Applied Microbiology and Biotechnology (2024) 108:459
https://doi.org/10.1007/s00253-024-13281-5

BIOTECHNOLOGICAL PRODUCTS AND PROCESS ENGINEERING

Enhanced ER protein processing gene expression increases rAAV yield and full capsid ratio in HEK293 cells

Qiang Fu[1] · Yongdan Wang[2] · Jiansong Qin[1] · Dongming Xie[2] · David McNally[2,3] · Seongkyu Yoon[1,2]

Received: 4 April 2024 / Revised: 9 August 2024 / Accepted: 14 August 2024

Abstract

The recombinant adeno-associated virus (rAAV) vector is among the most promising viral vectors in gene therapy. However, the limited manufacturing capacity in human embryonic kidney (HEK) cells is a barrier to rAAV commercialization. We investigated the impact of endoplasmic reticulum (ER) protein processing and apoptotic genes on transient rAAV production in HEK293 cells. We selected four candidate genes based on prior transcriptomic studies: *XBP1*, *GADD34* / *PPP1R15A*, *HSPA6*, and *BCL2*. These genes were stably integrated into HEK293 host cells. Traditional triple-plasmid transient transfection was used to assess the vector production capability and the quality of both the overexpressed stable pools and the parental cells. We show that the overexpression of *XBP1*, *HSPA6*, and *GADD34* increases rAAV productivity by up to 100% and increases specific rAAV productivity by up to 78% in HEK293T cells. Additionally, more prominent improvement associated with ER protein processing gene overexpression was observed when parental cell productivity was high, but no substantial variation was detected under low-producing conditions. We also confirmed genome titer improvement across different serotypes (AAV2 and AAV8) and different cell lines (HEK293T and HEK293); however, the extent of improvement may vary. This study unveiled the importance of ER protein processing pathways in viral particle synthesis, capsid assembly, and vector production.

Key points

- *Upregulation of endoplasmic reticulum (ER) protein processing (XBP1, HSPA6, and GADD34) leads to a maximum 100% increase in rAAV productivity and a maximum 78% boost in specific rAAV productivity in HEK293T cells*
- *The enhancement in productivity can be validated across different HEK293 cell lines and can be used for the production of various AAV serotypes, although the extent of the enhancement might vary slightly*
- *The more pronounced improvements linked to overexpressing ER protein processing genes were observed when parental cell productivity was high, with minimal variation noted under low-producing conditions*

Keywords Gene therapy · RAAV production · HEK293 cells · ER protein processing gene and apoptosis regulator gene · Capsid biosynthesis · Packaging efficiency

Introduction

Recombinant adeno-associated viruses (rAAVs) comprise an important vector for delivering transgenes in vivo (Fu et al. 2023b; Sha et al. 2021; Wang et al. 2019, 2024). However, due to increasing demand, the low productivity of rAAVs presents as a major challenge (Escandell et al. 2023; Fu et al. 2023b; Lee et al. 2022; Wang et al. 2019, 2024).

rAAV vector production is heavily dependent on host cells. A recent review paper provides a comprehensive summary of the role of host cells as manufacturing factories that support viral production. This same review also emphasizes the importance of cell line engineering as a future strategy

Qiang Fu and Yongdan Wang equally contributed to this paper.

✉ Seongkyu Yoon
seongkyu_yoon@uml.edu

[1] Department of Biomedical Engineering and Biotechnology, University of Massachusetts Lowell, Lowell, MA 01854, USA

[2] Department of Chemical Engineering, University of Massachusetts Lowell, Lowell, MA 01854, USA

[3] MassBiologics, University of Massachusetts Chan Medical School, Mattapan, MA 02126, USA

Published online: 04 September 2024

for improving manufacturing capability (Wang et al. 2024). Omics studies have demonstrated a series of responses in human embryonic kidney 293 (HEK293) cells that are triggered by viral production. These studies formed a hypothesis regarding how these responses might influence viral production (Chung et al. 2023; Lu et al. 2024; Wang et al. 2023). The transcriptomic data demonstrated that endoplasmic reticulum (ER) stress was upregulated when host cells produced vectors. The unfolded protein response (UPR) was then recruited to alleviate this ER stress and facilitate viral production (Wang et al. 2023). We observed extreme upregulation in (a) protein phosphatase 1 regulatory subunit 15 A (*GADD34* / *PPP1R15A*), (b) heat shock protein family member 6 (*HSPA6*), and (c) X-box binding protein 1 (*XBP1*). We identified these genes as important molecular features of viral production (Wang et al. 2023).

GADD34 is essential for the recovery of protein synthesis when UPR signaling (triggered by protein kinase R–like endoplasmic reticulum kinase (PERK)) temporarily attenuates translation (Harding et al. 2009). *HSPA6* is an ER chaperone regulated by UPR signaling through activating transcription factor 6 (*ATF6*). The expression of ER chaperones such as *HSPA6* can promote ER protein translocation, protein folding, protein secretion, and misfolded protein degradation (Hassler et al. 2015; Wu et al. 2007). *XBP1* is upregulated when UPR signaling is triggered by ER-to-nucleus signaling 1 (IRE1α), a protein kinase that exhibits ribonuclease (RNase) activity and can cleave a set of mRNAs (Hollien et al. 2009; Upton et al. 2012). Similarly, the activation of *XBP1* can also promote ER protein translocation, folding, and secretion (Kaufman 2002). Thus, we hypothesize that the overexpression of these protein processing genes can relieve the host cell stress induced by viral protein synthesis, mitigate the accumulation of misfolded or unfolded proteins, and enhance viral production. Furthermore, the overexpression of anti-apoptotic gene *B-cell lymphoma 2* (*BCL2*) increases lentiviral vector volumetric production by 53% in HEK293 cells (Formas-Oliveira et al. 2020). This demonstrates the essential role of ER protein processing and anti-apoptotic pathways in vector production capability and the overall performance of producer cells.

We mainly investigate how the rAAV vector yield and quality in HEK293T cells are influenced by the overexpression of ER protein processing genes (i.e., *GADD34*, *HSPA6*, *XBP1*) and *BCL2* during the transient transfection process. This overexpression was achieved by stably integrating these candidate genes into the safe harbor site (i.e., the adeno-associated virus integration site 1 (*AAVS1*) locus) in the HEK293T cells. These results demonstrate that the overexpression of ER protein processing genes can effectively boost rAAV yield in the high-producing condition but shows less impact in the low-producing condition. This same improvement can also be extended to different serotypes (e.g., AAV2 and AAV8) and different cell lines (e.g., HEK293T and HEK293).

Materials and methods

Plasmid construction

When constructing the *AAVS1*-sgRNA-Cas9 plasmid, the design of the sgRNA sequences that target the *AAVS1* locus was based on a prior publication (Shin et al. 2020). These sgRNA sequences were annealed and then integrated into the sgRNA-Cas9 backbone plasmid (pU6-(BbsI)_CBh-Cas9-T2A-mCherry) using Golden Gate assembly (Ran et al. 2013).

For donor plasmid construction, we amplified the *AAVS1* homology arms and selection marker via PCR from donor backbone plasmids (i.e., the *AAVS1*-eGFP donor, a gift from the Korea Advanced Institute of Science & Technology (KAIST)) (Hockemeyer et al. 2009; Shin et al. 2020). The four candidate genes (*BCL2*, *XBP1*, *HSPA6*, and *GADD34*) were then amplified via PCR using related plasmids. NEBuilder® HiFi DNA Assembly master mix (New England Biolabs, Ipswich, MA, United States) was used to assemble PCR fragments that contained homology arms sequences, selection genes, or inserts. This was followed by transformation in *Escherichia coli* DH5α competent cells (New England Biolabs, Ipswich, MA, United States). Plasmids were confirmed through sequencing and then extracted using an EndoFree® Plasmid Maxi Kit (Zymo, Irvine, CA, United States). Supplementary Tables S1 and S2 provide more information on the plasmids and primers used for cloning. SnapGene® (GSL Biotech LLC, San Diego, CA, United States) was used to document all DNA cloning.

Cell culture, stable transfection, and stable cell pool development

HEK293 (ATCC, CRL-1573) and HEK293T/17 (ATCC, CRL-11268) cells were grown in Dulbecco's Modified Eagle's Medium (DMEM) (Gibco, Waltham, MA, United States) supplemented with 10% fetal bovine serum (Gibco, Waltham, MA, United States). The cells were cultured in T-flasks (GenClone, EI Cajon, United States) with 5 mL working volume at 37 °C under 5% CO_2 and passaged every 4 days. To establish stable cell pools that expressed the four candidate genes at the *AAVS1* locus, we transfected the candidate genes donor and sgRNA-Cas9 vectors at a 1:1 (w/w) ratio with Lipofectamine™ 2000 (Invitrogen, Carlsbad, CA, United States) followed by 2 weeks of selection with 2 μg/mL of puromycin (Sigma-Aldrich, St. Louis, MO, United States). Stable cell pools were then denoted as "Gene of Interest (GOI) + ." Viable cell density (VCD) and

cell viability were assessed using a Cedex HiRes Analyzer (Roche Life Science, Indianapolis, IN, United States) and the trypan blue assay (Thermo Fisher Scientific, Waltham, MA, United States).

Genomic DNA extraction for junction PCR and RNA isolation for gene expression level

Cell pellets from the stable pools were collected for genomic DNA and RNA extraction. Genomic DNA was extracted from stable cell pools using a Quick-DNA™ Miniprep Plus Kit (Zymo, Irvine, CA, United States) following the manufacturer's protocol. The 5′/3′ junction PCR was performed using PrimeSTAR® HS Premix (Takara Bio, Shiga, Japan) by touchdown PCR (98 °C for 3 min (10×), 98 °C for 10 s, 65–55 °C (− 1 °C/cycle) for 30 s, 72 °C for 2 min (30×), 98 °C for 10 s, 55 °C for 30 s, 72 °C for 2 min, and 72 °C for 10 min) (Shin et al. 2020). The primer sequences used for the 5′/3′ junction PCR are listed in Supplementary Table S2. The gene expression assay (e.g., RNA isolation, cDNA reverse transcription, gene expression by RT-PCR) was described earlier (Wang et al. 2023). We used the comparative cycle threshold ($2^{-\Delta\Delta CT}$) method to analyze the relative gene expression changes between the stable pools and the parental cells. Supplementary Table S2 lists the primer sequences used to assess the expression levels of *BCL2*, *XBP1*, *HSPA6*, *GADD34*, and the gene for glyceraldehyde 3-phosphate dehydrogenase (*GAPDH*).

Vector production

Triple-plasmid transfection was performed for rAAV production. Three AAV-related plasmids were used, sourced from Addgene (Watertown, MA, United States): pAd-DeltaF6, pAAV2/2, and AAV-CMV-GFP. Supplementary Table S1 provides information about the plasmids used. Cells in the exponential growth phase were plated at a density of 5×10^5 cells per well in a 6-well plate. After a 24-h incubation, transfection was performed following the PEI-pro® manufacturer's instructions (Polyplus, Berkeley, CA, United States). Specific transfection conditions can be found in Supplementary Tables S3 and S4.

Analytical methods (genome titer, capsid titer)

The harvested cell cultures were aliquoted to 1 mL (including both cells and supernatant) in 1.5-mL centrifuge tubes 68-h post-transfection. All cultures were stored at − 80 °C for future analysis. The genome titer by quantitative polymerase chain rain (qPCR) and capsid titer by enzyme-linked immunosorbent assay (ELISA) followed the methods used in an earlier study (Fu et al. 2023a). The specific productivity can be calculated as

$$\mathrm{IVCD} = \frac{(\mathrm{VCD}_2 + \mathrm{VCD}_1) \times (t_2 - t_1)}{2} \quad (1)$$

$$Q_p = \frac{\mathrm{Titer}_2 - \mathrm{Titer}_1}{\mathrm{IVCD}} \quad (2)$$

Note: VCD denotes viable cell density in million cells/mL. VCD1 and VCD2 indicate viable cell density measured at time points t_1 and t_2, respectively. IVCD represents integral viable cell concentration in cell·h/mL. Titer stands for the viral genome titer in viral genome/L. Titer_1 and Titer_2 indicate viral genome titer measured at time points t_1 and t_2, respectively. Q_p stands for specific productivity in viral genome/cell*h.

Statistical analysis

GraphPad Prism 9 (GraphPad Software, LLC, San Diego, CA, United States) was used for raw data processing, statistical analysis, and to generate graphs. BioRender (Toronto, Canada) was used to generate figures.

Results

Workflow for the overexpression of ER protein processing and anti-apoptotic genes in HEK293 and HEK293T cells

Figure 1 shows the overall workflow of this study. To investigate the impact of various ER protein processing genes (*XBP1*, *HSPA6*, and *GADD34*) and anti-apoptotic genes (*BCL2*) on rAAV production in HEK cell lines, these genes were stably integrated into the *AAVS1* locus using the CRISPR gene editing technique. We constructed sgRNA-Cas9 plasmids targeting the *AAVS1* locus as well as donor template plasmids containing the gene of interest (GOI) (see Fig. 1, left panel). These constructed plasmid sequences were obtained via Sanger sequencing (or whole plasmid sequencing), and proper assembly was confirmed by aligning against appropriate references in SnapGene®. These two vectors were then transfected to host cells via Lipofectamine™ 2000 followed by two weeks of antibiotic selection. The cell pools with stably integrated candidate genes were passaged and maintained for later evaluations.

As shown in Fig. 1 (middle panel), the gene expression (mRNA) level was evaluated in these stable pools, followed by 5′ and 3′ junction PCR to confirm site-specific integration. We followed the sgRNA and the donor template (i.e., the *AAVS1* locus) design provided by previous literature (Shin et al. 2020). Due to the acceptable knock-in efficiency reported in the literature, we did not conduct a whole genome sequencing (Shin et al. 2020). The stable

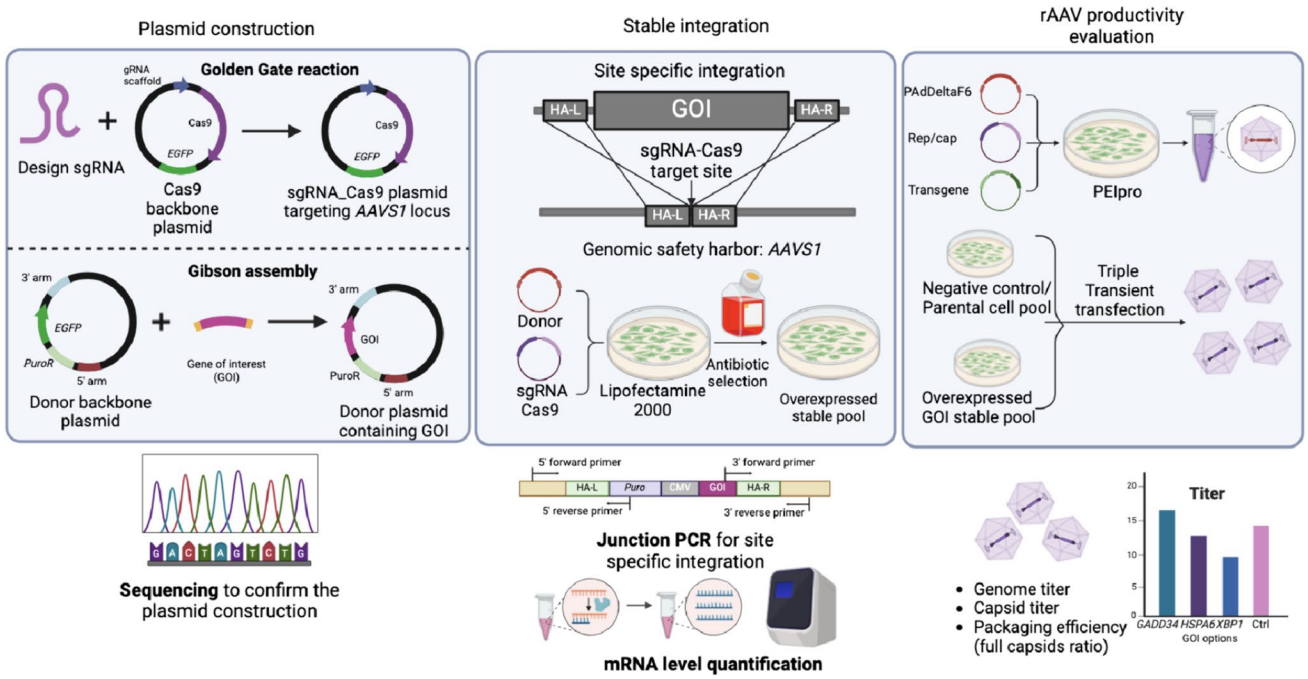


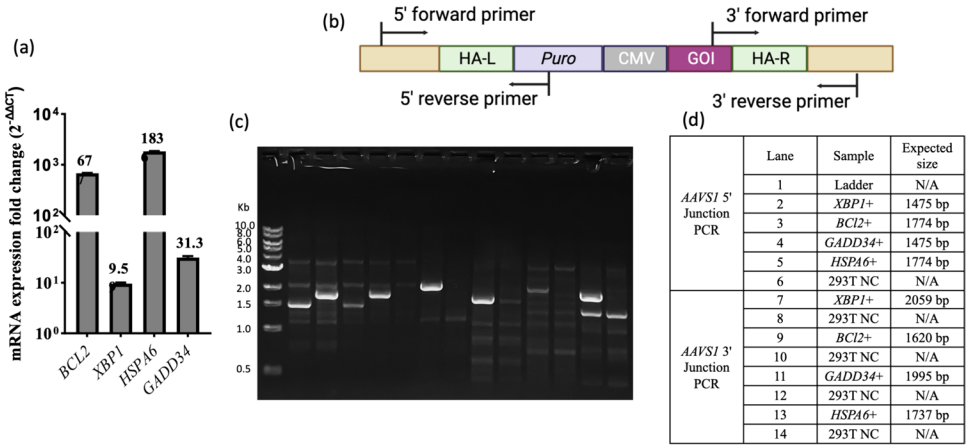
Fig. 1 Overall workflow of this overexpression study, including plasmid construction, stable integration, and evaluation of rAAV productivity in the overexpressed cell pools

pools containing the overexpressed candidate genes and the parental cells were then used to produce rAAVs via transient transfection with triple plasmids (Fig. 1). The effects of these ER protein processing genes on the genome and capsid titers were evaluated in HEK cell lines.

Evaluation of candidate gene overexpression and site-specific integration in HEK293T cells

HEK293T cells were selected and evaluated in detail because the same plasmid construction and the same stable integration process have been applied to both HEK293 and HEK293T cells. Each candidate gene, in its corresponding overexpressed stable cell, showed an expression level higher than the parental cells (Fig. 2a). More precisely, compared to the parental HEK293T cells, *BCL2* showed a 677-fold increase, *XBP1* showed a 9.6-fold increase, *HSPA6* showed a 1836-fold increase, and *GADD34* showed a 31.3-fold increase. The overexpression levels of each stably integrated gene were not the same. The difference could result from some random integrated genome copies of GOI. The effects of overexpression of ER-related genes are well-studied in stable production using CHO cells. Increasing evidence suggests overexpressing genes associated with UPR can serve to increase the productivity of CHO cell lines (Rutkowski and Kaufman 2004; Tigges and Fussenegger 2006). The outcomes depend on the overexpressed genes, target therapeutic proteins, and overexpression system (Kim et al. 2012). Due to the difference between viral vector and therapeutic protein production, as well as between cell lines, a more relevant study regarding lentiviral vector production in HEK cell lines was referenced. This study investigated different copy number integrations and mRNA level in detail (Formas-Oliveira et al. 2020). The transgene expression showed a trend to increase as the copy number of the candidate genes rose. In future study, altering the expression levels of these genes could serve as a way to fine-tune the degree of improvement.

Figure 2b shows our primer design strategy for the 5′ and 3′ junction PCR of the four candidate genes. In junction PCR, one primer is designed to anneal to the insert, and the other primer is designed to anneal to the sequence outside the homology arm. Primer designs outside the homology arms can be shared for all cases. In 5′ junction PCR, the inside (reverse) primer, which is annealed to the inserted puromycin sequence, was shared for all four overexpressed genes. In 3′ junction PCR, various inside (forward) primers were designed to target various inserted GOI sequences. The electrophoresis gel image (Fig. 2c and d) confirms that the genes were successfully integrated into the *AAVS1* locus. The expected size bands were shown in the overexpressed groups, while no band appeared in the control group.



	Lane	Sample	Expected size
AAVS1 5′ Junction PCR	1	Ladder	N/A
	2	*XBP1+*	1475 bp
	3	*BCl2+*	1774 bp
	4	*GADD34+*	1475 bp
	5	*HSPA6+*	1774 bp
	6	293T NC	N/A
AAVS1 3′ Junction PCR	7	*XBP1+*	2059 bp
	8	293T NC	N/A
	9	*BCl2+*	1620 bp
	10	293T NC	N/A
	11	*GADD34+*	1995 bp
	12	293T NC	N/A
	13	*HSPA6+*	1737 bp
	14	293T NC	N/A

Fig. 2 **a** Transcript level fold change of candidate genes (overexpressed stable cell pools/parental cells), normalized to housekeeping gene *GAPDH*. The data represent the mean and standard deviation of triplicate qPCR wells. **b** Primer design strategy for 5′/3′ junction PCR. **c** Verification of targeted integration at the *AAVS1* locus by 5′/3′ junction PCR. **d** Gel electrophoresis image, sample name, and corresponding expected band size. HA-L represents left homology arm. HA-R represents right homology arm. *Puro* represents puromycin resistance gene sequence. CMV represents cytomegalovirus promoter. N/A represents no specific defined band size

Vector productivity evaluation (AAV2) of HEK293T overexpressed stable pools via triple-plasmid transient transfection

Traditional transient transfection was conducted to produce rAAV vectors in both the overexpressed stable cells and the parental cells. We evaluated how the overexpression of candidate genes affects vector productivity under both low-producing and high-producing conditions, assuming that lower stress level is associated with low heterologous protein production, whereas elevated cell stress is linked to higher protein production. Two sets of transient transfection process parameters were selected: The first parameter was the low-producing condition (as shown in Supplementary Table S3), which followed the PEIpro® manufacturer protocol. The high-producing condition (as shown in Supplementary Table S4) was the optimal condition based on the literature (Grieger et al. 2016; Gu et al. 2018; Zhao et al. 2020). The genome titer and VCD were measured 68 h post-transfection (i.e., 68 HPT) for both conditions.

In the low-producing conditions (as depicted in Supplementary Fig. S1), overexpressed *HSPA6* stable cells displayed a genome titer comparable to that of the control, while the overexpression of the other three genes led to a decrease in the rAAV genome titer ranging from 30 to 60% in HEK293T. However, under high-producing conditions, all four genes showed significantly higher rAAV production than the control condition, with a titer improvement ranging from 48 to 80% (see Fig. 3a). Specifically, in the case of *BCL2* overexpression, there was a remarkable 79% boost in productivity, while *HSPA6* overexpression resulted in a 54% increase, and *GADD34* overexpression yielded an impressive 71% increase in productivity. Furthermore, we observed additional pronounced and significant improvements in specific productivity in all four overexpressed stable cells: *BCL2* + cells showed a 50% increase, *XBP1* + cells exhibited a substantial 91% increase, *HSPA6* + cells demonstrated a remarkable 116% increase, and *GADD34* + cells showed a notable 93% increase (see Fig. 3b). The VCD and viability for high-producing condition are shown in Supplementary Table S5. We also noticed that vector productivity improvement was not directly proportional to the rise in transcript levels of the candidate genes. This phenomenon may arise from variations in mRNA half-life in the cells and differences in the translation efficiency across individual genes (Metkar et al. 2024).

Packaging quality evaluation of AAV2 via capsid titer quantification for HEK293T overexpressed stable cells

The capsid titer is defined as the total number of synthesized capsid particles as quantified using ELISA. The packaging efficiency, expressed as the ratio of the genome titer (in vg/L) to the capsid titer (in capsid/L or abbreviated as cp/L) (i.e., the full capsid ratio), was assessed in HEK293T cells under high- and low-producing conditions using the methodology outlined in a prior study (Fu et al. 2023a). Figure 3c and d show the capsid titer and packaging efficiency under high-producing conditions. Supplementary Fig. S1 b and c show the same data for the low-producing conditions.

For the control groups, a capsid titer of 5×10^{13} cp/L was achieved under high-producing conditions, while a capsid titer of 5×10^{12} cp/L was observed under low-producing conditions. Notably, under both conditions, the capsid titer remained relatively constant between the overexpressed

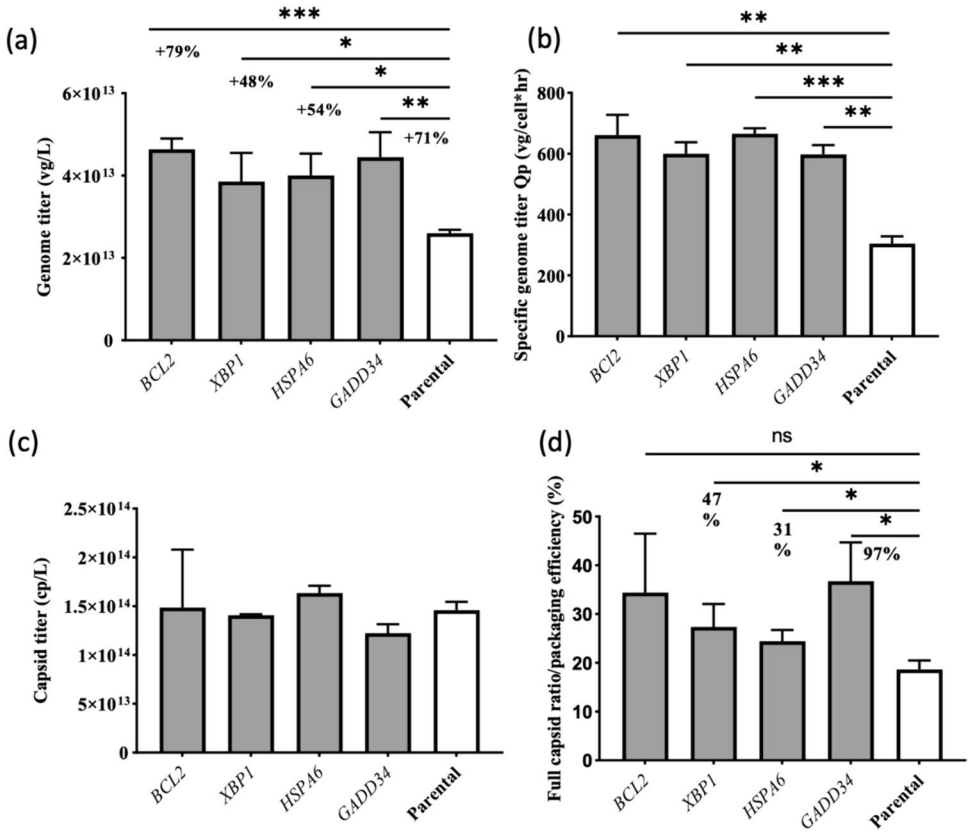


Fig. 3 Transient transfection for AAV2 production in HEK293T cells for both overexpressed stable pools and parental cells in the high-producing condition. **a** Genome titer in vg/L 68 HPT. **b** Specific genome titer Q_p in vg/cell × h at harvest. **c** Capsid titer in cp/L 68 HPT. **d** Packaging efficiency. The error bar indicates the mean and standard deviation derived from biological triplicates. To evaluate significant difference between treatment and parental groups, an unpaired two-tailed Student *t* test was conducted. *p* values < .05 were considered statistically significant. *$p < .05$, **$p < .01$, ***$p < .001$

stable cells and the parental control cells. One would expect that viral gene expression and viral protein production would be highly comparable because we employed an equal number of plasmids during transfection for both the overexpressed stable cell pools and the parental cells in rAAV production. The identical quantities of raw materials could explain the similar accumulation of capsids in all cases.

Packaging efficiency (the so-called full capsid ratio) is another important quality attribute for rAAV production, and we observed noticeable improvement in all four overexpressed stable cells in high-producing conditions. Among them, *XBP1* + (27.34%), *HSPA6* + (24.43%), and *GADD34* + (36.72%) stable cells demonstrated full capsid ratios significantly higher than the parental HEK293T cells (19.63%). This increased full capsid ratio in the overexpressed stable cells in high-producing conditions indicates the vital role these ER protein processing pathways and genes play in capsid protein biosynthesis, capsid assembly, and packaging. In addition to measuring capsid titer by ELISA, employing transmission electron microscopy (TEM) for direct visualization of viral vector particles might also be useful to evaluate the impact of the overexpression on vector quality. However, we find no significant difference in the full capsid ratios between the overexpressed groups and the control group in low-producing conditions. The relatively low ER stress in low-producing conditions might be effectively addressed by the inherent ER stress–processing genes in the control group; thus, the overexpression of these ER stress–processing genes is unlikely to provide additional stress relief. Moreover, the overexpression of ER protein–processing genes might play a more important role as ER stress escalates beyond the capacity of inherent stress-processing genes to alleviate it.

Assessment of alternate rAAV serotype (AAV8) production in HEK293T cells

To investigate whether the serotype determines the overexpression effects of ER protein processing genes on rAAV production, we conducted additional transient transfections for AAV8 production in the *XBP1* +, *HSPA6* +, *GADD*34 +, *BCL2* +, and parental HEK293T cells. The transfection parameters were identical to those used in the AAV2 high-producing condition. The genome titer and VCD were measured at 68 HPT. In general, under the same transfection conditions, a higher titer was achieved for AAV8 production (1.18×10^{14} vg/L) than that achieved for AAV2 production (2.60×10^{13} vg/L) in HEK293T parental cells. Figure 4 provides more detail about the genome titer and specific productivity, and Supplementary Table S6 shows the VCD and viability on harvest. The genome titers were higher in the overexpressed cells than in the parental cells. The result

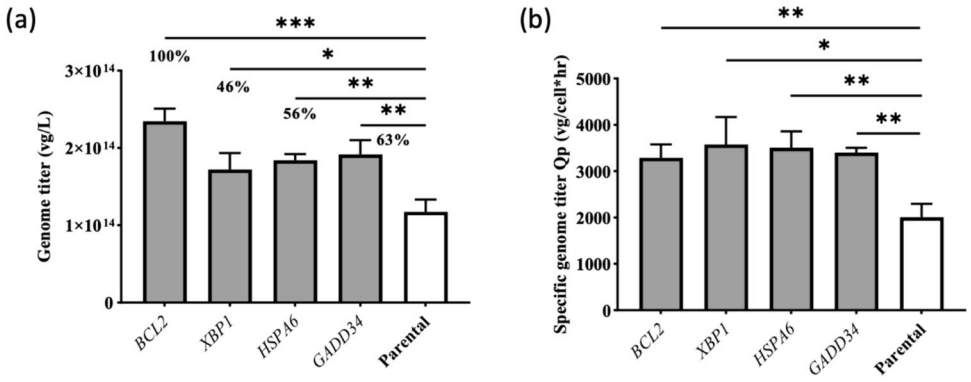


Fig. 4 Transient transfection for AAV8 production in HEK293T cells for both overexpressed stable pools and parental cells in the high-producing condition. **a** Genome titer in vg/L 68 HPT. **b** Specific genome titer Q_p in vg/cell × h at harvest. The error bar represents the mean and standard deviation for biological triplicates ($n = 3$). p values < .05 were considered statistically significant. *$p < .05$, **$p < .01$, ***$p < .001$

was a 46–100% increase in genome titer. The statistical p values further confirm the significance of the titer change. The improvement using both serotypes (AAV2 and AAV8) suggests the broad applicability of the overexpression of ER protein processing genes (*XBP1*, *HSPA6*, and *GADD34*) and anti-apoptotic genes (*BCL2*) as an engineering strategy in rAAV manufacturing.

Vector productivity evaluation (AAV2) in HEK293 overexpressed stable pools

In the production of viral vectors for gene therapies (e.g., rAAVs), the HEK293 and HEK293T host cell lines have been studied the most thoroughly. However, despite higher viral vector yields in general, the use of HEK293T cells still raises safety concerns due to the presence of an SV40 T antigen–encoding sequence (De et al. 2023; Zhao et al. 2020). We applied the same processes to HEK293 cells to evaluate whether the improvement associated with ER protein processing gene overexpression is cell line dependent. The duplicated processes included the stable integration of identified candidate genes to safe harbor sites, the evaluation of target gene expression levels in the stable pools, and transient transfection assessment for rAAV production in both high- and low-producing conditions for AAV2 production.

The transfection parameters were the same as those used for the HEK293T cells. As shown in Fig. 5 and Supplementary Fig. S2, the overexpression enhanced the genome titer and the specific productivity in the high-producing condition, but no significant impacts were observed in the low-producing condition in HEK293 cells. In high-producing cases, the overexpression of the anti-apoptotic gene *BCL2* led to a statistically significant increase (55%) in genome titer. The overexpression of the ER processing gene *HSPA6* resulted in a 102% increase in genome titer, the overexpression of *XBP1* contributed to a 70% increase, and the overexpression of *GADD34* resulted in a 91% boost in genome titer. The VCD and viability for high-producing condition in HEK293 are shown in Supplementary Table S7. This result confirms that ER protein processing gene overexpression is beneficial for viral vector production. However, the positive impacts of each specific gene on genome titer (which ranged from 50 to 100%) were cell line dependent.

Discussion

The therapeutic application and commercialization of rAAV gene therapies have been hindered by manufacturing difficulties and high production costs. Triple transient transfection in HEK293 cells is the common platform for

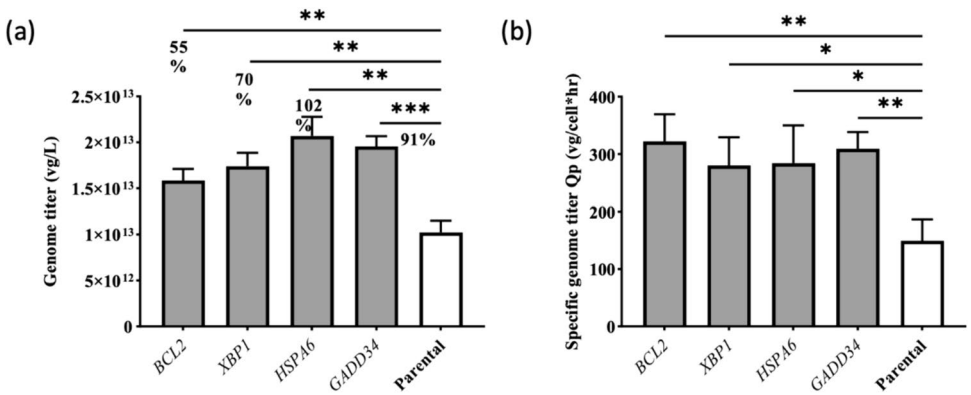


Fig. 5 Transient transfection for AAV2 production in HEK293 cells for both overexpressed stable pools and parental cells in the high-producing condition. **a** Genome titer in vg/L 68 HPT. **b** Specific genome titer Q_p in vg/cell × h at harvest. The error bar represents the mean and standard deviation for biological triplicates ($n = 3$). p values < .05 were considered statistically significant. *$p < .05$, **$p < .01$, ***$p < .001$

rAAV vector production. Thus, cellular features characterized during vector production provide relevant information for cell line engineering and productivity improvement.

A prior study using transcriptomics demonstrates that viral protein synthesis and replication trigger ER stress (Wang et al. 2023). Downstream UPR pathways were highly recruited to prohibit host cell stress, ensuring the survival of the cell and the production of the viral vector (Wang et al. 2023). *GADD34*, *HSPA6*, and *XBP1* were identified as molecular features for high producers, and these genes were involved in three respective UPR signaling pathways. In this study, we utilized a cell line engineering strategy to stably integrate candidate genes and increase their expression levels. Different strategies (e.g., medium supplementation targeting the specific pathway, transient gene expression) can also be applied to achieve target gene overexpression. However, stable overexpression is more favorable for long-term manufacturing processes where there is no need for additional small molecules or transfection using expensive plasmids. Thus, each gene was stably integrated into parental HEK cells to derive the overexpressed stable pools, which were later assessed for their vector production capability. All these evaluations were conducted as a proof-of-concept study to validate the essential roles of genes associated with ER protein processing during vector production.

ER protein processing pathways involved in vector production

Figure 6 provides details on the UPR signaling pathways and associated genes investigated in this study, including UPR signaling through PERK, IRE1α, and ATF6. The figure describes the pathways and provides the reason for selecting each candidate gene as well as its potential roles in vector productivity.

UPR signaling through PERK and *GADD34*: PERK is a kinase that phosphorylates eIF2α (a translation initiation factor) and temporarily attenuates protein synthesis when activated by ER stress (Kaufman 2002; Ron and Walter 2007). This translation attenuation is intended to reduce the ER protein load and alleviate ER stress. Simultaneously, phosphorylated eIF2α can trigger an *ATF4*-involved feedback loop to dephosphorylate eIF2α and restore normal protein synthesis by upregulating *GADD34* (Novoa et al. 2001). ER stress can also induce apoptosis if such UPR-adaptive responses are insufficient to relieve the stress and ensure cell survival (Ron and Walter 2007). Here, *GADD34* can restore protein translation when unfolded proteins accumulate as a result of viral protein synthesis and ER stress (Novoa et al. 2001). The overexpression of *GADD34* in this study could hypothetically restore viral protein synthesis and thereby promote capsid formation.

UPR signaling through IRE1α and *XBP1*: IRE1α is a protein kinase that can induce RNase activity under ER stress (Credle et al. 2005). This RNase activity can splice mRNA

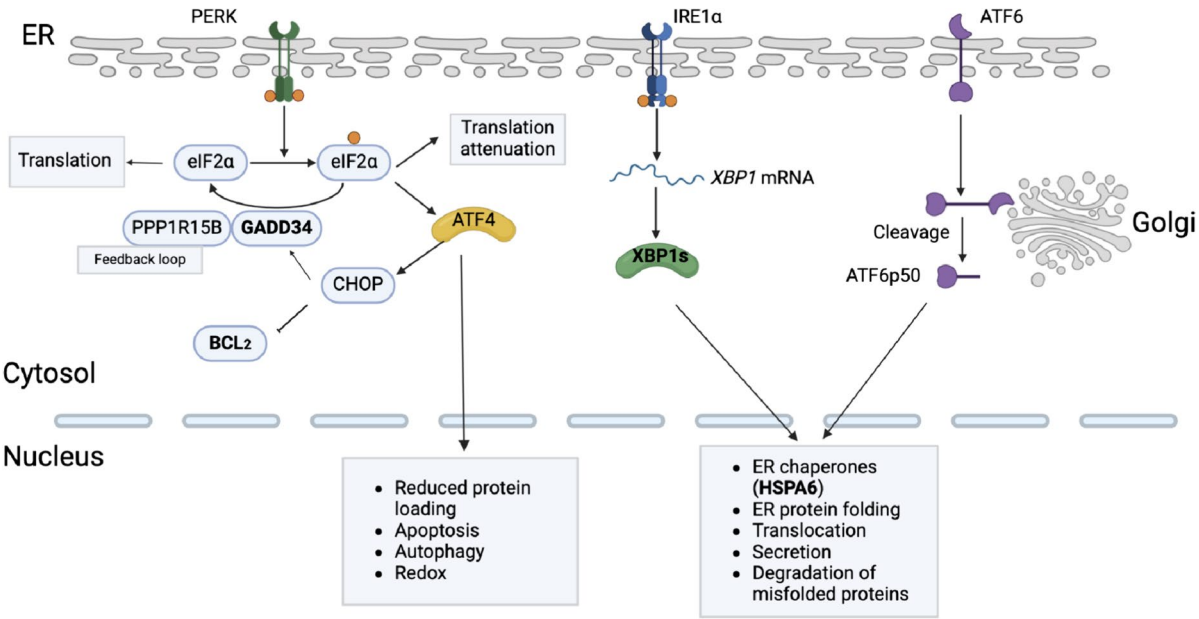


Fig. 6 ER protein processing pathways involved in this study, including UPR signaling through PERK, IRE1α, and ATF6. XBP1s represent multiple X-box binding proteins. CHOP represents C/EBP-homologous protein (CHOP). PPP1R15B represents protein phosphate 1 regulatory subunit 15B. All abbreviations shown in the figure represent encoded proteins by corresponding genes

that encodes *XBP1* and activate its transcription. This can further upregulate the genes involved in the expression of ER chaperones, ER protein translocation, ER protein folding, ER protein secretion, and the degradation of misfolded proteins (Calfon et al. 2002; Kaufman 2002; Lee et al. 2003; Yoshida et al. 2001). In this research, overexpressed XBP1 theoretically upregulates the transcription of genes that can facilitate ER protein folding, translocation, and secretion; which is beneficial for protein synthesis and capsid formation.

UPR signaling through *ATF6*: Acting parallel to *XBP1*, *ATF6* can also be cleaved and activated under ER stress, thus further upregulating the genes that encode ER chaperones and enzymes (Shoulders et al. 2013; Wu et al. 2007). Both genes can promote the proper folding, maturation, and secretion of proteins (Shoulders et al. 2013). HSPA6 is a heat shock protein that can function as an ER chaperone to facilitate proper folding and protect cells from stress (Munro and Pelham 1986). Heat shock proteins have been reported to participate in various stages of viral infection (e.g., viral gene expression, virus replication, virus assembly) (Wan et al. 2020). The overexpression of *HSPA6* should also benefit vector production. In this study, *BCL2* was overexpressed to decrease the apoptosis induced by the transfection process and accumulated ER stress, increase cell performance, and thus increase productivity.

Improvement of rAAV productivity and quality by overexpression of ER protein processing genes in HEK293T cells

HEK293T *XBP1*+, *GADD34*+, *HSPA6*+, and *BCL2*+ stable cells exhibited improved genome titer and packaging efficiency than the parental cells in the high-producing condition. No significant impact was observed for any candidate gene overexpression in the low-producing condition. We hypothesized that in the high-producing condition, more protein synthesis leads to more viral protein accumulation in the ER, which increases ER stress.

The overexpression of the ER protein processing genes involved in adaptive UPR pathways assumes a more essential role in alleviating ER stress, thereby improving protein folding for vector preparation and vector packaging. Another noticeable observation was that the capsid amounts produced in the overexpressed HEK293T cells were comparable to the parental cells. The marginal impact on capsid titer, combined with a remarkable improvement in packaging efficiency, suggests that (a) the improved genome titer was achieved without altering the total capsids produced and (b) the enhanced viral protein folding and vector packaging in the overexpression group improved packaging efficiency.

Additionally, cell growth in the overexpressed groups was lower than that of the negative control group. With decreased VCD in the overexpressed cells, significant improvements in specific productivity were observed (approximately one-fold), with the overexpression of *BCL2*, *HSPA6*, *GADD34*, and *XBP1*, respectively. With the stable integration of these ER protein processing genes, the continuous expression of these GOIs might consume nutrients and substrates, slowing down overall cell growth. In addition, this more effective protein folding might increase the ER burden and promote cell apoptosis. Also, the enhanced viral packaging process consumes extra energy, and this might explain the low performance in cell growth.

Broad applicability: various HEK293 cell lines and serotypes (AAV2 and AAV8)

To evaluate how parental cell line and AAV serotypes affect ER protein processing genes in terms of vector productivity and quality, we developed stable cell pools with overexpressed genes in both HEK293T and HEK293 cell lines and then compared the production of two serotypes (AAV2 and AAV8) in HEK293T cells. Similarly to what we observed in HEK293T, we find that the overexpression of all candidate genes in HEK293 cells significantly increased productivity in the high-producing condition, while the genome titer was either comparable to or lower than the low-producing condition. This overall consistency in titer improvement across both HEK293T and HEK293 cell lines confirms the importance of these ER processing genes across different cell lines. Additionally, although the extent of the improvement might vary, we observed a similar trend of rising AAV2 and AAV8 productivity under high-producing conditions in the HEK293T cells. We observed that the genes involved in ER processing play a role in enhancing rAAV production, and this effect is consistent across various cell lines and serotypes.

In the future, various overexpression combinations of ER processing genes may offer a much better improvement in vector productivity and overall product quality. Additionally, altering the expression levels of these genes could serve as a way to fine-tune the degree of improvement, similar to a previous study on lentiviral vector production (Formas-Oliveira et al. 2020). Additionally, a thorough investigation of cellular mechanisms might be necessary to fully elucidate the connection between ER stress, triggered UPR, and viral vector productivity. Cell stress assay that can produce fluorescent protein when the cells endure ER stress or UPR can be applied as the marker of cellular stress in the future (Harlen et al. 2019; Mohammed et al. 2023). Moreover, as more researchers devote their time and resources to constructing stable producers for vector manufacturing, any promising insights obtained from the transient transfection production process (e.g., enhancing ER processing genes for higher packaging efficiency) could offer valuable inputs and could

be applied to a stable production system, thus alleviating the existing bottleneck in rAAV biomanufacturing from a cellular physiology perspective (Escandell et al. 2023; Fu et al. 2023b; Jalšić et al. 2023; Lee et al. 2022).

Overall, this study unveils the significant role of ER protein processing pathways in rAAV capsid synthesis and vector preparation. The observed genome titer improvement is beneficial from improved protein formation, capsid quality, and enhanced packaging efficiency. The overexpression of *XBP1*, *HSPA6*, and *GADD34* in HEK293T cells increased rAAV productivity by up to 100%, increased the specific rAAV productivity by up to 78%, and increased the full capsid ratio value up to 37%. Furthermore, this productivity improvement was validated across different HEK293 cell lines (HEK293 and 293T cell lines) and various AAV serotypes (AAV2 and 8), although the extent of the enhancement might vary slightly. These findings further validate the molecular features identified from omics studies, and they suggest that cell line engineering can serve as a promising strategy to enhance vector yield and quality.

Supplementary Information The online version contains supplementary material available at https://doi.org/10.1007/s00253-024-13281-5.

Acknowledgements We would like to express our gratitude to all AMBIC member companies for their mentorship and their financial support.

Author contribution QF, YW, and SY conceived the manuscript. QF, YW, and JQ performed the experiments. QF and YW analyzed the data. QF and YW composed the manuscript. DM, DX, and SY provided input on data analyses. DM and SY edited the manuscript. SY secured the research fund and supervised the research project. All authors read and approved the final manuscript.

Funding This work was funded and supported by the Advanced Mammalian Biomanufacturing Innovation Center (AMBIC) through the Industry–University Cooperative Research Center Program under the U.S. National Science Foundation (Grant number 2100075). This work was also partially funded by the Massachusetts Life Science Center (MLSC).

Data availability The data that support the findings of this study are provided in the supplemental file.

Declarations

Conflict of interest The authors declare no competing interests.

Human and animal rights This article does not contain any studies with human participants or animals performed by any of the authors.

References

Calfon M, Zeng H, Urano F, Till JH, Hubbard SR, Harding HP, Clark SG, Ron D (2002) IRE1 couples endoplasmic reticulum load to secretory capacity by processing the *XBP-1* mRNA. Nature 415(6867):92–96

Chung C-H, Murphy CM, Wingate VP, Pavlicek JW, Nakashima R, Wei W, McCarty D, Rabinowitz J, Barton E (2023) Production of rAAV by plasmid transfection induces antiviral and inflammatory responses in suspension HEK293 cells. Mol Ther-Methods Clin Dev 28:272–283

Credle JJ, Finer-Moore JS, Papa FR, Stroud RM, Walter P (2005) On the mechanism of sensing unfolded protein in the endoplasmic reticulum. Proc Natl Acd Sci 102(52):18773–18784

De BP, Cram S, Lee H, Rosenberg JB, Sondhi D, Crystal RG, Kaminsky SM (2023) Assessment of residual full-length SV40 large T antigen in clinical-grade adeno-associated virus vectors produced in 293T cells. Hum Gene Ther 34(15–16):697–704

Escandell J, Moura F, Carvalho SB, Silva RJ, Correia R, Roldão A, Gomes-Alves P, Alves PM (2023) Towards a scalable bioprocess for rAAV production using a HeLa stable cell line. Biotechnol Bioeng 120(9):2578–2587

Formas-Oliveira AS, Basílio JS, Rodrigues AF, Coroadinha AS (2020) Overexpression of ER protein processing and apoptosis regulator genes in human embryonic kidney 293 cells improves gene therapy vectors production. Biotechnol J 15(9):1900562

Fu Q, Lee YS, Green EA, Wang Y, Park SY, Polanco A, Lee KH, Betenbaugh M, McNally D, Yoon S (2023a) Design space determination to optimize DNA complexation and full capsid formation in transient rAAV manufacturing. Biotechnol Bioeng 120(11):3148–3162

Fu Q, Polanco A, Lee YS, Yoon S (2023b) Critical challenges and advances in recombinant adeno-associated virus (rAAV) biomanufacturing. Biotechnol Bioeng 120(9):2601–2621

Grieger JC, Soltys SM, Samulski RJ (2016) Production of recombinant adeno-associated virus vectors using suspension HEK293 cells and continuous harvest of vector from the culture media for GMP FIX and FLT1 clinical vector. Mol Ther 24(2):287–297

Gu B, Bhat V, Dong W, Pham H, Pubill S, Kasaraneni N, Onishi E, Vitelli F, Seth A (2018) Establishment of a scalable manufacturing platform for in-silico-derived ancestral adeno-associated virus vectors. Cell Gene Ther Insights 4(S1):753–769

Harding HP, Zhang Y, Scheuner D, Chen J-J, Kaufman RJ, Ron D (2009) Ppp1r15 gene knockout reveals an essential role for translation initiation factor 2 alpha (eIF2α) dephosphorylation in mammalian development. Proc Natl Acd Sci 106(6):1832–1837

Harlen KM, Roush EC, Clayton JE, Martinka S, Hughes TE (2019) Live-cell assays for cell stress responses reveal new patterns of cell signaling caused by mutations in rhodopsin, α-synuclein and TDP-43. Front Cell Neurosci 13:535

Hassler JR, Scheuner DL, Wang S, Han J, Kodali VK, Li P, Nguyen J, George JS, Davis C, Wu SP (2015) The IRE1α/XBP1s pathway is essential for the glucose response and protection of β cells. PLoS Biol 13(10):e1002277

Hockemeyer D, Soldner F, Beard C, Gao Q, Mitalipova M, DeKelver RC, Katibah GE, Amora R, Boydston EA, Zeitler B (2009) Efficient targeting of expressed and silent genes in human ESCs and iPSCs using zinc-finger nucleases. Nat Biotechnol 27(9):851–857

Hollien J, Lin JH, Li H, Stevens N, Walter P, Weissman JS (2009) Regulated Ire1-dependent decay of messenger RNAs in mammalian cells. J Cell Biol 186(3):323–331

Jalšić L, Lytvyn V, Elahi SM, Hrapovic S, Nassoury N, Chahal PS, Gaillet B, Gilbert R (2023) Inducible HEK293 AAV packaging cell lines expressing Rep proteins. Mol Ther-Methods Clin Dev 30:259–275

Kaufman RJ (2002) Orchestrating the unfolded protein response in health and disease. J Clin Invest 110(10):1389–1398

Kim JY, Kim Y-G, Lee GM (2012) CHO cells in biotechnology for production of recombinant proteins: current state and further potential. Appl Microbiol Biotechnol 93:917–930

Lee A-H, Iwakoshi NN, Glimcher LH (2003) XBP-1 regulates a subset of endoplasmic reticulum resident chaperone genes in the unfolded protein response. Mol Cell Biol 23(21):7448–7459

Lee Z, Lu M, Irfanullah E, Soukup M, Hu W-S (2022) Construction of an rAAV producer cell line through synthetic biology. ACS Synth Biol 11(10):3285–3295

Lu M, Lee Z, Hu WS (2024) Multi-omics kinetic analysis of recombinant adeno-associated virus production by plasmid transfection of HEK293 cells. Biotechnol Prog e3428 https://doi.org/10.1002/btpr.3428

Metkar M, Pepin CS, Moore MJ (2024) Tailor made: the art of therapeutic mRNA design. Nat Rev Drug Discov 1:67–83

Mohammed NA, Lewis K, Hodges N, Michelangeli F (2023) Mechanisms of cell death induced by hexabromocyclododecane (HBCD) involves apoptosis, autophagy, and ER stress. J Biochem Molecular Tox e23397 https://doi.org/10.1002/jbt.23397

Munro S, Pelham HR (1986) An Hsp70-like protein in the ER: identity with the 78 kd glucose-regulated protein and immunoglobulin heavy chain binding protein. Cell 46(2):291–300

Novoa I, Zeng H, Harding HP, Ron D (2001) Feedback inhibition of the unfolded protein response by GADD34-mediated dephosphorylation of eIF2α. J Cell Biol 153(5):1011–1022

Ran F, Hsu PD, Wright J, Agarwala V, Scott DA, Zhang F (2013) Genome engineering using the CRISPR-Cas9 system. Nat Protoc 8(11):2281–2308

Ron D, Walter P (2007) Signal integration in the endoplasmic reticulum unfolded protein response. Nat Rev Mol Cell Biol 8(7):519–529

Rutkowski DT, Kaufman RJ (2004) A trip to the ER: coping with stress. Trends Cell Biol 14(1):20–28

Sha S, Maloney AJ, Katsikis G, Nguyen TN, Neufeld C, Wolfrum J, Barone PW, Springs SL, Manalis SR, Sinskey AJ (2021) Cellular pathways of recombinant adeno-associated virus production for gene therapy. Biotechnol Adv 49:107764

Shin S, Kim SH, Shin SW, Grav LM, Pedersen LE, Lee JS, Lee GM (2020) Comprehensive analysis of genomic safe harbors as target sites for stable expression of the heterologous gene in HEK293 cells. ACS Synth Biol 9(6):1263–1269

Shoulders MD, Ryno LM, Genereux JC, Moresco JJ, Tu PG, Wu C, Yates JR, Su AI, Kelly JW, Wiseman RL (2013) Stress-independent activation of XBP1s and/or ATF6 reveals three functionally diverse ER proteostasis environments. Cell Rep 3(4):1279–1292

Tigges M, Fussenegger M (2006) Xbp1-based engineering of secretory capacity enhances the productivity of Chinese hamster ovary cells. Metab Eng 8(3):264–272

Upton J-P, Wang L, Han D, Wang ES, Huskey NE, Lim L, Truitt M, McManus MT, Ruggero D, Goga A (2012) IRE1α cleaves select microRNAs during ER stress to derepress translation of proapoptotic Caspase-2. Science 338(6108):818–822

Wan Q, Song D, Li H, He M-l (2020) Stress proteins: the biological functions in virus infection, present and challenges for target-based antiviral drug development. Signal Transduct Target Ther 5(1):125

Wang D, Tai PW, Gao G (2019) Adeno-associated virus vector as a platform for gene therapy delivery. Nat Rev Drug Discov 18(5):358–378

Wang Y, Fu Q, Lee YS, Sha S, Yoon S (2023) Transcriptomic features reveal molecular signatures associated with recombinant adeno-associated virus production in HEK293 cells. Biotechnol Prog e3346 https://doi.org/10.1002/btpr.3346

Wang Y, Fu Q, Park SY, Lee YS, Park S-Y, Lee D-Y, Yoon S (2024) Decoding cellular mechanism of recombinant adeno-associated virus (rAAV) and engineering host-cell factories toward intensified viral vector manufacturing. Biotechnol Adv 108322 https://doi.org/10.1016/j.biotechadv.2024.108322

Wu J, Rutkowski DT, Dubois M, Swathirajan J, Saunders T, Wang J, Song B, Yau GD-Y, Kaufman RJ (2007) ATF6α optimizes long-term endoplasmic reticulum function to protect cells from chronic stress. Dev Cell 13(3):351–364

Yoshida H, Matsui T, Yamamoto A, Okada T, Mori K (2001) XBP1 mRNA is induced by ATF6 and spliced by IRE1 in response to ER stress to produce a highly active transcription factor. Cell 107(7):881–891

Zhao H, Lee K-J, Daris M, Lin Y, Wolfe T, Sheng J, Plewa C, Wang S, Meisen WH (2020) Creation of a high-yield AAV vector production platform in suspension cells using a design-of-experiment approach. Mol Ther-Methods Clin Dev 18:312–320