivalent to the one used in ThreeBac, repcapBV features both *Rep* and *Cap* open reading frames (ORFs). The *Rep* cassette expresses both *Rep78* and *Rep52* (in nearly equivalent levels) from a single transcript, while the *Cap* ORF, again from a single transcript, expresses AAV structural proteins (VP1, VP2, and VP3) in appropriate stoichiometric ratios to mimic the capsid composition of the wild-type AAV. Recently, rAAV production in insect cells using only one recombinant BV has been demonstrated with the OneBac and MonoBac systems. In the OneBac system,[21] an insect cell line with stably integrated *Rep* and *Cap* cassettes is infected by goiBV. In the MonoBac system,[22] Sf9 cells are instead infected by a recombinant BV carrying all cassettes necessary for rAAV production.

Although rAAV manufacturing with the BEVS is a promising technology, many challenges remain to be addressed. Mechanistic modeling is an invaluable tool to support (bio)pharmaceutical process development[25,26] because it allows a faster understanding of the process dynamics with less time-consuming and expensive experiments compared with traditional process development.[25,26] Recently, mechanistic modeling of rAAV production via transient transfection of mammalian cells identified bottlenecks in the process and was used to propose improvements to the system.[27] Mathematical models for BV infection and production of recombinant proteins and virus-like particles have been proposed in the past.[28–36] To the best of our knowledge, no mathematical model has been developed to describe and improve rAAV production with the BEVS, a process much more complex than standard manufacturing of recombinant proteins.

This work presents the first mechanistic model for BV infection and rAAV production from insect cells. The model is developed for two-wave synchronous infection; namely, it can simulate the process when BV multiplicity of infection (MOI) is large enough to infect all cells of the batch shortly after that infectious BV is added to the system (first wave) or, at most, after that budded BV is released from the first wave of infected cells (second wave). The model considers the intermediate steps of BV binding, BV transport to the nucleus and replication, release of budded BV, transcription and translation of AAV genes, rAAV capsid formation, Rep protein synthesis, vector genome amplification, and vector genome encapsidation. The parameters of the model are estimated from data available in the literature. The model is extensively validated for in-house and literature experimental datasets that report dynamic measurements of intermediates and rAAV titers collected during batch processing for TwoBac and ThreeBac. The model is also validated for literature data collected from stably transfected cell lines. An *in silico* process analysis is carried out, including model simulation, sensitivity analysis, and uncertainty analysis, to investigate the productivity bottlenecks for TwoBac and ThreeBac, the systems for which the model has been more extensively

Figure 1. TwoBac and ThreeBac constructs

(a) TwoBac delivers the genetic information for rAAV production through two BVs: goiBV and repcapBV. goiBV carries the ITR/GOI cassette, while repcapBV contains a *Rep* cassette from which both Rep52 and Rep78 are expressed and a *Cap* cassette from which the structural proteins are expressed. In ThreeBac, the same goiBV as in TwoBac is used, but *Rep* and *Cap* are delivered through two separate BVs: repBV and capBV, respectively. In repBV, Rep proteins are expressed through separate Rep52 and Rep78 cassettes. The promoters in all cassettes in TwoBac and ThreeBac are the strong *A. californica* very late promoters polh and p10, except for the Rep78 cassette in repBV, which contains the *O. pseudotsugata* weak immediate-early promoter ΔIE1 instead. (B) TwoBac: BV infection model. Uninfected cells duplicate with kinetic constant μ and present low death kinetics ($k_{death,T}$). Virions in the medium bind uninfected cells with kinetic constant k_{bind0}. Infected cells are infected by additional virus with decreased binding kinetics (k_{bind}). Coinfection from repcapBV and goiBV is necessary for rAAV production. Infected cells do not duplicate and experience an accelerated death kinetics ($k_{death,I}$). Budded virions are released in the very late infection stage with rate k_{rel}. Virions in the medium slowly degrade ($k_{d,V}$). (C) TwoBac: intracellular reaction-transport network. Receptor-bound BV is transported into the nucleus (k_{int}). Rerouting to lysosomes leads to degradation of several (η) BVs. In the late stage of the infection, the viral DNA replicates (k_{repl}) in the nucleus. Transcription from the promoters used in TwoBac (polh and p10) occurs during the very late infection stage. The two promoters present slightly different transcription kinetics (k_{p10} and k_{polh}). Rep52 and Rep78 are synthetized (k_{transl}) from the *Rep* transcript, while empty rAAV capsids form from the structural proteins translated from the *Cap* transcript. The transgene is expressed through the p10 promoter. Vector genome amplification (k_{ampl}) originates from the ITR/GOI cassette of goiBV. Rep78 and Rep52 play a role in vector genome amplification (K_{ampl}) and encapsidation (k_{encaps}), respectively. Non-negligible degradation is registered for Rep proteins ($k_{d,rep}$), for mRNA ($k_{d,mRNA}$), and for the non-encapsidated vector genome ($k_{d,DNA}$). Dashed lines indicate that reactants are not consumed in the reaction. (B) and (C) were created with BioRender (https://biorender.com/).

validated. Potential strategies are discussed for increasing productivity for TwoBac, currently the most widely used BEVS construct for rAAV production. Finally, a differential analysis is carried out between modeling rAAV production in insect and mammalian cell lines.

RESULTS

A model is developed for batch production of rAAVs in Sf9 cells using the BEVS. Taking as reference the TwoBac process, the formulation of the model is outlined in the "mathematical model formulation" section under materials and methods, which also summarizes the differences across TwoBac, ThreeBac, and stable cell line production modeling. The model features 35 parameters (Table 1), which are either retrieved from the literature or estimated from experimental data, as outlined in the "parameter estimation strategy" section under materials and methods. The TwoBac and ThreeBac constructs are shown in Figure 1A. The steps and species considered by the model for TwoBac and ThreeBac are shown in Figures 1 and S1, respectively. In the remainder of this manuscript, full rAAV capsids are interchangeably referred to as filled capsids or encapsidated vector genomes as opposed to empty capsids.

BV infection dynamics

Binding

A BV binding to insect cells is a fast process at MOIs (MOI between 0.01 and 100 plaque-forming units [PFU] per cell) and cell densities ($\gtrsim 5 \times 10^5$ cells/mL) of practical interest for the BEVS.[29,43] The viral uptake rate of infected cells decreases as the infection cycle progress and is eventually halted completely because of down-regulation of receptors induced by the virus.[29,32] Viral binding inhibition dynamics is important for the TwoBac and ThreeBac systems, where rAAV production requires host cell coinfection with at least, respectively, 2 and 3 different BVs (one per type). Parameter estimation on data reported by Nielsen[32] indicates that the binding kinetic constant for uninfected cells is equal to $k_{bind0} = 6.3 \times 10^{-7} \pm 1.1 \times 10^{-7}$ mL cell^{-1} h^{-1} and that the binding kinetic constant for infected cells (k_{bind}) starts decaying exponentially at the infection time $\tau_{bind} = 1.8 \pm 0.5$ h post infection (hpi), with decay constant $\beta_{bind} = 0.149 \pm 0.029$ h^{-1}. As a result, the BV uptake rate from infected cells decays rapidly with progress of infection, and infected cells do not practically bind any more virus by the time budded BV is released ($\tau_{rel,on} = 18 \pm 2$ hpi).[29,31,44] Hence, budded virions do not bind to the cells that produced them or to synchronously infected cells. The estimated parameters are consistent with fast BV binding, as found previously in the literature.[29] According to the model, for cell concentrations larger than approximately 1×10^6 cell mL^{-1}, all infective virions are internalized by the cells in 4–6 h from process onset (data not shown). As a result, for MOIs as low as 3 PFU/cell per BV type, more than 85% and 90% of cells achieve a productive infection configuration in, respectively, the ThreeBac and TwoBac systems (Table S1).

Transport to the nucleus and replication

Internalization and transport to the nucleus of receptor-bound BVs have been studied by Dee and Shuler,[29] who estimated the kinetic constants for endocytosis, release into the cytosol, and transport to the nucleus of BVs. All of the steps the receptor-bound BV undergoes until nuclear entry can be approximated with negligible error as a single step, with the kinetic constant corresponding to the rate-determining step; namely, the step presenting the lowest kinetic constant: release of the BV from the endosome into the cytosol ($k_{int} = 0.6 \pm 0.12$ h^{-1}).[29] BV degradation in lysosomes during trafficking to the nucleus is accounted for by considering that only $\eta = 50\%$ of the internalized BV reaches the nucleus.[29] The kinetic parameters for BV DNA replication are estimated from data collected in four batch experiments at different cell densities reported in the literature,[39] with the model fit shown in Figure S2. Consistent with the literature on BV infection,[13] BV DNA replication is found to start at $\tau_{repl,on} = 6 \pm 1$ hpi and last until $\tau_{repl,off} = 18 \pm 1$ hpi. The replication kinetic constant (k_{repl}) is estimated to be 0.732 ± 0.081 h^{-1}, corresponding to a BV DNA doubling time of about 1 h.

Validation of BV infection dynamics

The BV infection component of the model is validated by comparing the model prediction of the total intracellular and extracellular BV DNA copies with the experimental measurements during a batch infection reported by Vieira et al.[41] (Figure 2). In the first stage of the infection, BVs pass from the extracellular medium to the intracellular environment, where they are, in part, degraded in lysosomes. As a result, the total intracellular BV DNA level slightly decreases until DNA replication starts at $\tau_{repl,on}$. Soon after replication onset, the nascent BV DNA becomes predominant over the BV DNA that entered the nucleus following cell binding and trafficking. The BV DNA level increases during replication up to five orders of magnitude, increasing from the few units that primarily infected each cell up to 10^4–10^5 BV DNA copies per cell. The stochastic biological variability of $\tau_{repl,off}$ has a larger influence on the BV DNA level at the end of replication than the BV DNA level at replication onset (Figures 2 and S2). Replication stops at the onset of the very late infection stage, during which budding occurs. Only a small fraction (1%–10%) of the replicated BV DNA is released as budded virus,[13,29,39] but it is often sufficient to propagate the infection to (potentially remaining) uninfected cells in the culture medium.

AAV production dynamics

Transcription and translation

The model considers expression through the intermediate steps of transcription and translation of the AAV structural proteins and of the non-structural proteins Rep52 and Rep78, whose genes are carried in the recombinant BVs. If BV promoters are included in the ITR/GOI cassette of the goiBV,[19,20] then the model can account for expression of the transgene as an additional protein that competes with the structural and non-structural AAV proteins for the host macromolecular synthetic machinery. The promoters commonly used in the BEVS for expressing the AAV proteins for rAAV production (Figure 1A) are the *A. californica* polh and p10 promoters[10,19,20] and the IE1 and ΔIE1 promoters derived from a related BV, *Orgyia pseudotsugata* MNPV (OpMNPV).[19] While the polh and p10 promoters are active in the very late infection stage, the IE1 promoter

Table 1. Model parameters

Symbol	Value	95% CI	Unit	Parameter	Source
Cell growth and death dynamics					
$k_{death,cap}$	0	–	h^{-1}	infected cell death kinetics: effect of rAAV capsids	estimated from literature data[37]
$k_{death,DNA}$	2.9×10^{-3}	$(2.5 \times 10^{-3}, 3.3 \times 10^{-3})$	h^{-1}	infected cell death kinetics: effect of BV DNA	estimated from literature data[37]
$k_{death,rep}$	0	–	h^{-1}	infected cell death kinetic constant: effect of Rep78	estimated from literature data[37]
$k_{death,T}$	8×10^{-5}	$(6 \times 10^{-5}, 1 \times 10^{-4})$	h^{-1}	uninfected cell death kinetic constant	from literature[31]
μ	2.8×10^{-2}	$(2.7 \times 10^{-2}, 2.9 \times 10^{-2})$	h^{-1}	uninfected cell growth kinetic constant	from literature[31,38]
τ_{death}	24	(22, 26)	hpi	switch from uninfected to infected death kinetics	estimated from literature data[37]
Binding and transport to the nucleus					
k_{bind0}	6.3×10^{-7}	$(5.2 \times 10^{-7}, 7.4 \times 10^{-7})$	mL $cell^{-1}$ h^{-1}	uninfected cell binding kinetic constant	estimated from literature data[32]
k_{int}	0.60	(0.48, 0.72)	h^{-1}	BV internalization kinetic constant	from literature[29]
β_{bind}	0.149	(0.120, 0.178)	h^{-1}	infected cell binding decay kinetic constant	estimated from literature data[32]
η	0.50	(0.45, 0.55)	–	percentage of BV rerouted to lysosomes	from literature[29]
τ_{bind}	1.8	(1.3, 2.3)	h^{-1}	infected cell binding decay onset	estimated from literature data[32]
BV replication					
k_{repl}	0.732	(0.651, 0.8130)	h^{-1}	BV replication kinetic constant	estimated from literature data[39]
$\tau_{repl,on}$	6	(5, 7)	hpi	BV replication onset time	estimated from literature data[39]
$\tau_{repl,off}$	18	(17, 19)	hpi	end of BV replication	estimated from literature data[39]
BV release					
k_{rel}	9.8	(8.3, 11.3)	PFU $cell^{-1}$ h^{-1}	budded BV release kinetic constant	from literature[31]
$\tau_{rel,on}$	18	(16, 20)	hpi	budded BV release onset time	from literature[31]
$\tau_{rel,off}$	cell death	–	–	end of budded BV release	from literature[31]
Transcription					
$k_{\Delta IE1}$	60.79	(60.41, 61.17)	nt $cell^{-1}$ h^{-1}	ΔIE1 promoter transcription kinetic constant	estimated from literature data[40]
k_{polh}	607.90	(604.10, 611.68)	nt $cell^{-1}$ h^{-1}	polh promoter transcription kinetic constant	estimated from literature data[41]
k_{p10}	518.53	(518.02, 519.04)	nt $cell^{-1}$ h^{-1}	p10 promoter transcription kinetic constant	estimated from literature data[41]
$\tau_{\Delta IE1}$	8	(6, 10)	hpi	ΔIE1 promoter transcription onset	from literature[13]
τ_{polh}	18	(16, 20)	hpi	polh promoter transcription onset	estimated from literature data[41]
τ_{p10}	15	(14, 16)	hpi	p10 promoter transcription onset	estimated from literature data[41]
$\tau_{transcr,off}$	38	(36, 40)	hpi	time of transcription halt for all promoters	estimated from literature data[41]
Translation					
k_{transl}	2.94×10^{9}	$(2.78 \times 10^{9}, 3.11 \times 10^{9})$	nt $cell^{-1}$ h^{-1}	translation kinetic constant	estimated from literature data[37]
K_{transl}	2.74×10^{4}	$(2.61 \times 10^{4}, 2.87 \times 10^{4})$	mRNA $cell^{-1}$	MM constant for translation	estimated from literature data[37]
Vector genome amplification					
k_{ampl}	3.63×10^{7}	$(3.11 \times 10^{7}, 4.25 \times 10^{7})$	nt $cell^{-1}$ h^{-1}	vector genome amplification kinetic constant	estimated from literature data[37]
K_{ampl}	9.39×10^{6}	$(8.52 \times 10^{6}, 1.03 \times 10^{7})$	Rep78 $cell^{-1}$	MM constant for vector genome amplification	estimated from literature data[37]
Vector genome encapsidation					
k_{encaps}	831.09	(750.15, 920.76)	nt LS^{-1} h^{-1}	vector genome encapsidation kinetic constant	Estimated from literature data[37]
$K_{encaps,coeff}$	5	–	Rep52 $cell^{-1}$	coefficient for MM constant for encapsidation	fixed
Degradation					
$k_{d,V}$	7×10^{-3}	$(1 \times 10^{-3}, 1.3 \times 10^{-2})$	h^{-1}	extracellular BV degradation kinetic constant	from literature[31]
$k_{d,DNA}$	0.01	–	h^{-1}	non-encapsidated vector genome degradation kinetic constant	estimated from literature data[42]
$k_{d,GFP}$	3.2×10^{-3}	$(2.8 \times 10^{-3}, 3.6 \times 10^{-3})$	h^{-1}	GFP degradation kinetic constant	estimated from literature data[42]
$k_{d,rep}$	0.03	(0.02, 0.04)	h^{-1}	Rep proteins degradation kinetic constant	estimated from literature data[40,38]
$k_{d,mRNA}$	0.0610	(0.0597, 0.0623)	h^{-1}	mRNA degradation kinetic constant	estimated from literature data[41]

MM, Michaelis-Menten.

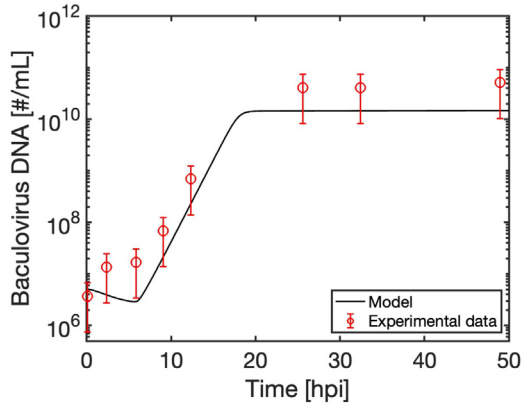


Figure 2. Model validation: BV infection dynamics

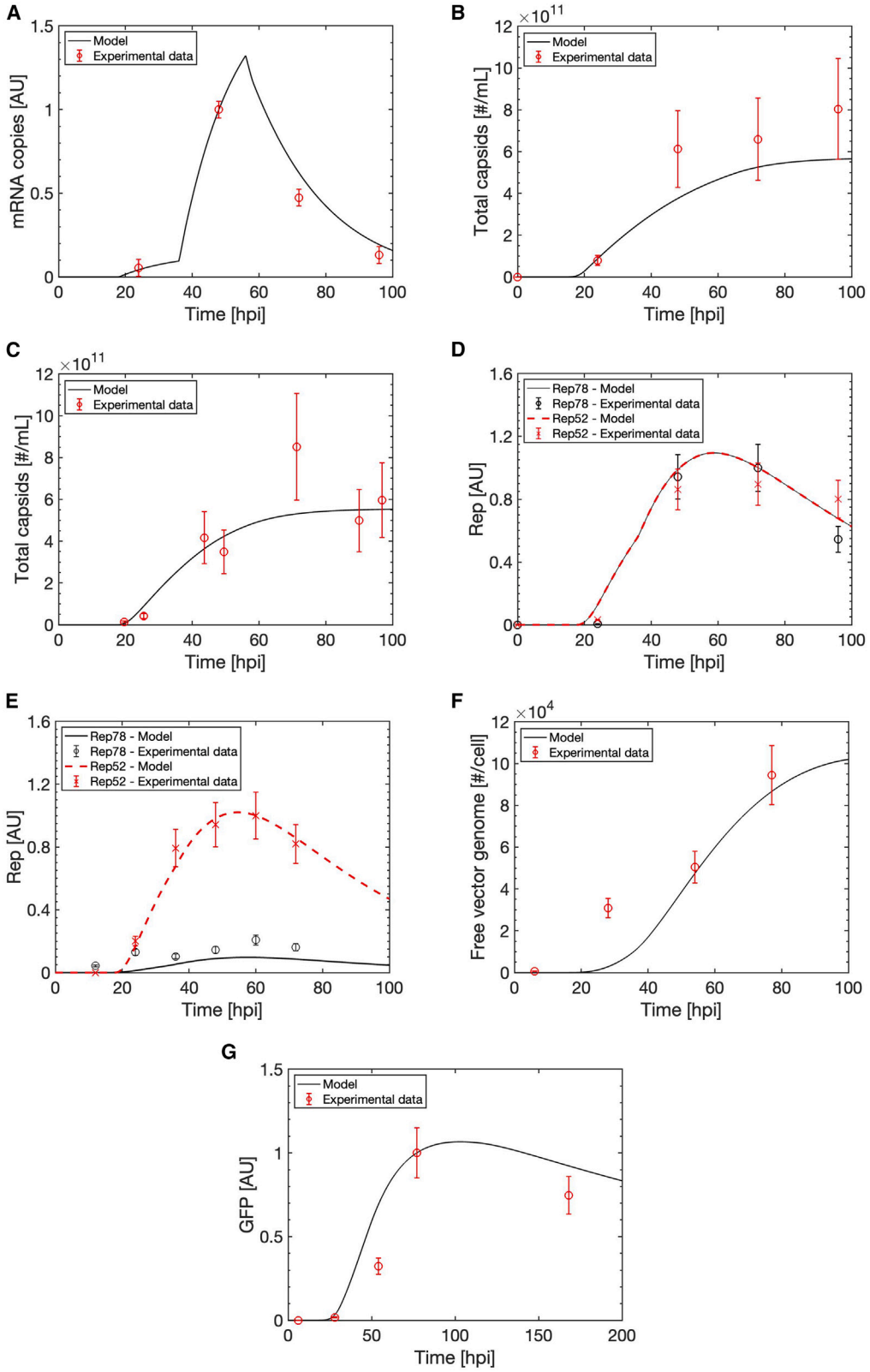
Shown is validation of cumulative extracellular and intracellular BV DNA concentration: model prediction vs. experimental measurements (data are from Figure 2 of Vieira et al.[41]).

and its partially deleted form, ΔIE1, present immediate-early transcription onset, as described in the literature.[13] The transcription kinetics for polh and p10 and the mRNA degradation kinetics are estimated by fitting the model predictions to experimental measurements of mRNA concentration in two experiments on virus-like particle production with the BEVS (Figure S3). Consistent with the literature,[13,20,30,45] the polh promoter (k_{polh} = 607.89 ± 3.79 transcribed nucleotides per hour per cell [nt h^{-1} $cell^{-1}$]) leads to slightly stronger transcription than the p10 promoter (k_{p10} = 518.53 ± 0.51 nt h^{-1} $cell^{-1}$), and transcription onset from polh (τ_{polh} = 18 ± 2 hpi) is delayed a few hours with respect to p10 (τ_{p10} = 15 ± 1 hpi). The polh and p10 promoters remain transcriptionally active until approximately $\tau_{transcr,off}$ = 38 ± 2 hpi, consistent with previous reports.[30] After that transcription halts, the mRNA degrades with a half-life of about 10 h ($k_{d,mRNA}$ = 0.061 ± 0.0013 h^{-1}). For the ΔIE1 promoter, the literature[45] reports that $\tau_{\Delta IE1}$ = 6 ± 2 hpi.[46] Based on an estimate from literature data of the relative expression rate between ΔIE1 and polh,[19,40] the kinetic constant of transcription from ΔIE1 is modeled as $k_{polh} = k_{\Delta IE1}/10$ = 60.79 nt h^{-1} $cell^{-1}$.[19,40] The model predictions are validated for mRNA concentration measurements from a dataset available in the literature[47] (Figure 3A). The data are from a BEVS in which a recombinant protein expression is transcriptionally regulated by the polh promoter. BV infection is carried out at low MOI, generating two infection waves. The model successfully predicts increasing target transcript concentration during the first wave. A much larger transcript concentration increase is registered in correspondence with the polh transcription phase of the second wave, from about 36 (= $\tau_{rel,on} + \tau_{polh}$) to about 56 (= $\tau_{rel,on} + \tau_{transcr,off}$) hours from process onset (considered the time when infectious BVs are added to the system).

Kinetic parameters for the translation rate of the *Rep* and *Cap* transcripts are estimated (simultaneous with vector genome amplification and encapsidation kinetic parameters; see "parameter estimation strategy" under materials and methods) using measurement of capsids and rAAV production from nine batch experiments with the ThreeBac system reported in the study by Aucoin et al.[37] The employed experimental dataset is particularly convenient for this parameter estimation exercise because the 9 batches are carried out at different MOI combinations for the three BVs, providing the model with orthogonal information. The dataset contains the transducing particle concentration, as determined by gene transfer assay, rather than the filled capsid concentration. Before model calibration, the infective particle concentration is converted into filled capsid concentration by assuming a filled-to-infective particle ratio equal to 1,000 ± 500.[20,48] The estimated parameters do not significantly change using a larger filled-to-infective particle ratio equal to 2,000 ± 1,000 (data not shown). The successful model validation for quantitative PCR measurements of filled capsid concentration discussed in the next sections (specifically in section "validation of rAAV production dynamics") supports the conversion ratio between infective and filled particles adopted for the calibration dataset. The estimated cumulative maximum translation rate for AAV gene transcripts is k_{transl} = 2.94 × 10^9 ± 1.67 × 10^8 nt h^{-1} $cell^{-1}$, while the estimated Michaelis-Menten constant for translation is K_{transl} = 2.74 × 10^4 ± 1.33 × 10^4 mRNA $cell^{-1}$. These parameters mean that, according to the model, in each infected cell, the sum of the translation rate of the transcripts of (if present) *Rep*, *Cap*, and the transgene cannot exceed about 2.93 × 10^9 nt h^{-1}. In the absence of competition (namely, in infected cells in which only one of the genes considered by the model is present), the translation rate of the relevant transcript will correspond to half of k_{transl} when the transcript concentration is equal to K_{transl}.

Capsid synthesis

The model provides a good fit of the total capsid measurements at 72 hpi in all experiments of the calibration dataset (Figure 4A), which are carried out at different MOI ratios for the three BVs of the ThreeBac system. The weakest fit is for the batch that starts with MOI = 1 PFU/cell for each BV, which is the only batch in which not all cells are infected during the first wave. For this experiment, slight measurement errors in the MOIs have a large effect on the model prediction. The experimental measurements validate the model prediction that the total capsid production is larger for batches carried out at a larger ratio between the MOI of capBV and the sum of the MOIs of repBV and goiBV (Figure 4A). The model explains this trend based on the competition between *Rep*, *Cap*, and the transgene for expression. For instance, in the four experiments starting with MOI = 9 for capBV, the largest capsid production is achieved when MOI = 1 PFU/cell for repBV and goiBV. Conversely, the lowest capsid production is achieved when MOI = 9 PFU/cell for repBV and goiBV. The capsid production of the two batches in which the sum of the MOIs of repBV and goiBV is 10 (i.e., MOI = 9 PFU/cell for repBV and MOI = 1 PFU/cell for goiBV and vice versa) lies between these two extremes. The model prediction of the capsid concentration is additionally validated by experimental data collected in house for rAAV-5 production with the TwoBac system (Figure 3B) and by an experimental dataset reported in the literature for rAAV-2 production in the ThreeBac system (Figure 3C). The MOI (PFU/cell)



(legend on next page)

used in both experiments is sufficiently large to infect all cells with one wave. The capsid concentration profile predicted by the model is aligned to experimental measurements. No capsids are present in the system until about 20 hpi; shortly after that, the polh promoter of the *Cap* cassette becomes active (τ_{polh} = 18 h). The capsid concentration rises quickly until about 40 hpi. At that point, the capsid synthesis rate decreases because of a halt in transcription ($\tau_{transcr,off}$ = 38 h) and loss of cell viability. Almost no capsids are produced after 72 hpi because of low mRNA levels (Figure 3A) and viability. No significant capsid degradation is registered, validating the model hypothesis.

Rep protein synthesis

The model predictions of Rep52 and Rep78 levels in TwoBac and ThreeBac systems are validated by experimental datasets from the literature (Figures 3D and 3E). The good agreement between model predictions and measurements is particularly interesting for TwoBac because the model parameters for *Rep* expression have been estimated on data collected in the ThreeBac system.[37] This result validates the model hypothesis that the dynamics of transcription from the *Rep52* polh promoter in ThreeBac and from the *Rep* polh promoter in TwoBac are the same and that the translation dynamics are equivalent in the two systems. In the TwoBac process (Figure 3D), Rep52 and Rep78 are present at similar levels.[20] The level of Rep proteins starts rising around 20 h from process onset, shortly after initiation of transcription from the polh promoter of the *Rep* cassette, and peaks at about 60 h from process onset, corresponding with the end of transcription for the second wave of infected cells. After that point, Rep proteins degrade with a half-life of about 20 h ($k_{d,rep}$ = 0.03 h^{-1}, estimated using data from the literature[38,40]). In the ThreeBac process, the Rep52 level follows the same trend as in TwoBac (Figure 3E) because *Rep52* is also expressed through the polh promoter in this case (Figure 1). The level of Rep78 is much lower because its expression is regulated by the weaker ΔIE1 promoter. The earlier onset of transcription from ΔIE1 with respect to polh is not enough to compensate for the difference in strength of the promoters. Actually, a low BV DNA level is registered before 18 hpi (Figure 2), when transcription is active from ΔIE1 but not active from polh.

Vector genome amplification

Rep78 activities are critical for vector genome amplification.[8] Hence, vector genome amplification cannot start until Rep78 has been expressed, and the overall process can be limited by a Rep78 deficiency.[27] To directly estimate the extent of this limitation in the BEVS, vector genome amplification is implemented in our model through Michaelis-Menten kinetics. From the calibration dataset,[37] the maximum vector genome amplification rate (in the absence of Rep78 limitation) is estimated to be k_{ampl} = 3.63 × 10^7 ± 5.72 × 10^6 nt $cell^{-1}$ h^{-1}. The Michaelis-Menten constant is equal to K_{ampl} = 9.38 × 10^6 ± 9.10 × 10^5 Rep78 $cell^{-1}$ (number of Rep78 proteins per cell). Experimental data[42] show that the non-encapsidated vector genome degrades with a half-life of about 70 h ($k_{d,DNA}$ = 0.01 h^{-1}). The vector genome amplification kinetics predicted by the model are validated for experimental measurements from the literature (Figure 3F).[42] Data were collected during repBV infection of insect cells stably transfected with an ITR-flanked GFP. No vector genome was encapsidated during the experiment because *Cap* was not introduced in the host. Even though the model parameters were estimated from ThreeBac experiments, the model accurately predicts the experimentally measured maximum level of vector genome in the stable cell line infection (≈8 × 10^4 copies per cell, achieved about 3 days post infection [dpi]). The (non-encapsidated) vector genome copy number starts growing at about 25 hpi, corresponding to the time when the level of Rep78 starts to increase during repBV infection (Figure 3E). The vector genome copy number peaks at about 80 hpi and then starts declining because of DNA degradation. The ITR-flanked GFP cassette stably transfected in the cell line used in the experiment contains a p10 promoter for transgene expression in insect cells. The measured concentration of GFP during the experiment is well predicted by the model (Figure 3G). When the vector genome is amplified in sufficient numbers (i.e., at about 25 hpi), GFP is detected in the system. Mimicking the vector genome amplification dynamics, the GFP synthesis rate peaks about 3 dpi. Additional validation of the vector genome amplification rate predicted by the model is shown in Figure S4.

Vector genome encapsidation

Vector genome encapsidation is the last step of rAAV particle formation. From the calibration dataset,[45] the encapsidation kinetic constant is estimated to be k_{encaps} = 831 ± 85 nt h^{-1} limiting species for encapsidation $(LS)^{-1}$, where the LS is the species between empty capsids and the non-encapsidated vector genome that has a lower concentration in the cell (Equation 36). The model provides a good fit of the filled capsid production at the end of the batch for the nine ThreeBac experiments from the literature used for model calibration (Figure 4B). The model simulation also resembles the intermediate data collected during the experiments, not used for model calibration (Figure S5), and the corresponding viability profile (Figure S6). The largest rAAV production is achieved in the experiment carried out at the highest MOI (= 9 PFU/cell) for all the three BVs, according to the model and the experimental measurements. All experiments for which an MOI = 9 PFU/cell was used for repBV achieved similar rAAV production levels, independent from the MOIs of capBV and goiBV, highlighting the important role of Rep proteins in production of filled capsids. The model predicts a small

Figure 3. Model validation: intermediates dynamics

(A) mRNA concentration. Data are from Figure 7 (AcNPV-Bgal) of Mitchell-Logean and Murhammer.[47] (B) Total rAAV capsids (TwoBac, rAAV-5). Data are from an in-house experiment. (C) Total rAAV capsids (ThreeBac, rAAV-2). Data are from Figure 5 of Meghrous et al.[40] (D) Rep (TwoBac, rAAV-2). Data are from Figure 2 of Smith et al.[20] (E) Rep (ThreeBac, rAAV-5). Data are from Figure 4 of Urabe et al.[46] (F) Non-encapsidated vector genome. Data are from Figure 3A of Li et al.[42] (G) GFP. Data are from Figure 3B of Li et al.[42]

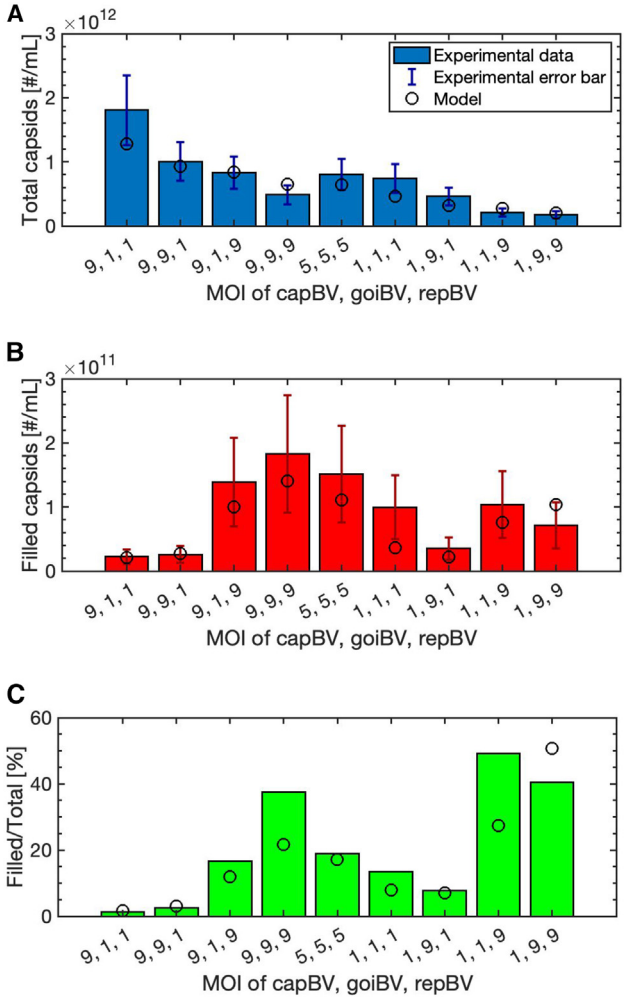


Figure 4. Model fit to the ThreeBac dataset used for model calibration

Shown is the model fit to the ThreeBac dataset (rAAV-2) used for estimation of k_{transl}, K_{transl}, k_{ampl}, K_{ampl}, and k_{encaps}: model predictions (circles) and experimental data (bars). Error bars refer to experimental data. (A) Total rAAV capsids at 72 hpi. Data are from Figure 3 of Aucoin et al.[37] (B) Filled rAAV capsids at 96 hpi. The reported experimental data are obtained by multiplying the infective viral particle titer reported in Figure 2A of Aucoin et al.[45] by a factor of 1,000. (C) Filled-to-empty capsid ratio.

increase in full rAAVs for experiments that have a larger MOI of goiBV (under the same MOI conditions for capBV and repBV). The percentage of filled particles (Figure 4C) increases significantly with the ratio between the MOI of repBV and the MOI of capBV. For instance, the two batches with the lowest ratio between repBV MOI and capBV MOI (= 1/9) have the lowest ratio of filled capsids, according to model and experimental measurements, even though they lead to the largest total capsid production. The two batches with the largest ratio between repBV MOI and capBV MOI (= 2) have the largest ratio of filled to total capsids. The percentage of filled particles also modestly increases with the MOI of goiBV (at fixed MOI for capBV and repBV).

Validation of rAAV production dynamics

Model validation for the BV infection compartment (Figure 2) and for the intermediates of rAAV production (Figure 3) is discussed above. This section describes the successful validation of the overall BV infection and rAAV production model by comparing the model prediction of the filled capsid concentrations with in-house data collected on the TwoBac system and with measurements from additional ThreeBac,[38] TwoBac,[12,20] and OneBac[10] literature datasets. In all considered datasets, the filled capsid concentration is accurately quantified through PCR.

First, consider a study carried out in the ThreeBac system.[38] The model simulation is well aligned with the experimental measurements (Figure 5A). A low total MOI (0.3 PFU/cell, corresponding to 0.1 PFU/cell for each BV), compatible with a two-wave infection, was used in the experiment. At this low MOI, the likelihood of any cell co-infecting a cell with the three different BVs is below 0.1% (Table S1); thus, rAAV produced from the first wave of infected cells was below the limits of detection. The filled capsid concentration starts to grow about 45–50 h from process onset, 25–30 hpi for cells infected in the second wave, with a delay of about 10 h from onset of capsid synthesis (Figures 3B and 3C). Then the filled capsid concentration increases steadily until it stabilizes about 4.5 days from process onset because of loss of viability.

All data discussed so far with respect to model calibration and validation refer to experiments in the ThreeBac system (except for validation of Rep proteins profiles in TwoBac; Figure 3C). Without any change to the estimated parameters, TwoBac implementation of the model predicts the filled capsid concentration measured in an in-house experiment of rAAV-5 production with the TwoBac system (Figure 5B). The model predictions are well aligned with the corresponding total capsid concentration (Figure 3C) and viability measurements (Figure S7). The model is further validated for the filled capsid concentration (Figure 5C) measured in two additional studies from the literature (Kurasawa et al.[12] and Smith et al.[20]). Despite the TwoBac system was used under similar experimental conditions (including initial MOIs [PFU/mL]) in these two studies,[12,20] the filled capsid production per cell achieved by Smith et al.[20] is lower than that obtained by Kurasawa et al.[12] The main difference between the two experimental procedures lies in the goiBV construct. While Smith et al.[20] used the goiBV reported in Figure 1A, the goiBV used by Kurasawa et al.[12] had a 5% shorter ITR-flanked sequence and included only the cytomegalovirus (CMV) promoter (no p10 promoter). Different viability profiles (not fully reported by the authors) or lab-to-lab variability, especially in determining BV titers, might play a role in the production difference between the two studies. The model explains the increase in rAAV production based on the difference between the goiBVs used in the two studies. According to the model, the absence of the strong p10 promoter in the ITR-flanked sequence of goiBV leads, by itself, to about a 20% increase in filled capsid production because of an increase in *Rep* and *Cap* expression as a consequence of reduced competition with transgene expression (as discussed for capsid production; Figure 4A). In

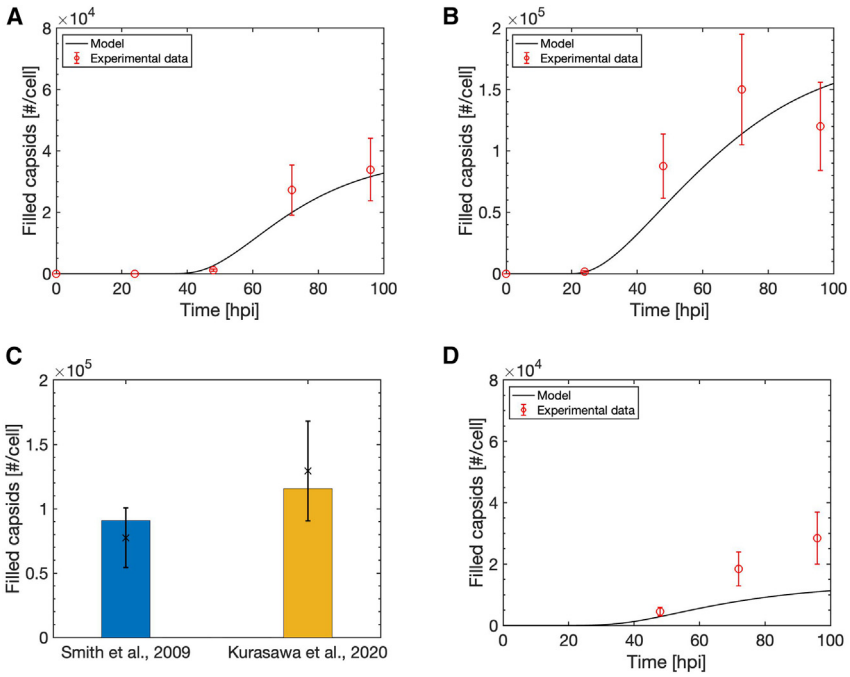


Figure 5. Model validation: rAAV production

(A) ThreeBac (rAAV-2). Data are from Figure 5 of Mena et al.[38] (B) TwoBac (rAAV-5). Data are from an in-house experiment. (C) TwoBac (rAAV-2). Data are from Table 1 of Smith et al.[12,20] and from Figure 3A of Kurasawa et al.[12] Shown are model predictions (bars) and experimental data (markers). (D) OneBac (rAAV-5). Data are from Figure 4A (bioreactor) of Joshi et al.[10] In all plots, rAAV concentration [#/cell] is reported as rAAV concentration in the system [#/mL] normalized by the cell density at inoculation [cell/mL].

addition, the model also estimates that the shorter ITR-flanked sequence can be amplified faster, leading to a corresponding production increase of 5%.

Differently from TwoBac and ThreeBac, the *Rep* and *Cap* cassettes in the OneBac system[10] are stably integrated in the genome of the insect cells, and rAAV is produced upon infection using only one BV that carries the ITR-flanked transgene.[21] Recently, Joshi et al.[10] achieved rAAV-5 titers (2.5–3.5 $\times$ 10^{11} filled capsids per mL) with OneBac higher than demonstrated previously with TwoBac in 1-L-scale bioreactors for several rAAV serotypes.[11,49] Figure 5D compares the prediction of OneBac implementation of the model for measurements of filled capsid concentration for a representative experiment from the study of Joshi et al.[10] Although the total rAAV production is underestimated, the model successfully describes the main rAAV production dynamics and predicts that about 20% of the produced capsids are filled, as found experimentally.[10] This validation, carried out without any re-estimation of the model parameters, is particularly challenging because the *Rep* and *Cap* integrated in the cells of OneBac have many differences with respect to those in the experiments used for model calibration.[50] For instance, the Kozak consensus sequences are different, and *Cap* comes from AAV-5 rather than AAV-2 as in the experiments considered for model calibration. Interestingly, the rAAV production per cell of the considered OneBac experiment is comparable with that of ThreeBac (Figure 5A) and is 3–4 times lower than the production per cell achieved with TwoBac (Figures 5B and 5C). The high cell density ($\approx 1 \times 10^7$ cell/mL) to which cells have been grown in the study by Joshi et al.[10] is the main factor that led to the high titer obtained with OneBac, rather than a greater productivity per cell compared with TwoBac.

Cell growth and death dynamics

Uninfected Sf9 cells have a growth kinetic constant μ=2.8 $\times$ 10^{-2} h^{-1}[31,38] and a low death kinetic constant ($k_{death,T}$) of about 8 $\times$ 10^{-5} h^{-1}.[31] In turn, BV-infected Sf9 cells do not undergo mitosis and have an increased death rate that leads to death of the whole population of synchronously infected cells in 3–5 dpi.[13] It has been suggested that the death rate of infected cells is correlated with the concentration in the host of BV DNA,[30] of Rep proteins (known to cause cytotoxicity in mammalian cells[51]) and possibly of rAAV capsids (because of a potential cytolytic effect). Viability measurements (Figure S6) across the nine batch experiments with the ThreeBac system carried out by Aucoin et al.[37] are used to estimate the correlation of the death rate of infected cells with the model estimations of the levels of BV DNA, Rep78, and total rAAV capsids in the cell. The infected cell viability shows no significant decay until a certain time after infection (τ_{death}), corresponding to 24 ± 2 hpi. After that, the death rate is correlated with the logarithm of the BV DNA concentration, proportionally to $k_{death,DNA}$ = 2.9 $\times$ 10^{-3} ± 4 $\times$ 10^{-4} h^{-1}. Interestingly, the parameter estimation indicates that the death rate is not directly influenced by the concentration of Rep78 ($k_{death,rep}$ = 0 h^{-1}). The dependency of the death rate on the rAAV capsid concentration is also found to be negligible ($k_{death,cap}$ = 0 h^{-1}) because it does not improve the model fit to the cell viability in a statistically significant way, according to the Akaike information criterion.[52] The estimated parameters are generally representative of the death rate dynamics for BV-infected insect cells[13,31] and lead to a good prediction of the viable cell density concentration measured during an in-house experiment (Figure S7). These results indicate that the cell death rate is directly correlated with the BV DNA concentration, while no direct correlation emerges between cell death rate and the concentration of Rep78 and capsids. Hence, the accelerated death rate experienced by infected cells appears to be mainly due to expression of BV proteins, while rAAV production does not diminish cell viability.

In silico analysis of the process

The validated model for BV infection and rAAV production is used to gain insights into the process dynamics. The focus is on TwoBac and ThreeBac, for which the model has been exhaustively validated, also

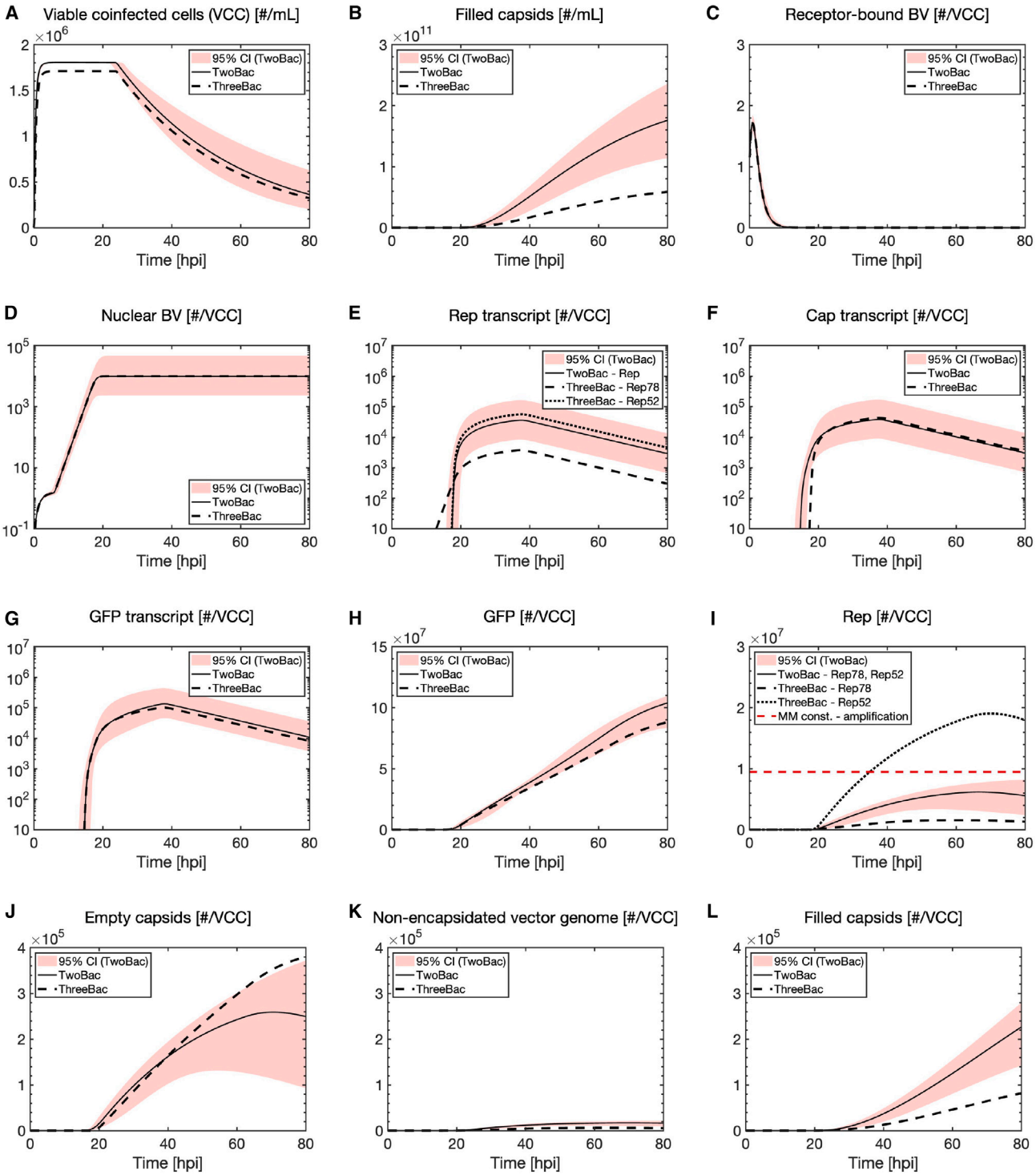


Figure 6. *In silico* analysis of the process: sample plots

Shown is the *in silico* analysis of the process: sample plots for TwoBac and ThreeBac, generated with MOI = 3 PFU/cell for each BV and cell density at infection equal to 2×10^6 cell mL^{-1}. The 95% confidence interval (CI) of the model prediction for TwoBac is also reported. (A and B) Total concentration in the system: (A) viable coinfected cells

(legend continued on next page)

with respect to the intermediates. Sample simulation plots for MOI = 3 PFU/cell for each BV and a cell density of 2×10^6 cells mL^{-1} for TwoBac and ThreeBac are shown in Figure 6. The plots also report the 95% confidence interval on the model predictions for TwoBac, obtained through forward uncertainty propagation by Monte Carlo sampling (see "Sensitivity analysis and uncertainty propagation" under materials and methods). The corresponding rates for the intracellular reactions that lead to rAAV formation in cells presenting a productive coinfection are reported in Figure 7A. A sensitivity analysis of the model parameters, expressing their relative impact on rAAV production for TwoBac, is shown in Figure 7B.

The total concentration in the system of viable cells presenting productive coinfection and of filled rAAV capsids are reported, respectively, in Figures 6A and 6B. Most (>80%) cells achieve productive coinfection for TwoBac and ThreeBac (Figure 6A). A significant viability decrease is registered starting from 24 hpi, and less than 25% are still viable 3 dpi. As expected, the production of filled rAAV capsids is significantly larger for TwoBac than for ThreeBac (Figure 6B).

Figures 6C–6L show the intracellular concentration of the intermediates for rAAV production in viable cells having productive coinfection. The concentration of BV bound to host receptors (Figure 6C) peaks at process onset, caused by fast BV binding, followed by internalization and transport to the nucleus. The copy number of BV DNA in the nucleus grows moderately in the first phase of the process because of internalization of the extracellular BV and exponentially in the replication phase (6–18 hpi; Figures 6D and 7A). From 24 hpi, the cumulative BV DNA copy number in coinfected viable cells decreases because of loss of cell viability, although the copy number in cells that are still viable remains approximately constant.

Transcript concentration profiles (Figures 6E–6G) are closely related to the transcription dynamics of the associated promoters. The concentration of all transcripts starts decaying significantly after the halt of transcription ($\approx$ 40 hpi) because of mRNA degradation and loss of cell viability. *Cap* transcript concentration dynamics in TwoBac and ThreeBac are similar (Figure 6E), with slightly earlier transcription onset in TwoBac (p10 promoter) than in ThreeBac (polh promoter). A single *Rep* transcript is produced in TwoBac, while *Rep52* and *Rep78* are transcribed separately in ThreeBac (Figure 1). Transcription of *Rep78* (ΔIE1 promoter) in ThreeBac is much weaker than *Rep* transcription (polh promoter) in TwoBac (Figures 6E and 7A). The GFP transcript reaches a higher level in TwoBac than in ThreeBac because of the larger amount of template available in TwoBac as a consequence of stronger vector genome amplification.

Protein levels (Figures 6H–6J) start to rise as soon as transcription of the corresponding genes is initiated (Figure 7A). Rep52 is less abundant in TwoBac than in ThreeBac (Figure 6I) because of the lower transcript concentration and the weaker translation caused by the leaky scanning mechanism. On the other hand, the level of Rep78 is much higher in TwoBac than in ThreeBac. The capsid synthesis rate and the concentration of empty capsids across the batch are similar for the two systems (Figures 6J and 7A).

The level of free vector genome available for encapsidation is much lower than the empty capsid levels for TwoBac and ThreeBac (Figures 6J and 6K). By overlapping the rates of capsid synthesis, vector genome amplification, and encapsidation in TwoBac (Figure 7C), the model indicates that the encapsidation rate is controlled by the amplification rate and that both are significantly lower than the capsid synthesis rate. Hence, the non-encapsidated vector genome is an LS for rAAV formation, and vector genome amplification is the limiting step within the reaction-transport network. Comparing the Rep78 level in coinfected cells for the TwoBac system with the estimated Michaelis-Menten constant for amplification, the model indicates that Rep78 is limiting for vector genome amplification, also when the uncertainty of the model prediction is considered (Figure 6I). In TwoBac, encapsidation appears to be limited only by the availability of a non-encapsidated vector genome. Model simulation shows that about 90% of the synthesized vector genome copies are successfully packaged, with most of the remaining genomes not being encapsidated because of loss of cell viability (Figure S8). For ThreeBac, the non-encapsidated vector availability is even more limiting for rAAV production (Figure 6K) because of the lower Rep78 level (Figure 6I) and the consequent lower vector genome amplification and encapsidation rates (Figure 7C) compared with TwoBac.

The sensitivity analysis for TwoBac also indicates that vector genome amplification is the limiting step for rAAV production (Figure 7B). Of all model parameters, the maximum vector genome amplification rate (k_{ampl}) affects rAAV production the most, whereas the Michaelis-Menten constant for amplification (K_{ampl}) is the sixth parameter in the ranking. The time of transcription onset from the polh promoter (τ_{polh}), controlling *Rep* transcription, is also significantly affecting production, together with the parameters related to viability (τ_{death} and $k_{death,DNA}$). The maximum translation rate (k_{transl}) is the fifth parameter in the ranking, while the rest of the parameters that have a non-negligible effect on production are related, in order of decreasing importance, to BV DNA replication ($\tau_{repl,off}$, k_{polh}, and $\tau_{repl,on}$), transcription (k_{polh} and k_{p10}), encapsidation (k_{encaps}), and Rep protein degradation ($k_{d,rep}$).

DISCUSSION

Achieving high productivity in rAAV manufacturing is of utmost importance for successfully delivering gene therapy treatments

and (B) filled rAAV capsids. (C–L) Intracellular concentration in viable coinfected cells: (C) BV bound to the cell surface (for each type of BV of the system), (D) BV DNA in the nucleus (for each type of BV of the system), (E) *Rep* transcripts, (F) *Cap* transcript, (G) transgene transcript, (H) GFP, (I) Rep proteins (the Michaelis-Menten [MM] constant for Rep78 limitation to vector genome amplification is reported), (J) empty capsids, (K) non-encapsidated vector genome, and (L) filled capsids. VCC, viable coinfected cell (i.e., a viable cell that has a coinfection capable of rAAV production).

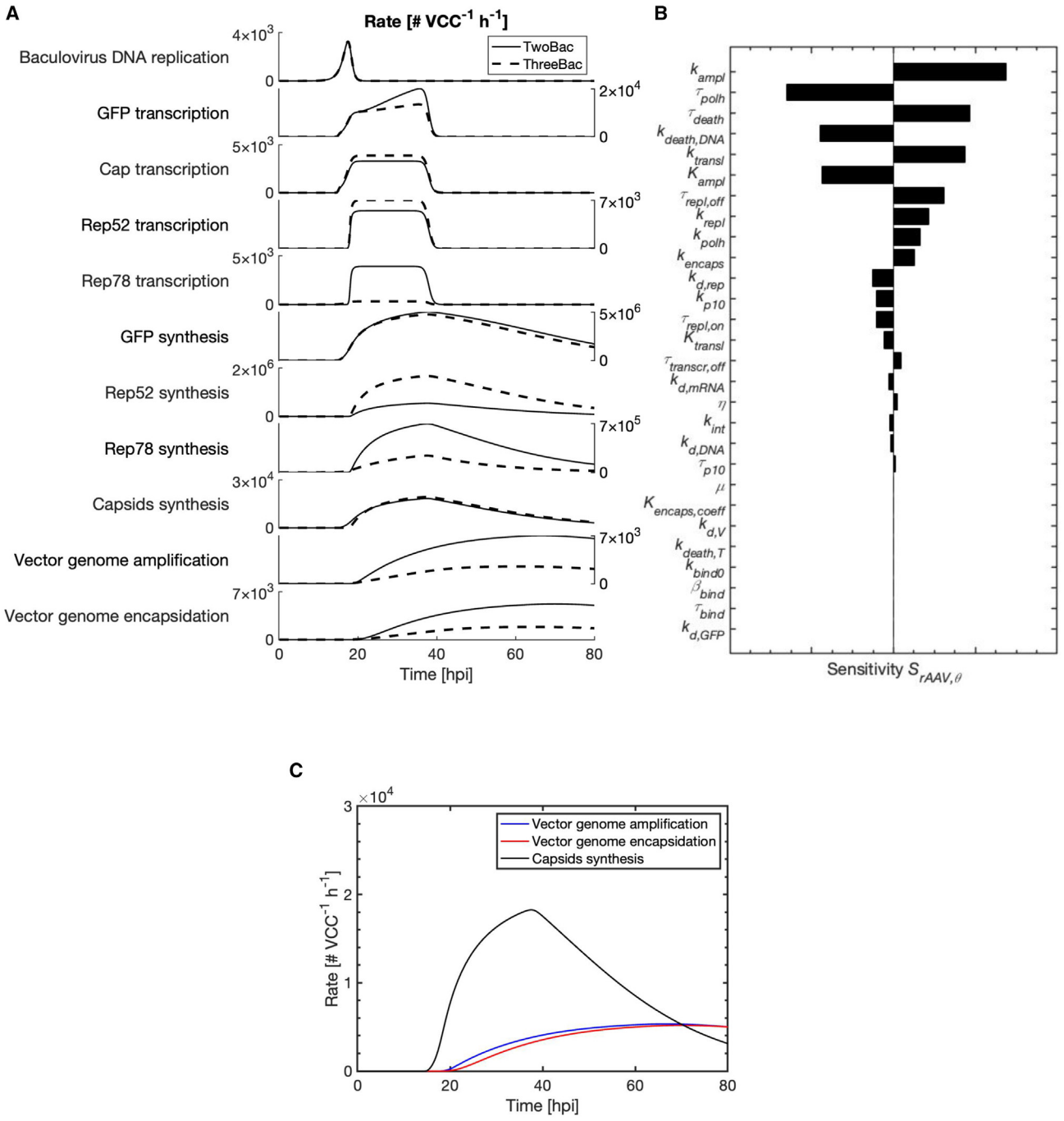


Figure 7. *In silico* analysis of the process: reaction rates and sensitivity analysis

Shown is the *in silico* analysis of the process: reaction rates and sensitivity analysis, generated with MOI = 3 PFU/cell per BV and cell density at infection equal to 2×10^6 cell mL^{-1}. (A) Dynamic trends of reaction rates for TwoBac (continuous line) and ThreeBac (dashed line) in viable cells presenting productive coinfection. (B) Sensitivity of rAAV production to TwoBac model parameters (Equation 52). (C) Comparison among rates of vector genome amplification, vector genome encapsidation, and rAAV (empty) capsid synthesis in TwoBac simulation.

to patients. However, current rAAV manufacturing processes with mammalian and insect cells present severe limitations in terms of productivity. In this work, the first model for rAAV production in the Sf9 cell/BEVS is presented and used to investigate the bottlenecks that limit rAAV production in state-of-the-art processes.

A validated quantitative model for BV infection and rAAV production

The model predictions of rAAV and of intermediate concentrations are in excellent accordance with experimental measurements collected in our lab and with experimental datasets reported in the most significant studies from the literature of rAAV production with the BEVS. Successful model validation is demonstrated for TwoBac (Figures 3B, 3D, 5B, 5C, and S7) and ThreeBac experiments (Figures 3C,3E, 4, 5A, and S4–S6) and, without any additional parameter estimation, for experiments using stable cell lines with integrated ITR/GOI (Figures 3G and 3H) or *Rep* and *Cap* (Figure 5D) cassettes. All parameters pass the 95% confidence t test for parameter precision of Equation 49 (scores not shown), except for $k_{d,V}$, which does not significantly affect rAAV production (Figure 7C). The confidence intervals of the model predictions (Figure 6) correspond to the experimental variability measured in the datasets used for model calibration and validation (Figures 2, 3, 4, 5, and S2–S7) and are generally aligned with the intrinsic biological variability found in the literature.[12,20,23,24,37,38,40] The concentrations of certain intracellular species, such as empty capsids in viable coinfected cells (Figure 6J), have apparently large 95% confidence intervals at an infection age greater than 50–60 hpi. However, few cells remain viable at this infection stage (Figure 6A). Nonetheless, when of interest for specific applications, the confidence interval of the model predictions having higher uncertainty could be reduced by re-fitting the model parameters on larger experimental datasets specifically designed for this purpose.[53]

Regarding the BV compartment, the model shows good agreement with experimentally measured BV DNA concentration during infection (Figures 2 and S2). According to the model, productive coinfection is achieved in the vast majority of cells for MOI per BV as low as 3 PFU/cell in TwoBac and ThreeBac (Table S1; Figure 6A). Experimental measurements and model analysis show that, in the replication phase ($\approx$6–18 hpi), the BV DNA level grows from the few copies that infected the host up to $\approx 10^5$ copies per cell (Figures 2 and 6D). The final DNA level at the end of replication seems to be more dependent on the biological variability of $\tau_{repl,off}$ (Table 1) rather than on infection conditions such as MOI or cell density (Figure S2). Hence, experimental and modeling results indicate that, with MOI per BV as low as 2–3 PFU/cell, rAAV production is not significantly impacted by unsuccessful or unproductive coinfection or low BV level.

The rAAV production compartment of the model accounts for the expression of AAV genes through the intermediate steps of transcription and translation, for the synthesis of rAAV capsids and of Rep52 and Rep78 proteins, and, finally, for vector genome amplification and encapsidation. The model is conceived with a plug-and-play structure, which can simulate gene expression through different gene-promoter arrangements, such as used in TwoBac and ThreeBac (Figure 1A). The model simulation is successfully validated for mRNA concentration measurements collected during BEVS experiments (Figures 3A and S3). The model predictions are also aligned with respect to experimental concentrations of proteins related to rAAV production, including the total rAAV capsid concentration (Figures 3B, 3C, and 4A), Rep proteins (Figures 3D and 3E), and transgene protein (Figure 3G).

Model simulation and experimental results highlight competition between *Cap*, *Rep*, and the transgene for the host translation machinery. Accordingly, the model estimates that stronger transgene expression results in reduced rAAV production, as found previously experimentally (Figures 4 and 5B).[37] Experimental measurements also validate the model prediction for the vector genome amplification (Figures 3F and S4) and encapsidation dynamics, including the predicted rAAV titer achieved in TwoBac and ThreeBac (Figures 5A–5C). The model prediction shows a good alignment with measurements of rAAV titer from one experiment with the OneBac construct reported in the literature as well (Figure 5D).

Limiting steps for rAAV production

Model simulation for TwoBac shows that, in coinfected cells, the concentration of free vector genome available for encapsidation (Figure 6K) is generally much lower than the concentration of empty capsids (Figure 6J). Hence, the non-encapsidated vector genome is the LS for rAAV formation, up to the point where the encapsidation rate is essentially controlled by the amplification rate (Figure 7C). The model indicates that vector genome amplification is, in turn, limited by Rep78 (Figures 6I and 7B). According to the model, the vector genome limitation is stronger in ThreeBac than in TwoBac (Figures 6J, 6K, and 7A) because of weaker *Rep78* expression (Figures 6E and 6I) and consequent weaker vector genome amplification (Figure 7A). Experimentally, rAAV production in ThreeBac (Figure 5A) is typically lower than in TwoBac (Figures 5B and 5C). The literature attributes the greater rAAV production in TwoBac to a larger coinfected cell population, producing rAAVs for a longer period of time prior to cell death.[23] Instead, the model-based analysis indicates that use in TwoBac of the polh promoter for expressing *Rep78*, rather than the weaker ΔIE1 as in ThreeBac (Figure 1A), plays a main role in achievement of larger rAAV productivity. Indeed, *Rep78* expression has been found experimentally to be stronger in TwoBac (Figure 3D) than in ThreeBac (Figure 3E). Additional experiments reported in the literature confirm that stronger *Rep78* expression is correlated with larger vector genome amplification.[37,42]

It has been well established that the AAV non-structural Rep proteins are essential for AAV propagation.[8,54,55] The single-stranded linear DNA virus genome undergoes "rolling hairpin" replication where the palindromic 3′ ITR forms an energetically stable secondary structure that functions as primer for DNA polymerase extension. The intramolecular replicative intermediate consists of the input template strand covalently linked to the nascent strand through the ITR. The AAV p5 Rep proteins (Rep78/68) bind to a repetitive motif (GCTC) in the A stem of the ITR[56,57] and, through ATPase/helicase activity, exposes the terminal resolution site (trs), a short motif (5′AGT*TGG)[58] in a small hairpin at the boundary of the A/D′ region of the ITR. The tyrosine (Y156)[59] within the rolling-circle replication catalytic pocket forms a phosphodiester with the 5′ thymidine

(Rep-*T).[60] The p5 Rep helicase activities unwind the ITR, transferring the ITR-Rep nucleoprotein to the nascent strand that now presents the 3′-thymidine as the primer for DNA polymerase extension. As a result of rolling hairpin replication and terminal resolution, each complementary strands contain input and newly synthesized DNA. While the p5 Rep proteins are primarily involved in replication functions as described above, the p19 Rep proteins (Rep52/40) are required for encapsidation of vector genomes. Mutating the methionine initiation codon to glycine for Rep52/40 did not affect AAV DNA replication or capsid assembly but resulted in a 100-fold reduction of infectious virus.[61] In the insect-cell-based processes considered in this work, only Rep78 and Rep52 proteins were present in the system because the corresponding spliced versions of Rep78 and Rep52, Rep68 and Rep40, respectively, do not form.[19,20]

The model-based finding that transgene expression limits rAAV production because of competition for the host translation machinery is also supported by experimental evidence (Figure 4A). Hence, reduction of transgene expression by removing insect cell promoters from the ITR/GOI cassette in the BEVS can lead to a significant increase in filled rAAV capsids (although a proportional increase of empty capsids is also expected). For rAAV manufacturing in mammalian cells, instead, it is challenging to reduce transgene expression without also losing the therapeutic effect. Following the same rationale, productivity in the BEVS can be enhanced by reducing or suppressing expression of BV genes not useful for production of rAAV or budded BV.

The model indicates that encapsidation and Rep52 deficiency are not limiting rAAV production in TwoBac and ThreeBac. Model simulation for TwoBac, in which amplification is more effective (Figure 7A) and *Rep52* is more strongly expressed (Figure 6I), shows that roughly 90% of the vector genomes are effectively encapsidated, with the remaining 10% of vector genomes not being packaged for loss of viability (Figure S8). The Michaelis-Menten coefficient expressing Rep52 limitation to packaging ($K_{encaps,coeff}$) is also found to have a negligible sensitivity on rAAV production (Figure 7B). Experimentally, repBV infection in a stable cell line with integrated ITR/GOI leads to vector genome amplification (Figure 3F) up to levels similar to filled capsid production in ThreeBac (Figure 5A) when the competition between *Rep* and *Cap* for expression (occurring in ThreeBac but not in the stable cell line experiment) is accounted for. This experimental finding further supports the model-based result that encapsidation and Rep52 are not limiting filled capsid production in TwoBac and ThreeBac.

Regarding OneBac, the model suggests that the low filled capsid production per cell (Figure 5D) is also due to low vector genome amplification in this case. Rep78 limitation to vector genome amplification also emerges for OneBac. The ratio between the integrated *Cap* and *Rep* gene copies per cell in the experiment has been measured to be about 10:1,[10] which, the model indicates, is not optimal for achieving a large filled-to-empty capsid ratio (Figure 4). However, model validation has not been carried out for OneBac as exhaustively as for TwoBac and ThreeBac, and a more detailed analysis would be needed to draw reliable conclusions on production bottlenecks in OneBac.

Maintaining high cell viability for as long as possible during BV infection is crucial for achieving large rAAV titers (Figure 7B). Model fit to viability measurements (Figure S6) indicates that the cell death rate is correlated with the number of BV DNA copies in the cell with a logarithmic dependence. Rep78 concentration does not appear to significantly affect viability, in contrast to what has been found in mammalian cells.[51] The cytolytic effect of rAAV capsids is also found to be negligible in the considered experimental datasets. The result that neither *Rep* nor *Cap* expression contributes significantly to the death rate of infected Sf9 cells is consistent with the literature finding that, following passage amplification, the titers of repBV, capBV, and goiBV are similar.[62]

Model limitations

Although the model accurately reproduces several experimental datasets from the literature, the following limitations must be considered. The model has been constructed and calibrated based on the recombinant BV constructs introduced in Urabe et al.[19] and Smith et al.[20] Model accuracy is not guaranteed for potential constructs that present a substantially different expression profile for the BV proteins that are not explicitly considered by the model. Nonetheless, the model can be rearranged to simulate constructs presenting the same BV backbone of those considered in this work, but a different configuration of foreign genes expressed through the polh, p10, or ΔIE1 promoters. Promising validation results have been obtained for data collected on a OneBac system.[10] A more complete model validation, including additional datasets and measurements of intermediates concentration, is needed to assess whether the model predictions in OneBac are consistently as accurate as proved for TwoBac and ThreeBac. Additionally, the parameters of the model have been estimated for production of rAAV-2. Different serotypes might present different dynamics for the steps of the reaction-transport network that involve the intervention of Rep and Cap proteins from rAAV serotypes other than AAV-2. Although the model was demonstrated to be able to track production dynamics for rAAV-5 in our in-house experiment (Figures 3B and 5B) and in experiments from the literature (Figures 3E and 5D), re-estimation of the parameters associated with vector genome amplification and encapsidation might be necessary to improve the accuracy of model predictions for other serotypes. Substrate and metabolite limitations because of depletion are not considered in the model, and neither are the effects of noxious metabolic products and metabolite pool depletion, under the assumption that the cellular macromolecular synthesis is shut off soon after process onset. Although results show that vector genome encapsidation is not limiting for the considered implementations of the BEVS, encapsidation could become limiting with different arrangements of the BV construct. To simulate these scenarios with the model, a more accurate estimation of packaging kinetics should be conducted, improving the estimation of Rep52 limitation and, potentially, considering ATP limitation to encapsidation.[63] While the model was validated for data coming from different laboratories, lab-to-lab variability is expected to affect the accuracy of the model predictions. Re-fit of model parameters, especially of the viability-related parameters ($k_{death,DNA}$ and τ_{death}), would be needed to produce the most accurate predictions

in a specific setup. Finally, for poorly agitated systems, such as hollow-fiber bioreactors,[64] it might be necessary to implement partial differential equations in the model to describe the spatial gradients of extracellular species (i.e., cells and virions) in the medium.

rAAV production modeling: Insect cells vs. mammalian cells

Triple transient transfection (here referred to as triple transfection) of pDNA into mammalian cells (especially human embryonic kidney 293 [HEK293] cells) is the most common alternative to rAAV manufacturing via the BEVS.[7] The reaction-transport network of processes based on mammalian and insect cells share several steps[27] but also have many differences that should be accounted for during process modeling. First and foremost, insect cell-based processes involve an infection, which halts cell growth and leads to cell death within a few days. In contrast, the cell cycle is not initially disrupted following transient transfection, with only modest loss of viability under suitable transfection conditions (plasmid concentrations, reagents, etc.). This difference is probably the reason why Rep78 cytotoxicity affects the viability of mammalian cells[51] more than of BV -infected insect cells, as estimated in this work. Other differences between infection and transient transfection processes involve DNA replication and infection propagation. BV DNA replicates to high levels upon infection, whereas the pDNA that enters cells during transient transfection does not replicate, potentially providing fewer templates available for transcription. At the same time, BV infection can self-propagate to uninfected cells through budding, while subsequent transient transfection waves can only occur upon introduction of fresh pDNA into the culture medium. The genes in the plasmids used for transient transfection and the genes introduced in recombinant BVs in the BEVS have inherent differences. While *Rep*, *Cap*, and ITR/GOI cassettes are present in the construct in both cases, transient transfection of mammalian cells requires introduction of a helper plasmid, carrying genes of a helper virus that is needed for rAAV production. In the BEVS, instead, the BV itself serves as helper virus, avoiding the need to introduce additional helper genes into the host. Furthermore, only *Rep52* and *Rep78* are expressed during rAAV manufacturing in the BEVS, whereas the spliced proteins Rep40 and Rep68 are also produced in transient transfection. Genes that lead to rAAV production are expressed in the BEVS through the BV promoters, while in transient transfection, the promoters of wild-type AAV are conventionally used. Finally, empty and filled capsids secrete into the cytosol and into the extracellular matrix in mammalian cells,[65] facilitated by specific egress factors.[66] No significant secretion is registered in insect cells, especially for empty capsids,[67] although some vectors are found in the medium with the passing of time as a result of loss of cell viability. Mammalian and insect cell processes have different efficiencies for rAAV production. In transient transfection, the typical filled-to-empty particle ratio is 5%–30%.[8,9] In the BEVS, filled particle ratios as high as 50%–80% have been achieved.[11,20,49] Recently, a comprehensive mechanistic model for rAAV manufacturing via transient transfection of HEK293 cells has been presented in the literature.[27] The model indicated that a temporal lag between capsid synthesis and vector genome amplification plays a main role in the low efficiency of transient transfection in mammalian cells. This lag does not occur in insect cell processes (Figure 7A), potentially explaining the difference in efficiency between the two systems. However, given the many differences between mammalian and insect cells processes, a more thorough differential model-based analysis should be conducted to understand what the most critical steps are that originate the efficiency gap.

Conclusions

The current interest in achieving large production of rAAV viral vectors motivates the construction of a mechanistic model suitable for increasing understanding and to support process optimization. In this work, a first mechanistic model for rAAV production in the BEVS is developed. The model quantitatively summarizes datasets and information from key contributions published in the literature on the BEVS and on rAAV manufacturing in the BEVS. Successful validation is achieved for in-house experimental data and for multiple datasets from the literature of rAAV manufacturing via the BEVS with the TwoBac and ThreeBac constructs. Validation for one literature dataset indicates that the model can also explain the dynamics occurring during rAAV manufacturing with the OneBac construct, although a more exhaustive validation should be carried out to confirm the general validity of this result. The model is applied to analyze the production bottlenecks for TwoBac, currently the most widely used BEVS construct for rAAV production. The analysis indicates that larger rAAV production can be achieved by inducing stronger vector genome amplification. Stronger *Rep* expression, especially for *Rep78*, can increase the rate of vector genome amplification, according to the model estimation. Incidentally, ThreeBac is found to be even more affected by Rep78 limitation to rAAV production than TwoBac. Reducing expression of the transgene and of BV genes non-essential for the process can also lead to increased rAAV production by diminishing the competition for *Rep* and *Cap* expression. On the other hand, the model indicates that vector genome encapsidation and Rep52 limitations do not significantly limit rAAV production in the considered TwoBac and ThreeBac constructs. The model also suggests that productive coinfection, in which binding is not limiting production, is achieved for an MOI as low as 2 or 3 PFU/cell for each type of BV. Maintaining large cell viability for as long as possible during infection is crucial to achieve high rAAV titers. The model-based analysis indicates that the cell death rate is correlated with the number of BV DNA copies in the cell with a logarithmic dependence. Differently from transiently transfected mammalian cells, the viability dynamics of Sf9 cells expressing Rep78 are not different from the viability dynamics of other BV -infected cells that do not express Rep78. In future work, the model can be used to design and test *in silico* different BV constructs with novel promoter-cassette arrangements, which can potentially lead to enhanced rAAV production.

MATERIALS AND METHODS

Mathematical model formulation

Overview

For conciseness, the model formulation is first described in detail with reference to the TwoBac system. The ThreeBac and stable cell lines implementations of the model used for generating the results reported

in this article are discussed in dedicated sections below. The model equations reproduce the BV infection dynamics and the reaction-transport network sketched in Figure 1. The model is implemented in MATLAB (MathWorks, Waltham, MA, USA). In all simulations, the model equations are solved using the ordinary differential equation solver ode45 available in MATLAB. The model is developed for two-wave synchronous infection and considers 9 viable extracellular species (Figure 1B). In addition to the concentrations in the culture of the repcapBV virion (V_{rc}), the goiBV virion (V_{goi}), and the viable uninfected cells (T), the model tracks the concentrations of viable infected and coinfected cells I_j, considering cells infected by repcapBV (first wave, $j = rc1$; second wave, $j = rc2$), by goiBV (first wave, $j = goi1$; second wave, $j = goi2$), and of viable coinfected cells (first wave, $j = co1$; second wave, $j = co2$). The concentration of all extracellular species in the system is assumed to be homogeneous because the characteristic time for BV binding to cells is significantly larger than the characteristic time for macroscopic transport in the systems considered in this work.[29,68] The model accounts for 13 intracellular species X_j or each type j of viable infected and coinfected cells. The intracellular species X_j are (Figure 1C) as follows: receptor-bound repcapBV ($B_{rc,j}$), receptor-bound goiBV ($B_{goi,j}$), repcapBV DNA in the nucleus ($vDNA_{rc,j}$), goiBV DNA in the nucleus ($vDNA_{goi,j}$), Rep transcript ($mRNA_{rep,j}$), Cap transcript ($mRNA_{cap,j}$), transgene transcript ($mRNA_{goi,j}$), Rep52 ($Rep52_j$), Rep78 ($Rep78_j$), non-encapsidated vector genome (GOI_j), transgene protein (GFP_j), rAAV empty capsids ($Caps_j$), and rAAV filled capsids ($rAAV_j$). The transgene protein is denoted here as GFP because it is the most commonly used transgene in experiments for process development. The extension to any other transgene is straightforward. The concentration of all intracellular species is calculated for each infected and coinfected cell type, even though not all infections lead to production of the whole set of intracellular species (e.g., Rep proteins are produced only upon repcapBV infection). This implementation, although not optimal for speed, is designed to simulate more easily the different BV constructs and experimental setups analyzed in this study (TwoBac, ThreeBac, and stably transfected cell lines). Additional equations are set up for tracking the concentration of nonviable cells and of the respective intracellular species content. In these equations, not detailed for sake of conciseness, the only phenomenon that is accounted for is the degradation of intracellular species (with the same kinetics as in viable cells). Denoting with X_j the cumulative concentration in the system of an intracellular species across a given class j (#/mL), the total concentration in the system of a given intracellular species X is calculated as

$$X = \sum_j X_j + X_{nonviable} \qquad \text{(Equation 1)}$$

where $X_{nonviable}$ is the cumulative concentration of the considered species in nonviable cells. The phenomena considered by the model are: BV binding, BV transport to the nucleus and replication, release of budded BV, transcription and translation of AAV genes, rAAV capsid formation, Rep protein synthesis, vector genome amplification, and vector genome encapsidation (Figure 1C). BV infection presents an early (0–6 hpi), a late (6–18 hpi), and a very late (>18 hpi) stage, characterized by activation of different reactions and mechanisms within the infected cell. For instance, phenomena such as DNA replication and budding occur only in the late and very late stages, respectively (Figure 1C). Because the model considers synchronous infection, the infection age of the first wave of infected and coinfected cells is set to the time that passed from process onset, whereas the infection age of the second wave is the time from onset of budding for the first wave (τ_{rel}). For simplicity, the remainder of this discussion drops the two-wave notation, and the model is presented for a one-wave scenario without any loss of generality. The notation for infected and coinfected cells I_j and for the concentration of the respective intracellular species X_j becomes $j \in \{rc, goi, co\}$.

Cells, virions, and binding

The balance for uninfected cells is

$$\frac{dT}{dt} = \mu T - k_{bind0} T\left(V_{rc} + V_{goi}\right) - k_{death,T} T \qquad \text{(Equation 2)}$$

where μ, k_{bind0}, and $k_{death,T}$ are, respectively, the growth, binding, and death kinetic constants of uninfected cells. Baculovirus-infected cells do not undergo mitosis,[13] and their balances are

$$\frac{dI_{rc}}{dt} = k_{bind0} T V_{rc} - k_{bind} I_{rc} V_{goi} - k_{death,rc} I_{rc} \qquad \text{(Equation 3)}$$

$$\frac{dI_{goi}}{dt} = k_{bind0} T V_{goi} - k_{bind} I_{goi} V_{rc} - k_{death,goi} I_{goi} \qquad \text{(Equation 4)}$$

$$\frac{dI_{co}}{dt} = k_{bind}\left(I_{rc} V_{goi} + I_{goi} V_{rc}\right) - k_{bind} I_{goi} V_{rc} - k_{death,co} I_{co} \qquad \text{(Equation 5)}$$

where $k_{death,j}$ is the equivalent death kinetic constant of infected cells (for $j \in \{rc, goi, co\}$), and k_{bind} is the equivalent binding kinetic constant for infected cells, which decays with the infection progress as[13,69]

$$k_{bind} = \begin{cases} k_{bind0}, & \text{for } t_{inf} < \tau_{bind} \\ k_{bind0} exp\left(-\beta_{bind}\left(t_{inf} - \tau_{bind}\right)\right), & \text{for } t_{inf} \geq \tau_{bind} \end{cases} \qquad \text{(Equation 6)}$$

where t_{inf} is the infection age, β_{bind} is the binding decay kinetic constant, and τ_{bind} is the binding decay onset. The death rate of infected cells is faster than for uninfected cells, and most cells become nonviable between 3 and 5 dpi.[13,31] The equivalent death kinetic constant of infected cells is modeled by

$$k_{death,j} = \begin{cases} k_{death,T}, & \text{for } t_{inf} < \tau_{death} \\ k_{death,DNA} \log\left(vDNA_{rc,j} + vDNA_{goi,j}\right) + k_{death,rep} \log\left(Rep78_j\right) + k_{death,cap} \log\left(Caps_j + rAAV_j\right) & \text{for } t_{inf} \geq \tau_{death} \end{cases} \qquad \text{(Equation 7)}$$

where τ_{death} is the infection age of switching from uninfected to infected cell death kinetics, and the parameters $k_{death,DNA}$, $k_{death,rep}$, and $k_{death,cap}$ are kinetic constants relating $k_{death,j}$ to intracellular concentrations of BV DNA, Rep78, and rAAV capsids. Literature studies suggest that the concentration of these intracellular species might be correlated with the accelerated death rate of infected cells.[30,51] However, following the parameter estimation findings reported under results (specifically in section "cell growth and death dynamics"), the effect of *Rep* and *Cap* expression on $k_{death,j}$ can actually be neglected ($k_{death,DNA} = k_{death,rep} = 0$), and Equation 7 is simplified as

$$k_{death,j} = \begin{cases} k_{death,T}, & \text{for } t_{inf} < \tau_{death} \\ k_{death,DNA} log\left(vDNA_{rc,j} + vDNA_{goi,j}\right) & \text{for } t_{inf} \geq \tau_{death} \end{cases} \quad \text{(Equation 8)}$$

The balances for the BV virions are

$$\frac{dV_{rc}}{dt} = r_{rel,rc} - k_{bind0} V_{rc} T - k_{bind} V_{rc}\left(I_{rc} + I_{goi} + I_{co}\right) - k_{d,V} V_{rc} \quad \text{(Equation 9)}$$

$$\frac{dV_{goi}}{dt} = r_{rel,goi} k_{bind0} V_{goi} - k_{bind} V_{goi}\left(I_{rc} + I_{goi} + I_{co}\right) - k_{d,V} V_{goi} \quad \text{(Equation 10)}$$

where $k_{d,V}$ is the degradation kinetic constant for BV, and $r_{rel,rc}$ and $r_{rel,goi}$ are the cumulative release rates of repcapBV and goiBV from infected cells, respectively. The release rates are given by

$$r_{rel,i} = \begin{cases} 0, & \text{for } t_{inf} < \tau_{rel,on} \\ k_{rel} \sum\limits_{j \in \{rc,goi,co\}} \dfrac{vDNA_{i,j}}{\sum\limits_{k \in \{rc,goi\}} vDNA_{k,j}}, & \text{for } t_{inf} \geq \tau_{rel,on} \end{cases} \quad \text{(Equation 11)}$$

for $i \in \{rc, goi\}$, where k_{rel} is the release kinetic constant, and τ_{rel} is budding onset time. The formulation of Equation 11 assumes that infected cells release a fixed number of virions per unit of time during the very late stage and that, in the case of coinfection, virion production is split between repcapBV and goiBV, proportional to the respective DNA copies in the cell. Although the literature approximately agrees on the values of k_{rel} = 9.8 ± 1.5 PFU cell^{-1} h^{-1} and $\tau_{rel,on}$ = 18 ± 2 hpi,[13,31] different works suggest different endpoints for budded virus release, ranging from ≈35 hpi[29] up to cell death occurrence.[31] For our model, the latter option is chosen (Equation 11). Because the model developed in this work considers only systems in which all cells are infected during the first or second synchronous infection waves, the budding end time does not practically affect the model simulation.

Transport to nucleus and replication

As for all intracellular species, the balances for the BV bound to the receptors of viable infected cells are formulated in terms of cumulative concentration in the system:

$$\frac{dB_{rc,rc}}{dt} = V_{rc}\left(k_{bind0} T + k_{bind} I_{rc}\right) - B_{rc,rc}\left(k_{int} + k_{death,rc} + k_{bind} V_{goi}\right), \quad \text{(Equation 12)}$$

$$\frac{dB_{goi,goi}}{dt} = V_{goi}\left(k_{bind0} T + k_{bind} I_{goi}\right) - B_{goi,goi}\left(k_{int} + k_{death,goi} + k_{bind} V_{rc}\right). \quad \text{(Equation 13)}$$

where k_{int} is the kinetic constant for internalization. Equations 12 and 13 include a source term, accounting for BV binding, and a sink term, accounting for internalization, loss of viability, and for the amount of receptor-bound BV that passes from the infected to the coinfected cells balance consequently to cell coinfection. The balances for the BV bound to the receptors of viable coinfected cells present analogous contributions:

$$\frac{dB_{rc,co}}{dt} = k_{bind}\left(V_{rc}\left(I_{goi} + I_{co}\right) + B_{rc,rc} V_{goi}\right) - B_{rc,co}\left(k_{int} + k_{death,co}\right), \quad \text{(Equation 14)}$$

$$\frac{dB_{goi,rc}}{dt} = k_{bind}\left(V_{goi}\left(I_{rc} + I_{co}\right) + B_{goi,goi} V_{rc}\right) - B_{goi,co}\left(k_{int} + k_{death,co}\right). \quad \text{(Equation 15)}$$

The balances for the viral DNA in the nucleus are

$$\frac{dvDNA_{rc,rc}}{dt} = \eta k_{int} B_{rc,rc} - vDNA_{rc,rc}\left(k_{bind} V_{goi} + k_{death,rc}\right) + k_{repl} f_{repl} vDNA_{rc,rc}, \quad \text{(Equation 16)}$$

$$\frac{dvDNA_{goi,goi}}{dt} = \eta k_{int} B_{goi,goi} - vDNA_{goi,goi}\left(k_{bind} V_{rc} + k_{death,goi}\right) + k_{repl} f_{repl} vDNA_{goi,goi}, \quad \text{(Equation 17)}$$

$$\frac{dvDNA_{rc,co}}{dt} = \eta k_{int} B_{rc,co} + k_{bind} V_{goi} vDNA_{rc,rc} - k_{death,co} vDNA_{rc,co} + k_{repl} f_{repl} vDNA_{rc,co}, \quad \text{(Equation 18)}$$

$$\frac{dvDNA_{goi,co}}{dt} = \eta k_{int} B_{goi,co} + k_{bind} V_{rc} vDNA_{goi,goi} - k_{death,co} vDNA_{goi,co} + k_{repl} f_{repl} vDNA_{goi,co}, \quad \text{(Equation 19)}$$

where η is a coefficient accounting for BV degradation in lysosomes during trafficking to the nucleus, k_{repl} is the BV DNA replication kinetic constant, and f_{repl} is the activation function for BV replication. Following the BV infection dynamics:[13]

$$f_{repl} = \begin{cases} 0, & \text{for } t_{inf} < \tau_{repl,on} \\ 1, & \text{for } t_{inf} \geq \tau_{repl,off} \end{cases} \quad \text{(Equation 20)}$$

Beside DNA replication, the source contribution of Equations 16–19 accounts for internalization. As for receptor-bound BV (Equations 12–15), Equations 16–19 present a contribution that factors out of the balance the BV DNA in cells that become nonviable and a contribution that moves the viral DNA content of infected cells that, at a given time, become coinfected from the infected cell balance to the coinfected cell balance. The latter contribution is necessary for closure because binding, even though weak, can still occur at replication onset, and a cell infected by only one of the two BVs might obtain coinfection status after having already started BV replication. This contribution does not appear in the balances of the other intracellular species that are presented in the following discussion because, when they form, BV binding is weak or not happening anymore, preventing the flux of intracellular species from infected to coinfected cell balance.

Transcription and translation

The model includes equations for expression of Rep52, Rep78, viral proteins, and the transgene. Transgene expression is explicitly considered because it can potentially reduce the total host expression capability available for AAV structural and non-structural proteins. All other proteins that are expressed during BV infection are not featured in the model because they are not expected to directly affect rAAV production. In case their expression might indirectly affect rAAV production (e.g., by sharing limited transcription factors with *Rep* or *Cap* promoters), such an effect would not significantly vary across the considered experimental conditions, and it would already be accounted for in the regressed model parameters.

The transcription of a generic gene g located in the DNA of a BV of type i in the nucleus of infected cell of type j is modeled by

$$\frac{dmRNA_{g,j}}{dt} = \frac{k_l}{nt_g} f_l vDNA_{i,j} - \left(k_{d,mRNA} + k_{death,j}\right) mRNA_{i,j} \quad \text{(Equation 21)}$$

where k_l is the transcription kinetic constant for promoter l associated with gene g, f_l is the activation function for transcription from promoter l, nt_g is the number of nucleotides of gene g, and $k_{d,mRNA}$ is the degradation kinetic constant for mRNA. As in the balances for receptor-bound BV (Equations 16–19) and nuclear viral DNA (Equations 16–19), Equation 21 accounts for the decrease of the total level in the system of mRNA available for translation because of loss of cell viability. For TwoBac, Equation 21 becomes

$$\frac{dmRNA_{rep,j}}{dt} = \frac{k_{polh}}{nt_{rep78}} f_{polh} vDNA_{rc,j} - \left(k_{d,mRNA} + k_{death,j}\right) mRNA_{rep,j} \quad \text{(Equation 22)}$$

$$\frac{dmRNA_{cap,j}}{dt} = \frac{k_{p10}}{nt_{cap}} f_{p10} vDNA_{rc,j} - \left(k_{d,mRNA} + k_{death,j}\right) mRNA_{cap,j}, \quad \text{(Equation 23)}$$

$$\frac{dmRNA_{goi,j}}{dt} = \frac{k_{p10}}{nt_{goi}} f_{p10} vDNA_{goi,j} - \left(k_{d,mRNA} + k_{death,j}\right) mRNA_{goi,j}. \quad \text{(Equation 24)}$$

The activation functions for the polh and p10 promoters are

$$f_{polh} = \begin{cases} 0, & \text{for } t_{inf} < \tau_{polh} \text{ or } t_{inf} > \tau_{transcr,off} \\ 1, & \text{for } \tau_{polh} \le t_{inf} \le \tau_{transcr,off} \end{cases} \quad \text{(Equation 25)}$$

$$f_{p10} = \begin{cases} 0, & \text{for } t_{inf} < \tau_{p10} \text{ or } t_{inf} > \tau_{transcr,off} \\ 1, & \text{for } \tau_{p10} \le t_{inf} \le \tau_{transcr,off} \end{cases} \quad \text{(Equation 26)}$$

The translation rate $r_{transl,g,j}$ for transcript $mRNA_{g,j}$ in cells j is calculated, in terms of $\left[\#\ h^{-1}mL^{-1}\right]$, with Michaelis-Menten kinetics, accounting for saturation of the host translation machinery:

$$r_{transl,g,j} = \frac{k_{transl}}{nt_g} \frac{mRNA_{g,j}}{K_{transl} + mRNA_{g,j}} f_{sat}, \quad \text{(Equation 27)}$$

where k_{transl} is the translation kinetic constant corresponding to the maximum translation rate, K_{transl} is the Michaelis-Menten constant for translation, and f_{sat} accounts for competition for translation among different transcripts:

$$f_{sat} = \begin{cases} 1, & for\ R_{sat} \le 1, \\ 1/R_{sat}, & \text{for } R_{sat} > 1, \end{cases} \quad \text{(Equation 28)}$$

$$R_{sat} = \sum_g \frac{mRNA_{g,j}}{K_{transl} + mRNA_{g,j}}. \quad \text{(Equation 29)}$$

With the formulation of Equation 27, translation spontaneously halts with mRNA degradation.

The corresponding balances for the Rep proteins for the TwoBac system are

$$\frac{dRep52_j}{dt} = \frac{k_{transl}}{0.5\left(nt_{rep52} + nt_{rep78}\right)} I_j \frac{mRNA_{rep,j}}{K_{transl} + mRNA_{rep,j}} f_{sat} - k_{d,rep} Rep52_j - k_{death,j} Rep52_j \quad \text{(Equation 30)}$$

$$\frac{dRep78_j}{dt} = \frac{k_{transl}}{0.5\left(nt_{rep52} + nt_{rep78}\right)} I_j \frac{mRNA_{rep,j}}{K_{transl} + mRNA_{rep,j}} f_{sat} - k_{d,rep} Rep78_j - k_{death,j} Rep78_j, \quad \text{(Equation 31)}$$

where $k_{d,rep}$ is the degradation kinetic constant for Rep. In Equations 30 and 31, nt_g is taken as the average number of nucleotides between *Rep52* and *Rep78*, following the experimental finding that the leaky scanning mechanism of TwoBac leads to formation of Rep52 and Rep78 at a ratio approximately equal to 1:1.[20]

The balance for the transgene protein is

$$\frac{dGFP_j}{dt} = \frac{k_{transl}}{nt_{goi}} I_j \frac{mRNA_{goi,j}}{K_{transl}+mRNA_{goi,j}} f_{sat} - k_{d,GFP}\, GFP_j - k_{death,j} GFP_j, \quad \text{(Equation 32)}$$

where $k_{d,GFP}$ is the degradation kinetic constant for the protein. Under the hypothesis of fast capsid assembly,[27,70,71] the viral protein formation and assembly are lumped into a single step. The empty capsid balance becomes

$$\frac{dCaps_j}{dt} = \frac{k_{transl}}{60\, nt_{cap}} I_j \frac{mRNA_{cap,j}}{K_{transl}+mRNA_{cap,j}} f_{sat} - k_{d,GFP}\, GFP_j - k_{death,j} Caps - r_{encaps,j}, \quad \text{(Equation 33)}$$

where $r_{encaps,j}$ is the vector genome encapsidation rate in infected cells of type j. This equation neglects capsid degradation, following experimental evidence from the literature,[37,40] and considers that capsid proteins VP1, VP2, and VP3 have the same relative synthesis rate, as assumed in previous models.[27] Equation 33 does not account for empty capsid secretion into the cytosol, following experimental findings.[67] With the passing of time from inoculation, empty and full capsids are found in the medium as a result of loss of cell viability, although intracellular capsids remain predominant.[11]

Vector genome amplification and encapsidation

Rep78 is essential for vector genome amplification.[8,72] Vector genome amplification is modeled through Michaelis-Menten kinetics, accounting for potential Rep78 limitation. The balance for the free vector genome available for encapsidation is

$$\frac{dGOI_j}{dt} = \frac{k_{ampl}}{nt_{vg}} f_{ampl} I_j \frac{Rep78_j}{K_{ampl}+Rep78_j} - \left(k_{d,DNA} + k_{death,j}\right) GOI_j - r_{encaps,j}, \quad \text{(Equation 34)}$$

where k_{ampl} is the kinetic constant for amplification, K_{ampl} is the Michaelis-Menten constant, nt_{vg} is the number of nucleotides in the ITR/GOI cassette, f_{ampl} is the activation function for amplification, equal to 1 when at least one template is available for vector amplification in the cell and equal to 0 otherwise, and $k_{d,DNA}$ is the degradation kinetic constant for the non-encapsidated vector genome. Limitations induced by low levels of template are not considered, following experimental evidence.[42]

Vector genome encapsidation is mediated by Rep52, which forms stable intermediate complexes with the empty capsids, and intervenes in genome packaging.[8,73] Rep52 limitation is modeled through Michaelis-Menten kinetics, imposing the Michaelis-Menten constant in cells j ($K_{encaps,j}$) equal to

$$K_{encaps,j} = LS_j K_{encaps,coeff}, \quad \text{(Equation 35)}$$

where $K_{encaps,coeff} = 5$, and LS_j is the LS for encapsidation in cells j:

$$LS_j = \min\left\{Caps_j, GOI_j\right\}. \quad \text{(Equation 36)}$$

This modeling choice is based on the 1:1 mapping between Rep52 and capsids in formation of intermediate complexes during encapsidation,[8] assuming that Rep52 is not significantly limiting packaging when it is present in the cell in concentrations of one or more orders of magnitude larger than LS_j. The resulting encapsidation rate is

$$r_{encaps,j} = \frac{k_{encaps}}{nt_{vg}} LS_j \frac{Rep52_j}{K_{encaps,j}+Rep52_j}. \quad \text{(Equation 37)}$$

The corresponding balance for the filled rAAV capsids is

$$\frac{drAAV_j}{dt} = r_{encaps,j} - k_{death,j} rAAV_j. \quad \text{(Equation 38)}$$

ThreeBac modeling

The BV infection dynamics and the reaction-transport network for ThreeBac are shown in Figure S1. ThreeBac modeling is carried out with the same rationale adopted for TwoBac, but accounting for the presence of a third BV for gene delivery and for the different arrangements and promoters of the *Rep*, *Cap*, and ITR/GOI cassettes (Figure 1A). In ThreeBac, there are 6 types of infected and coinfected cells I_j for each of the two synchronous waves (Figure S1A). Balances for uninfected and infected cells and for the virions are analogous to the balances for TwoBac (Equations 2–11). The ThreeBac model accounts for 16 intracellular species, which is three more than for TwoBac. Two additional intracellular species are related to the presence of one more type of receptor-bound BV and BV DNA in the nucleus, whereas the third additional species originates from the separate transcription of *Rep52* and *Rep78* in ThreeBac (Figure S1B). Balances for receptor-bound BV and BV DNA in the nucleus are developed accordingly to the TwoBac model (Equations 12–20). The balances for the transcripts are derived from Equation 21, accounting for the promoter-cassette arrangement of ThreeBac:

$$\frac{dmRNA_{rep78,j}}{dt} = \frac{k_{\Delta IE1}}{nt_{rep78}} f_{\Delta IE1} vDNA_{rep,j} - \left(k_{d,mRNA} + k_{death,j}\right) mRNA_{rep78,j}, \quad \text{(Equation 39)}$$

$$\frac{dmRNA_{rep52,j}}{dt} = \frac{k_{polh}}{nt_{rep52}} f_{polh} vDNA_{rep,j} - \left(k_{d,mRNA} + k_{death,j}\right) mRNA_{rep52,j}, \quad \text{(Equation 40)}$$

$$\frac{dmRNA_{cap,j}}{dt} = \frac{k_{polh}}{nt_{cap}} f_{polh} vDNA_{cap,j} - \left(k_{d,mRNA} + k_{death,j}\right) mRNA_{cap,j}, \quad \text{(Equation 41)}$$

$$\frac{dmRNA_{goi,j}}{dt} = \frac{k_{p10}}{nt_{cap}} f_{p10} vDNA_{goi,j} - \left(k_{d,mRNA} + k_{death,j}\right) mRNA_{goi,j}, \quad \text{(Equation 42)}$$

where $k_{\Delta IE1}$ is the transcription kinetic constant for the $\Delta IE1$ promoter, and $f_{\Delta IE1}$ is the promoter's activation function:

$$f_{\Delta IE1} = \begin{cases} 0, & \text{for } t_{inf} < \tau_{\Delta IE1} \text{ or } t_{inf} > \tau_{transcr,off}, \\ 1, & \text{for } \tau_{\Delta IE1} \leq t_{inf} \leq \tau_{transcr,off}, \end{cases} \quad \text{(Equation 43)}$$

Translation; Rep, Cap, and transgene protein synthesis; capsid formation; vector genome amplification; and encapsidation are modeled as in the equations for TwoBac (Equations 27–38).

Stably transfected cell line modeling

The model considers two different configurations for rAAV production through BV infection of stably transfected cell lines. The first configuration involves infection with repBV of a cell line with an integrated ITR/GOI cassette.[42] This scenario is directly simulated using the same implementation of the model used for ThreeBac, imposing the concentration of goiBV DNA in the nucleus equal to one copy per cell and the concentration of capBV and goiBV virions equal to zero. The second configuration involves the OneBac process, in which goiBV is used to infect a cell line stably transfected with *Rep* and *Cap* cassettes.[50] The OneBac process is simulated using the model implementation for TwoBac, imposing the concentration of the repcapBV virions equal to zero. In the OneBac experiment considered in this work, weak Rep and Cap amplification is registered within the cell during BV infection.[10] Hence, in Equation 21, the number of viral DNA templates per cell available for transcription is set equal to the Rep and Cap copy numbers measured during the experiment (about 50 and about 450 copies per cell, respectively).

Parameter estimation strategy

The model has 35 parameters (Table 1). The parameters are estimated through a six-step (steps 1–6; Table S2) approach, as further detailed in the "Additional information on the parameter estimation strategy" section in the supplemental information. In all steps except for step 6 (discussed below), parameter estimation is carried out through maximum-likelihood estimation in the log space of the parameters:[74]

$$\widehat{\boldsymbol{\theta}} = \arg\min_{\boldsymbol{\theta}} \Phi(\boldsymbol{\theta}) . \quad \text{(Equation 44)}$$

$$\Phi(\boldsymbol{\theta}) = (\mathbf{y} - \tilde{\mathbf{y}}(\boldsymbol{\theta}))^{\top} \mathbf{V}_{\mathbf{y}}^{-1} (\mathbf{y} - \tilde{\mathbf{y}}(\boldsymbol{\theta})) \quad \text{(Equation 45)}$$

where $\widehat{\boldsymbol{\theta}}$ is the vector of parameter estimates, $\mathbf{y}$ is the vector of measurements, $\tilde{\mathbf{y}}(\boldsymbol{\theta})$ is the vector of model prediction for the measured variables, and $\mathbf{V}_{\mathbf{y}}$ is the measurement covariance matrix. The covariance matrix of the estimated parameters ($\mathbf{V}_{\boldsymbol{\theta}}$) is computed through the Hessian approximation:[74]

$$\mathbf{V}_{\boldsymbol{\theta}} = \mathbf{H}(\Phi(\widehat{\boldsymbol{\theta}}))^{-1}. \quad \text{(Equation 46)}$$

In each step (again, except for step 6), a multi-start approach with a sequential quadratic algorithm is used to solve the optimization (Equations 44 and 45) by minimizing the objective function (Equation 45) locally from 1,000 random starting points. The retained estimated parameters correspond to the parameter set that leads to the lowest objective function across the 1,000 parameter estimation runs. For this parameter set, the Hessian for computing $\mathbf{V}_{\boldsymbol{\theta}}$ (Equation 46) is evaluated numerically.[75] From $\mathbf{V}_{\boldsymbol{\theta}}$, 100 $(1 - \alpha)$% confidence intervals ($CI_{i,100(1-\alpha)\%}$) for each estimated parameter $\widehat{\theta}_i \in \widehat{\boldsymbol{\theta}}$ are calculated as[74]

$$CI_{i,100(1-\alpha)\%} = \left(\widehat{\theta}_i - t_{1-\frac{\alpha}{2},N_y-N_\theta} \sqrt{V_{\theta_{ii}}}, \widehat{\theta}_i + t_{1-\frac{\alpha}{2},N_y-N_\theta} \sqrt{V_{\theta_{ii}}} \right), \quad \text{(Equation 47)}$$

where $t_{1-\frac{\alpha}{2},N_y-N_\theta}$ is the Student's t value with a significance level α and $(N_y - N_\theta)$ degrees of freedom, and $V_{\theta_{ii}}$ is the variance of estimated parameter $\widehat{\theta}_i$, equal to the diagonal entry of $\mathbf{V}_{\boldsymbol{\theta}}$ corresponding to $\widehat{\theta}_i$.

The parameters obtained in steps 1–5 are required to run the parameter estimation of step 6 (Table S2). Hence, in step 6, maximum a posteriori estimation is used to propagate the uncertainty in the parameters obtained in steps 1–5 to the new set of estimated parameters. At the same time, the parameters obtained in steps 1–5 are updated within the maximum a posteriori estimation procedure by leveraging the experimental information from the dataset considered in step 6. A multi-start approach with a sequential quadratic algorithm is used to solve the maximum a posteriori estimation problem (Equations 44 and 48), by minimizing locally, from 1,000 random starting points, the objective function:[74,76]

$$\Phi(\boldsymbol{\theta}) = (\mathbf{y} - \tilde{\mathbf{y}}(\boldsymbol{\theta}))^{\top} \mathbf{V}_{\mathbf{y}}^{-1} (\mathbf{y} - \tilde{\mathbf{y}}(\boldsymbol{\theta})) + (\boldsymbol{\theta} - \boldsymbol{\theta}_{prior})^{\top} \mathbf{V}_{prior}^{-1} (\boldsymbol{\theta} - \boldsymbol{\theta}_{prior}), \quad \text{(Equation 48)}$$

where $\boldsymbol{\theta}$ includes all the model parameters, $\boldsymbol{\theta}_{prior}$ contains the values of the parameters available at the end of step 5, and $\mathbf{V}_{prior}$ is the corresponding covariance matrix. The entries of $\mathbf{V}_{prior}$ and $\boldsymbol{\theta}_{prior}$ corresponding to the parameters that are still unknown at step 6 (Table S2) are appropriately set so that their contribution to the second term on the right-hand side of Equation 48 is null. The estimated parameters that lead to the lowest objective function across the 1,000-parameter estimation runs is used to initialize a bootstrap analysis[77] (1,000 samplings), to better characterize the estimated parameter distribution. The parameters estimated through the bootstrap analysis are reported in Table 1. The values of the parameters previously estimated in steps 1–5 do not present any significant variation in step 6. The bootstrap analysis provides the 95% confidence interval reported in Table 1 for the new estimated parameters (Table S2). For the parameters that had been previously estimated in steps 1–5, the (wider) confidence interval previously calculated through Equation 47 are conservatively retained and reported in Table 1, rather than the tighter confidence interval obtained through bootstrapping in step 6. The statistical significance of all estimated

parameters is verified with the following t test for parameter precision:[53]

$$\frac{\widehat{\theta}_i}{t_{1-\frac{\alpha}{2},N_y-N_\theta}\sqrt{V_{\theta_{ii}}}} > t_{1-\alpha,N_y-N_\theta}. \quad \text{(Equation 49)}$$

The set of 35 parameters obtained through parameter estimation are used in the simulations for model validation (Figures 2, 3, 5, S4, and S7) and for *in silico* analysis of the process (Figures 6, 7, and S8). The only exception are the parameters for the infected cell death kinetics ($k_{death,DNA}$ and τ_{death}), which, for model validation, are re-estimated to match the viability measurements in each experiment when available. The resulting parameters (not reported) present a maximum deviation of 50% from the values reported in Table 1, indicating that viability can be much affected by lab-to-lab variability.

Data are extracted from the figures of the cited literature using WebPlotDigitizer or ImageJ.

Sensitivity analysis and forward uncertainty propagation

The sensitivity matrix $\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}}$ of the model states $\mathbf{x}$ with respect to the model parameters $\boldsymbol{\theta}$ is computed by integrating the sensitivity equations alongside the model equations:[72]

$$\frac{d\mathbf{x}(t)}{dt} = f(\mathbf{x}(t), \boldsymbol{\theta}) \quad \text{(Equation 50)}$$

$$\frac{d}{dt}\left(\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}}\right) = \frac{\partial \mathbf{f}}{\partial \mathbf{x}}\frac{\partial \mathbf{x}}{\partial \boldsymbol{\theta}} + \frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}} \quad \text{(Equation 51)}$$

The Jacobians $\frac{\partial \mathbf{f}}{\partial \mathbf{x}}$ and $\frac{\partial \mathbf{f}}{\partial \boldsymbol{\theta}}$ are calculated using the automatic differentiation toolbox ADiGator.[78] The normalized cumulative sensitivity S_{x_i,θ_i} across the batch of a state $x_i \in \mathbf{x}$ to a parameter $\theta_i \in \boldsymbol{\theta}$ is computed from

$$S_{x_i,\theta_i} = \frac{\widehat{\theta}}{t_f}\int_0^{t_f}\frac{\partial x_i}{\partial \theta_i}(t)dt \quad \text{(Equation 52)}$$

where t_f is the batch duration. The 95% confidence interval of the model prediction reported in Figure 6 is computed through a forward Monte Carlo approach with 5×10^4 realizations, sampling, for each realization, a different set of model parameters from the distribution defined by the parameter estimate and variance matrix (Table 1).

Shake flask production of rAAVs

Cell culture and rAAV production

Sf9 cells were cultured in 30 mL serum-free medium (SFM4Insect, HyClone, UT, USA) in a 125-mL vented Erlenmeyer flask (Fisher Scientific, NH, USA) placed in an orbital shaker incubator at 27°C and 135 rpm. When the Sf9 cell density reached 1.5 million cells/mL, the infection was initiated by inoculating the medium, at 1:100 volumetric dilution, with insect cells previously infected with goiBV and insect cells previously infected with repcapBV (obtained as in Cecchini et al.[11] and stocked at 20 million cells/mL). Cell suspensions were collected at 24, 48, 72, and 96 hpi. Cell diameter, percentage viability, and viable cell density were determined by incubating the cell suspension with an equal part of 0.4% trypan blue (Thermo Fisher Scientific, MA, USA), followed by quantification using an automated cell counter (Countess Automated Cell Counter, Thermo Fisher Scientific).

Vector characterization

Aliquots of the sample 24, 48, 72, and 96 after inoculation were treated with lysis buffer (20 mM Tris HCl, 1 mM $MgCl_2$ hexahydrate, and 0.5% Tween 20) supplemented with 20 U/mL Turbonuclease (Millipore Sigma, MA, USA) for 4 h at 37°C. The cell debris was removed by centrifugation at 15,000 × *g* for 20 min at 4°C. The supernatant containing AAV5 particles was collected to quantify the encapsidated AAV-5 genome and capsids using droplet digital PCR (ddPCR) and an enzyme-linked immunosorbent assay (ELISA), respectively. Samples were diluted with nuclease-free ultrapure water (Ampliqon, Odense, Denmark) to attain the desired concentration before setting up the reaction mix. The primers used for the AAV5 vector quantification were as follows: forward, 5′ GCAAAGACCC CAACGAGAAG-3′; reverse, 5′TCACGAACTCCAGCAGGACC-3′. Samples (2 μL) were mixed with EvaGreen Supermix (Bio-Rad, CA, USA) along with forward and reverse primers in ddPCR 96-well plates (Bio-Rad), followed by droplet generation (Qx200 Droplet Generator, Bio-Rad) into a clean 96-well plate. The plates were sealed at 180°C for 5 s (PX1 PCR Plate Sealer, Bio-Rad). The PCR reaction was performed with a thermal cycler (C1000 Touch Thermal cycler, Bio-Rad), using three steps: enzyme activation (°C, 5 min, 1 cycle), DNA denaturation (95°C, 30 s, 40 cycles), and primer annealing (60°C, 1 min, 40 cycles). After thermal cycling, samples were quantified using a droplet reader (QX200, Bio-Rad) using QuantaSoft software. The AAV5 Xpress Titration ELISA Kit (Progen Biotechnik, Heidelberg, Germany) was utilized to detect the assembled AAV5 capsids. The samples were diluted in the sample buffer supplied in the kit, and the manufacturer's protocol was followed to quantify AAV5 capsids. In brief, the diluted samples were incubated with anti-AAV5 capture antibodies pre-coated on 96-well plate strips for 20 min. Following incubation, biotinylated anti-AAV5 detection antibody, streptavidin-peroxidase conjugate, and tetramethylbenzidine substrate were added sequentially to quantify the AAV-5 capsids. The absorbance was measured using a standard well plate reader at 450 nm. AAV-5 capsid titer was determined from the standard curve using a 4-parameter logistic fit (4PL) based on the manufacturer's guidelines.

DATA AND CODE AVAILABILITY

A MATLAB implementation of the model presented in this manuscript can be downloaded at the following link: https://github.com/francescodestro/rAAV_BEVS.

SUPPLEMENTAL INFORMATION

Supplemental information can be found online at https://doi.org/10.1016/j.omtm.2023.05.019.

ACKNOWLEDGMENTS

This work was done in Cambridge, MA, USA. This work was supported by the US Food and Drug Administration through contract 75F40121C00131.

AUTHOR CONTRIBUTIONS

Conceptualization, F.D., C.N., P.W.B., S.L.S., A.J.S., R.M.K., and R.D.B.; formal analysis, F.D.; investigation, F.D., J.J., P.S., and J.M.K.; methodology, F.D. and R.D.B.; software, F.D.; validation, F.D. and J.J.; visualization, F.D.; writing – original draft, F.D. and R.D.B.; writing – review & editing, F.D., J.J., P.S., C.N., J.M.W., P.W.B., S.L.S., A.J.S., S.C., R.M.K., and R.D.B.; funding acquisition, C.N., J.M.W., P.W.B., S.L.S., A.J.S., R.M.K., and R.D.B.; project administration, P.W.B.; supervision, P.W.B., S.L.S., A.J.S., S.C., R.M.K., and R.D.B.

DECLARATION OF INTERESTS

R.M.K. is an inventor on patents related to recombinant AAV technology and owns equity in gene therapy-related companies. Portions of the recombinant AAV technology studied in this report are covered by United States and European patents assigned to the Secretary of the U.S. Department of Health and Human Services. A fraction of the licensing fees and royalty payments made to the National Institutes of Health is distributed to the inventors (R.M.K.) in accordance with U.S. Government and National Institutes of Health policy.

REFERENCES

1. Smalley, E. (2017). First AAV gene therapy poised for landmark approval. Nat. Biotechnol. *35*, 998–999. https://doi.org/10.1038/nbt1117-998.
2. Ylä-Herttuala, S. (2012). Endgame: glybera finally recommended for approval as the first gene therapy drug in the European Union. Mol. Ther. *20*, 1831–1832. https://doi.org/10.1038/mt.2012.194.
3. U.S. National Library of Medicine (2022). ClinicalTrials.gov. https://www.clinicaltrials.gov.
4. Evaluate Ltd (2022). Evaluate Pharma 2022. https://www.evaluate.com/thought-leadership/pharma/world-preview-2022-report.
5. Rininger, J., Fennell, A., Schoukroun-Barnes, L., Peterson, C., and Speidel, J. (2019). Capacity Analysis for Viral Vector Manufacturing: Is There Enough?. https://bioprocessintl.com/manufacturing/emerging-therapeutics-manufacturing/capacity-analysis-for-viral-vector-manufacturing-is-there-enough/.
6. Elverum, K., and Whitman, M. (2020). Delivering cellular and gene therapies to patients: solutions for realizing the potential of the next generation of medicine. Gene Ther. *27*, 537–544. https://doi.org/10.1038/s41434-019-0074-7.
7. Clément, N., and Grieger, J.C. (2016). Manufacturing of recombinant adeno-associated viral vectors for clinical trials. Mol. Ther. Methods Clin. Dev. *3*, 16002. https://doi.org/10.1038/mtm.2016.2.
8. 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., and Braatz, R.D. (2021). Cellular pathways of recombinant adeno-associated virus production for gene therapy. Biotechnol. Adv. *49*, 107764. https://doi.org/10.1016/j.biotechadv.2021.107764.
9. Adamson-Small, L., Potter, M., Falk, D.J., Cleaver, B., Byrne, B.J., and Clément, N. (2016). A scalable method for the production of high-titer and high-quality adeno-associated type 9 vectors using the HSV platform. Mol. Ther. Methods Clin. Dev. *3*, 16031. https://doi.org/10.1038/mtm.2016.31.
10. Joshi, P.R.H., Cervera, L., Ahmed, I., Kondratov, O., Zolotukhin, S., Schrag, J., Chahal, P.S., and Kamen, A.A. (2019). Achieving high-yield production of functional AAV5 gene delivery vectors via fedbatch in an insect cell-one baculovirus system. Mol. Ther. Methods Clin. Dev. *13*, 279–289. https://doi.org/10.1016/j.omtm.2019.02.003.
11. Cecchini, S., Virag, T., and Kotin, R.M. (2011). Reproducible high yields of recombinant adeno-associated virus produced using invertebrate cells in 0.02-to 200-liter cultures. Hum. Gene Ther. *22*, 1021–1030. https://doi.org/10.1089/hum.2010.250.
12. Kurasawa, J.H., Park, A., Sowers, C.R., Halpin, R.A., Tovchigrechko, A., Dobson, C.L., Schmelzer, A.E., Gao, C., Wilson, S.D., and Ikeda, Y. (2020). Chemically defined, high-density insect cell-based expression system for scalable AAV vector production. Mol. Ther. Methods Clin. Dev. *19*, 330–340. https://doi.org/10.1016/j.omtm.2020.09.018.
13. Rohrmann, G.F. (2019). Baculovirus Molecular Biology, 4th edition (National Center for Biotechnology Information (US)).
14. Smith, G.E., Summers, M.D., and Fraser, M.J. (1983). Production of human beta interferon in insect cells infected with a baculovirus expression vector. Mol. Cell Biol. *3*, 2156–2165. https://doi.org/10.1128/MCB.3.12.2156-2165.1983.
15. Yee, C.M., Zak, A.J., Hill, B.D., and Wen, F. (2018). The coming age of insect cells for manufacturing and development of protein therapeutics. Ind. Eng. Chem. Res. *57*, 10061–10070. https://doi.org/10.1021/acs.iecr.8b00985.
16. Wickham, T.J., Davis, T., Granados, R.R., Shuler, M.L., and Wood, H.A. (1992). Screening of insect cell lines for the production of recombinant proteins and infectious virus in the baculovirus expression system. Biotechnol. Prog. *8*, 391–396. https://doi.org/10.1021/bp00017a003.
17. Von Drygalski, A., Giermasz, A., Castaman, G., Key, N.S., Lattimore, S., Leebeek, F.W.G., Miesbach, W., Recht, M., Long, A., Gut, R., et al. (2019). Etranacogene dezaparvovec (AMT-061 phase 2b): normal/near normal FIX activity and bleed cessation in hemophilia B. Blood Adv. *3*, 3241–3247. https://doi.org/10.1182/bloodadvances.2019000811.
18. FDA (2022). FDA Approves First Gene Therapy to Treat Adults with Hemophilia B. https://www.fda.gov/news-events/press-announcements/fda-approves-first-gene-therapy-treat-adults-hemophilia-b.
19. Urabe, M., Ding, C., and Kotin, R.M. (2002). Insect cells as a factory to produce adeno-associated virus type 2 vectors. Hum. Gene Ther. *13*, 1935–1943. https://doi.org/10.1089/10430340260355347.
20. Smith, R.H., Levy, J.R., and Kotin, R.M. (2009). A simplified baculovirus-AAV expression vector system coupled with one-step affinity purification yields high-titer rAAV stocks from insect cells. Mol. Ther. *17*, 1888–1896. https://doi.org/10.1038/mt.2009.128.
21. Mietzsch, M., Grasse, S., Zurawski, C., Weger, S., Bennett, A., Agbandje-McKenna, M., Muzyczka, N., Zolotukhin, S., and Heilbronn, R. (2014). OneBac: platform for scalable and high-titer production of adeno-associated virus serotype 1–12 vectors for gene therapy. Hum. Gene Ther. *25*, 212–222. https://doi.org/10.1089/hum.2013.184.
22. Galibert, L., Jacob, A., Savy, A., Dickx, Y., Bonnin, D., Lecomte, C., Rivollet, L., Sanatine, P., Boutin Fontaine, M., Le Bec, C., and Merten, O.W. (2021). MonoBac system–A single baculovirus for the production of rAAV. Microorganisms *9*, 1799. https://doi.org/10.3390/microorganisms9091799.
23. Kotin, R.M., and Snyder, R.O. (2017). Manufacturing clinical grade recombinant adeno-associated virus using invertebrate cell lines. Hum. Gene Ther. *28*, 350–360. https://doi.org/10.1089/hum.2017.042.
24. Joshi, P.R.H., Venereo-Sanchez, A., Chahal, P.S., and Kamen, A.A. (2021). Advancements in molecular design and bioprocessing of recombinant adeno-associated virus gene delivery vectors using the insect-cell baculovirus expression platform. Biotechnol. J. *16*, e2000021. https://doi.org/10.1002/biot.202000021.
25. Hong, M.S., Severson, K.A., Jiang, M., Lu, A.E., Love, J.C., and Braatz, R.D. (2018). Challenges and opportunities in biopharmaceutical manufacturing control. Comput. Chem. Eng. *110*, 106–114. https://doi.org/10.1016/j.compchemeng.2017.12.007.
26. Destro, F., and Barolo, M. (2022). A review on the modernization of pharmaceutical development and manufacturing-Trends, perspectives, and the role of mathematical modeling. Int. J. Pharm. *620*, 121715. https://doi.org/10.1016/j.ijpharm.2022.121715.
27. 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., and Braatz, R.D. (2021). Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. Mol. Ther. Methods Clin. Dev. *21*, 642–655. https://doi.org/10.1016/j.omtm.2021.04.006.

28. Saxena, A., Byram, P.K., Singh, S.K., Chakraborty, J., Murhammer, D., and Giri, L. (2018). A structured review of baculovirus infection process: integration of mathematical models and biomolecular information on cell–virus interaction. J. Gen. Virol. *99*, 1151–1171. https://doi.org/10.1099/jgv.0.001108.

29. Dee, K.U., and Shuler, M.L. (1997). A mathematical model of the trafficking of acid-dependent enveloped viruses: application to the binding, uptake, and nuclear accumulation of baculovirus. Biotechnol. Bioeng. *54*, 468–490. https://doi.org/10.1002/(SICI)1097-0290(19970605)54:5<468::AID-BIT7>3.0.CO;2-C.

30. Roldão, A., Vieira, H.L.A., Charpilienne, A., Poncet, D., Roy, P., Carrondo, M.J.T., Alves, P.M., and Oliveira, R. (2007). Modeling rotavirus-like particles production in a baculovirus expression vector system: infection kinetics, baculovirus DNA replication, mRNA synthesis and protein production. J. Biotechnol. *128*, 875–894. https://doi.org/10.1016/j.jbiotec.2007.01.003.

31. Power, J.F., Reid, S., Radford, K.M., Greenfield, P.F., and Nielsen, L.K. (1994). Modeling and optimization of the baculovirus expression vector system in batch suspension culture. Biotechnol. Bioeng. *44*, 710–719.

32. Nielsen, L.K. (2003). Virus production from cell culture, kinetics. In Encyclopedia of Cell Technology (John Wiley & Sons, Inc.). https://doi.org/10.1002/0471250570.spi106.

33. De Gooijer, C.D., Koken, R.H., Van Lier, F.L., Kool, M., Vlak, J.M., and Tramper, J. (1992). A structured dynamic model for the baculovirus infection process in insect-cell reactor configurations. Biotechnol. Bioeng. *40*, 537–548. https://doi.org/10.1002/bit.260400413.

34. Licari, P., and Bailey, J.E. (1992). Modeling the population dynamics of baculovirus-infected insect cells: optimizing infection strategies for enhanced recombinant protein yields. Biotechnol. Bioeng. *39*, 432–441. https://doi.org/10.1002/bit.260390409.

35. Jang, J.D., Sanderson, C.S., Chan, L.C., Barford, J.P., and Reid, S. (2000). Structured modeling of recombinant protein production in batch and fed-batch culture of baculovirus-infected insect cells. Cytotechnology *34*, 71–82. https://doi.org/10.1023/A:1008178029138.

36. Canova, C.T., Inguva, P.K., and Braatz, R.D. (2023). Mechanistic modeling of viral particle production. Biotechnol. Bioeng. *120*, 629–641. https://doi.org/10.1002/bit.28296.

37. Aucoin, M.G., Perrier, M., and Kamen, A.A. (2006). Production of adeno-associated viral vectors in insect cells using triple infection: optimization of baculovirus concentration ratios. Biotechnol. Bioeng. *95*, 1081–1092. https://doi.org/10.1002/bit.21069.

38. Mena, J.A., Aucoin, M.G., Montes, J., Chahal, P.S., and Kamen, A.A. (2010). Improving adeno-associated vector yield in high density insect cell cultures. J. Gene Med. *12*, 157–167. https://doi.org/10.1002/jgm.1420.

39. Rosinski, M., Reid, S., and Nielsen, L.K. (2002). Kinetics of baculovirus replication and release using real-time quantitative polymerase chain reaction. Biotechnol. Bioeng. *77*, 476–480. https://doi.org/10.1002/bit.10126.

40. Meghrous, J., Aucoin, M.G., Jacob, D., Chahal, P.S., Arcand, N., and Kamen, A.A. (2005). Production of recombinant adeno-associated viral vectors using a baculovirus/insect cell suspension culture system: from shake flasks to a 20-L bioreactor. Biotechnol. Prog. *21*, 154–160. https://doi.org/10.1021/bp049802e.

41. Vieira, H.L.A., Estêvão, C., Roldão, A., Peixoto, C.C., Sousa, M.F.Q., Cruz, P.E., Carrondo, M.J.T., and Alves, P.M. (2005). Triple layered rotavirus VLP production: kinetics of vector replication, mRNA stability and recombinant protein production. J. Biotechnol. *120*, 72–82. https://doi.org/10.1016/j.jbiotec.2005.03.026.

42. Li, L., Dimitriadis, E.K., Yang, Y., Li, J., Yuan, Z., Qiao, C., Beley, C., Smith, R.H., Garcia, L., and Kotin, R.M. (2013). Production and characterization of novel recombinant adeno-associated virus replicative-form genomes: a eukaryotic source of DNA for gene transfer. PLoS One *8*, e69879. https://doi.org/10.1371/journal.pone.0069879.

43. Power, J.F., Reid, S., Greenfield, P.F., and Nielsen, L.K. (1996). The kinetics of baculovirus adsorption to insect cells in suspension culture. Cytotechnology *21*, 155–163.

44. Vanarsdall, A.L., Okano, K., and Rohrmann, G.F. (2005). Characterization of the replication of a baculovirus mutant lacking the DNA polymerase gene. Virology *331*, 175–180. https://doi.org/10.1016/j.virol.2004.10.024.

45. Grose, C., Putman, Z., and Esposito, D. (2021). A review of alternative promoters for optimal recombinant protein expression in baculovirus-infected insect cells. Protein Expr. Purif. *186*, 105924. https://doi.org/10.1016/j.pep.2021.105924.

46. Urabe, M., Nakakura, T., Xin, K.-Q., Obara, Y., Mizukami, H., Kume, A., Kotin, R.M., and Ozawa, K. (2006). Scalable generation of high-titer recombinant adeno-associated virus type 5 in insect cells. J. Virol. *80*, 1874–1885. https://doi.org/10.1128/jvi.80.4.1874-1885.2006.

47. Mitchell-Logean, C., and Murhammer, D.W. (1997). bcl-2 expression in Spodoptera Frugiperda Sf-9 and Trichoplusia Ni BTI-Tn-5B1-4 insect cells: effect on recombinant protein expression and cell viability. Biotechnol. Bioeng. *56*, 380–390. https://doi.org/10.1002/(SICI)1097-0290(19971120)56:4<380::AID-BIT4>3.0.CO;2-K.

48. Aucoin, M.G., Perrier, M., and Kamen, A.A. (2008). Critical assessment of current adeno-associated viral vector production and quantification methods. Biotechnol. Adv. *26*, 73–88. https://doi.org/10.1016/j.biotechadv.2007.09.001.

49. Cecchini, S., Negrete, A., and Kotin, R.M. (2008). Toward exascale production of recombinant adeno-associated virus for gene transfer applications. Gene Ther. *15*, 823–830. https://doi.org/10.1038/gt.2008.61.

50. Kondratov, O., Marsic, D., Crosson, S.M., Mendez-Gomez, H.R., Moskalenko, O., Mietzsch, M., Heilbronn, R., Allison, J.R., Green, K.B., Agbandje-McKenna, M., and Zolotukhin, S. (2017). Direct head-to-head evaluation of recombinant adeno-associated viral vectors manufactured in human versus insect cells. Mol. Ther. *25*, 2661–2675. https://doi.org/10.1016/j.ymthe.2017.08.003.

51. Schmidt, M., Afione, S., and Kotin, R.M. (2000). Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53. J. Virol. *74*, 9441–9450. https://doi.org/10.1128/jvi.74.20.9441-9450.2000.

52. Burnham, K.P., and Anderson, D.R. (2004). Model Selection and Multimodel Inference (Springer).

53. Franceschini, G., and Macchietto, S. (2008). Model-based design of experiments for parameter precision: state of the art. Chem. Eng. Sci. *63*, 4846–4872. https://doi.org/10.1016/J.CES.2007.11.034.

54. Mccarty, D.M., Pereira, D.J., Zolotukhin, I., Zhou, X., Ryan, J.H., Muzyczka, N., Mccarty, D.M., Ryan, J.H., Zolutukhin, S., Zhou, X., et al. (1994). Identification of linear DNA sequences that specifically bind the adeno-associated virus Rep protein. J. Virol. *68*, 4988–4997. https://doi.org/10.1128/JVI.68.8.4988-4997.1994.

55. Carter, B.J. (2004). Adeno-associated virus and the development of adeno-associated virus vectors: a historical perspective. Mol. Ther. *10*, 981–989. https://doi.org/10.1016/j.ymthe.2004.09.011.

56. Chiorini, J.A., Yang, L., Safer, B., and Kotin, R.M. (1995). Determination of adeno-associated virus Rep68 and Rep78 binding sites by random sequence oligonucleotide selection. J. Virol. *69*, 7334–7338.

57. Chiorini, J.A., Wiener, S.M., Owens, R.A., Kyöstió, S.R., Kotin, R.M., and Safer, B. (1994). Sequence requirements for stable binding and function of Rep68 on the adeno-associated virus type 2 inverted terminal repeats. J. Virol. *68*, 7448–7457.

58. Im, D.-S., and Muzyczka, N. (1990). The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. Cell *61*, 447–457.

59. Smith, R.H., and Kotin, R.M. (2000). An adeno-associated virus (AAV) initiator protein, Rep78, catalyzes the cleavage and ligation of single-stranded AAV ori DNA. J. Virol. *74*, 3122–3129.

60. Chejanovsky, N., and Carter, B.J. (1990). Mutation of a consensus purine nucleotide binding site in the adeno-associated virus rep gene generates a dominant negative phenotype for DNA replication. J. Virol. *64*, 1764–1770.

61. Chejanovsky, N., and Carter, B.J. (1989). Mutagenesis of an AUG codon in the adeno-associated virus rep gene: effects on viral DNA replication. Virology *173*, 120–128. https://doi.org/10.1016/0042-6822(89)90227-4.

62. Negrete, A., Yang, L.C., Mendez, A.F., Levy, J.R., and Kotin, R.M. (2007). Economized large-scale production of high yield of rAAV for gene therapy applications exploiting baculovirus expression system. J. Gene Med. *9*, 938–948. https://doi.org/10.1002/jgm.1092.

63. Chemla, Y.R., Aathavan, K., Michaelis, J., Grimes, S., Jardine, P.J., Anderson, D.L., and Bustamante, C. (2005). Mechanism of force generation of a viral DNA packaging motor. Cell *122*, 683–692. https://doi.org/10.1016/j.cell.2005.06.024.

64. Jardin, B.A., Zhao, Y., Selvaraj, M., Montes, J., Tran, R., Prakash, S., and Elias, C.B. (2008). Expression of SEAP (secreted alkaline phosphatase) by baculovirus mediated

transduction of HEK 293 cells in a hollow fiber bioreactor system. J. Biotechnol. *135*, 272–280. https://doi.org/10.1016/J.JBIOTEC.2008.04.006.

65. Wolfisberg, R., Kempf, C., and Ros, C. (2016). Late maturation steps preceding selective nuclear export and egress of progeny parvovirus. J. Virol. *90*, 5462–5474. https://doi.org/10.1128/jvi.02967-15.

66. Elmore, Z.C., Patrick Havlik, L., Oh, D.K., Anderson, L., Daaboul, G., Devlin, G.W., Vincent, H.A., and Asokan, A. (2021). The membrane associated accessory protein is an adeno-associated viral egress factor. Nat. Commun. *12*, 6239. https://doi.org/10.1038/s41467-021-26485-4.

67. Gallo-Ramírez, L.E., Ramírez, O.T., and Palomares, L.A. (2011). Intracellular localization of adeno-associated viral proteins expressed in insect cells. Biotechnol. Prog. *27*, 483–493. https://doi.org/10.1002/btpr.565.

68. Petricevich, V.L., Palomares, L.A., González, M., and Ramírez, O.T. (2001). Parameters that determine virus adsorption kinetics: toward the design of better infection strategies for the insect cell - baculovirus expression system. Enzym. Microb. Technol. *29*, 52–61. https://doi.org/10.1016/S0141-0229(01)00323-4.

69. Shuler, M.L., Cho, T., Wickham, T., Ogonah, O., Kool, M., Hammer, D.A., Granados, R.R., and Wood, H.A. (1990). Bioreactor development for production of viral pesticides or heterologous proteins in insect cell cultures. Ann. N. Y. Acad. Sci. *589*, 399–422.

70. Myers, M.W., and Carter, B.J. (1980). Assembly of adeno-associated virus. Virol *102*, 71–82. https://doi.org/10.1016/0042-6822(80)90071-9.

71. Grosse, S., Penaud-Budloo, M., Herrmann, A.-K., Börner, K., Fakhiri, J., Laketa, V., Krämer, C., Wiedtke, E., Gunkel, M., Ménard, L., et al. (2017). Relevance of assembly-activating protein for adeno-associated virus vector production and capsid protein stability in mammalian and insect cells. J. Virol. *91*, 011988-17–e1217. https://doi.org/10.1128/jvi.01198-17.

72. Hong, M.S., Velez-Suberbie, M.L., Maloney, A.J., Biedermann, A., Love, K.R., Love, J.C., Mukhopadhyay, T.K., and Braatz, R.D. (2021). Macroscopic modeling of bioreactors for recombinant protein producing Pichia pastoris in defined medium. Biotechnol. Bioeng. *118*, 1199–1212. https://doi.org/10.1002/bit.27643.

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

74. Bard, Y. (1974). Nonlinear Parameter Estimation, 4th ed. (Academic Press).

75. D'Errico, J. (2023). Adaptive robust numerical differentiation. MATLAB Central File Exchange. Retrieved March 23, 2023. tps://www.mathworks.com/matlabcentral/fileexchange/13490-adaptive-robust-numerical-differentiation.

76. Beck, J.V., and Arnold, K.J. (1977). Parameter Estimation in Engineering and Science (John Wiley & Sons).

77. Efron, B., and Tibshirani, R.J. (1994). An Introduction to the Bootstrap (CRC Press).

78. Weinstein, M.J., and Rao, A.V.((2017). Algorithm 984: ADiGator, a toolbox for the algorithmic differentiation of mathematical functions in MATLAB using source transformation via operator overloading. ACM Trans. Math Software *44*, 1–25. https://doi.org/10.1145/3104990.