Development of an Improved Transient Triple Transfection System with Focus on Engineering the Helper Plasmid for Enhanced rAAV Production

Doctoral Thesis

by

Thilo Pohle

Department of Chemical and Biological Engineering

University of Sheffield

Sheffield

August 2023

Development of an Improved Transient Triple Transfection System with Focus on Engineering the Helper Plasmid for Enhanced rAAV Production

Doctoral Thesis

by

Thilo Pohle

Department of Chemical and Biological Engineering

University of Sheffield

Sheffield

August 2023

1st Supervisor: Prof David C James

2nd Supervisor: Dr Adam J Brown

Registration Number: 190235268

Email Address: thpohle1@sheffield.ac.uk

Date of Submission: 18th August 2023

Table of Contents

2 Abstract

Recombinant adeno-associated virus (rAAV) has established itself as the predominant *in vivo* gene therapy vector in clinical applications. AAV's outstanding biological traits and efficacy for *in vivo* gene transfer, numerous successful clinical trials, and multiple approved products are driving increasing vector demand. However, low production yields remain a major challenge in rAAV manufacturing and contribute to the enormous costs per dose of rAAV-based gene therapies. As the well-established HEK293-based transient triple transfection systems, involving Rep/Cap, Helper, and ITR/GOI plasmids, remain the most used manufacturing methods in academia and industry, a key aspect of vector production is the molecular design of rAAV expression plasmids. Therefore, an improved plasmid system was designed based on the industry standard expression system of REGENXBIO. The utilisation of the novel Rep/Cap and Helper plasmids with individually controllable genes yielded a 1.8-fold increase in virus genome (VG) titre.

This thesis focuses primarily on the Helper plasmid, which provides essential Adenovirus genes that orchestrate cellular and viral gene expression, manipulating the host cell environment for efficient AAV virus replication. Systematic removal of redundant sequences and molecular genetic changes streamlined and modularized the plasmid, optimising its structure and simplifying future genetic engineering. The detailed analysis of the Helper plasmid genes provided novel insights into their influence on rAAV production and resulted in (i) optimised E4 open reading frame subsets increasing VG titres, (ii) unveiling the in *E2A*'s 5' UTR located *L4-33K/22K* as a new beneficial Helper function, and (iii) nearly halving the plasmid size to 8.6 kb.

A variety of genetic engineering and synthetic biology tools were then employed to assess the relative expression demands of individual components for high-titre rAAV production. Separation of the Helper genes into distinct plasmids, determination of E2A and E4 mRNA levels, the utilisation of heterologous promoters and bicistronic Helper gene expression provided valuable insights into the essential roles of Helper genes and foundational knowledge about their transcription requirements. Sufficiently strong *E2A* transcription emerged as a pivotal prerequisite for high rAAV titres. The results highlighted the complex molecular interactions of AAV replication, underpinning that precise balancing and regulation of components are imperative for an optimised plasmid system.

Enhancement of the Helper plasmid composition was achieved through iterative genetic engineering, involving reconfiguration of gene positions and orientations, along with the integration of CMV-derived synthetic promoters. These innovative plasmid designs combined positive outcomes of this research, culminating in a significant 2.9-fold increase in rAAV yields compared to the initial Helper plasmid.

Taken together, this project underscored the preeminent role of Helper plasmid functions in determining rAAV yields, improved both rAAV production titres and the plasmid system itself, and established a solid foundation for further research and development of rAAV expression systems.

3 Introduction

Gene therapeutics hold remarkable potential to not only alleviate but cure a multitude of genetic and acquired diseases and therefore help countless patients worldwide. Presently, they represent biotechnologies most exceptional tools and promises of breakthrough innovations to advance modern medicine into a new era. Efficient delivery of gene therapeutic nucleic acids, which have the potential to introduce advantageous alterations to cells and heal patients, requires a vector capable of effectively shuttling the genetic payload into target cells. The adeno-associated virus (AAV) possesses outstanding properties to be utilised as such a vector for *in vivo* gene therapies. However, pressing need for meeting current high demands and addressing production costs necessitates comprehensive research into enhanced production technologies.

3.1 AAV's Field of Application: Gene Therapy

Within the realm of modern medicine, gene therapeutics have emerged as cutting-edge solutions, offering promising remedies for a wide range of genetic disorders. These disorders including haemophilia A and B, retinal dystrophy and spinal muscular atrophy (SMA). Four previously life impairing or life-threatening inherited conditions are now treatable with FDA- and/or EMA-approved AAV-based gene therapeutics which are now in the clinic (ASGCT & Citeline, 2023; FDA, 2023).

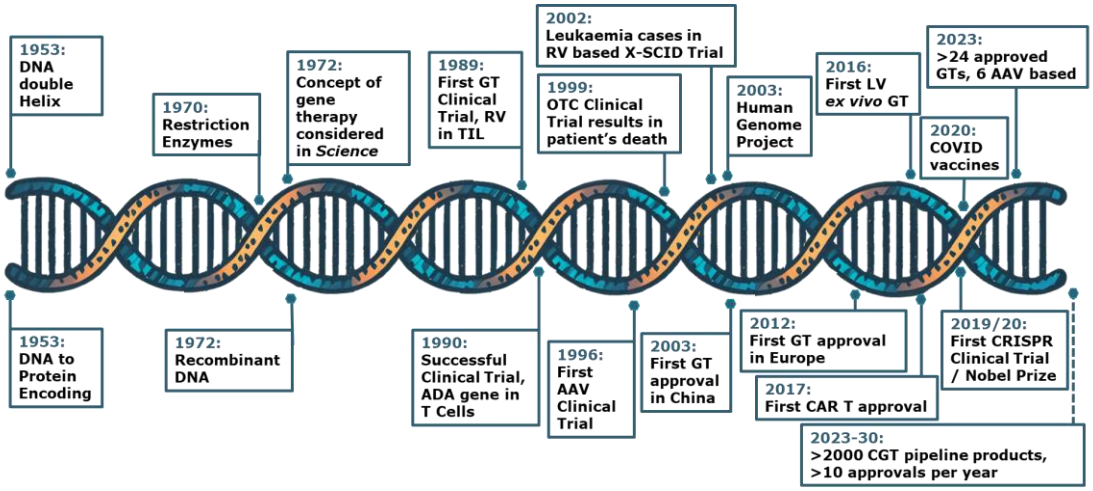


Figure 1: Milestones of the gene therapy field.

Gene therapy is generally defined as the therapeutic delivery of nucleic acids into cells of a patient to cure a disease, but there are numerous different mechanisms how the inserted gene can act and deploy its therapeutic effect in the cell (Scheller & Krebsbach, 2009). The very first ideas and attempts to use viruses for gene delivery date back as far as 1960s/1970s (Nirenberg, 1967). Progress in ethical considerations, virology, and advancements in biological and genetic engineering during the 1970s and 1980s led to the conceptualisation and exploration of viral vectors for gene delivery, as well as first experiments in cell cultures and animal models (Friedmann, 1992; Miller, 1992; Mulligan, 1993; Nicholl, 2023). This pursuit marked a significant milestone in the field of gene therapy when, in 1990, the first approved gene therapy procedure successfully cured a four-year-old patient afflicted with severe combined immunodeficiency (SCID) due to adenosine deaminase deficiency (Anderson et al., 2008; Blaese et al., 1995). However, despite these early achievements, reports of adverse effects and fatalities dealt severe blows to the field, culminating in the halt of all clinical trials in the United States. In Europe, four boys treated against X-SCID died of leukaemia, while in the USA an 18-year-old patient with ornithine transcarbamylase (OTC) deficiency received treatment as part of a study in 1999, using an adenoviral vector carrying the OTC gene and succumbed to a severe immune reaction four days later, potentially induced by antibody-dependent enhancement (Arabi et al., 2022; Howe et al., 2008). Although such tragedies reduced euphoria and posed a challenge to the field, slowing it down for a while, they also prompted regulatory reforms and motivated further research and technological advancements, leading to improved vector design and safety in gene therapy.

Early studies in gene therapy primarily focused on addressing single-gene defects such as cystic fibrosis, sickle cell anaemia, and muscular dystrophy by replacing or disrupting the defective gene. However, the field has significantly evolved over the past three decades, leading to the development of various therapeutic strategies. Scientific and technological advancements have expanded the applications of gene therapeutics beyond monogenetic disorders, of which many often are rare orphan diseases, to encompass more complex, systemic, and prevalent disorders such as central nervous system conditions and cancer. In addition to gene augmentation, the scientific community has pioneered gene inhibition, suicide gene therapy, as well as gene editing (Bulaklak & Gersbach, 2020; Dunbar et al., 2018; Lapteva et al., 2020). Regarding gene editing, recent breakthroughs in nuclease editing tools like CRISPR/Cas and TALEN are currently opening opportunities for novel therapy approaches, which add to the already comprehensive gene therapy toolbox (Raguram et al., 2022). It is worth noting that all currently considered gene

therapeutics in the western world are solely targeted at somatic cells, and germline editing is prohibited in most jurisdictions, including the US, China, and Europe. The dawn of CRISPR systems and the claimed gene editing of human embryos by a Chinese group, however, underscore impact, significance, but also risks of advances in gene therapy and therefore open up several ethical questions and problems once again (Raposo, 2019).

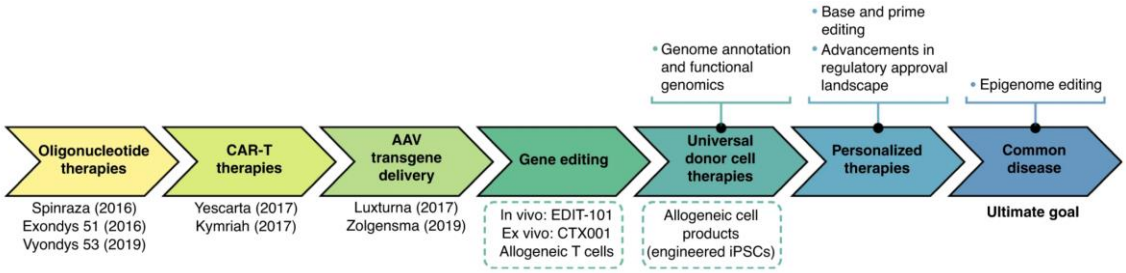


Figure 2: Timeline illustrating gene therapy milestones and advancement in treatment options towards approaches for prevalent diseases, from (Bulaklak & Gersbach, 2020).

Innovations and efforts in the field yielded over 3600 clinical trials to date. Among these, more than 1000 trials are currently ongoing, and over 2000 gene therapies are in various stages of development, from preclinical to pre-registration (Journal of Gene Medicine, 2023). Thus far resulting in 24 approved gene therapies (including genetically modified cell therapies, excluding RNA therapies). However, with the FDA's goal of approving 15 to 20 cell and gene therapies (CGTs) a year and the prediction of 13 new CGTs in 2023 the age of cell and gene therapy medicines is still just beginning (ASGCT & Citeline, 2023; Shahryari et al., 2019; Stephen Kemler & Adam Lohr, 2022).

Generally, gene therapy treatments are broadly categorized into two different administration approaches: *in vivo* and *ex vivo* procedures. *Ex vivo* approaches involve genetic modification of cultured heterologous or autologous cells outside the patient, which are later administered to the patient. Many gene therapies currently approved by the FDA, particularly those used in cancer treatment, fall under this category (Arabi et al., 2022; Lapteva et al., 2020). This includes all FDA-approved CAR T (chimeric antigen receptor T-cell) therapies, although ongoing pre-clinical trials may result in the emergence of *in vivo* CAR-T therapies (Michels et al., 2022; Xin et al., 2022). In contrast *ex vivo* approaches, *in vivo* gene therapies involve the direct delivery of therapeutic nucleic acids to the patient. For improved delivery efficiency of the genetic material into the cells viral and non-viral vectors are used. While both types have advantages and disadvantages, viral vectors have historically been the preferred choice due to their higher efficiency, despite their inferior safety

profile. Due to viruses' natural abilities to transfer genetic material into host-cells, viral vectors are naturally well equipped to overcome the barriers to deliver their payload into the targeted cells' nucleus. Distinctions are made between integrative vectors, like lentivirus or retroviruses, and non-integrative vectors, like the Adenovirus (AdV). The ideal vector would be inexpensive and easy to produce in large amounts, efficiently transduce exclusively targeted cells, show long-term expression of the gene of interest (GOI), and cause no negative adverse effects such as inflammation, cytotoxicity, or any kind of immune responses. Unfortunately, the ideal vector has yet to be developed (Kaji and Leiden, 2001). Nonetheless, the adeno-associated virus (AAV) possesses several advantages over other viral vector systems and fulfils many of the aforementioned criteria (Atchison et al., 1965; Dunbar et al., 2018; Naso et al., 2017).

Among the extraordinary characteristics of adeno-associated virus for gene therapeutic purposes is its replication incompetence in absence of a Helper virus, which is a significant safety feature. AAV is inherently non-pathogenic, non-cytotoxic and causes a generally low immune response, which sets it apart from other viruses such as Adenoviruses. Unlike many other viruses, it can infect dividing and also non dividing cells in G0-phase. With its spectrum of serotypes AAV offers a broad range of tissues that can be infected with different specificities, and some can even cross the blood brain barrier. Wild-type (wt) AAV shows beneficially low random integration events. It even comes with the unique ability of site directed integration into the AAVS1 locus on chromosome 19, although this has little relevance for recombinant AAV vectors (rAAV) due to their lack of the *rep* gene. However, the inverted terminal repeats (ITRs) at both ends of AAV's genome enable it to persist long-term as an episomal concatemer (Hastie & Samulski, 2015; Naso et al., 2017; Samulski & Muzyczka, 2014; D. Wang et al., 2019).

Nevertheless, there are also certain disadvantages in using AAV in gene therapy applications. Firstly, a considerable proportion of the population already possesses pre-existing immunity against wild-type AAVs. Up to 50% to 96% of the population have been previously infected with AAV serotype 2 and a significant portion of the population carries capsid-specific neutralizing antibodies and memory T cells against various serotypes (Büning et al., 2008; Calcedo et al., 2009; Klamroth et al., 2022). Despite the existence of at least 14 natural serotypes and over 100 variants, these statistics indicate the need for capsid modifications (Gao et al., 2004, 2005). Such modifications can be beneficial for expanding the range of diseases and target cell types that can be effectively addressed, with the potential of more precise targeting and higher efficiency through modern protein engineering (Büning & Srivastava, 2019). Secondly, AAV's naturally small size and genome limit the payload

capacity to only 4.7 kb, making it unsuitable for delivering large genes unless this limitation is overcome by employing dual vectors that are annealed and spliced after transduction (Chamberlain et al., 2016; McClements & MacLaren, 2017). Further, concerns exist regarding the stability or better longevity of transgenes persisting as concatemers, particularly in rapidly dividing cell types, as well as the immunogenicity of rAAVs, especially in higher dosages. Additionally, production-related issues also persist, including the presence of undesired sequences such as host genome fragments and bacterial plasmid DNA during the production process. Another quality issue involves the large proportion of empty capsids that can be prevalent in rAAVs production processes. Lastly, vector titres of current rAAV productions still require to be increased to meet demands and reduce therapeutic price tags (Merten, 2016; Naso et al., 2017; Qu et al., 2019).

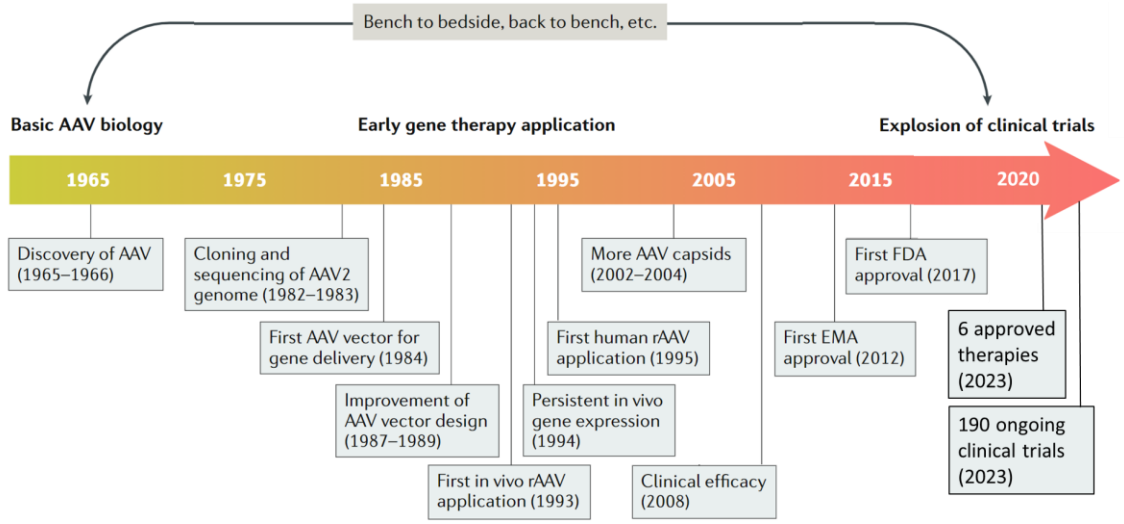


Figure 3: Historical development of rAAV gene therapy vectors from bench to bedside, amended illustration from Gao et al. (2005).

Regardless of the mentioned downsides and thanks to all its upsides, including its relatively simple manufacturing, AAV remains the most utilized *in vivo* gene therapy vector worldwide. rAAVs have been employed in 350 clinical trials, making up 9.5 % of all gene therapy clinical trials (Journal of Gene Medicine, 2023). Milestones on its way becoming the preferred vector of choice for *in vivo* gene therapies are shown in Figure 3. The fact that the cell and gene therapy space is a very diverse space, often considered as a single discipline, can lead to an underestimation of AAV's therapeutic importance. To give an overview of this market and AAVs impact, a brief excerpt of a marketing exercise that was performed during this thesis is presented.

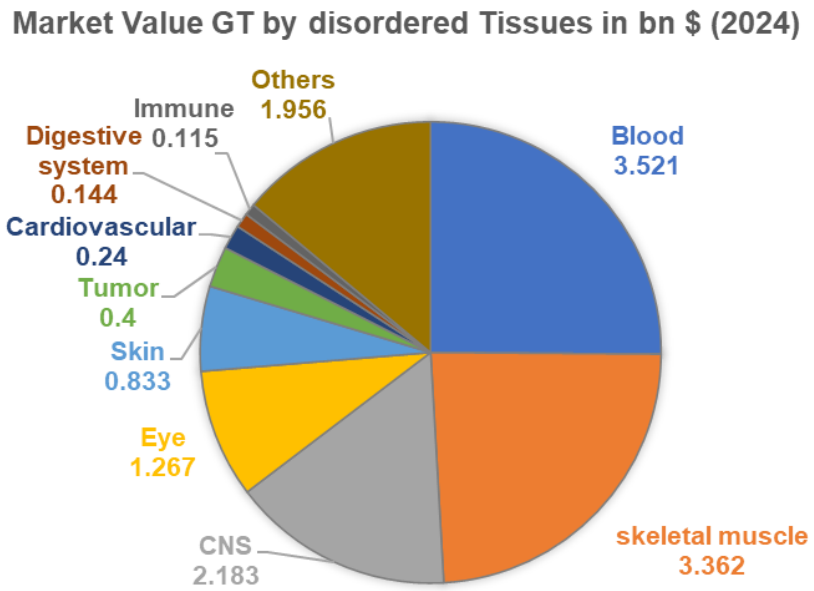
It should be note that financial figures of many resources regarding CGT should be interpreted with caution due to variations resulting from the field's heterogeneity, differing categorisations by financial providers, as well as undisclosed and

estimated data, which can lead to significant discrepancies. Nevertheless, it is certain that the CGT market, as a whole, is experiencing substantial growth, with an estimated value of $1.9 billion in 2019 which is projected to reach $42.6 billion by 2030. As of 2019, 38 % of this market value belongs to gene modified cell therapies (*ex vivo* gene therapies). *In vivo* gene therapies represent 45 % (bcc Research, 2022). The referenced market report values non-modified cell therapies with only 17 % of the CGT market value, demonstrating not only the importance of gene modifying therapies, but also their extraordinary price tags. Even in 2023, despite over 20 additionally gene therapies being approved, less than half of all cell and gene therapies are genetically modified. However, their compound annual growth rate (CAGR) for the next years (2021-2026) has a high evaluation with 31.9% for gene modified cell therapies and 34.1% for gene therapies, showing trust in upcoming pipeline products (bcc Research, 2022). Looking at those therapeutic pipelines, meaning products that are in between pre-clinical and pre-registration, currently 72 % belong to *ex vivo* therapies. Cancer and rare diseases, of which most are also oncology based, being the predominantly targeted disorders (ASGCT & Citeline, 2023). This demonstrates the current dominance of CAR-T's in CGT pipelines.

Our own research of 50 selected cell and gene therapy companies' pipelines draws a similar picture in many ways but shows that only about 40% are ex vivo products. Several factors may contribute to this slightly shifted distribution, including the maturity and progression of clinical studies for different types of products. Among other things industry acquisitions result in a higher proportion of products in advanced clinical stages compared to non-industry sponsored products. Additionally, a higher percentage of gene therapeutic products already advanced to phase II and III studies than cell-based immune-oncology products (ARM, 2022). Notably, as of Q1 2023, RNA therapies represent 30% of all gene therapy pipeline products, a 3% increase from ASGCT's previous report in Q1 2021, likely reflecting the field's development since the successful introduction of Moderna and Pfizer/BioNTech's COVID-19 vaccines (ASGCT & Citeline, 2023; ASGCT & PharmaIntelligence, 2021). A 3% increase in the proportion of all gene therapy trials compared to ASGCT's first landmark report in Q1 2021 is likely an indication of the

field's development since the success of Moderna and Pfizer/BioNTech's COVID-19 vaccines.

A



B

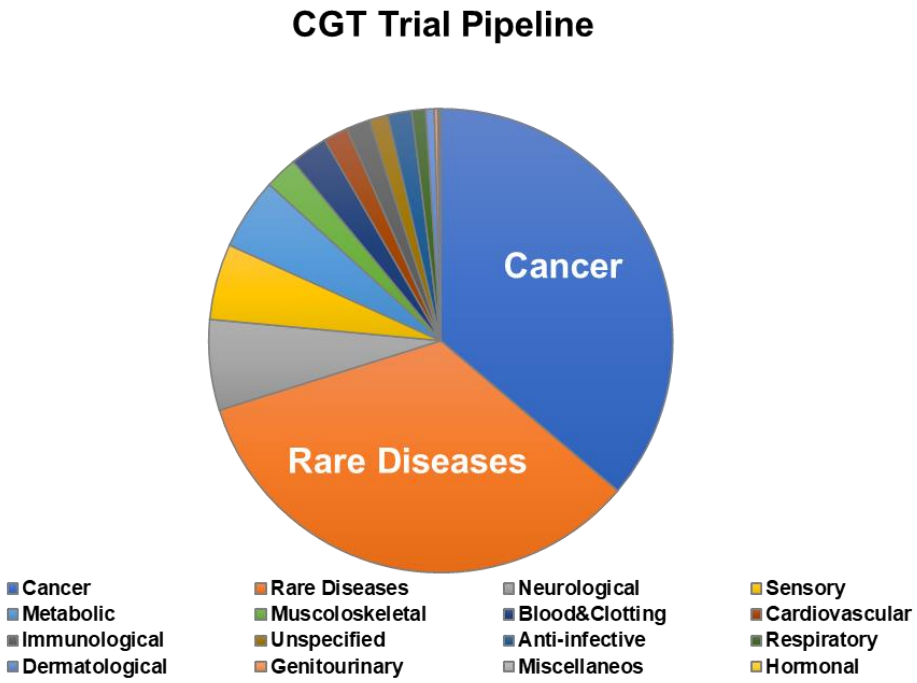


Figure 4: Gene therapy market overview. A) Market value by disease area in 2024 estimated (EveluatePharma & BCG, 2020). B) Distribution of cell and gene therapy pipeline products (pre-clinical to pre-registration) (ASGCT & Citeline, 2023).

Our own analysis of gene therapy pipelines reveals that while nearly all tissues and organs are targeted, the top five targets, namely T-cells, liver, central nervous system (CNS), hematopoietic stem cells (HSC), and eye, account for 76% of all analysed products. The described emphasise on oncogenic gene therapy targets is

evident, as 22% of the considered products target T-cells. Among the diverse range of vectors that can be utilized for gene therapeutics, including mRNA-carrying lipid nanoparticles and other viruses, like Herpes simplex, Adeno- and retro viruses, lentivirus vectors are predominantly used for CAR-T and other ex vivo approaches. As a result, lentivirus vectors have a usage rate of at least 48% for products in ongoing clinical trials (see Figure 5).

The current focus on lentivirus for ex vivo applications and adeno-associated virus (AAV) for *in vivo* applications is not as pronounced in total clinical trial figures, given to the 30-year history of gene therapy clinical trials, in which many studies in past and present have been conducted employing alternative vectors. Hence, the total usage of lentivirus is only 9.9% of all studies and AAV vectors also have a relatively low overall usage rate of only 9.5% (Journal of Gene Medicine, 2023). Nevertheless, AAV remains the most widely used *in vivo* vector in ongoing trials, accounting for at least 46% according to the Alliance for Regenerative Medicine's 2021 analysis (ARM, 2021). A graphic of their 2022 report illustrates best the predominance of lentivirus as ex vivo (cell-based immune-oncology) and AAV as *in vivo* gene therapy vectors (ARM, 2022). Further, AAV's influence and usage rate might increase in the future. More prevalent diseases will be treated with *in vivo* GT's and there is an anticipated transition towards *in vivo* applications for numerous cell and gene therapies, specifically in the context of lymphocytic adoptive cell therapies. Part of this development will be novel, emerging strategies using AAVs in diverse cancer treatments, such as the innovative, universal AAV immune-gene therapy approach of the recently founded company Siren Biotechnology (Grinstein, 2023; Nawaz et al., 2021).

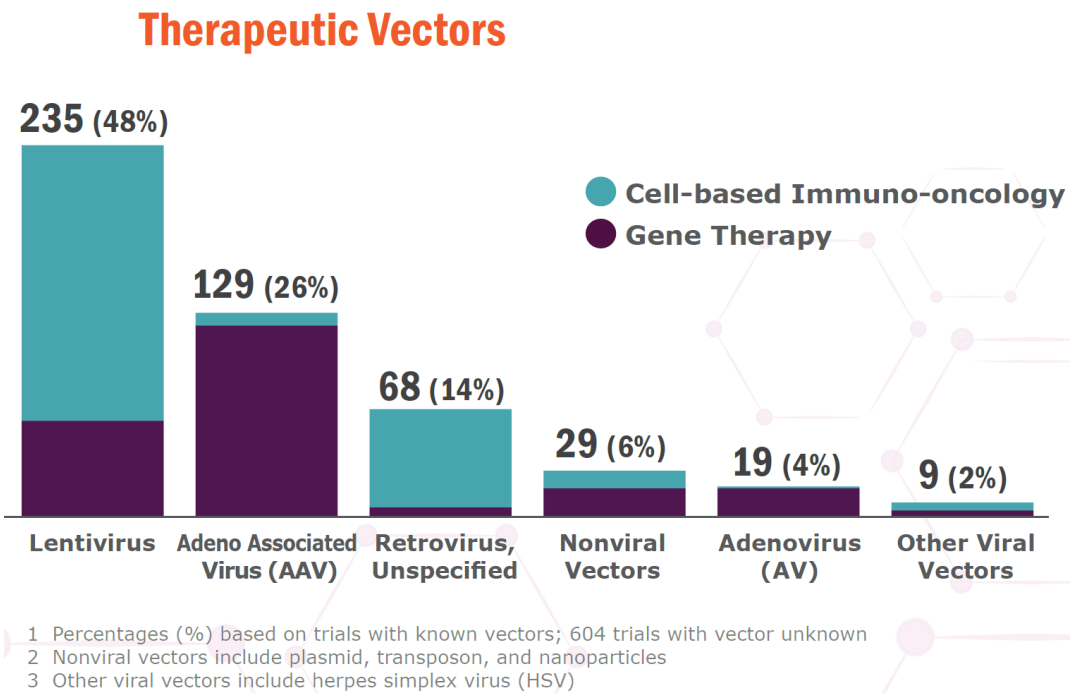


Figure 5 – Vectors used in gene therapy clinical trials. Graphic taken from the Alliance for Regenerative Medicine's 2022 report (ARM, 2022)

AAV is not only the vector of choice for most *in vivo* gene therapeutics in clinical trials, such as genetic diseases like Duchenne muscular dystrophy, Wilson disease and Parkinson's Disease, butit is also the vector used for the first two FDA approved *in vivo* gene therapies. Luxturna (voretigene neparvovec), is an AAV2-based medication containing the RPE65 gene to cure retinitis pigmentosa and Leber's congenital amaurosis, two forms of the rare inherited retinal dystrophy. It was the first *in vivo* gene therapeutic approved by the FDA in 2017. Subsequently, Zolgensma (onasemnogene abeparvovec), an AAV9-based treatment targeting the rare genetic disease spinal muscular atrophy (SMA) type 1, received approval in 2019. Like all other approved AAV based therapeutics it only requires a one-time application. Instead of facing paralysis over time and a fatal prognosis, administration of one dose of 1×10^{14} vector genomes/kg body weight, containing an intact SMA1 copy, is intended to cure the infants for their lifetime. However, a single dose also comes with a price tag of $2.1 million, making it the world's most expensive drug until recently. It was surpassed by CSL Behring's AAV5-based haemophilia therapy Hemgenix (etranacogene dezaparvovec), costing $3.5 million. While these costs are significant, these medications bring tremendous improvements in quality of life for patients and have the potential to achieve complete cures. In contrast, traditional treatments often only alleviate symptoms and can be more expensive in the long run. Nonetheless, the high price of over 1 million euros for Europe's first AAV-based *in vivo* gene therapeutic Glybera

(alipogene tiparvovec) was part of its downfall and withdrawal from the market only five years after approval by the EMA.

AAV's success as a gene therapy vector is undeniable, positioning it as one of the most promising and the most popular *in vivo* vector in the current praxis. However, numerous challenges persist. Among these are the aforementioned high drug prices, which represent a significant hurdle for gene therapy applications. Addressing these challenges requires improved production strategies to help reduce costs and enhance the accessibility of AAV-based therapeutics for a larger patient population.

3.2 Virus Biology

The production of recombinant adeno-associated virus (rAAV) involves a complex interplay between viral and cellular genes, their expression, and interactions. An integral aspect of this process is the natural transformation of the host cell into a virus manufacturing facility that is orchestrated by the viruses to enable efficient expression of viral genes and the formation of functional viral particles. Therefore, it is imperative to gain a comprehensive understanding of the biology of AAV and its helper virus, Adenovirus, is essential for effectively engineering this intricate production system.

3.2.1 Adeno-associated Virus

Adeno-associated viruses belong to the Parvoviridae family, specifically the genus Dependovirus. The term “depend” refers to one of its most extraordinary and for its safe biotechnological use most important features. It is dependent on a Helper virus coinfection to replicate and proliferate. Accordingly, it was rather an indication of AAV’s fundamental biology than coincidence that, in 1965, AAV was first discovered as a contaminant in a preparation of Adenovirus (AdV) (see Figure 6A) (Atchison et al., 1965). The approximately 4-fold larger eponymous Adenovirus (~100 nm) is one of AAV’s natural Helper viruses. Other viruses that can function as Helper virus for AAV are for example herpes simplex virus (HSV), vaccinia virus, or human papillomavirus (Georg-Fries et al., 1984; Grieger & Samulski, 2012; Ogston et al., 2000; Schlehofer et al., 1986; Walz et al., 1997). AAV requires a coinfection with a Helper virus mainly for the replication of its genome and to enhance transcription of its own genes, both characterizing the entrance into AAVs lytic pathway of its biphasic life cycle (reviewed in 3.2.1.2). The distinct Helper virus functions of the necessary Ad proteins are reviewed in chapter 3.2.2.

AAV possesses a small (~22 nm), non-enveloped, icosahedral-shaped virion containing a 4.7 kb large genome of single stranded DNA (ssDNA). The capsid is assembled of 60 protein entities, composed of the three capsid proteins VP1, VP2

and VP3, in a 1:1:10 ratio. Albeit recent research by Wörner et al. (2021) demonstrated that virus capsid assembly is stochastic, leading to considerable deviations from this paradigmatic VP-composition ratio. The resulting capsid is icosahedral with a T = 1 symmetry, comprising two-, three- and five-fold symmetry axes (Balakrishnan & Jayandharan, 2014; Gonçalves, 2005; Samulski & Muzyczka, 2014; Weitzman & Linden, 2012). At each of the twelve five-fold symmetry axes a pore with a diameter of 1.2 nm is located (see Figure 6 B). Packaging of the genome into the assembled capsid takes palace through this pore (Bleker et al., 2005).

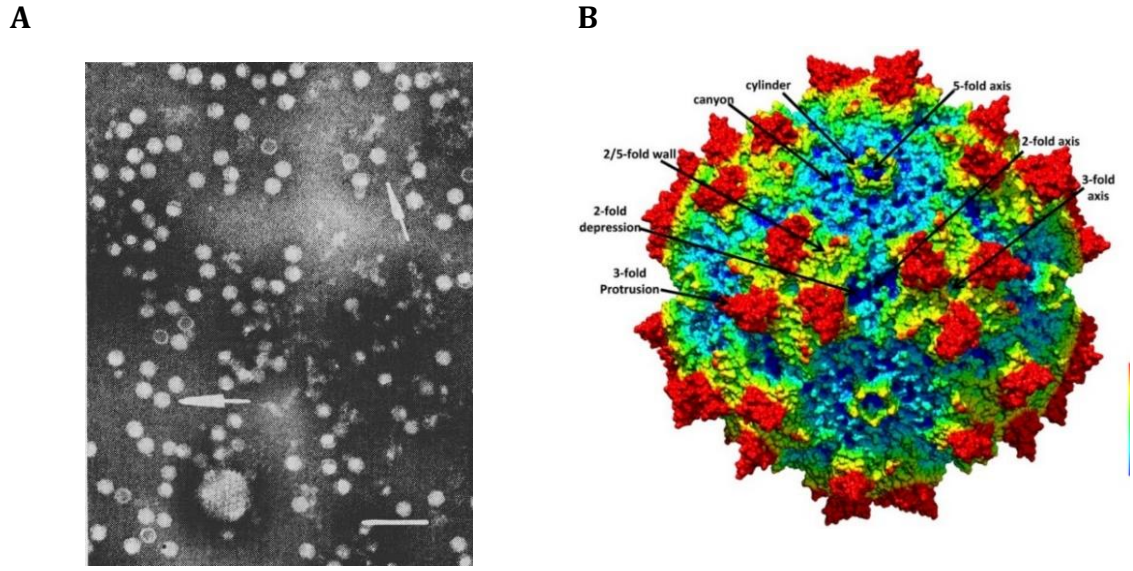


Figure 6 – (A) First report of AAV from (Atchison et al., 1965).Electron microscopy photography of a slightly, by differential centrifugation, purified Adenovirus (AdV) cell culture preparation. Smaller AAV particles are indicated by arrows, surround a much larger AdV particle. (B) AAV1 generated by Chimera from 60 VP monomers (RCSB PDB # 3NG9), representing the general capsid structure of AAVs. Radially colours show the distance from the capsid centre in Å. Symmetry axes as well as capsid surface features are indicated by labelled arrows. Graphic from Tseng and Agbandje-McKenna (2014).

All structural proteins of AAV share a common C-terminal sequence, which includes VP3 (62 kDa). Additionally, VP2 (73 kDa) and VP1 (87 kDa) contain unique N-terminal sequences. While the assembly of capsids can be achieved with VP3 alone, VP1 plays a crucial role in infection due to its conserved phospholipase A2 (PLA2) sequence. This sequence is essential for endosomal escape and overall infectivity (Balakrishnan & Jayandharan, 2014; Girod et al., 2002; Grieger & Samulski, 2005). On the other hand, VP2 is dispensable for infectivity, allowing for the insertion of large peptides at its N-terminus without affecting capsid assembly (Warrington et al., 2004). The serotype of an AAV is determined by its capsid structure, as the capsids affinity to specific cell surface receptors is decisive for the virion's entry into the cell and therefore its infectivity of a certain tissue. So far, 14

serotypes from both human and non-human primates have been identified, along with over 100 different variants (Gao et al., 2005). Additionally, there are AAV serotypes of different other species, e.g. avian AAVs that can also infect humans (Yates et al., 1981). The wide range of serotypes exhibited by AAV contributes to its broad tissue tropism. Moreover, through the use of recombinant capsid modifications, such as retargeting strategies, the tropism of AAV can be expanded or tailored to specifically target tumour cells (Büning et al., 2008; Gao et al., 2004, 2005).

The discovery of post-transcriptional modifications (PTMs) on viral proteins (VPs) and their presence on the mature capsid surface has shed light on an additional crucial aspect of AAV's tropism, since capsid PTMs are likely to influence viral infectivity (Giles et al., 2018; Jin et al., 2017; Mary et al., 2019). Moreover, the significance of the production process on product quality has been underscored by Rumachik et al. (2020), who revealed expression system-dependent differences in PTMs. Post-translational modifications of structural and non- structural proteins of AAV are likely to be made and be involved in various stages of AAV's life cycle. These modifications include co-translational modifications for cellular transport and overcoming cellular barriers, as well as during packaging, virion maturation, and fine-tuning of receptor binding for optimal viral infectivity (Fattaey & Consigli, 1989; Mary et al., 2019; Ribet & Cossart, 2010; Weger et al., 2004).

3.2.1.1 AAV's Genome and Encoded Proteins

The approximately 4.7 kb short, linear single-stranded DNA (ssDNA) genome of AAV consists of two genes, *rep* and *cap*, flanked by a 145 bases long inverted terminal repeat (ITR) at each end.

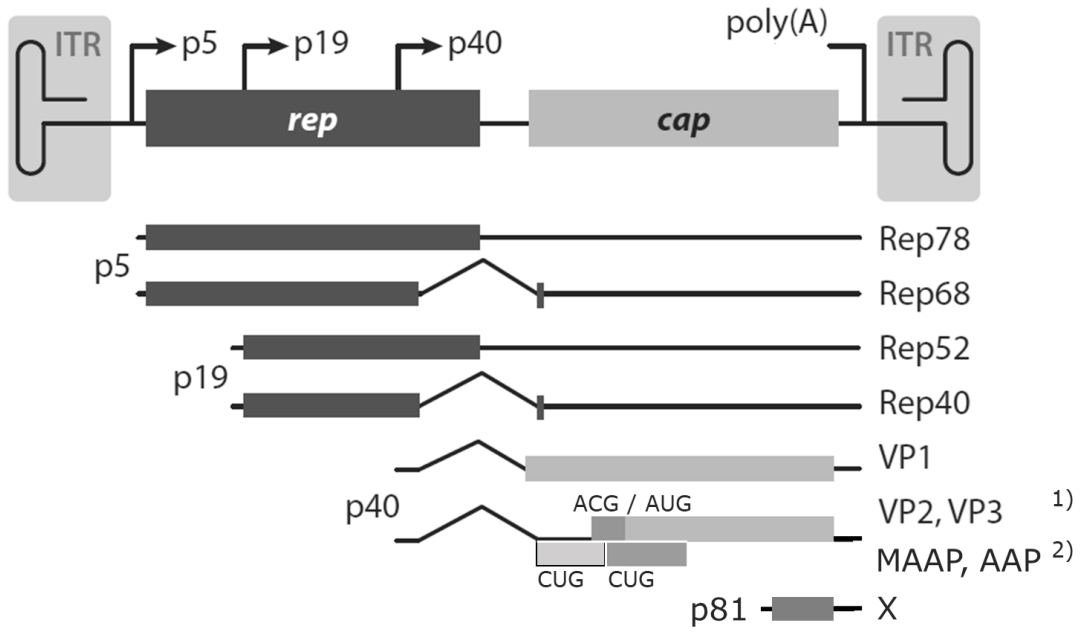
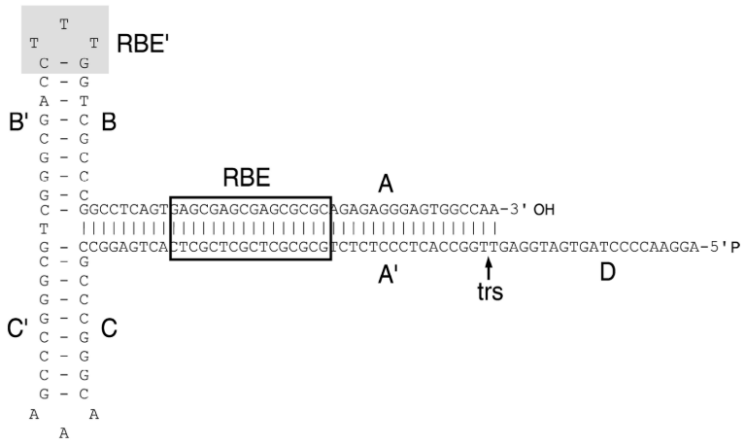


Figure 7 - Genetic map of wtAAV. Rep mRNAs and gene in dark grey, *cap* and its mRNA products in light greys. Additionally, alternate reading frames of VP (1), AAP, MAAP (2) are distinguished by numbers and the start codons of VP2 (AUG), VP3 (AUG), MAAP and AAP (CUG) are inscribed. The graphic is modified referring to Samulski and Muzyczka (2014).

The ITRs

The ITRs serve as the only essential *cis*-elements for replication, integration, and packaging. This feature allows for the packaging of a gene of interest (GOI) in recombinant AAV by replacing *rep* and *cap* with the GOI for therapeutic applications. ITRs form a T-shaped hairpin structure, facilitated by a 125-base palindromic sequence forming a complementary base-pairing. The ITRs consists of two arms, B-B' and C-C', as well as a stem palindrome, A-A'. The inner 20 base pairs are unpaired and referred to as the D-sequence (Gonçalves, 2005; Koczot et al., 1973; Rose et al., 1969). The free 3'-OH of the ITR serves as the origin of replication for AAV, enabling its self-priming strand-displacement or rolling hairpin replication mechanism (see Figure 8 B). This mechanism requires the binding of Rep78/68 proteins at the Rep binding element (RBE), a 22-base pair sequence, as well as the 5-base RBE' located at the tip of the hairpin. Rep protein-mediated nicking at the 7-base terminal resolution site (trs) within the ITR allows for replication of the ITR and subsequent genome amplification (Brister & Muzyczka, 2000; Ryan et al., 1996; Snyder et al., 1993).

A



B

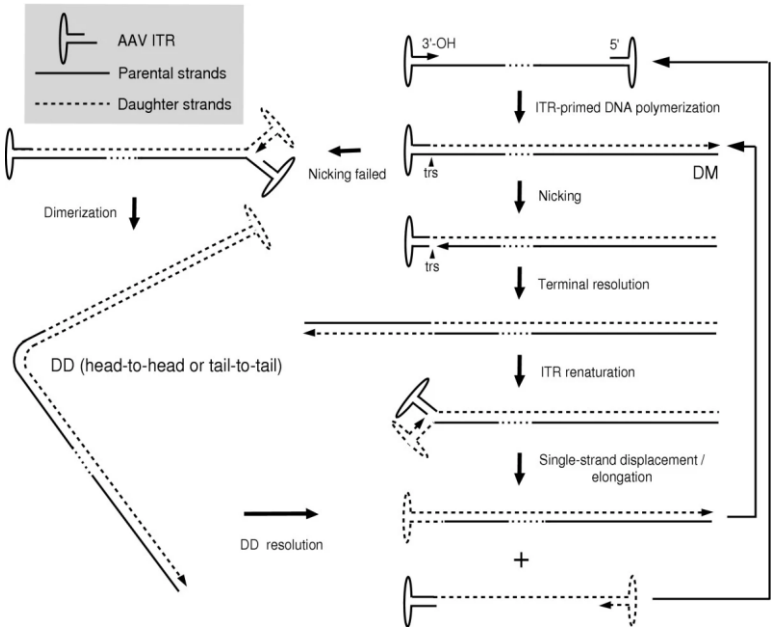


Figure 8 – (A) AAV2 ITR secondary structure. (B) Schematic illustration of AAV's strand-displacement DNA replication model. Illustrations taken from Gonçalves (2005).

The *rep* Gene

A vast coding capacity expansion is realized through alternate splicing, different reading frames and alternative start codons. As a result, only two genes are responsible for encoding a total of nine proteins. The *rep* gene encodes the four non-structural replicases, namely Rep78 (78 kDa), Rep68 (68 kDa), Rep52 (52 kDa), and Rep40 (40 kDa). mRNAs for the large Reps (Rep78 and Rep68) originate from the p5 promoter upstream of the gene. The transcription of Rep52/40 is driven by the p19 promoter located within *rep* gene. For the maturation of Rep 68 and Rep40 their

pre-mRNAs undergo splicing at the same splice site, resulting in the production of Rep68 instead of Rep78 and Rep40 instead of Rep52. The splice acceptor site is situated within the second gene, the *cap* gene. The larger Rep proteins, in conjunction with the ITRs, play a crucial role in integrating the AAV genome into the host genome, facilitating genomic rescue, and regulating AAV gene expression. They are also essential for the replication of AAV's single-stranded genome (Weitzman & Linden, 2012). Notably, Rep78 has been demonstrated to arrest the host cell in the S-phase (Berthet et al., 2005). Its overexpression is known to be anti-proliferative, as e.g. shown in NIH3T3 and MEF cells, or even cytotoxic (Saudan et al., 2000; Schmidt et al., 2000). Rep78 encompasses all three primary functional Rep-domains: the N-terminal DNA binding and endonuclease activity, the central domain with its ATPase and helicase activity, plus a nuclear localisation sequence (NLS) and the C-terminal zinc-finger domain, which interacts with a multitude of cellular factors. The other Reps, being translated from shorter mRNAs of the *rep* gene, lack some of these domains, but therefore provide different functions. All four Rep proteins possess a central helicase domain belonging to the superfamily 3 class, sharing similarities with AAA+ domain proteins. Rep68 acts as a site-specific endonuclease *cap*able of unwinding DNA, while the two small Reps (Rep52 and Rep40) are 3'-5' DNA helicases crucial for packaging the viral genome into the capsid (Balakrishnan & Jayandharan, 2014; Collaco et al., 2003; James et al., 2003; Yoon-Robarts et al., 2004; Zarate-Perez et al., 2013). Both the small and large Reps regulate the p5 promoter, as well as other viral and cellular promoters, which is elucidated in more detail in 3.2.2.6 (Kyöstiö et al., 1994, 1995; Murphy et al., 2007; Pereira et al., 1997).

The *cap* Gene

The *cap* gene is responsible for encoding the three viral capsid proteins, namely VP1, VP2, and VP3, as well as three non-structural proteins: the assembly activating protein (AAP), the proposed X-protein and the recently discovered membrane associated accessory protein (MAAP). Notably, all three of these proteins are situated in a different reading frame than the VPs. Pre-mRNA derived from the p40 promoter is spliced into a minor splice (VP1) or the major splice (VP2, VP3, MAAP, AAP, X) (see Figure 7). Among the capsid proteins, VP3 is the most abundant, initiated by the common and strong start codon AUG. In contrast, translation of VP2 from the same spliced mRNA initiates with the upstream positioned and weaker start codon ACG, which is often skipped by ribosomes due to leaky scanning (Daya & Berns, 2008).

Within the alternate reading frame, translation of both, MAAP and AAP, commence with the weak start codon, CUG (Ogden et al., 2019; Sonntag et al., 2010). The X gene is believed to have its own promoter, p81, and is transcribed with an AUG start

codon (Cao et al., 2014). The X-protein is thought to be involved in replication and enhances the replication of vector DNA, although its precise function remains incompletely understood and has been poorly investigated thus far (Cao et al., 2014, 2015).

AAP is involved in viral assembly and guides VP3 proteins to the nucleus (see 3.2.1.2). AAP is essential for many serotypes and is suggested to function as a scaffolding protein for capsid assembly, as well as to stabilize VPs and protect them from degradation (Grosse et al., 2017; Sonntag et al., 2010, 2011).

Similar to AAP, MAAP participates in the intracellular trafficking of capsid proteins. Furthermore, it appears to contribute to AAV replication, while also aiding in the secretion of certain serotypes. As its name implies, MAAP primarily localizes to cytoplasmic and nuclear membranes, where it is believed to interact with these membranes, playing an essential role in the progression of infection, proliferation, and potentially the egress of the virus. Like other non-structural proteins of AAV, MAAP exhibits multiple functions. It is involved in AAV's control of Helper virus proliferation in the host cell and is suggested to be associated with viral packaging and selectivity of packaging based on ITR specificity (Galibert et al., 2021; Timpe et al., 2006). Furthermore, it contributes to the cellular distribution of VPs while also influencing their degradation. Truncations and mutations of the 13 kDa protein have been shown to result in increased levels of VPs and capsids, reduced degradation, altered packaging efficiency and specificity, and increased virus genome titres. These findings suggest that MAAP holds promise as a potential target for engineering purposes, as highlighted by Galibert et al. (2021). In summary, while the VP proteins form the capsid, the non-structural *cap* proteins have distinct roles in the life cycle of AAV and are crucial for its replication.

3.2.1.2 AAV’s biphasic life cycle

The presence or absence of a Helper virus determines the pathway of AAV infection. In the absence of a Helper virus, AAV follows the lysogenic pathway, where its goal is to persist in the infected cell. When a Helper virus co-infects the cell, the lytic pathway and replication become available.

Lysogenic pathway and genomic integration

This infection pathway is characterized by a site-specific integration of the AAV genome into the host genome. The most frequent site for these Rep-dependent, specific integration events is the long arm of chromosome 19 (19q13.3-qter) termed AAVS1 (Adeno-Associated Virus Integration Site 1). The Rep-binding element or site (RBE / RBS) present in the AAV ITR is also found in AAVS1 (Kotin et al., 1990, 1991).

Both are bound simultaneously by Rep78/68 for integration. Integrations can also occur at other genomic locations near consensus RBS', such as chromosome 5p13.3 named AAVS2 and AAVS3 on chromosome 3p24.3 (Hüser et al., 2010). The exact mechanism of AAV's targeted genome integration is complex and, in some parts, still not fully resolved. However, it is certainly mediated by the large Reps and requires functional RBE and trs motifs (D. M. McCarty et al., 2004; Weitzman et al., 1994). Following integration, AAV gene expression is suppressed by Rep proteins, resulting in a dormant state of the virus in the host genome. Upon infection with a Helper virus, the provirus can be reactivated, leading to genomic rescue.(Beaton et al., 1989; Hamilton et al., 2004; Kyöstiö et al., 1995). Genomic rescue and AAV replication can also be induced in the absence of Helper virus by activating cellular stress response genes with genotoxic stimuli e.g. UV radiation or chemical carcinogens (Berns & Linden, 1995; Yakobson et al., 1987; Yalkinoglu et al., 1988).

RT-qPCR experiments have shown that the frequency of site-specific genomic integration at the AAVS1 site of wild-type AAV is approximately 0.1% per infectious virus particle (previously reported as 0.1-0.5% in earlier studies), which should not be confused with transduction efficiencies that can reach up to 80% in vitro, depending on the serotype and cell type (Hamilton et al., 2004; Hüser et al., 2002; D. M. McCarty et al., 2004). AAV genome integrations are more likely than those of non-integrating viruses like AdV, but less likely than those of retroviruses, which in contrast to AAV require integration for gene expression (D. M. McCarty et al., 2004). The site-specific integration of AAV is a unique feature among known eukaryotic viruses, making it highly attractive for gene therapy applications at the beginning of its journey as a therapeutic vector (Kotin, 1994). However, with rAAV used in gene therapeutics, site-specific integration cannot occur due to the absence of Rep proteins. Instead, rAAV integration into the genome is random and rare, as shown by Schnepp et al. (2003), who found that > 99.5% of transduced vector DNA was not integrated.

Instead, rAAVs persist as, mostly multimeric, circularized, episomal concatemers (Nakai et al., 2000; J. Yang et al., 1999). These episomal concatemers enable long-term persistence and gene expression. For example, a study by Penaud-Budloo et al. (2008) proved episomal persistence and gene expression in non-human primate muscles for up to 22 months. Canine studies promise long-term persistence, showing unimpaired expression in the eye twelve years after administration (Hauswirth, unpublished) (Berns & Muzyczka, 2017). However, it is important to consider the age and life expectancy of animal models as critical challenges for conducting long-term studies and statements about maximum persistence periods. Additionally, the persistence of episomes is likely influenced by the rate of cell

division, suggesting that non-dividing or slowly dividing cells may retain AAV episomes for an extended period. The exploration of this rAAV feature has diminished the relevance and research efforts of Rep-mediated integration in therapeutic application, which could, for example, be realised through co-expression of Rep in the targeted tissue.

Infection, lytic pathway and genome replication

AAV can infect both dividing and non-dividing cells with varying efficiency depending on tissue and serotype (Podsakoff et al., 1994; Zincarelli et al., 2008). The process of AAV infection involves multiple steps, encompassing receptor binding, cell entry, intracellular trafficking, endosomal escape, nucleus entry, viral uncoating, and, in absence of Helper virus functions, implantation of viral DNA (Nonnenmacher & Weber, 2012). The virion enters the cell through receptor mediated endocytosis. Hence, the virus has to binding to a receptor on the cell surface. While AAV2's primary receptor is heparin sulphate proteoglycan (HSPG), other serotypes engage different receptors due to variations in their capsid structures (Büning et al., 2008; Lux et al., 2005; Summerford & Samulski, 1998). Once the reversible capsid structural rearrangement takes place, it facilitates co-receptor mediated endocytosis and migration towards the nuclear area (Asokan et al., 2006; Sanlioglu et al., 2000). Although the precise mechanisms responsible for AAV's retrograde transport and endosomal escape remain incompletely resolved, acidification within the endosome exposes hidden capsid regions, such as the N-terminus of VP1. These regions contain important sequences for endosomal escape, particularly PLA2, as well as possible nuclear localisation sequences (NLS) like BR1-4 (Girod et al., 2002; Grieger et al., 2006, 2007). This allows for nuclear import, enabling the entire capsid containing the virus' DNA to enter the nucleus. However, this process takes time, with the majority of viral particles still located in a perinuclear compartment 16-20 hours after infection (Bartlett et al., 2000).

The presence of AAV capsids in the nucleus, as demonstrated by Sonntag et al. (2006) and as well as super resolution imaging by Kelich et al. (2015), suggests that the slow uncoating process occurs in the nucleus and not in the cytosol (Bartlett et al., 2000; J. S. Johnson & Samulski, 2009; Nicolson & Samulski, 2014; Sanlioglu et al., 2000). To enter the nucleus AAV particles hijack cellular transport mechanisms and get shuttled through the nuclear pore complex (NPC). The interactions required for this process may vary among serotypes (Junod et al., 2021; Nicolson & Samulski, 2014). Within the nucleus, AAV accumulates in nucleoli (Görlich et al., 1995).

Once the genome is released, the initial step, independent of the presence or absence of a Helper virus, involves the synthesis of the second strand and basal expression

of Rep proteins (Agbandje-McKenna & Kleinschmidt, 2012). This step is known to be a rate-limiting and was therefore optimized in some rAAV applications by the use of self-complementary genomes (scAAV) (Ferrari et al., 1996; D. McCarty et al., 2001). However, the use of scAAV is limited, as it halves the genome size. After the synthesis of the second strand of AAV, its replicative life cycle in the presence of AdV infection can be divided into four major steps, as follows: 1. Double-strand synthesis, 2. Expression of *rep* and *cap* genes, 3. DNA replication, and 4. Packaging of synthesized single-stranded DNA (ssDNA) into pre-assembled capsids.

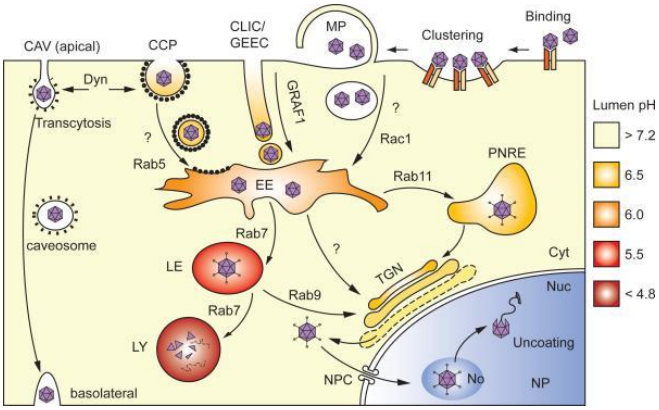


Figure 9: Schematic model of entry and intracellular trafficking of AAVs (Nonnenmacher & Weber, 2012). Abbreviations: CCP=clathrin-mediated endocytosis, CAV=caveolar endocytosis, EE=early endosome, LE=late endosome, LY=lysosome, PNER=perinuclear recycling endosome, TGN=trans-Golgi network, ER=endoplasmic reticulum, NPC=nuclear pore complex, NP=nucleoplasm, No=nucleolus

AAV DNA replication follows a unidirectional, self-priming strand-displacement, or rolling hairpin mechanism, requiring the AAV-encoded Rep78 and Rep68 proteins synthesized earlier, host cell DNA polymerase δ (DNApol-δ), its accessory proteins, replication factor C (RFC), proliferating cell nuclear antigen (PCNA), and the minichromosome complex maintenance (MCM). The folded left ITR with its free 3'-OH serves as the replication primer. Initially, the host cell's replication machinery synthesizes the secondary strand unidirectionally. Subsequently, a duplex monomer (DM) with a covalently closed hairpin structure is formed. The Rep78 and Rep68 proteins bind at the Rep binding elements (RBE and RBE') and their activated endonuclease function cleaves the bottom strand site-specifically at the trs (at nt 125), creating a new 3'-OH for replication of the ITR. Thereby, Rep78/68 stays covalently bound to the newly created 5'-end of the parental ITR. Afterwards, the palindromic sequences of the linear duplex ITRs can renaturate into their terminal hairpin structure, providing new 3'-OHs for single-strand displacement syntheses of both strands. In case the nicking at the trs fails, a double-length double-stranded molecule (DD) is formed with either a head-to-head or a tail-to-tail confirmation. DDs can be resolved to DMs through the ITR sequences at the symmetry axis, which can then also be used as templates for single-stranded displacement replication. The single-strand displacement synthesis uses the cellular complexes pol δ, MCM, and their associated proteins for amplification. The result is

one single-stranded genome plus one DM that can both be used to generate more DMs for replication. Alternatively, the ssDNA genome can be packed into capsids (Gonçalves, 2005; Samulski & Muzyczka, 2014). AAV DNA strands of each polarity are produced and packaged with equal frequency, resulting in half of the particles containing plus strands and the other half containing minus strands (Berns & Adler, 1972; Berns & Labow, 1987; Mayor et al., 1969; Rose et al., 1969).

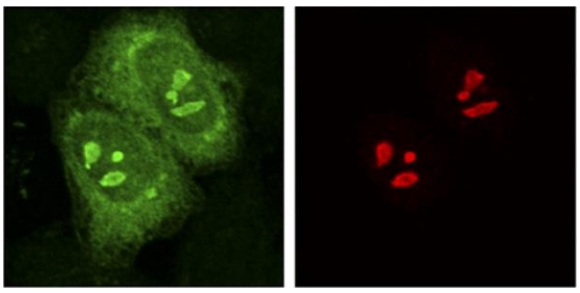


Figure 10 - Intracellular localisation of AAV2 VPs (green, left picture) and assembled VP3 only capsids (red, right picture). Indirect double immunofluorescence of transfected HeLa cells. Immunostaining was obtained with VP antiserum VP51 (for VPs) and mAb A20 (for assembled capsids) from (Sonntag et al., 2010).

Newly synthesized ssDNA genomes are then packaged into preassembled capsids. Capsid assembly occurs within the nucleus, necessitating the entry of viral proteins after their ribosomal synthesis in the cytosol. VP1 and VP2 accumulate rapidly in the nucleus, while VP3 remains partially in the cytoplasm (Grieger et al., 2006; Ruffing et al., 1992; Wistuba et al., 1997). This discrepancy arises from the presence of nuclear localisation signals (NLSs) in the N-terminal regions of VP1 and VP2. Hansen et al. (2000) revealed the relation between nuclear import and the BR3 sequence PAKKRL/N at amino acids 166 to 171/172, which does only exist in VP1 and VP2. So far, no functional NLS has been reported for VP3. In contrast, import of VP3 into the nucleus is mediated by AAP, as shown by Sonntag et al. (2010). Their work also revealed that fusing an NLS to VP3 alone is insufficient for viral assembly of AAV2. Consequently, AAP not only directs VPs to the nucleus but also to nucleoli, where it contributes to the actual assembly process. However, this dependence on AAP is serotype-specific, as elucidated by Maurer et al. (2018), showing AAP independent VP3 nuclearisation and capsid assembly for AAV4, AAV5 and AAV11. Regardless of the serotype, VPs initially localize to the nucleoli and accumulate there (see Figure 5). This localisation may accelerate capsid assembly, as a high concentration of VPs favours capsid formation. The icosahedron is assembled from VP oligomers, particularly pentamers that likely from already in the cytoplasm. Capsids show co-localisation with Rep proteins, but their detection in the absence of Rep proteins suggests that Rep proteins are not necessary for the assembly process (Myers & Carter, 1980; Wistuba et al., 1995, 1997). Furthermore, the specific role of nuclear proteins, such as nucleolin and nucleophosmin, in association with VPs and capsids, remains unresolved (Bevington et al., 2007; Qiu &

Brown, 1999). However, the presence of AAP as a scaffolding protein is essential for most serotypes for capsid assembly (Sonntag et al., 2010).

The capsid formation process occurs rapidly, while packaging takes several hours (Myers & Carter, 1980). Packaging is facilitated through the 1.2 nm pore of the capsids, utilizing the helicase activity of Rep52/40. During the replication Rep78/68 seem to be bound to the 5'-ITR and serve as packaging signal. The experiments Nony et al. (2003) performed support this theory, showing that genomes without ITRs can be packaged when upstream of rep a p5 promoter containing an RBE sequence is present. The DNA-Rep78/68 complex can then bind via Rep78/68 to Rep52/40. Now, Rep52 and Rep40 interact with the ITR's D-sequence and bind to the capsid. Subsequently, the helicase activity transfers the ssDNA, led by its 3'-end, into the capsid (J. A. King et al., 2001). The efficiency of packaging is influenced by genome size, with optimal packaging occurring between 4.1 and 4.9 kb (Dong et al., 1996). Genomes smaller than half the wild-type size have been observed to be packaged with two copies per virion, suggesting a "head full" mechanism for packaging (Agbandje-McKenna & Kleinschmidt, 2012). The whole proliferation process of AAV is enabled by the virus' use of the host cell's protein synthesis machinery. This includes post transcriptional modifications of non-structural and structural AAV proteins which accompany the virus' amplification.

Finally, with regards to cell exit, AAV is likely reliant on its Helper virus, as no active egress mechanism has been identified for AAV. In contrast, lysis of the host cell in the lytic pathway is believed to be dependent on the cytopathogenic effects induced by the Helper virus. While there is no known dedicated egress pathway for AAV, there are variations in cell release efficiencies among different serotypes. Consequently, during rAAV production, some serotypes are predominantly found in the culture supernatant, while others require cell lysis as an initial downstream step. These variances are likely influenced by the tropism specific to each capsid. For instance, the heparin-binding motif present in AAV2 appears to contribute to its retention, as mutations in this motif have been shown to result in increased supernatant titres (Vandenberghe et al., 2006, 2010). Interestingly, MAAP has been shown to play a role in the release of virions, at least partially in association with extra cellular vesicles (Elmore et al., 2021).

3.2.2 AAV's Prominent Helper – Adenovirus – Helper Functions, Gene Regulation and Interactions on Molecular Level

In general, Helper viruses and their proteins/products function as transcription factors (TFs) and/or interact with cellular proteins to activate, enhance, and regulate viral transcription and translation, as well as manipulate the host's antiviral

response and apoptosis (Meier et al., 2020). Viruses generally manipulate host cells to serve as ideal virus factories. Since AAV is very small and has limited encoding capacity on its genome, it partially utilizes the abilities of other virus for these purposes. Among various Helper viruses for AAV, AdV is particularly significant. In rAAV production, AdV5 genes are commonly employed to provide the necessary Helper functions (Clément & Grieger, 2016). Therefore, this chapter will focus on describing the functions of human Adenovirus 5 genes essential for AAV production, as well as the interactions and gene regulations involving AAV, AdV, and host cell components.

The minimum set of required Ad5 Helper genes has been identified as the non-structural proteins E1A (E = AdV early gene), *E1B55K*, *E2A*, and *E4orf6*, as well as the VA (viral-associated) RNA (Carter et al., 1983; Grimm et al., 1998; M. M. Huang & Hearing, 1989a; Janik et al., 1981, 1989; Laughlin et al., 1982; Matsushita et al., 1998; Muzyczka, 1992; Richardson & Westphal, 1981, 1984; Samulski & Shenk, 1988; Tratschin et al., 1984; Xiao et al., 1998). Notably, the absence of Ad replication proteins (AdV polymerase, AdV terminal protein) and their minimal impact on AAV replication highlight that AAV replication is primarily executed by cellular replication proteins in the presence of Ad coinfection.

E1A	General transcription factor, AdV early promoter activation, oncogene	AAV promoter activation, drives cells to S-phase
E1B19K	Inhibits proapoptotic Bcl-2 homologs (Bax and Bak), induces autophagy, oncogene	Enhances AAV vector titers
E1B55K	In complex with E4orf6 it prevents E1A mediated p53 stabilization, oncogene	Involved in AAV mRNA export, promotes AAV second-strand synthesis
E2A	ssDNA binding protein, viral DNA replication & mRNA processing	AAV promoter regulation, AAV genome replication, Rep splicing, capsid protein production
E4orf6	In complex with E1B55K it prevents E1A mediated p53 stabilization, supports viral DNA replication and RNA processing	Promotes AAV second-strand synthesis, inhibits the MRN complex
VA RNA	inhibits the eIF-2 protein kinase, promotes viral protein translation	Prevents E4orf6/E1B mediated degradation of AAV capsids & Rep52

Figure 11 - Summary of AdV5 Helper functions from Meier et al. (2020).

3.2.2.1 E1A

The AdV early region 1A (E1A) gene produces five mRNA splice transcripts with two major splices, the translated proteins (here referred to as E1A) possess similar sequences and characteristics, depending on the presence of more or less of its numerous short linear interaction motifs (SLiMs). E1A acts as a kind of “hub”-protein that interacts with a multitude of cellular factors (at least 32 primary protein interactions and 4000 unique associations) to reprogram nearly the entire cell for ideal virus production conditions (Ferreon et al., 2009; Horwitz et al., 2008; C. R. King et al., 2018; Pelka et al., 2008). By hijacking the cellular nuclear import, it localizes in the nucleus and functions as a kind of general-purpose transcription factor, to enable and enhance viral gene expression, as well as to suppress antiviral host cell mechanisms (Bayley & Mymryk, 1994; Köhler et al., 2001; Lyons et al.,

1987; Nevins et al., 1979; Spindler et al., 1985). Since E1A is the first protein that is expressed after AdV infection, many of its functions are necessary to enable further protein expression and pathways of AdV and AAV (Matsushita et al., 2004; Nevins et al., 1979). These functions rely on E1A's protein-protein interactions as E1A does not directly bind to DNA. Therefore, E1A's TF-like functions and other DNA interactions are mediated by other proteins and only indirectly by E1A itself (Frisch & Mymryk, 2002).

One important initial feature of E1A's host cell preparation for virus production is its stabilisation of the preinitiation complex (PIC) and the recruitment of cofactors to initiate and sustain transcription of viral and specific host cell genes. Therefore, E1A interacts with a multitude of TFs and cellular proteins associated with transcriptional pathways, including the TATA-binding protein (TPB), a subunit of TFIID for transcription initiation, but also p300, CBP, RB1 and many more (Avantaggiati et al., 1996; Egan et al., 1989; Geisberg et al., 1994; Lundblad et al., 1995; Song et al., 1995). In this manner, E1A enhances particularly the transcription of promoters containing the TFRE (transcription factor regulatory element) CRE (cAMP-responsive element) by interacting with CREB (CRE-binding protein) and ATF-1 (activating transcription factor 1), both of which bind to CRE. This effect is mediated through the cAMP/PKA (cyclic AMP/protein kinase A) signalling pathway, which regulates a diverse set of cellular functions, many of which are through activation of TFs like CREB. The phosphorylation of CREB by PKA is required for it to interact with the transcriptional coactivator CREB-binding protein (CBP) that then initiates the transcription (Chrivia et al., 1993a, 1993b; Edwards & Scott, 2000; Sassone-Corsi, 2012). This mechanism is particularly significant for AdV and AAV because most AdV promoters, including the ones of *E2A* and *E4*, contain CRE elements (Sassone-Corsi, 1988). Additionally, CREB interactions are also likely for all three AAV promoters (according to an analysis with https://tfbind.hgc.jp/) (Tsunoda & Takagi, 1999). Therefore, E1A is an activator for all viral proteins required for AAV production. Increase of AdV protein expression due to E1A is about 50-100 fold (Berk et al., 1979; Jones & Shenk, 1979). Another known mechanism through which E1A activates transcription is by releasing the transcription factor E2F from its binding partners pRb, p107, and p130. E2F containing promoters showing increased activity in E1A presence include the immediate early promoter of human cytomegalovirus (CMV), c-myc and AdV E2 early promoter (Bagchi et al., 1990; Cockett et al., 1991; Hiebert et al., 1989; Kovesdi et al., 1987; Metcalf et al., 1994). However, E1A is also known to repress transcription from several promoters including the E2 late promoter and SV40A. Excess copies of the affected enhancer can counteract the repression caused by E1A, indicating that the repression is

caused by the activation of a negatively acting repressor product in the presence of E1A (Borrelli et al., 1984; Rossini, 1983; Sogawa et al., 1989; Velcich & Ziff, 1985; Webster et al., 1988). Notably, there are functional differences and even antagonistic effects of E1A products. The E1B promoter, as well as the early E2, E3 and the E4 promoter depend on activation by E1A's largest protein product, whereas its second largest product (243R) represses them (Dery et al., 1987).

E1A also induces and even arrests S-Phase in cells that are in resting phase (G0), likely through interactions with p300/CBP, p107 and pRb triggering DNA replication, which even continues during the cell cycle arrest (Grand et al., 1998). Even though AdV provides its own replication machinery, it seems logical that a cell in a state of active DNA replication provides more required substrates and is more beneficial for virus production than a cell in basic survival mode. Another aspect for this cell preparation is E1A's interaction with histone-directed deacetylases (HDACs), histone-directed acetyltransferases (HATs), and other nuclear factors. Through these interactions, E1A remodels the chromatin of host cells, thereby regulating both host cell and viral gene transcription (Gallimore & Turnell, 2001). In addition to E1A's efforts to promote transcription and expression of viral genes, it also interferes with the host cell's expression of its own proteins, as well as its immune response (Olanubi et al., 2017; Zemke & Berk, 2017). Importantly, it changes PTMs of host cell proteins by altering the PTM-machinery itself, e.g. by altering SUMOylation pathways. As one way to reduce cellular immune activity against the virus intrusion, E1 interferes with MHC-I dependent antigen representation by alteration and inhibition of the immunoproteasome (C. R. King et al., 2018). Another of these inhibitions is targeting p53, the product of tumour suppressor gene TP53, a key protein of cell cycle, apoptosis and genetic stability. The stabilisation of p53 by E1A leads to the cell's arrest in S-phase, but also drives the cell into apoptosis (Steegenga et al., 1996; Turnell, 2000). To counter this, E1B55K and E4orf6 form a complex with p53, which leads to ubiquitination of and consequent degradation of p53 (Lowe et al., 1993; Querido et al., 1997). Further apoptosis induction is countered by E1B's second gene product its 19K protein (Rao et al., 1992). The fine balance between E1A and E1B is central to Adenovirus' oncogenic properties, as it leads to cell immortalisation instead of apoptosis. As a result, the stable integration of approximately five E1 copies transformed human embryonic kidney cells into the primary cell line used for rAAV production, HEK293 (Blackford & Grand, 2009; Graham et al., 1977; Hidalgo et al., 2019; Lin et al., 2014; White, 2001a). Due to E1's presence in the cell line, it can be omitted from the plasmids used for rAAV production. E1A is essential for AAV production because it relieves the repression of the AAV p5 promoter by binding to the cellular

transcription repressor YY1. This binding converts YY1 from a repressor to a transcription activator. Additionally, upstream of the YY1 binding site, there is a MLTF (major late transcription factor) site that acts similarly to E1A as an activator (L. S. Chang et al., 1989; Shi et al., 1991).

3.2.2.2 E1B

The E1B gene encodes two proteins, namely E1B55K (55 kDa) and E1B19K (19 kDa), which play a crucial role in inhibiting apoptosis induced by E1A. Thereby, E1B19K interacts with Beclin-1, while E1B55K binds to p53. E1B's 19 kb protein prevents degradation of the antiapoptotic protein MCL1, induced by E1A, through mimicry and binding to BAK, which effectively halts the apoptotic cascade initiated by ML1's binding partner BAK (Cuconati et al., 2003; White, 2001b). Although both proteins enhance AAV titres, only E1B55K is considered indispensable. The redundancy of E1B19K can be attributed to its cellular homologues, Bcl-2 and Bcl-xL, which likely compensate for its functions (Matsushita et al., 2004; Subramanian et al., 2015).

The activity of E1B55K is regulated through phosphorylation and SUMOylation, and it also functions as a SUMO-1 ligase itself, leading to the polyubiquitylation of p53 (Hidalgo et al., 2019). SUMO (small ubiquitin-like modifier) proteins comprise a protein family that covalently attaches to other proteins, thereby modulating various cellular processes such as nuclear-cytosolic transport, transcriptional regulation, protein stability, stress responses, cell cycle progression, and apoptosis (Geiss-Friedlander & Melchior, 2007).

Moreover, the 55K protein facilitates the export of viral mRNA from the nucleus by providing a leucine-rich nuclear export signal (NES), while simultaneously inhibiting the export of cellular mRNA via the CRM-1 pathway. This mRNA export is further downregulated by E4orf6 (Endter et al., 2005; Pilder et al., 1986). E1B55K is also involved in recruiting essential proteins to viral replication centres. However, the execution of numerous vital functions by E1B55K occurs within its complex with E4orf6, making it challenging to distinguish between the functions of E1B55K alone and the E4orf6/E1B55K complex, especially in older studies. As a result, the reported functions of E1B55K, E4orf6, and the E4orf6/E1B55K complex often overlap (Hidalgo et al., 2019). The same assumption can be made regarding numerous Helper functions, as it can be challenging to differentiate between a direct function and an observation resulting from a cascade of functions, such as at the transcriptional level, depending on the type of study.

3.2.2.3 *E4* and the Proteins of its Open Reading Frames

The *E4* gene of the Adenovirus encodes at least 18 distinct mRNAs, which in turn give rise to seven different proteins. Deletion mutant studies have shown that six of these proteins and their according open reading frames (ORFs) can be removed without inhibiting Adenovirus proliferation. However, the deletion of Orf6 leads to reduced Adenovirus yields. *E4orf6* is consequently the only *E4* ORF capable of providing sufficient functions for replication in the absence of all other *E4* ORFs. Interestingly, in the absence of Orf6, Adenovirus replication remains possible with different ORFs as long as *E4orf3* is present, suggesting a compensatory function between Orf3 and Orf6, where the loss of one can be partially compensated by the other (Halbert et al., 1985; Leppard, 1997; Virtanen et al., 1984). Nevertheless, studies by Huang and Hearing (1989) demonstrated that Orf6 is more efficient. Furthermore, Ferrari et al. (1996) and Samulski and Shenk (1988) established that the 34 kDa protein expressed from *E4orf6* is essential for AAV production, alongside E1B55K. This finding served as a crucial step towards the development of Helper plasmids. One of the pioneering Helper plasmids was created by Matsushita et al. (1998), who demonstrated that rAAV production without Adenovirus could surpass the traditional method by relying solely on transient plasmid transfection. Additionally, their group established that E4orf6 alone represents the only indispensable component of *E4* for rAAV production with transient transfection of plasmids. Consequently, regarding the *E4* gene, the main focus of this review, but also the majority of research, is on its essential product, E4orf6.

All E4 mRNAs are transcribed from the E4 promoter and then spliced accordingly. Activation of the E4 promoter occurs early during infection, primarily driven by the E1A transcriptional activation. Transcriptional activity of the E4 promoter persists throughout the late phase of infection. However, transcription decreases due to the repression exerted by the *E2A* gene product. Additionally, the E1A-mediated transactivation is subjected to feedback inhibition by the E4orf4 protein, further contributing to the decrease in E4 mRNA. Furthermore, levels of individual E4 mRNAs are regulated post-transcriptionally through alternative splicing.

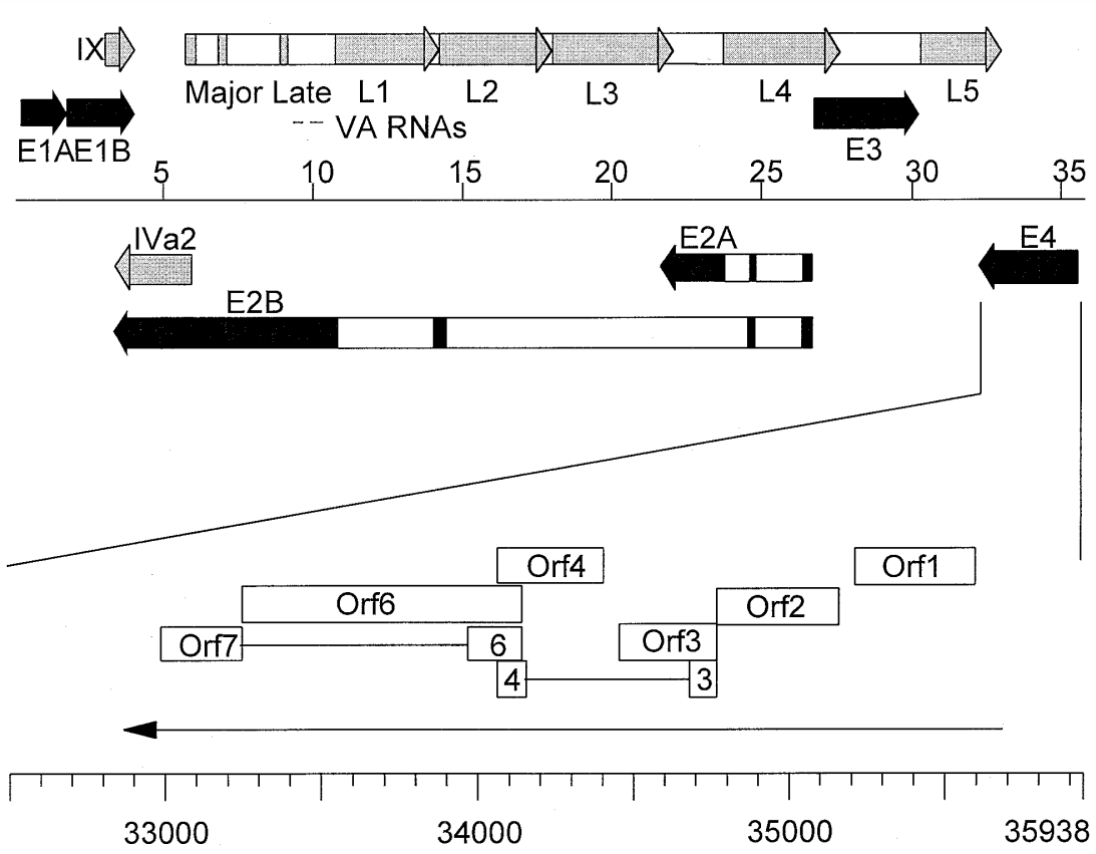


Figure 12 – Illustration of the AdV5 genome with focus on the *E4* gene from Leppard (1997).

E4orf6

E4orf6 plays a pivotal role in various aspects AAV proliferation. Experiments by Allen, Halbert, and Miller (2000) even demonstrated that E4orf6 alone is sufficient as the only required AdV Helper function in HEK293 cells when only heterologous promoters are used. This is in line with the findings of Richardson and Westphal (1981), who showed that with microinjected mRNA only E4 is sufficient and required for AAV proliferation. Additionally, based on findings of their experiments, the investigators postulated a AdV Helper gene regulation cascade of *E1*→*E2A*→*E4*.

E4orf6 is involved in mRNA processing, mRNA transport, and the suppression of host cell synthesis, including proteins involved in viral DNA replication, double-strand repair, and apoptosis. The majority of E4orf6's Helper functions are executed in conjunction with E1B55K. Inhibition of p53 can be facilitated by E4orf6 when complexed with E1B55K and independently. Similarly, it can enhance second strand synthesis on its own, although this effect is attenuated compared to the impact of E4orf6/E1B55K (Ferrari et al., 1996; Fisher et al., 1996). In the study by Ferrari et al. (1996) it is shown that second strand synthesis is a limiting step in AAV production that can be improved by E4orf6. Additionally, the study suggests that the use of genotoxic chemicals can mimic this effect. Although the precise mechanism

by which E4orf6 promotes AAV second strand synthesis remains incompletely understood, it appears that the accumulation of essential replication proteins in viral replication centres (VRCs) is a critical function of both E4orf6 alone and E4orf6/E1B55K. Hence, genotoxic stress may facilitate replication in a similar manner (Leppard, 1997). As previously mentioned, distinguishing the individual functions of E4orf6 and E1B55K proves difficult because many of these functions are performed by the complex formed by these two adenoviral proteins. These include promoting the degradation of newly assembled capsids and Rep52 through the ubiquitin pathway, which actually can be mitigated by co-expression of other Helper genes, particularly VAI, in stoichiometrically favourable amounts (Meier et al., 2020). Therefore, many functions of E4orf6 are better described as the mechanisms of the E4orf6/E1B55K complex.

E4orf6/E1B55k

The E4orf6/E1B55k complex presents crucial Helper functions for AAV, providing various essential mechanisms for AAV's replication. Among the many functions of this complex are the facilitation of AAV mRNA export, stabilizing the viral genome during replication, promoting translation, and preventing DNA damage response. Many of its cell altering actions, are achieved by the creation of an E3 ubiquitin ligase. For the formation of this ubiquitin ligase the cellular proteins elongin B and C, Cul5, and Rbx1 are recruited by E4orf6 and E1B55K functions as the substrate. Ubiquitinated targets, including p53, DNA ligase IV, integrin α3, and ATRX are then marked for subsequent proteasomal degradation (Cheng et al., 2013). However, some cellular proteins can be degraded by subcomplexes of the ligase, formed either with E4orf6 or E1B55K alone.

The nuclear localisation and recruitment of replication factors to AAV viral replication centres are likely executed through the E1B55K and E4orf6 complex (Blackford & Grand, 2009). The same goes for the shutdown of cellular mRNA export. The E4orf6/E1B55k complex exhibits the ability to ubiquitinate and degrade messenger RNPs (ribonucleoproteins) or other proteins involved in cellular mRNA nuclear export. By disrupting the host cell mRNA export pathway, it promotes the viral mRNA export (Berk, 2005; Blackford & Grand, 2009). Notably, the complex employs ubiquitination as a means to inhibit p53. Moreover, both E1B55K and E4orf6 independently possess the capability to suppress p53 activity (Weitzman, 2005).

The interference of the E4orf6/E1B55K complex with host cell proteins is exemplified prominently by its degradation of MRE11, a homologous double-strand break repair endonuclease. This interference disrupts the MRN

(MRE11/Rad50/NBS1) complex through ubiquitination and subsequent proteasomal degradation. The MRN multi protein complex plays an important role in detecting and signalling double-strand breaks and is involved in the repair pathways of homologous recombination (HR) and non-homologous end joining (Lamarche et al., 2010; Stracker et al., 2002). Moreover, previous studies have demonstrated that the MRN complex limits AAV transduction and replication (R. A. Schwartz et al., 2007). Therefore, the degradation of the MRN complex explains one aspect of the ability of E4orf6/E1B55K to enhance AAV titres. The complex also contributes to rAAV production through its capacity to enhance second strand synthesis in recombinant AAV production (Fisher et al., 1996). On the other hand, EB55K/E4orf6's E3 ligase also ubiquitinates Rep52 and VPs, which are consequently degraded. It is not clear, however, if this is an evolved AdV mechanism against AAV, a mechanism of AAV for regulation of component amounts or just a byproduct of the complex' limited target specificity (Nayak et al., 2008).

The complex' has a large role in the often-quoted arms race between virus and host. An example is the competition against the nuclear protein SPOC1 (survival-time associated PHD (plant homeodomain) protein in ovarian cancer 1). SPOC1 is part of cellular DNA damage response (DDR) associated with condensing of the chromatin. SPOC1 overexpression reduces AdV titres, and the protein was also detected interacting with E1B55K and *E2A*'s gene product the AdV DBP (DNA binding protein), showing that it is associated with a direct response to the adenoviral infection. But the virus' E1B55K/E4orf6 complex acting as a E3 ubiquitin ligase is able to counter this by deSUMOylation of the SPOC1's interaction partner KAP1 and proteasomal degradation of SPOC1 (Bürck et al., 2016; Schreiner et al., 2013). Next to making sure the chromatin stays accessible, the complex' destructive ability also guarantees the transcription of viral genes by degrading the death domain associated protein (Daxx), which inhibits basal expression of viral genes. However, Daxx can also be degraded in the absence of E4orf6 in the case of AdV5. Whereas for AdV12 (but not AdV5), E4orf6 alone promotes the degradation of TOPBP1. Demonstrating that some cellular proteins can be degraded by subcomplexes of the ligase, formed either with E4orf6 or E1B55K alone (Blackford et al., 2010; Cheng et al., 2013; Forrester et al., 2011).

E4orf1

All open reading frames of E4orf1-4, particularly E4orf1, have been reported to exert immunomodulatory effects on the host (Sangare et al., 2021, 2022). The E4orf1 protein shares structural similarities with dUTPases, although no enzymatic activity has been detected (Leppard, 1997; Weiss et al., 1997). Instead, E4orf1 activates the PI3-kinase/Akt pathway through protein-protein interactions,

specifically by associating with PDZ domain-containing proteins. These proteins primarily function as scaffold proteins for assembling signalling complexes at the plasma membrane (Sheng & Sala, 2001). By engaging the PI3-kinase, E4orf1 triggers the activation of Akt, mTOR, and p70S6K, consequently modulating protein synthesis and promoting cell survival. Moreover, it induces the phosphorylation of NF-κB, a heterodimer composed of p50 and RelA subunits, which functions as a transcription factor regulating cytokine release, cellular immune responses, and cell survival (Biancalana et al., 2021; Sangare et al., 2022; Thomas et al., 2009). Hence, E4orf1 exhibits oncogenic characteristics in AdV9, although its impact for other serotypes seems to be inferior. Other modes of interaction of E4orf1, leading to similar alterations in protein synthesis and cell survival, have also been reported (Thomas et al., 2009).

E4orf2 and Eorf3/4

The functional characteristics of two proteins encoded by *E4* have yet not been fathomed. The Orf2 protein, which is detected in infected HeLa cells during the early phase of infection, exists as a soluble cytoplasmic component. No evidence has been found to suggest its involvement in the formation of complexes with viral or cellular proteins (Weitzman, 2005). On the other hand, the existence of E4orf3/4 has been predicted based on AdV2 mRNA structure analysis, but experimental confirmation is still lacking (Dix & Leppard, 1993, 1995; Virtanen et al., 1984).

E4orf3

E4orf3 and E4orf6 are functionally redundant for AdV's lytic pathways and independently sufficient for efficient viral DNA replication, late viral protein synthesis, host protein synthesis shut-off, and virus production (Shepard & Ornelles, 2004). E4orf3's association with E1B55K leads to its relocalisation into the nucleus, but no complex formation has been detected between E4orf3 and E1B55K. Both, E4orf3 and E4orf6 can prevent concatemer formation, likely by targeting required cellular proteins involved in non-homologous recombination, such as ligase IV and DNA-PK, as well as the Mre11 complex, through alternative mechanisms (Araujo et al., 2005; Stracker et al., 2002).

E4orf3 is the most abundant E4 ORF on mRNA level and one of the earliest proteins being expressed after infection. It then acts in the nucleus, more specifically, it associates with the nuclear matrix and introduces the reorganisation of PML oncogenic domains (PODs). PODs are large nuclear structures involved in multiple cellular functions, including transformation, genomic stability, DNA repair, transcriptional control, apoptosis, and the interferon response. Reorganisation of PODs is thought to be advantageous for viral replication, as many viruses target and

rearrange these structures (Carvalho et al., 1995; Donovan-Banfield et al., 2020; Hoppe et al., 2006). PML and MRN relocalisation and reorganisation by E4orf3 is an important control mechanism of AdV replication (Evans & Hearing, 2005). By association with E1A, E4orf3 is part of the virus' chromatin remodelling and enhances viral genome expression (Soriano et al., 2019). Similarly, E4orf6 also interacts with E1A through its complex with E1B55K, in this case enhancing the relief of E2F and therefore enhancing the transactivation of E2F-regulated promoters (Dallaire et al., 2016). The interchangeability of E4orf6 is apparent, but the individual mechanistic are distinct. Inactivation of p53 for instance can also be executed by Orf3. However, in contrast to E4orf6, Orf3's mechanism is E1B55K independent and based on heterochromatin formation at p53 target promoters (Soria et al., 2010). Generally, despite its compensational functions, it seems that the ability of cell alteration by heterochromatin formation could be E4orf3's core function for AdV proliferation.

E4orf4

The E4orf4 protein is dispensable for AdV production, but deletion mutants lacking this ORF are more cytotoxic. E4orf4 serves as an antagonist to other AdV functions and acts to inhibit the cellular immune response (Halbert et al., 1985; Müller et al., 1992). Although it was shown that E4orf4 protects against apoptosis, it is also known to induce cell death through other pathways, a mechanism that may be relevant during the late stages of the AdV life cycle (Pechkovsky, Lahav, et al., 2013; Pechkovsky, Salzberg, et al., 2013). E4orf4 primarily localizes in the nucleus, particularly in the viral replication centres (VRCs), and exerts its main effects by regulating protein phosphorylation through interaction with its binding partner the protein phosphatase 2A (PP2A). By dephosphorylating the otherwise hyperphosphorylated c-Fos and E1A proteins, E4orf4 counteracts the transactivation ability of the E1A protein on AP1-containing promoters, thereby attenuating their activity (Kleinberger & Shenk, 1993; Müller et al., 1992). Furthermore, E4orf4 downregulates the E4 promoter, forming a feedback loop, as well as the E2 promoter through dephosphorylation (Mannervik et al., 1999). These interferences and down regulations of other AdV proteins are likely a balancing mechanism of the virus, preventing premature cell death. Next to its main partner PP2A, E4orf4 interacts with various cellular proteins and participates in the virus-mediated inhibition of the host cell's DDR. E4orf4 enhances viral production while simultaneously perturbing the cell cycle by targeting the mTOR pathway. It downregulates transcription factors, including JunB and c-Myc, to reduce host cell gene expression, and is involved in chromatin remodeling processes employed by AdV. Furthermore, E4orf4 modulates splicing patterns of AdV mRNAs. However,

most of these functions seem redundant with the expression of other AdV genes and therefore its main function is likely function as a regulator of AdV gene expression. In this function it is known to reduce E1A's upregulation of the E4 promoter. Additionally, E4orf4 downregulates the E2 promoter activation by E1A and to lesser extent the one by E4orf6/7 (Kleinberger, 2020).

E4orf6/7

E4orf6/7 is a fusion product of Orf6's N-terminus and Orf7. Its most prominent feature is the activation of the transcription factor E2F. Notably, E4orf6/7 can substitute for E1A in E2F-dependent transactivation or, more effectively, enhance the transactivation process through its unique binding mechanism. E4orf6/7 achieves this by stabilizing the DNA binding capacity of E2F through self-dimerization and dimerisation of two E2F molecules (M. M. Huang & Hearing, 1989b; Marton et al., 1990; Raychaudhuri et al., 1990). While the activation of the E2 promoter and increased production of the DNA binding protein (DBP) are the most extensively studied aspects of E4orf6/7 function, it is plausible that the activation of E2F (and DP) by E4orf6/7 is also vital for establishing the favourable environment required for efficient viral production within the host cell. Numerous cellular genes possess E2F transcription factor-binding sites in their promoters, including critical genes involved in S-phase induction, such as E2F-1 (D. G. Johnson et al., 1994; Leppard, 1997). The regulation of E4orf6/7 appears to be mediated through SUMOylation by E1B55K. Although mutation studies have not demonstrated reduced DBP levels, the findings suggest that the expression of cellular proteins controlled by E2F-containing promoters may be affected (Melling, 2018).

3.2.2.4 E2A

The gene product of *E2A*, known as AdV's 72 kDa DNA binding protein (DBP), functions as a single-stranded DNA binding protein and binds to the ssDNA genome of AAV, thereby enhancing AAV DNA replication. The Rep proteins facilitate the localisation of DBP to AAV VRCs, thereby promoting replication. DBP plays multiple roles in supporting AAV proliferation, including mRNA export, stability, and processing (Rep splicing), as well as transcription activation of VPs (Carter et al., 1992; Matsushita et al., 1998; Stracker et al., 2004; Ward et al., 1998).

Both E2A and E4orf6 have known post-transcriptional effects on AAV mRNAs. Although the underlying mechanism is unclear, the DBP is believed to enhance transcript levels and translation of AAV RNAs together with VAI RNA (Janik et al., 1981). An important function of E2A is its role as a transcription activator. It can interact with TFs such as the nuclear factors I and III (NFI and NFIII), who's DNA

binding is enhanced drastically by DBP's binding to the NF proteins (Stuiver & van der Vliet, 1990). Thereby, DBP stimulates DNA polymerase binding, thus promoting viral DNA replication (van Breukelen et al., 2003). The activation of the p5 promoter by E2A is particularly relevant for AAV production. There may be a potential feedback regulation between E2A and Rep, as the E2A promoter contains a Rep binding site, and all four Rep proteins inhibit *E2A* gene expression through their purine nucleotide binding site (Casper et al., 2005; Jing et al., 2001; Nada & Trempe, 2002). As mentioned above, the E2A promoter also contains a CRE element, which can be bound and repressed by CREB. Consequently, E1A controls the expression of E2A through the cAMP/PKA pathway (Fax, Lehmkühler, et al., 2000; Fax, Lipinski, et al., 2000).

Contrary to E2A's promotion of AAV production, Carter et al. (1992) demonstrated that production of AAVs is possible without DBP. However, DNA replication, Rep and VP synthesis, and consequently viral titres, were significantly decreased. Matsushita et al. (1998) even stated that among all Helper functions in their plasmid system, E2A has the most significant impact on rAAV titres. This is likely attributed to DBP's transcriptional activation of the p5 promoter and AdV's early promoters, including E1A and its own, as indicated by increased expression rates. In contrast, DBP slightly inhibits the *E4* (L. S. Chang & Shenk, 1990). In contrast, Richardson and Westphal (1981) showed with microinjected nucleic acid experiments, that E2A mRNA is required for *E4* activation and thereby AAV production. Ward et al. (1998) demonstrated that the Adenovirus DBP can be substituted with the homologous human protein RPA, further supporting the findings of Carter et al. (1992). Despite general findings that *E2A* significantly improved rAAV titres, this indicates the possibility of developing an *E2A*-less plasmid system for rAAV production, provided that cultivation conditions and other factors such as gene regulation cascades are appropriately adjusted.

Another function of E2A is its promotion of AAV genome circular concatemerisation, while the presence of E4orf6 favours a linearized form (Duan et al., 1999). However, this concatemerisation might be a general effect from ssDNA binding proteins on the AAV genome. Therefore, the existence of multimeric, circularized concatemers of AAV and rAAV can be explained by the presence of substitute cellular ssDNA binding proteins in the absence of a Helper virus or its functions. Conversely, during lytic infection, when AdV functions including E4orf6 are present, having easily readable and replicable linearized DNA would be advantageous.

3.2.2.5 Virus Associated RNA (VA RNA)

RNA polymerase III transcribes the *VA RNA* gene, resulting in two virus associated RNAs (VA RNAs): VA RNAI (VAI) and VA RNAII (VAII), with VAI being the predominant species. *VA RNA* is indispensable for AAV production, but it has been demonstrated that only VAI is sufficient for AAV production. The abundance of *VA RNA* transcripts is remarkably high, with greater than 10^8 VAI copies per cell, whereas the minor species, VAII, is only expressed at about 5×10^6 copies per cell. The higher abundance of VAI is attributed to its stronger promoter outcompeting VAII transcription during infection (Bhat & Thimmappaya, 1984; Söderlund et al., 1976; Vachon & Conn, 2016).

VAI is a multifunctional non-coding RNA that is essential for AdV and AAV replication due to its remarkable functions in interfering with the host cell's anti-viral defence and cellular processes, including nuclear RNA export, protein synthesis and miRNA generation. VAI's single stranded RNA forms into a conserved secondary structure as a double stranded RNA, containing a terminal stem, a central domain, and is basically folded over in half at the apical stem. VA RNA lengths vary and are typically around 160 nt (~150-200 nt). The variations origin in multiple termination sites and VA RNA's interactions with nuclear factor 1 (NF1) that is involved in RNA Pol III termination. Some readthroughs can reach up to 690-950 nucleotides (Akusjärvi, 1985; Punga et al., 2020). Although the general structure is conserved, VA RNA sequences vary greatly between serotypes, certain regions, including the Box A and Box B RNA Pol III promoter sequences and the complementary sequences GGGU and ACCC in the central domain, being conserved (Ma, And and Mathews, 1996; Vachon and Conn, 2016).

VAI's crucial role in AAV replication primarily stems from its capacity to enhance viral protein translation (M. H. P. West et al., 1987). The most extensively studied action of VA RNA is its inhibition of the cellular protein kinase (PKR), which is activated by dsRNA. Since most viruses express dsRNA at some point due to their dsRNA nature, RNA genome replication, or in the case of AdV and AAV, due to overlapping gene sequences, this is an effective way of virus detection by the cell. By inhibiting PKR, VAI prevents the shutdown of cellular transcription mediated by PKR's phosphorylation of eukaryotic translation initiation factor 2α (elF2α). The mechanism by which VAI acts as an inhibitor rather than an activator of PKR is not fully understood, but it is hypothesized that the abundance of wobble base pairs in VAI allows it to bind to PKR without activating it. Structural features of VAI, particularly the pseudoknot structure in the central domain, play a critical role in determining its inhibitory or activating properties towards PKR. With its inhibition

of PKR, VAI thwarts a cellular defence mechanism against viral takeover and facilitates viral gene expression (Dzananovic et al., 2017; Hood et al., 2019; Kitajewski et al., 1986; O'Malley, 1986; Punga et al., 2020).

The viral-associated RNA exerts inhibitory effects on the host's antiviral defences, particularly by targeting the cellular microRNA (miRNA) and miRNA-related processes. VA RNAs influence miRNA biogenesis through several mechanisms. Firstly, it competes with cellular RNAs for Exportin 5-mediated export from the nucleus, thereby impeding the transportation of miRNAs. VA RNAs are present in cytosol and nucleus, but accumulate in the nucleus during infection. Exp5 transports dsRNA-binding proteins and RNAs into the cytosol, among these are tRNAs, pre-miRNAs and Dicer mRNA (Gwizdek et al., 2003; Lu & Cullen, 2004; Yi et al., 2003). Secondly, VA RNAs saturate the pre-miRNA processing enzyme Dicer, preventing efficient miRNA maturation (Andersson et al., 2005; Vachon & Conn, 2016). Thirdly, the VA RNA molecule directly interferes with the assembly and function of the RNA-induced silencing complex (RISC), which is essential for miRNA-mediated gene silencing. Dicer-processed VA RNAs are incorporated into RISC, enabling them to selectively target cellular genes (Aparicio et al., 2010; Bellutti et al., 2015).

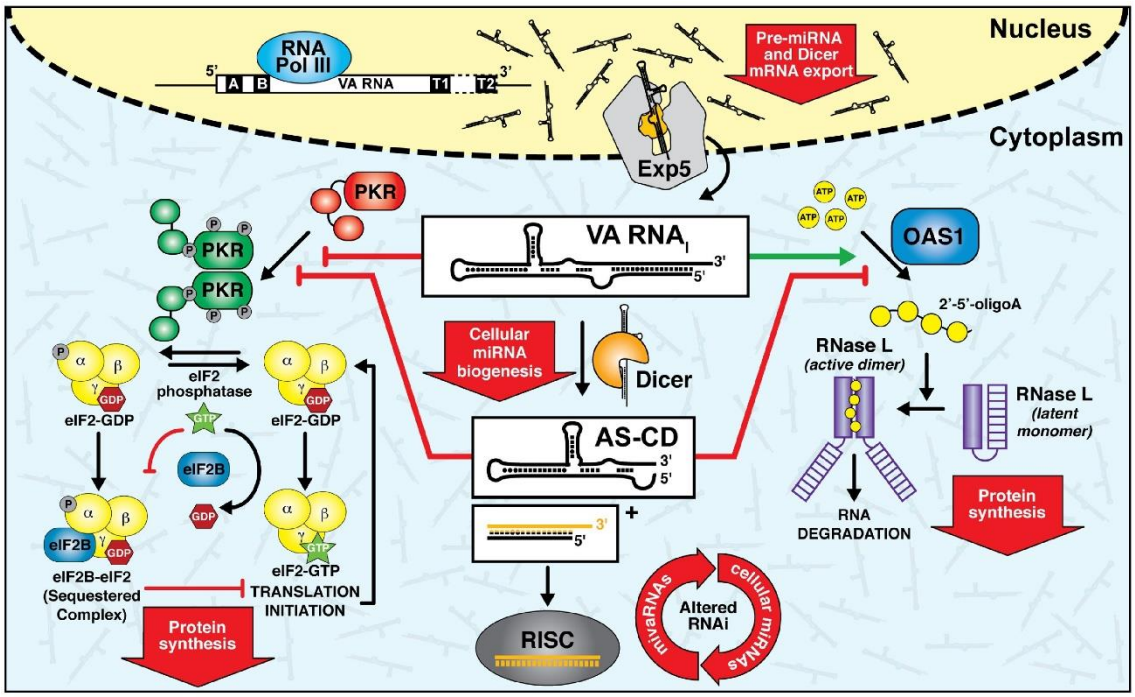


Figure 13: VA RNA interactions and actions against host cell mechanisms (Vachon & Conn, 2016).

In addition to these mechanisms, in which VA RNA subverts the cell's own defence mechanisms, VAI has also non-miRNA-based ways to counter the cell's immune response against viral infections. Notably, VAI interacts with various immune proteins, including OAS1, which, like PKR, is involved in dsRNA detection (Kondo et al., 2014; S. L. Schwartz & Conn, 2019; Vachon et al., 2015). Despite its actions

against the cellular immune system VA RNA still activates cellular inflammatory cytokines and type I interferon expression. Activation of the RIG-I and MDA-5 receptors initiates a signalling cascade leading to the phosphorylation and activation of IRF-3 and NF-κB (Yoneyama & Fujita, 2010). Additionally, during virus infection, the host's inflammasome is activated. However, VA RNA's interaction with PKR enables it to inhibit the inflammasome and suppress cytokine and inflammatory response cascades (Darweesh et al., 2019; Punga et al., 2020). Furthermore, VA RNA has been shown to protect capsids and Rep52 from degradation, mediated by E1B55K/E4orf6 (Meier et al., 2020). These functions elucidate the findings of Matsushita et al. (1998) and other research groups, who have demonstrated that VA RNA promotes capsid production and AAV proliferation.

Overall, the intricate interactions between VA RNAs and host cell proteins, as well as their modulation of antiviral and immune response pathways, underscore their critical involvement in Ad and AAV replication. By enhancing viral protein translation, inhibiting PKR-mediated translation shutdown, protecting viral components from degradation, and modulating miRNA export and synthesis, VA RNAs establish a favourable cellular environment for viral propagation. The multifaceted functions of VA RNAs highlight their significance and warrant further investigation to fully elucidate their comprehensive mechanisms and explore their potential therapeutic applications in the context of viral infections.

3.2.2.6 Rep Interactions

AAV's Rep proteins not only facilitate AAV replication and packaging but also act as transcription, translation, and host cell modulators. Reps regulate various processes, including their own transcription, host cell protein (HCP) transcription and activity, as well as Helper and AAV gene expression. These regulations involve multiple complex mechanisms of positive and negative regulation. Regarding HCP interference, there are at least 188 cellular proteins of all sorts of cellular functionalities interacting with Rep, including proteins from the MCM complex, Ku helicase (DNA replication), RCN1 (membrane transport), SMC2 (chromatin dynamics), PKA, EDD1 (ubiquitin ligase), IRS4 (signal transduction), and FUS (splicing) (Chiorini et al., 1998; Nash et al., 2009). Given the extensive nature of these regulations, this chapter will present only a subset of the commonly observed and relevant Rep functions.

In general, Rep proteins' interactions with host cells can be categorized as Helper functions, as they serve a similar pro-viral purpose as the Ad Helper genes. These interactions modulate host cell functions to ensure enhanced viral proliferation, downregulation of HCP production, and suppression of anti-viral responses. Many

of these interactions exhibit similarities or redundancy with the actions of the Helper functions, while some may counteract them. This counteracting effect may be a result of Rep proteins' additional role in minimizing resource competition with Adenovirus proliferation (Timpe et al., 2006).

One primary mechanism by which Rep prepares the cell for virus production is through cell cycle arrest, particularly in the S-phase. Rep78 achieves this by binding to p53 and preventing adenoviral degradation of the cell cycle and tumour suppressor gene. As a result, Rep78 inhibits cell growth and cell cycle progression. However, Saudan et al. (2000) report that both Rep78 and Rep68 induce cell cycle arrest. Rep68 seems to arrest cells more in the G1 and G2 phases, while Rep78 in S-phase. Cells expressing Rep78 and arrested in the S-phase accumulate hypophosphorylated pRb. The retinoblastoma protein (pRb) is a tumour suppressor protein that inhibits cell cycle progression at the G1-checkpoint by blocking S-phase entry by repressing the transcription of genes required for cell cycle progression. Phosphorylation of pRb inactivates it, relieves repression, and allows the cell to enter the S-phase. Repression is then achieved through chromatin remodeling and binding to E2F sites in the promoters of these genes (Giacinti & Giordano, 2006). Interference with these cellular mechanisms is familiar from the essential AdV Helper genes, indicating Reps collaboration with these, but also the difficulty of differentiation between individual functions and their effects. Notably, additional mechanisms are suggested by which Rep78 induces S-phase arrest. One of them involves Rep78 binding to cell cycle regulatory phosphatases, while another involves Rep78 causing DNA damage response (DDR)-mediated cell cycle arrest by nicking cellular chromatin (Berthet et al., 2005; Saudan et al., 2000). DDR pathways are activated through Rep's binding to and unwinding of DNA. However, DDR can also be provoked by the inverted terminal repeats (ITRs) and the single-stranded nature of the AAV genome. Interference of AAV with DDR may serve to facilitate second-strand synthesis and promote AAV replication in general (Vogel, 2013).

The indirect transcriptional regulation of Rep78, achieved through pRb inhibition and activation of E2F-containing promoters, is one of many mechanisms by which AAV regulates its own, Helper, and cellular protein expression. Rep78 also interacts with the transcription cofactor PC4, resulting in transcriptional interference. However, the relief of this effect in the presence of Adenovirus indicates the involvement of Helper functions in these regulatory processes (Weger et al., 1999). The interplay between Rep proteins and Helper functions, such as the downregulation of E2A by Reps, has been mentioned earlier. Rep78 is further known to repress the promoters of *E1A* and *E4*, but transactivates *E1B* without E1A expression. Co-expression of E1A causes Rep78 to repress all these promoters (Jing

et al., 2001; Timpe et al., 2006). Studies exploring the inhibition of heterologous and endogenous promoters by Rep have revealed that among the Rep proteins, Rep78 demonstrates the highest level of inhibition. The degree of inhibition decreases as the size of Rep increases, and it becomes nearly negligible with Rep40. Moreover, the results suggest that the inhibition of heterologous promoters is mainly mediated through protein-protein interactions, rather than Rep directly binding to DNA. (Hörer et al., 1995). Through interactions with transcription factors, cofactors, and other cellular proteins, Rep can profoundly alter the transcriptional landscape and repress many promoters even in the absence of E1A, such as those of SV40, HIV, and CMV (Antoni et al., 1991; Heilbronn et al., 1990; Hermonat, 1994; Hörer et al., 1995; J. Li et al., 1997). Strong or overexpression of the large Rep proteins, particularly Rep78, not only disrupts the cell cycle but also exhibits cytotoxicity, leading to apoptosis and drastic reduction of recombinant AAV (rAAV) titre (J. Li et al., 1997; Schmidt et al., 2000).

One of the most relevant regulatory functions of Rep for rAAV production is its impact on the AAV promoters. In the absence of Adenovirus Helper functions, Rep78/68 represses protein expression from all three confirmed AAV promoters (p5, p19, and p40) through different mechanisms. During latent infection, these promoters exhibit low basal transcription, repressed by all four Reps. Repression of p5 occurs despite the involvement of YY1 and MLTF. This repression can then be relieved by E1A and E2A, respectively. YY1 acts as a repressor, but in its presence, it becomes an activator of the p5 promoter. YY1 repression is in fact facilitated by Rep68 through binding to the RBE of p5. Thus, E1A indirectly relieves the auto-repression exerted by the large Reps. Down-regulation of p19 is also Rep mediated and just as with p5, this inhibition is relieved in the presence of Helper virus (Beaton et al., 1989; Kyöstiö et al., 1995). The p40 gene products are also repressed by the large Rep proteins, but in this case, the repression occurs at the translational level and not transcriptionally. These effects are not implemented through Rep binding to a regulatory element of the promoter itself but are mediated by cellular or Helper virus genes (Tratschin et al., 1986).

In addition to their inhibitory functions, Rep78/68 can also act as activators (Pereira et al., 1997). Both, ITR and p5 RBEs have an activating function for p19 and p40 in the presence of Adenovirus Helper functions. The mechanism by which Rep78/68 activates p19 transcription involves the interaction of Rep78/68 bound to the RBE with the SP1 transcription factor. SP1 can bind to p19 at positions -50 and -130 relative to the transcription start site (TSS). Additionally, interaction with a CArG-like element at position -140 is required. Activation of p19 requires Rep78/68, Sp1 and either the ITR or p5 RBE for large Rep binding. In the absence of ITRs, the large

Reps and TF Sp1 can form a complex that acts as a scaffold, positioning the p5 YY1 complex next to p19 and thereby delivering 5'-enhancer elements for increased p19 expression (Lackner & Muzyczka, 2002). In contrast, p5 is only activated by the ITR RBE, while its own RBE prepresses p5 in the presence of Helper functions. However, small Reps exhibit the ability to alleviate p5 inhibition, indicating certain feedback loop regulations between the Reps by the formation of a Rep78/68-Rep52 complex (Pereira et al., 1997). The transactivation mechanism of p40 is very similar to that of p5. It also requires an upstream RBE, either from p5 or the ITR, as well as the CArG-140 element in the p19 promoter and the SP1-50 element in the p40 promoter (Pereira et al., 1997; Pereira & Muzyczka, 1997a, 1997b). These transactivations are critical for AAV proliferation and, consequently, for rAAV production, as they are required to initiate maximal transcriptional activity for the Rep52/40 proteins, but also for regulation of adequate amounts of all other AAV proteins.

In orchestration with all other cellular and viral proteins and mechanisms, Rep-mediated transcriptional and translational (splicing) control specific regulates AAV proteins to specific stoichiometries. This leads to p5 transcripts being expressed the least, followed by p19 transcripts, which are increased due to Rep enhancement, and p40 transcripts as the most abundant AAV mRNA species. Additionally, spliced variants of all AAV mRNAs increase during infection of Adenovirus due to adenoviral stimulation of splicing and mRNA export. In the late stage of infection up to 90 % of all p40 mRNAs and up to 50 % of all p5 and p19 transcripts are spliced. AAV mRNAs possess relatively high stability (>6h persistence after actinomycin D treatment) and their ratio in late stage transfection is regulated to approximately 1:3:18 (p5:p19:p40), reflecting the individual promoter strength in presence of AdV Helper functions and Rep. Interestingly, expression levels of components change over time, with the evident logic of p5 mRNAs being transcribed first and predominant p40 in the late stage when capsids are required (Mouw & Pintel, 2000).

3.3 Recombinant Production of AAV

As the interest in recombinant adeno-associated viruses (rAAVs) as delivery vectors for gene therapy continues to grow and an increasing number of AAV gene therapies enter pre-clinical and clinical trials, there is growing demand for efficient and scalable platforms for high-titre AAV production (Clément & Grieger, 2016; Smith et al., 2018; Srivastava, 2016).

3.3.1 Transient Triple Transfection

Initially, plasmid transfection-based rAAV manufacturing was enabled by Jude Samulski cloning the AAV genome into a pBR322 plasmid (pSM620) (Samulski et al., 1982). Very early rAAV productions used transfection of adherent cells, first with plasmids exchanging parts of the AAV genome, resulting in the first transduced cells using AAV as a gene transfer vector (Tratschin et al., 1984, 1985). It was discovered that the ITRs are the only *cis* elements required for packaging and virus production. All other viral genes could be provided in trans and the AAV genome between the ITRs can therefore be replaced by the GOI. This concept, together with the construction of the first plasmid allowing to swap the content between ITRs (psub201) enabled exchange of the full AAV genome and elimination of replication competent, wildtype AAV contamination (Samulski et al., 1987, 1989). The result: separated Rep/Cap and an ITR/GOI plasmids, which were transfected into mammalian cells like HeLa or Detroit 6, e.g. by DEAE-dextran method, along with co-infection of Adenovirus (AdV5) to supply the required Helper functions. However, this method still resulted in a mixture of rAAV and AdV in the producing cells, due to the use of replication competent AdV as Helper virus. To remove the AdV, laborious purification techniques such as heat treatments and CsCl density gradients for ultra-centrifugation were necessary (Hermonat & Muzyczka, 1984; McLaughlin et al., 1988; Samulski et al., 1982, 1983, 1989; Samulski & Muzyczka, 2014). Unfortunately, toxicity caused by contaminating Ad components was still reported frequently.

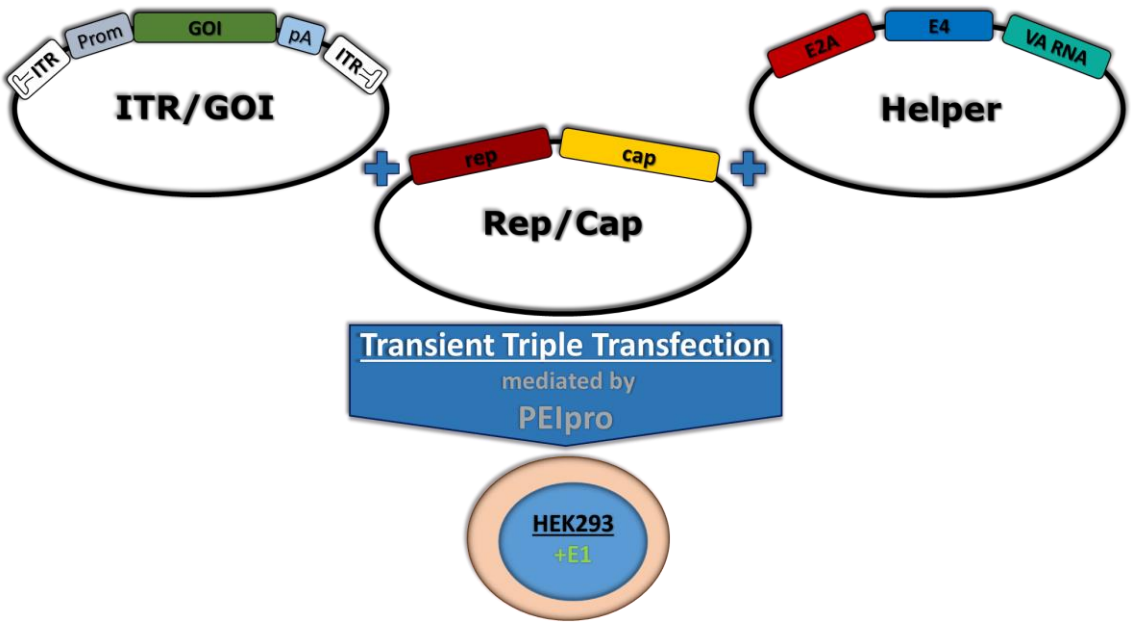


Figure 14: Schematic overview of the used transient triple transfection system, including the three plasmids ITR/GOI, Rep/Cap and Helper. PEIpro mediated transfection is carried out in HEK293 suspension cells, which stably express the Helper gene E1.

A major breakthrough in rAAV production came with the definition of essential AdV-genes for effective AAV replication and their subsequent cloning into a bacterial plasmid. Ferrari et al. (1996) produced the first rAAV without AdV5 co-infection, whilst investigating *E4orf6*'s accelerating effect on rAAV transduction through increased second strand synthesis. However, this approach isolated the DNA for transfection from previously produced dl309 AdV. About two years later, the same group performed the cloning of plasmid pXX5 and pXX6, the first AdV Helper plasmids (Xiao et al., 1998). This eliminated AdV contaminants for the first time. Additionally, an improved Rep/Cap plasmid (pXX2) was created and sequences that could potentially be used for homologous recombination were removed from the plasmids, greatly reducing the risk of wildtype-like, replication-competent AAV contaminations (Aucoin et al., 2008; Matsushita et al., 1998; Natsoulis, 1997; Samulski et al., 1989; X.-S. Wang et al., 1998). This all plasmid-based approach, known as transient triple transfection, is typically performed using HEK293 cells, which already carry the E1 Adenovirus Helper genes due to their immortalisation process, making it obsolete to include E1 in the AdV Helper plasmid. To the present day, the transient triple transfection method remains one of the most commonly employed techniques for rAAV production, particularly in research, but also numerous pre-clinical and clinical trials, due to its flexibility and convenience (Aucoin et al., 2008; Colella et al., 2018; J. F. Wright, 2008).

Over the years, rAAV titres from transfection of adherent cells could be raised from 100-1000 vg/cell (virus genome copies per cell) in the early attempts, to 2×10^5 vg/cell, matching the improved specific titres achieved with wild-type AAVs (Samulski & Muzyczka, 2014; J. F. Wright, 2008). Triple transfection of adherent HEK293 cells is a well-established laboratory approach for relatively quick small-scale production, allowing screening of various rAAV constructs by simply changing one or two plasmids. Albeit this method is used for production of material for many clinical trials, adherent cell cultures suffer from poor scalability due to the two-dimensional growth limitation, as opposed to suspension cultures that can expand in three dimensions. For instance, just a single preclinical study for Duchenne muscular dystrophy (DMD) would require 1250-25000 Corning 10-CellSTACK devices for production, equivalent to up to 2 million 10 cm-dishes, assuming three cohorts of three patients and the currently tested DMD doses ranging from $5x10^{13}$ to 5×10^{14} vg/kg (Deng et al., 2022; Merten, 2016). Other AAV based drugs that are applied intravenously, are similarly dosed, e.g. Zolgensma $1.1x10^{14}$ vg/kg. Notably, downstream processes for these GMP grade, highly purified doses have a limited of up to efficiency 65-80 %, but due to requirement of multiple purification steps, recovery rates are often only between 5-25 %. That means with a current upstream process which can achieve a crude titre of up to 1×10^{14} vg/L and a 5 kg weighting infant, about 10 to 100 L cell culture would be required per treatment (Clément, 2019; Florea et al., 2023; Zhao et al., 2020). These calculation examples demonstrate scale-up difficulties and limitations of adherent systems, but also the general demand for scalable production (and purification) methods and their improvement even with current suspension cell-based bioreactor processes.

The utilisation of suspension cell cultures for rAAV production is one step towards achieving better scalability. Triple transfection of suspension cultures offers similar advantages and disadvantages to transient transfection of adherent cultures, such as versatility and the ability to quickly exchange plasmids. However, plasmids are required, which can be a significant cost factor in large-scale current Good Manufacturing Practice (cGMP) quality production (Cameau et al., 2019; Merten, 2016). However, unlike adherent systems, suspension-based approaches can be scaled up instead of scaled out. For example, Durocher et al. (2007) demonstrated the scalability of rAAV2 production using 293F cells in bioreactors, which allowed for process regulation and increased titres. Grieger, Soltys, and Samulski (2016) also showed the feasibility of large-scale production with transient triple transfection of 293F cells in a 20 L wave bag reactor. Compared to adherent triple transfection, suspension cultures generally use serum-free, chemically defined medium and polyethylenimine (PEI) as a transfection agent, instead of media containing

animal-originated serum and calcium phosphate as transfecting agent for the plasmids, which are often applied in an approximately equal molar ratio. Initially, the achieved titres per cell in suspension cultures were only a tenth of those obtained in adherent production, at 1.4×10^4 vg/cell (Park et al., 2006). However, Grieger et al. (2016) developed a HEK293 suspension cell line, cultivation and transfection protocols that increased specific production to 1×10^5 vg/cell, matching the specific titres achieved with adherent triple transfection. Moreover, the volumetric titres in suspension cultures are even higher due to the higher cell densities (Clément, 2019; Kotin, 1994). With their optimized cell lines and procedures, crude titres exceeding 1×10^{11} vg/mL can be achieved today. Extensive research is carried out to optimize transfection agents, media, cells, PEI to DNA ratios, plasmid ratios and many more variables of the transient transfection process. However, high production costs, moderate yields, and limited scalability still remain challenges for transient transfection systems, particularly for the production of future mainstream therapeutics. Nevertheless, the inherent flexibility of the manufacturing method makes it potentially appealing for individual therapies and less prevalent diseases, which includes many genetic diseases. Further, as of today, it has high relevance due to its long establishment and current use in a multitude of ongoing studies (Clément, 2019; Merten, 2016; Zhao et al., 2020).

For the sake of completeness, it is worth mentioning non-transfection-based alternatives for large-scale production. These alternatives can be broadly divided into two approaches. The first approach involves infection with other viruses, such as herpes simplex virus or baculovirus, as transducing vectors, effectively replacing the use of polyplexed plasmids. In this case, the baculovirus itself serves as the Helper function. Baculovirus systems, such as OneBac2.0, typically employ insect cells (e.g., Sf9) with separate baculovirus vectors encoding *cap*, *rep*, and ITR/GOI sequences. These systems are currently capable of large-scale production up to 200 L and yield high titres of $>10^{12}$ vg/mL (Cecchini et al., 2011; Merten, 2016; Mietzsch et al., 2015). Another remarkable approach is the recently presented TESSA/TESSA2.0 production system, employing AdV as the vector to transduce HEK293 cells for AAV production. Whereas quality and infectivity of rAAV produced by baculovirus systems are often much worse than HEK-based ones, the TESSA2.0 system does not only show up to 30-fold higher titres than triple transfection, but also increased infectivity with the additional advantage of better scalability (Clément, 2019; Su et al., 2022). However, the second general approach offers even better scalability and involves the use of stably transfected cell lines, such as HeLa or BHK, that stably express the Helper and/or AAV genes. In these cases, the integrated genes are likely controlled by inducible promoters due to the cytotoxic

nature of the large Rep proteins. Cell lines with only the *rep* and *cap* genes integrated are referred to as packaging cell lines, which still require plasmids for delivering ITR/GOI and Helper functions or a Helper virus. Adding a stably integrated sequence for ITR/GOI transforms a packaging cell line into a producer cell line that only requires the infection of a Helper virus or transfection of Helper virus genes (Aucoin et al., 2008; Clément & Grieger, 2016; Merten, 2016). Another approach, which is even more versatile and advanced, involves the stable integration of inducible *rep* and Helper genes in a cell line, followed transient transfection or stable integration of *cap* and ITR/GOI genes for rAAV production. An example of such an approach is the ELEVECTA® Alpha Cell Line developed by Cevec, which can yield titres of up to 10^{11} vg/mL (Cevec, 2020).

3.3.2 Genetic Engineering of Triple Transfection Plasmids

Viruses can be turned into viral vectors by removing unessential and unwanted sequences of the virus' genome until only the essential parts for the production of this therapeutic trojan horse remain (Bouard et al., 2009). The development of the plasmid-based, Helper virus free, transient transfection approach was mainly based on the work of the Samulski lab, creating the three plasmids ITR/GOI, Rep/Cap and Helper for rAAV manufacturing. Over the course of time remarkably little changes were made to the basic foundations and compositions of these plasmids. Initially, in the first decade of rAAV production via transient triple transfection, most molecular cloning efforts focused on modifying sequences between the ITRs. This was done to exchange different GOIs and optimize the expression of the therapeutic DNA by altering the sequences themself, their promoters, and 5'- and 3'-UTRs. Recent years have witnessed an increased emphasis on the development of new, synthetically designed, and modified capsids, making *cap* a very relevant and prominent target for molecular engineering. This focus on capsid engineering aims to enhance tissue specificity, making it a prominent area for molecular engineering. However, this work primarily focuses on optimizing rAAV production at the plasmid level to increase vector genome (VG) titres. Therefore, previously performed changes to plasmids for production enhancement will be presented in the following.

One of the most important improvements to the Rep/Cap plasmid for virus titre augmentation was made very early after the first Rep/Cap plasmids were created, by (J. Li et al., 1997). Their exchange of promoters for *rep* lead to the realisation that Rep overproduction is counterproductive and even cytotoxic. Consequently, the start codon was changed from ATG to ACG to reduce Rep78/68 expression. Additionally, a p5 promoter was placed downstream of the genome's poly-adenylation (pA) signal. The addition of a second p5 aimed to enhance the expression of other AAV promoters, as the deletion of p5 had previously led to their

downregulation. These modifications resulted in a 15-fold increase in rAAV titre (J. Li et al., 1997; Xiao et al., 1998). Most Rep/Cap plasmids used in laboratories today closely resemble the resulting pXX2 plasmid. However, many incorporate an additional change to reduce large Rep expression, namely the deletion of most of the promoter including its TATA box. Reconstruction of the origin of this modification, when and by who it was carried out first, turns out challenging. It is uncertain if plasmids patented by Wilson and Xiao (1998) were the first incorporating this change or if the idea is originated in the work of Colosi and colleagues, who also filed patent applications for p5 deletions in 1998 (Surosky et al., 1997; Georges Natsoulis, Gary Kurtzman and Peter Colosi, 2005; Peter Colosi, 2008). Other groups, including Wustner et al. (2002) and Grimm et al. (2003) explored similar ideas.

Grimm and colleagues destroyed the TATA box and used a 5'-UTR frt for large Rep expression in one experiment and used a MMTV-LTR promoter instead of p5 in another. The first of these experiments was performed along their attempt of cross-packaging, meaning that *rep* and *cap* origin from different serotypes. Something that became the default solution for all rAAV, using different *cap* serotypes, because standardised use of serotype 2 *rep*, for packaging of ITRs of serotype 2, enables very convenient GOI exchanges. The second experiment, employing a MMTV-LTR promoter instead of p5, was performed to eliminate Rep-mediated replication, packaging and homologous recombination and therefore prevent wtAAV contaminations (Grimm et al., 1998, 2003). A concern that was previously addressed by flipping the direction of the AAV genes by Allen et al. (1997). In their experiment, Allen et al. (1997) additional exchanged promoters of *rep* and *cap*, an approach that was performed by several groups at the time (Ogasawara et al., 1998; Vincent et al., 1997). Although promising results were shown, particularly with the exchange of the p40 promoter to a CMV promoter, these approaches did not become the preferred methods for AAV production. As a result, interest and research in promoter exchanges for production enhancement were only revived recently.

For the Helper plasmid, three initial designs were suggested. The 21.8 kb large pDG plasmid, which contains the AAV(2) genome with its p5 substitute to the MMTV-LTR promoter. The genes are arranged in the order: *rep* and *cap*, *VA RNA*, followed by *E2A* and *E4* on the complementary strand. The adenoviral genes of another of these three first Helper plasmids, pXX6, which was already mentioned above, are arranged the same, but the plasmid does not contain the AAV genes. The third publication presenting a Helper virus for Helper contaminant free rAAV production, however, has a different sequence of genes and explored the individual genes more in depth. Matsushita et al. (1998) created individual fragments of the three essential genes and combined them to full Helper plasmids of 13.9 kb and 11.6 kb of AdV5

and AdV2, respectively. Fragments were combined as such that the promoters of *E2A* and *E4* for are next to each other and the genes are facing opposite directions, the *VA RNA* gene was placed at the 3' end of *E4* and transcription was directed towards the *E4* gene. By using subsets of the three essential genes, they could also demonstrate that only *E2A* or a combination of *E2A* and *VA RNA* or *E4* are efficient for rAAV production in 293 cells, but only the combination of all three resulted in titres equivalent to AdV co-infections. Further, individual genes on individual plasmids, controlled by CMV showed that *E4orf6* was equivalent to the full *E4* gene. Heterologous promoters to control the AdV Helper genes were also used by Allen et al. (1997). Contrary to the Matsushita et al. (1998) publication it was found that E4orf6, transcribed from a CMV promoter, was the only required AdV gene for rAAV production. Notably, *VA RNA* and *E2A* independence could only be achieved by the use of heterologous promoters for *rep* (MT promoter) and *cap* (CMV), whereby p40 mediated VP expression was not possible with *E4orf6* alone. Furthermore, the use of *rep* and *cap* split onto two individual plasmids was beneficial, which might have been due to the proximity of MT and CMV promoter in the previous Rep/Cap plasmid. Nevertheless, the approaches of heterologous promoters. Matsushita et al. (1998) and Allen et al. (1997) did not prevail. Instead, the two most used Helper plasmids for clinical trials are pXX6 (pXX6-80, now pALD-X80) and the commercially available pADΔF6. Both possess the same gene sequence *E4*, *E2A*, followed by *VA RNA* on the non-coding strand. Likely pADΔF6 is thereby a derivative, excluding further redundant sequences and reducing the plasmid size by 3 kb to 15.7 kb (Hildinger et al., 2001; Xiao et al., 1998). Among some other Helper plasmids that have been utilized in experiments and studies, the Avigen/Agilent pHelper plasmid is of particular relevance, as the same or very similar plasmids are currently available from different sources. It offers a further size reduction to 11.6 kb and practically resembles the Helper plasmid described by Matsushita et al. (1998), already patented by the same researchers (Colosi, 1996).

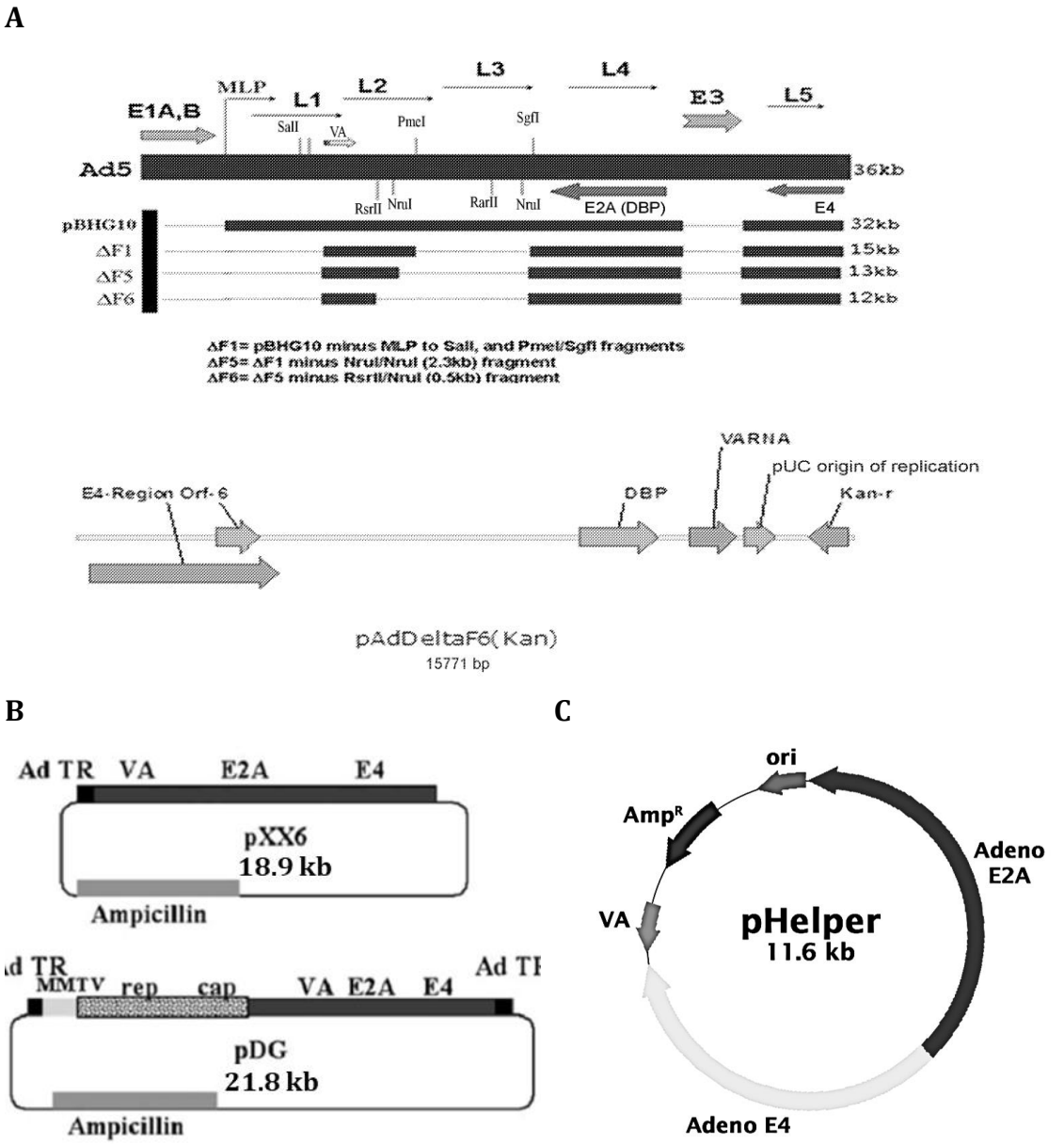


Figure 15 : Initial AdV Helper designs for Helper virus free rAAV production. (A) Construction of pADΔF6 by Hildinger et al. (2001). (B) Helper plasmid designs of pXX6 (Xiao et al., 1998) and pDG (Grimm et al., 1998) modified schematic from Chadeuf et al. (2005). (C) Plasmid map of Avigen's pHelper, similar to design of Matsushita et al., (1998) (Colosi, 1996).

The lack of optimisation of these plasmids, which have been in use for decades, bares opportunities for improvement of rAAV titre increases and vector quality. This is for example shown by the size reduced Helper plasmid of Emmerling et al. (2016), which achieved a >10-fold increase in vg/cell to the pDG plasmid. The research of Emmerling et al. (2016) also rediscovered the idea of previous approaches of separating *rep* and *cap* for individual control, cloning them into different plasmids while also eliminating Rep78. Other groups revisited the idea of dual instead of triple vector systems for transient transfection (Tang et al., 2020; van Lieshout et al.,

2023). One such design resembles the design of the pDG plasmid developed by Grimm et al. (1998).

Additional avenues for improvement include the ITR plasmid, which has seen various changes to optimize promoter, GOI sequence, 5' and 3' UTRs for transduction efficiencies. However, the overall structure of the ITR plasmid has remained largely unchanged from the first cis plasmids, except for the addition of MCSs for easier exchange of GOI, promoter, as well as 5' and/or 3'-UTRs, for example by Yue and Duan (2002). Other studies have investigated altered ITRs, enabling self-complementary AAV and thereby improving transduction efficiency by eliminating the need for second strand synthesis. Optimized ITRs could further increase packaging and genome replication during production, particularly for pseudotyped rAAV. Moreover, Rep and capsid modifications have the potential to enhance titres, while newly discovered proteins like MAAP and AAP offer opportunities for productivity improvement through molecular engineering. Recent research by Galibert et al. (2021) demonstrated that MAAP variants increased AAV2 titres by 3.5-fold while reducing contaminant packaging. Brimble et al. (2022) made changes to p5 promoter, involving introductions of spacer sequences, to prevent RBS dependent packaging of contaminant sequences into capsids. Genuinely, the promoters of rAAV systems bare hold additional potential beyond vector payload tissue specificity, as demonstrated earlier by Allen et al. (2000).

3.3.3 Synthetic Promoters and their Use for rAAV Production

The utilisation of heterologous promoters for the production of recombinant proteins in mammalian cells is common practice. For instance, monoclonal antibodies (mAbs) are typically produced in CHO cells, where foreign promoters like the hCMV-IE promoter (referred to as CMV) or the simian virus 40 early promoter (SV40) are used to express the heavy and light chains of the protein (Romanova & Noll, 2018). The most commonly used promoters are constitutive and derived from viruses. While these promoters are generally strong, they can be susceptible to epigenetic silencing and cell-cycle dependency (Prösen et al., 1996; Brightwell et al., 1997; Brooks et al., 2004; Kim et al., 2011). Furthermore, both endogenous and heterologous promoters used for recombinant protein production can be relatively large, which poses challenges for rAAV vector capacities, as well as also transient transfection due to limited transfection efficiency of larger plasmids.

In recent times, synthetic promoters have emerged as a viable alternative for recombinant protein production, offering advantages such as smaller size, resistance to silencing, predictable and increased transcriptional activity, and even inducibility (Brown et al., 2014; Johari et al., 2019; Romanova & Noll, 2018).

Typically, synthetic promoters are generated either by screening randomized DNA sequences or by assembling publicized cis-regulatory elements upstream of a minimal core promoter (Brown et al., 2014; Ogawa et al., 2007; Schlabach et al., 2010). Although the binding sites of transcription factors (TFs) or transcription factor regulatory elements (TFREs) of popular promoters like CMV, widely used for recombinant expression in mammalian cell lines, are well known, the complexity of this promoter in an organism that evolved to reproduce in various species and tissues is immense (Coulon et al., 2013; Sinzger et al., 2008; Stinski & Isomura, 2008). Hence, CMV's activity in different cells and tissues varies (Qin et al., 2010). Synthetic promoters offer the ability for increased transcriptional control in a certain cell type by analysis of available TFs or the activity of TFREs. Based on this concept, our laboratory has conducted bioinformatic analysis and screening of active TFREs, leading to the discovery of novel compositions of TFRE blocks. This has resulted in the creation of synthetic promoter libraries that encompass relatively small promoters with variable activity, up to 2.2 or 2.5 times higher than CMV (Brown et al., 2014; Johari et al., 2019).

To screen TFRE activity, heterotypic promoters with 6 or 7 copies of the same TFRE are employed upstream of a minimal core promoter, usually derived from CMV. Pre-selection of TFREs for testing is done through various bioinformatic analyses, usually based on RNA sequencing combined with TF binding prediction upstream of relevant genes. Transcriptional strength can then be measured by detection of a reporter protein such as GFP or the secreted alkaline phosphatase (SEAP). The use of homotypic promoters in different cell lines enables screening for specificity. By combining TFREs into heterotypic promoters, synthetic promoters with various properties can be created (Brown et al., 2014, 2017; Johari et al., 2019; Johari, Mercer, et al., 2021; A. O. Johnson et al., 2022).

Different heterologous and synthetic promoters are regularly used for defined and even tissue-specific expression of the therapeutic gene packaged in rAAV (Chai et al., 2023; Greig et al., 2021; Nieuwenhuis et al., 2023; Skopenkova et al., 2021). However, the use of heterologous promoters to enhance rAAV production is currently limited and has mainly been explored in the past or with alternative expression systems, such as the baculovirus system and stable cell lines, rather than for transient triple transfection of HEK293 cells (J. M. Allen et al., 2000; Cevec, 2020; Grimm et al., 1998; Mietzsch et al., 2015; Ogasawara et al., 1998; Qiu & Pintel, 2002; Reed Clark et al., 1996; Su et al., 2022). Only recently, there has been renewed interest in modifying AAV promoters for transient transfection systems, as demonstrated by the modifications to p5 by Brimble et al. (2022) and the use of a tetracycline-regulated capsid expression system by Ohba et al. (2023). Therefore,

the consideration of synthetic promoters for regulated gene expression of *rep* and *cap*, as well as Helper genes, holds promises for enhancing rAAV production.

3.3.4 Use of Small Molecules for rAAV Manufacturing

Small molecule chemical additives have been recognized for their ability to enhance recombinant protein production in biopharmaceutical manufacturing processes. The addition of chemicals can significantly improve a cell's biopharmaceutical production capabilities, although the effectiveness varies depending on the specific cell line, process, and product (M. J. Allen et al., 2008; Johari et al., 2015; W. C. Yang et al., 2014). For example, a study by Yang et al. (2014) demonstrated that the addition of valproic acid (VPA) to antibody-producing CHO cells resulted in significant increases in one cell line but had no effect on another. It is important to note that the concentration and timing of chemical administration are critical variables that need to be optimized.

Other well-known additives and their intended functions include the use of sodium butyrate (NaBu) as histone deacetylase inhibitor, chemical chaperones like tauroursodeoxycholic acid (TUDCA), or proteasome inhibitors like MG132 (Jiang & Sharfstein, 2008; Kusaczuk, 2019; Mitchell & Samulski, 2013). However, often the cause of effect for many of these additives remain unknown, especially since large screening panels of small molecules are in use (J. Chang et al., 2020). For instance, the addition of different peptones is believed to not only influence the availability of amino acids or peptides but also affect the regulation of transcriptional and translational processes in various manners (Pham et al., 2005). The popular cryoprotectant dimethyl sulfoxide (DMSO) can also be used as an additive to enhance protein production. However, the cellular mechanisms underlying its effects are diverse, similar to those of many other small molecule additives. These mechanisms include effects on glycolysis enzymes, induction of G0/G1 phase, chaperone proteins for PTMs, and more (J. Li et al., 2006; C.-H. Liu & Chen, 2007).

It has long been recognized that chemicals can also impact the production of AAV. As early as 1988, Yalkinoglu et al. (1988) demonstrated that the addition of metabolic inhibitors causing genotoxic stress could substitute for the Helper virus in AAV's genomic rescue and replication. In 1996, Ferrari et al. showed that hydroxy urea stimulates second strand synthesis and therefore AAV transduction und production by its effects on the cells. Further, rAAV small molecule studies were carried out searching for chemicals to increase the efficacy of the vector. Mitchell and Samulski (2013) studied proteasome inhibitors and Nicolson et al. (2016) performed high-throughput screenings, identifying vector specific transduction enhancements of different chemical groups, with epipodophyllotoxins being found

to be most effective. Additionally, attempts have also been made to use small molecule additives for production. Yu et al. (2021) demonstrated up to 10-fold increased VG titres can be achieved by adding salts like NaCl and KCl. Interestingly, the effects were not due to changes in osmolality, but were rather thought to be caused by cellular changes creating a favourable cellular environment for AAV production. However, these beneficial results were specific to their HSV-based platform and could not be replicated in a transient context, highlighting once again the individuality and process specificity of chemical additives. In the context of transient, PEI-mediated transfections for rAAV production, Hildinger et al. (2007) demonstrated that, in addition to the large benefits of media optimisation, the addition of soy peptone could increase AAV yields by 30 %.

Although small molecules show encouraging effects for rAAV production, the use in manufacturing is not as established as it is for recombinant protein production. Nonetheless, many process development challenges are transferable to the more complex rAAV production. Therefore, employing chemical screenings to modulate cells into optimized virus factories, akin to the functions of Helper viruses, represents a sensible step for implementing and optimizing a manufacturing process for high virus titre rAAV production.

4 Aim

The primary objective of this research project was to enhance the production of recombinant adeno-associated viruses (rAAV). To achieve this goal, a novel plasmid system was to be developed for use in the widely employed transient triple transfection methodology within rAAV manufacturing processes.

The main task of this work was the genetic engineering the Rep/Cap and Helper plasmids, utilizing tools and principles derived from synthetic biology. The design criteria for these plasmids included two crucial aspects (i) enhanced flexibility to facilitate subsequent genetic modifications and (ii) the ability to independently manipulate their constituent genes. These prerequisites should lay the foundation for a majority of the work and the overarching strategy of the project: exerting control over the individual genes of the system through the utilisation of synthetic promoters. As an integral part of this endeavour, these promoters were to be engineered to optimize the relative quantities of the system's components. Ultimately, the system was supposed to be fine-tuned and balanced on transcriptional level, aiming to reduced manufacturing costs by yielding higher quantities and potentially enhancing the quality of rAAV production.

As part of a larger research effort, this PhD project involved two additional researchers for a portion of its duration. Consequently, this thesis focuses primarily on the optimisation of the Helper plasmid.

5 Material and Methods

Experiments were conducted with the following methods if not indicated otherwise.

5.1 Bacterial and Molecular Biological Methods

Molecular genetic work and required DNA engineering is commonly performed in the bacterial organism *Escherichia coli* (*E. coli*). Cultivation, cloning, and other methods involving this bacterium and the manipulation of DNA are described in the following.

5.1.1 Cultivation of *E. coli* Cultures

Amplification of plasmids was performed with cultures of Escherichia coli. For Helper plasmids the strain NEBstable (NEB) was used, all other plasmids were amplified in DH5α (Invitrogen).

Cells were grown on LB agar medium supplemented with the appropriate antibiotic in 10 cm plastic petri dishes for single colony selection after plasmid transformation into the bacterial cell. Plates were incubated at 37°C overnight, about 16 h for DH5α cells, and about 20 h for NEBstable cells.

Bacterial suspension cultures for plasmid isolations were grown in 50 mL centrifuge tubes in 10 mL of LB broth supplemented from 1000× stock solutions with the corresponding antibiotic for small scale (mini-prep) or 30 mL in a 250 mL glass Erlenmeyer flask for medium scale plasmid preparations (Midi-prep). Cultures were incubated at 37°C for 15 to 17 h (DH5α) or 20 to 24 h (NEBstable) in a shaking incubator (Infors) at a speed of 200 rpm.

5.1.2 Creation of Chemically Competent *E. coli* Cells

Transformation competent *E. coli* of the strain DH5α were purchased from Invitrogen. However, cells from the strain NEBstable were purchased only once and afterwards propagated and in house chemically prepared to make them competent for plasmid uptake. Therefore, cells were taken from the purchased cryo-culture and an overnight pre-culture was grown in 5 mL LB broth at 37°C, 200 rpm orbital shaking. For the main culture five times 100 mL LB broth in glass 1L Erlenmeyer flasks were inoculated with 1 mL starter culture each. Incubation was performed as per usual. OD_{600} was measured hourly up to an OD_{600} of 0.2. Afterwards the culture density was determined every 20 min. At an OD_{600} of 0.35 to 0.4 the cultures were put on ice immediately and cooled for at least 20 min. Cells were afterwards spun down at 4°C and 3000 ×g for 15 min in 50 mL centrifugation tubes. After discarding the supernatant, cell pellets of each tube were washed with 20 mL sterile, ice cold 100 mM MgCl2 solution and combined into eight tubes afterwards (50 mL each). A

spin at 4°C and 2000 ×g for 15 min was carried out to pellet the cells again, the supernatant was discarded and the cells in each tube resuspended in 50 mL sterile, ice cold 100 mM CaCl2 solution. Another spin at 4°C and 2000 ×g for 15 min was carried out and cells resuspended in a total of 50 mL of sterile, ice cold 85 mM CaCl2 solution with 15 % glycerol. The suspension was split into two 50 mL centrifuge tubes and spun at 4°C and 1000 ×g for 15 min. The resulting supernatant was decanted, and the cell pellet resuspended in 2 mL of the previous solution. 50 µL aliquots of the cell suspension were transferred into 1.5 mL reaction tubes and immediately frozen in a small vessel with liquid nitrogen. Competent cells were stored at -80°C.

5.1.3 Sterilisation of Bacterial Media and Laboratory Equipment

Media and solutions that needed to be aseptic were sterilised in an autoclave for 30 min at 121 °C and a pressure 1.05 bar. Pipette tips, plate lids and other laboratory equipment that required sterility were treated the same way for 20 min. The autoclave was run by a technician, who also placed any goods that were required to be dry, e.g. pipette tips, in a drying cabinet overnight.

5.1.4 Isolation of Plasmid DNA

Plasmid DNA from E. coli was isolated with Qiagen Plasmid Plus Midi or Maxi kits for medium scale preparations and Qiagen's QIAprep Spin Miniprep kits for small scale preparations. The procedures were carried out according to the manufacture's protocols. For molecular cloning purposes Miniprep plasmid isolations were used, which were eluted in 30 µL of the kits EB buffer. Plasmid preparations for transfections of mammalian cells were always performed with Plasmid Plus Midi or Maxi kits. For the elution step 100 µL and 400 µL of the kits EB buffer were used. The incubation step of the elution buffer was prolonged in all cases to 5 min for higher yields. Dilutions of the plasmid solutions for transfections were prepared with TE buffer (pH 7.4).

5.1.5 Agarose-Gel Electrophoresis

DNA fragments were separated via agarose-gel electrophoresis. Gels were prepared with 1 to 2.5 % w/v of agarose in 70 mL TAE buffer (8.3 pH). Analytical gels were loaded with 20 µL of a mixture of DNA containing solution and 6x loading dye. Preparative gels with 60 µL of this mixture. The agarose TAE solution was heated in a microwave until boiling. For DNA staining 5 µL of ethidium bromide were added to the gel after the solution had slightly cooled down. The gel was poured into a plastic tray with detachable rubber ends tray for agarose gel casting. To create larger and smaller well sizes, different combs were used during the gel casting. Polymerisation proceeded for a minimum of 30 min at room temperature.

Afterwards, the gel was placed in the running chamber filled with TAE buffer and the samples as well as 4 µL of the DNA size marker (1kb or 100bp HyperLadder, Bioline Reagents) were loaded into the wells. The separation proceeded for at least 40 min at a constant voltage of 120 V. Ethidium bromide-stained DNA bands were visualised on a blue light table DNA fragment excision or in a UV transilluminator (ImageQuant-RT ECL Imager, GE Healthcare) for image capturing.

5.1.6 DNA Extraction from Agarose Gels

DNA fragments were extracted from gels with the QIAquick Gel extraction kit from Qiagen. Therefore, the respective bands were cut out with a scalpel from the gel, the gel piece was then transferred into a 2 mL reaction tube and weighted. All following steps were carried out according to the manufacturers protocol and the DNA eluted in 20 µL of the kits EB buffer.

5.1.7 PCR

For DNA amplifications by polymerase chain reactions Q5 high-fidelity polymerase (NEB) was used. Table 1 lists the reaction mixture concentrations and volumes.

Table 1: Ingredients for 50 µL Q5 HF polymerase chain reaction.

Component	Concentration	Volume / µL
Q5 2x MM	2×	25
Forward primer	10 µM	2.5
Reverse primer	10 µM	2.5
DMSO	100 %	1-2.5
Template DNA	Variable (0.002-0.2 ng/µL)	Variable
HPW	-	Add to 50 µL total volume

DMSO (final concentration up to 5%) was only used if previous optimisation PCR attempts were not useful. Table 2 shows the utilised thermal cycler programs.

Table 2: Temperature profile of Q5 PCR programs of regular and touchdown (TD) gradient variants. $T_m$anneal = annealing temperature of primer w/o overhang.

Step		Temperature / °C		Duration / s	
		regular	TD gradient	regular	TD gradient
Lid heating		115	115	Permanent	Permanent
Initial denaturation		98	98	60	60
Denaturation	1) TD: 10 × -1°C per cycle 2) TD: 25-30 × at Tm	98	98	6	6
1) Annealing 2) Annealing		$T_m$anneal	1) 10°C >Tm 2) $T_m$anneal	20	20
Elongation		72	72	20 per kb	20 per kb
Final extension		72	72	120	420
Storage		4	4	∞	∞

5.1.8 PCR Cleanup

To remove buffer, di-nucleotides, primers, and the polymerase enzyme from the PCR reaction solution Qiagen's QIAquick PCR Purification Kit was used. The manufacturer's instructs were followed and the elution performed with 20 µL of the kit's EB buffer.

5.1.9 Determination of DNA Concentrations

DNA purities and concentrations are determined by the UV-Vis-spectrometer NanoDrop One$_c$ with 1 µL of sample. DNA concentrations were measured at 260 nm, where DNA shows an absorbance peak due to its base pairs. In contrast, proteins show an absorbance maximum at 280 nm due to their aromatic amino acids. Thus, the ratio of the absorbance at 260 and 280 nm can be used to verify the purity of a DNA sample.

5.1.10 DNA Restriction Digest

Restriction enzymes to cut DNA plasmids and fragments were purchased from NEB. Reactions were carried out with up to 1 µg or up to 4 µg DNA in volumes of 20 µL for analytical digests or 50 µL for preparative digests, respectively. The reaction mix contained corresponding volumes of the for the specific enzymes required buffer (normally rCutSmart Buffer), ddH2O and no more than 10 % of enzyme solution. Reaction times were chosen according to the enzymes either 20 min or 60 min. Digests were performed at 37 °C if not stated otherwise by the enzymes manufacturer.

5.1.11 5'-Dephosphorylation of Vector DNA

To reduce self-ligation of vector backbones, especially of blunt end cut constructs, some DNA fragments were dephosphorylated prior to ligation. This 5'-dephosphorylation was performed by adding buffer and enzyme of NEB's Antarctic Phosphatase to the restriction digest mix. After an incubation of 30 min to 60 min at 37 °C the reaction was stopped with a heat-inactivation at 80 °C for 2 min.

5.1.12 DNA Ligation

For ligation of two purified DNA fragments NEB's T4 ligation kit was used. In a total reaction volume of 20 µL, 1 µL of enzyme and 2 µL of the T4 DNA ligase reaction buffer (10×) were utilised. The reaction was performed with a total of 100 ng DNA in a ratio 3:1 insert to backbone. Necessary DNA amounts are calculated with the NEBioCalculator online tool. Incubation was carried out at room temperature for 10 min for sticky end reactions or for blunt end and difficult constructs at 16 °C overnight. Preparation of the reaction mix was always prepared in 0.2 mL or 0.5 mL reaction tubes on ice.

5.1.13 Chemical Transformation of DNA into *E. coli*

To transform plasmid DNA in *E. coli* cells, 2 µL of the ligation mix or in case of retransformation 100 pg – 100 ng (usually 10 ng) of purified plasmid DNA were added to 50 µL chemically competent cells in a 1.5 mL tube. After gentle mixing cells were incubated for 20 min on ice. Subsequently, the cells were heat shocked at 42 °C for 45 seconds. Directly afterwards 450 µL of SOC or NEBstable outgrowth medium were added, and the mixture was incubated for 1 h at 37 °C and 850 rpm in a thermo mixing block. Then 100 µL of the culture were plated out on an agar plate containing the required antibiotic for overnight incubation. For difficult transformations the transformation culture was spun down at 5000 ×g for 5 min and resuspended in 100 µL to use all cells for plating.

5.1.14 DNA Sequencing

Plasmid sequences were confirmed by DNA sequencing. The service was either carried out by Eurofins or Genewiz (Azenta). The received sequencing data was compared with the digitally planned and created plasmid map in the software program SnapGene.

5.1.15 In Silico Creation of DNA Sequences

Plasmids, primers, and synthetic DNA fragments were designed in the software SnapGene. For the creation of primer sequences the NEB Tm Calculator tool was used, as well as Thermo Fisher Scientific's Multiple Primer Analyzer to minimise the presence of primer dimers, self-binding, and secondary structures. Primers used for

SDM reactions were created with the tool NEBaseChanger. Primers for Gibson Assembly like reactions (NEB HiFi Assembly) were planned with the tool NEBuilder. Primer sequences were subsequently listed and saved in an Excel sheet, as well as entered in the according plasmid maps in the SnapGene software.

5.1.16 Molecular Cloning with NEBuilder HiFi DNA Assembly and Site-Directed Mutagenesis

The NEBuilder HiFi DNA assembly kit is used for Gibson Assembly cloning approaches. NEB's Q5 Site-Directed Mutagenesis kit was used for altering DNA with an SDM cloning approach. The procedures were carried out according to the manufacturer's protocols. DNA amounts for the HiFi DNA assemblies and ratios of the fragments varied. The required DNA amounts were calculated with the online tool NEBiocalculator.

5.2 Mammalian Cell Culture and Analytical Methods

For the production of rAAVs and its optimisation transiently transfected HEK293 cells were used. Promoters were tested in HEK as well as in CHO cultures, which were also used to produce SARS-COV2 spike protein.

All mammalian cell cultivations were carried out at 37 °C with 5 % CO2. Transfected cultures in 24-well plates or in 50 mL TubeSpin bioreactor tubes were cultivated with humidification to 80 % relative humidity and shaken at 230 rpm (25 mm throw). The shaking speed for larger cultures in Erlenmeyer flasks was 200 rpm with a 25 mm throw or 140 rpm with a 50 mm throw.

5.2.1 Cultivation of HEK Cells

Two different types of HEK293 suspension cells were provided by REGENXBIO. A polyclonal cell line (HEK SKMB), and a monoclonal cell line (HEK 4B9-11A4). Both REGENXBIO HEK cell lines were cultivated in Dynamis medium (Gibco) supplemented with 6 mM Glutamine (Gibco). HEK 293-F cells (Thermo Fisher) were cultivated in Freestyle medium. Expi293F cells (Thermo Fisher) were cultivated in Expi293 expression medium (Gibco). CHO K1 cells (Lonza) were cultivated in CD-CHO medium with 8 mM Glutamine. All cell lines were cultivated as described in 5.2 and split on a three-day basis. For the routine passaging cells were diluted to 0.2 or 0.3×10^6 vc/mL. The regular culture and vessel volumes are listed in Table 3.

Table 3: Tissue culture vessels working volumes and growth areas.

Cultivation vessel / -	Volume medium / mL	Cultivation vessel / -	Volume medium / mL
24-well plate	0.77	E125 Flask	30
50 mL TubeSpin	5	E250 Flask	60
50 mL TubeSpin	10	E500 Flask	120

5.2.2 Creation and Thawing of Cryo-Stocks

For long term storage of mammalian cells, the viable cell density of cultures in exponential growth-phase (three days after passaging) was measured and the cultures then spun down at 200 ×g for 5 min. The supernatant was discarded, and the cell pellet resuspended in 4 °C cold culture medium supplemented with 10 % v/v DMSO. The volume of the of the DMSO containing medium was chosen to adjust the viable cell density to 1×10^6 cell/mL. Aliquots of 1 mL of the resulting suspension were prepared in cryovials, which were placed into an isopropanol filled Mr. Frosty Freezing container that guarantees a continues temperature decrease of -1 °C/min. The filled freezing container was then placed in a -80 °C freezer for 24 h.

Afterwards, the cryovials were transferred into a liquid nitrogen tank for long term storage.

To thaw cells taken from the cryostat container, the vial was placed into a 37 °C warm water bath until the cell suspension was fully thawed. Further, the cryopreserved culture was transferred quickly into a 125 mL Erlenmeyer flask containing 30 mL pre-heated culture medium.

5.2.3 Determination of Cell Density and Viability

Ordinary measurements of cell density and viability of mammalian cultures were carried out with a ViCell automated cell counter (Beckmann). The device detects dead cells by trypan blue staining and distinguishes cells from debris and other impurities according to parameters set by the manufacturer. If necessary, samples were diluted with medium or PBS up to a factor of 1:4. HEK cells that needed to be dissociated for cell counting due to heavy clumping were sampled into a 1.5 mL reaction tube and mixed with an equal volume of Accumax (Thermo). The mixture was incubated for 20 min at 37 °C and 140 rpm before the cell count. The final volume of diluted and undiluted samples loaded into the cell counter was 550 µL.

5.2.4 Transfection Efficiency Measurement

To get an idea of how effective the uptake of transfected plasmids was, GFP expression of cells was determined either by flow cytometry or with a Countess 3 FL automated cell counter (Thermo Fisher). Flow cytometry was carried out by the university's flow cytometer core with an Attune acoustic focusing cytometer (Applied Biosystems) until its abolition. For GFP measurements with the Countess cell counter, 10 µL of sample, if necessary diluted with PBS up to 8-fold, were injected into the device's single use cartridges. The fluorescence gain was adjusted either by the device automatically or manually with the reference being a sample of non-transfected cells.

5.3 rAAV Production in HEK Suspension Cells

Recombinant production of AAVs in HEK cells was performed in different culture volumes by PEIpro mediated transient transfection.

5.3.1 Transient Transfection

Cultivated HEK cells were seeded in a new vessel one day before transfection at 2.2×10^6 vc/mL (viable cells per millilitre). On the day of transfection, cells grown to $4.5\text{-}5\times10^6$ vc/mL were transfected with a total of 0.63 µg DNA per 10^6 vc if not mentioned otherwise. The regular plasmid ratio was 2:1:0.1 pHelper : Rep/Cap : ITR/GOI. However, plasmid ratios were changed during the course of this work. Changes are mentioned in the result chapters accordingly and a

final ratio of 2:1:0.5 was used for the final sets of experiments. The DNA to PEI ratio was 1:1.75 in all cases. PEIpro (Polyplus) and the DNA mix were mixed with Dynamis medium (Gibco) in separate tubes, before combining DNA premix and PEI premix. The combined transfection mix had a volume of 10 % of the cultivation volume. The PEI premix was always added to the DNA premix. Afterwards the solution was mixed by vortexing for 3-5 seconds, followed by an incubation of 10 to 15 min at RT. The incubated mix was then added to the cells, which were cultivated for three days.

Transfection conditions for GFP measurements to determine promoter strengths deviated. Here the premix of DNA and the premix of PEI were created with 150 mM NaCl (Polyplus). The transfection mix was only incubated for 4 min, 0.8 µg DNA per 10^6 vc and a DNA to PEI ratio of 1:3 was used, in accordance to Johari et al., (2022).

5.3.2 Small Molecule Addition

The addition of small molecules to transfected cells was performed as stated in (Scarrott et al., 2023). If not otherwise stated, supplementation was carried out 4 h post transfection. In DMSO dissolved chemicals M344 (50 mM) and Nocodazole (10 mM) were diluted 1:100 to 500 µM and 1 µM in Dynamis medium prior to adding them to the cultures for final concentrations of 2.5 µM and 4 µM, respectively.

5.3.3 Cell Harvest and Lysis

For rAAV production, cultivation was terminated three days after transfection. To lyse the cells and keep the viruses released to the medium, samples of 450 µL were taken from the cell suspension and transferred into a 1.5 mL tube containing 50 µL of 10× lysis buffer. The lysis was carried out at 37 °C and 230 rpm for 1 h with the tubes horizontally placed in a tube rack in the incubator (25 mm throw). Tubes were vortex heavily before and after incubation. Afterwards, the cell debris was separated by centrifugation at 4 °C, 13000 ×g for 10 min.

5.3.4 DNase I Digest of rAAV Samples

DNA not protected by AAV capsids was digested with DNase I prior to ddPCR measurement of genomic rAAV titres. Therefore, 5 µL of the lysed samples separated from cell debris were mixed into 45 µL of DNaseI reaction master mix. The master mix contained 2 µL of DNaseI (Roche), 5 µL 10× DNaseI reaction buffer, 5 µL of 1 % Pluronic F-68 (Gibco) and 33 µL of nuclease-free water per sample. The reaction mix was incubated for 1 h at 37 °C in a PCR machine. Overnight storage was performed at 4 °C if necessary, otherwise samples were stored long term at -20 °C.

5.3.5 ddPCR for Determination of Viral Genomic Copy Numbers

Genomic rAAV titres were determined by droplet digital PCR (ddPCR) with BioRad's Supermix for probes. At first, lysed and consequently DNase I digested samples were diluted 1:4000 with a PCR buffer mix containing one part 10× PCR buffer (Applied Biosystems), one part 1 of 1 % Pluronic F-68 (Gibco), and eight parts nuclease-free water, as well as 0.2 µg per mL single-stranded salmon sperm DNA. 2 µL of diluted sample were then pipetted into 18 µL of ddPCR master mix. The master mix contained 10 µL ddPCR mix for probes (BioRad), 5 µL nuclease-free water and 5 µL of a primer probe mix. This primer probe mix constituted of 40 µL forward and 40 µL revers primers (100 µM), as well as 10 µL of the probe solution (100 µM) and 910 µL nuclease free water per mL. Primers and probes amplify or bind to the polyA signal sequence of the rAAV genome, respectively. Droplets were generated in the QX200 droplet generator (BioRad) using the whole 20 µL of the sample in the middle well and 70 µL of droplet generation oil for probes (BioRad). After the liquids were sucked through the microfluidics of the cartridge, 40 µL of the oil solution containing the aqueous micro droplets, were transferred into a ddPCR 96-well plate (BioRad). After all samples were loaded into the plate the plate was heat sealed and placed into a C1000 Touch Thermal Cycler PCR machine (BioRad). The PCR program was run as described in Table 4. Next, the plate was put into the droplet reader CX200 (BioRad) and the samples were processed by it. For analysis, a manual threshold was set at the fluorescence value 850 for all samples, all droplet events below were regarded as negative.

Table 4: ddPCR thermal cycling conditions. for quantification of genomic copy numbers with the BioRad supermix for probes.

Step	Cycles	Temperature / °C	Duration / min
Enzyme activation	1	95	10
Denaturation	40	95	0:30
Annealing/extension		57	1
Enzyme deactivation	1	98	10
Hold cooling	1	12	>30

All titre measurements were performed with crude lysates to avoid sample to sample variations due to purification yield differences.

5.3.6 AAV8 Capsid ELISA

Measurement of capsid titres was performed with the AAV8 Titration ELISA from Progen. Samples were diluted 1:5000 to adjust their virus content to be in the linear range of the test's standard curve. The manufactures protocol was strictly followed and the horseradish peroxidase catalysed colour reaction of TMB, halted by sulfuric

acid, was then measured in a SpectraMax iD5-3024 plate reader (Molecular Devices) at 450 nm, as well as at 650 nm. For correction of the TMB's peak absorbance at 450 nm, the absorbance value at 650 nm was subtracted, as advised by the manufacturers protocol.

5.3.7 Determination of Promoter Strength through GFP Expression

The method was conducted as described in Johari et al. (2022). To summarise, cell densities were adjusted to 1×10^6 vc/mL the day before transfection in 10 mL of fresh medium. The transfection was then carried out with 8 µg plasmid DNA and 24 µL PEImax. 48 h post-transfection GFP expression was measured with an iD5 microplate reader as described.

For the HEK promoter constructs instead of 10 mL, 5 mL culture were transfected with PEIpro instead of PEImax, DNA and PEI concentrations and ratios were unchanged.

5.3.8 Measurement of mRNA Transcript Levels

mRNA transcripts of the adenoviral *E2A* and *E4* (orf6+orf6/7) genes were measured by ddPCR. First, cell pellets of 650 µL culture 72 h post-transfection were created through a 5 min spin at 300 ×g. The pellets were then resuspended in 150 µL RNAlater (Thermo Fisher Scientific), stored overnight in at 4 °C and then transferred for long term storage into a -80 °C freezer. The stored cells thawed on ice, diluted with 250 µL PBS and then spun down for 5 min at 600 ×g. The resulting pellet was resuspended with RLT buffer of Qiagen's RNAeasy plus kit supplemented with 2-mercaptoethanol as recommended by the manufacturer. Cell lysis was performed with a QIAshredder. Therefore, the resuspended cells were transferred to the QIAshredder column and spun at 21000 ×g for 3 min. Thereupon, the protocol of the Qiagen RNAeasy plus RNA extraction kit was followed strictly. For the elution step 50 µL RNase-free water were used. The elution was then further purified with a NEB Monarch RNA Cleanup Kit (50 µg). The manufactures protocol was followed with the addition of a second spin to dry the column. The eluted RNA was diluted to 100 ng/µL and then aliquoted and stored at -80°C.

For measurement of the respective genes' mRNA copy numbers, the to 100 ng/µL total RNA adjusted samples were thawed and diluted 1:1000 with RNAse-free water. 2 µL of the dilutions were then used with BioRad's One-Step RT-ddPCR kit. Reverse-transcribed cDNA copy numbers of E2A were measured with a FAM-labelled probe, *E4* (orf6+6/7) copy numbers with a HEX-labelled probe. Thermocycler conditions were used as shown in Table 5.

Table 5: One-Step RT-ddPCR thermal cycling conditions.

Step	Cycles	Temperature / °C	Duration / min
Reverse transcription	1	50	60
Enzyme activation	1	95	10
Denaturation	40	95	0:30
Annealing/extension		55	1
Enzyme deactivation	1	98	10
Cooling	1	12	>30

5.4 Statistical Analyses

Statistical Analyses were performed with the software Graph Pad Prism. Comparison of means of two populations were performed with unpaired, two-tailed t tests. Multiple means were compared by ordinary one-way ANOVA with Tukey post-hoc analysis for multiple comparison of means. Significances in figures are indicated as asterisks with P < 0.05 = *; 0. 005 = **; 0.0005 = ***; 0.0001 = ****.

5.5 Materials

Lists of the used materials and equipment can be found in 8.3.

6 Results and Discussion

6.1 Examination and Modification of the Helper Plasmid and its Components for Improved rAAV Production

A primary objective of this work was the creation of an innovative Helper plasmid with enhanced characteristics for rAAV production. Through genetic engineering, the aim was to analyse the current Adenovirus (AdV) Helper plasmid, modernise, simplify, and modularise it, in pursuit of improved rAAV production qualities, with a specific focus on greater recombinant virus genome titres. The starting point was the widely used pADΔF6, which has remained unchanged for over two decades. Despite being 3 kb smaller than the other highly popular Helper plasmid, pXX6, pADΔF6's large size of 15.8 kb poses challenges for plasmid uptake in transient transfection (Hildinger et al., 2001; Kreiss, 1999; Xiao et al., 1998). Consequently, the intention was to streamline the plasmid's structure, reducing it to its essential components for high-titre rAAV production, while also enhancing its functionality.

6.1.1 Assembly of Essential Helper Sequences to Create an Advanced and More Versatile Basis Plasmid

To maximize virus production and optimise the Helper plasmid, essential genes and genetic elements of the pADΔF6 Helper plasmid for rAAV production needed to be determined. Existing literature established *VA RNA*, *E2A*, and *E4* as essential AdV5 Helper genes (Ferrari et al., 1996; M. M. Huang & Hearing, 1989a; Muzyczka, 1992). Despite this knowledge, the commonly used Helper plasmids still contain additional AdV5 regions beyond these essential genes (Colosi, 1996; Grimm et al., 1998; Hildinger et al., 2001; Matsushita et al., 1998; Xiao et al., 1998). These extraneous Adenovirus genes and gene fragments are likely just artifacts from the creation of the first Helper plasmids and result from cloning the AdV genome into a plasmid vector. Some of these genes were excluded, while others remained, possibly due to technical limitations, even though they were known or believed to be irrelevant for rAAV production. Thus, efforts were made to eliminate as much of this non-essential and non-coding DNA as possible. The targeted DNA sequences for elimination from pADΔF6 included: (i) the L5 (fibre) gene, positioned downstream of the *E4* gene, (ii) the UXP protein, with exons located upstream of the E2E promoter, inside *E2A*'s second non-coding exon, and at the start of *E2A*'s CDS, (iii) gene fragments of E2B coded pTP, AdV's protease, a fragment of VAII, and non-coding DNA, flanking *E2A* and VAI, (iv) parts of the plasmid backbone and AdV ITR positioned upstream of *E4*, (v) the *L4* coded 100K, 33K and 22K proteins located in the *E2A*'s 5 UTR, and (vi) potentially unrequired *E4* ORFs (see Figure 16 B).

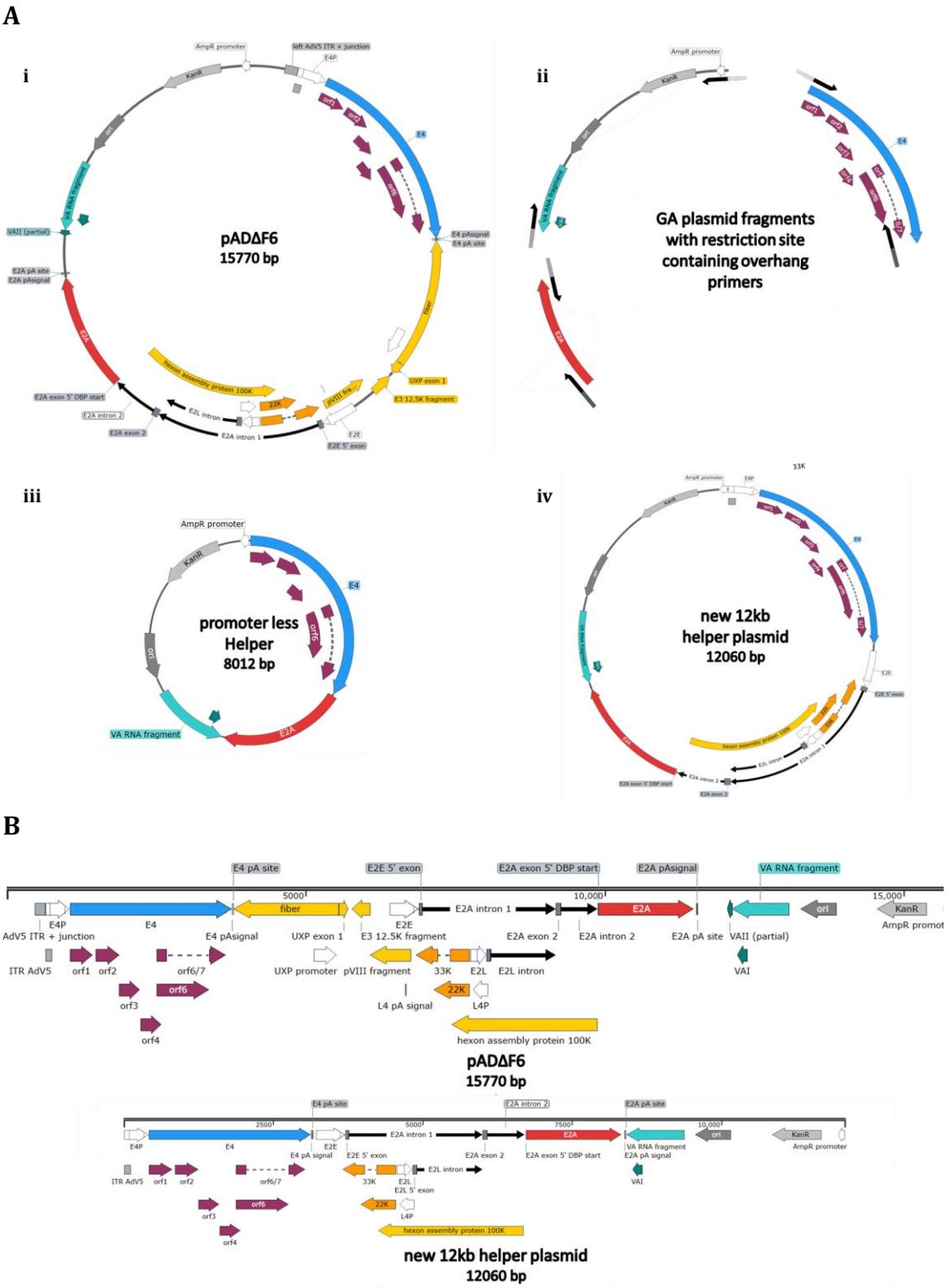


Figure 16: Creation of the new 12 kb Helper plasmid (Helper 1.0). (A) Plasmid map of (i) the original, parental plasmid pADΔF6 (Hildinger et al., 2001). (ii) Plasmid fragments amplified by PCR with overhang primers containing new restriction sites for subsequent introduction and exchange of promoters, poly A sequences and other genetic elements. (iii) Assembled Helper plasmid without promoters and poly A sequences. (iv) New 12 kb Helper plasmid containing VAI, as well as the full *E4* gene, *E2A* and its 5'-UTR with the *L4* products 33K/22K, together with their endogenous promoter and poly A sequences. (B) Linearized plasmid maps of pADΔF6 (Helper 0.1) and the new 12 kb Helper plasmid (Helper 1.0).

The construction strategy of the new Helper plasmid is shown in Figure 16 A. PCR amplification of fragments from pADΔF6 was performed using overhang primers containing unique restriction sites, enabling straightforward exchanges of genetic parts and future plasmid modifications, as well as assembly of the plasmid by Gibson Assembly (Gibson et al., 2009). This cloning strategy made it possible to exclude all unwanted sequences in one cloning step, instead of excision in sequential steps. The introduced restriction sites up- and downstream of the *E4* and *E2A* genes were then used to reinsert the genes' endogenous promoters but can also be used for further plasmid optimisations in the future providing a further novel feature to the new Helper plasmid. After the reintroduction of the endogenous promoters a size reduction of 3.8 kb was achieved. Virus genome titres measured with the new, 12 kb large, Helper plasmid were slightly lower with 0.86-fold of the yield obtained with pAdΔF6 (Figure 17). This decrease was regarded as acceptable and recoverable with further optimisations, considering that no plasmid ratio or other optimisation that were previously performed for the pAdΔF6 plasmid were conducted for the new Helper plasmid.

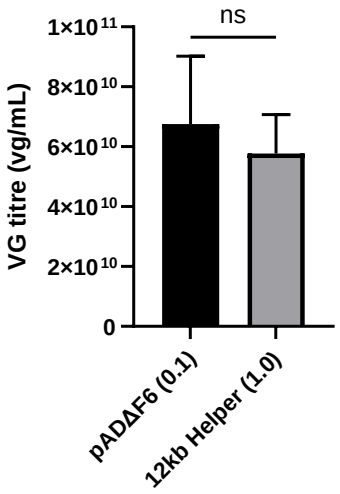


Figure 17: Virus genome titre per mL of the parental plasmid pADΔF6 (plasmid version 0.1) and the newly assembled 12kb Helper plasmid (1.0). Display of mean and standard deviation, statistical analysis was performed as an unpaired t-test of biological replicates (n=19).

6.1.2 Confirmation of *L5 (fiber)* Gene Redundancy

The largest fragment excluded in the new Helper plasmid contained the *L5* (fiber) gene. *L5* was of particular concern, because it is known that the adenoviral fiber protein can present one of the major contaminations for downstream processing. As pAdDeltaF6 and pXX6 lack the Major Late Promoter (MLP) responsible for driving the expression of all late Adenovirus proteins, including *L5*, this should be less of a concern (Farley et al., 2004; Leppard, 2014). Nevertheless, minor fiber protein contaminations were still present in transfections with pXX6, likely due to a cryptic promoter upstream of the gene (Xiao et al., 1998). Consequently, the *L5* gene was excluded from the new plasmid construct, providing a size reduction as well as the potential functional improvement.

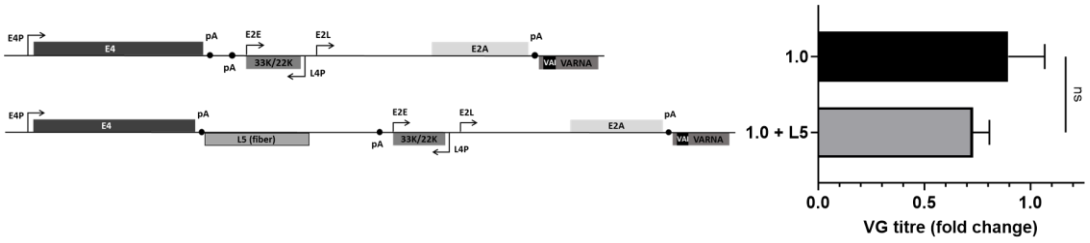


Figure 18: Analysis of the L5 (fiber) gene including 2.6 kb fragment, which was excluded from the 12kb Helper plasmid. Virus genome titre fold change related to the transfection of control Helper plasmid 0.1. Schematic depiction of 1.0 (12kb Helper) and the same plasmid with the reinsertion of the previously eliminated sequence between the *E4* 3' end and *E2A*'s promoter E2E. Arrows denote promoters, black dots poly A sequences. Genes, 5' and 3'-UTRs, but not promoters and polyA signals are drawn to approximate scale. Error bars display standard deviation, an unpaired t-test was performed for statistical analysis of the means of biological replicates of 1.0 (n=19) and 1.0+L5 (n=4).

To investigate whether the reduction in titre was caused by the removal of the fiber gene and whether this gene played a role in improved rAAV production in the plasmid, the *L5* gene was reinserted into the 12 kb Helper plasmid 1.0 (Figure 18). A non-significant decrease in VG titre indicated that the previous decline was not related to the removed *L5* gene. Instead, the slight titre increase of the smaller plasmid without the 2.6 kb fragment containing the *L5* gene might suggest moderately increased transfection efficiency. This marginal increase was further confirmed by the addition of L5 to another plasmid construct (see appendix Figure 48). Besides *L5*, a fragment of *pVIII* and the nearly full sequence for *E3 12.5K* were still present. The results confirmed that these gene sequences, which were already missing their upstream major late promoter, had no positive effect on rAAV production. The reinserted sequence also contained the first exon and start codon of the UXP protein located upstream of *L5*, as well as its promoter located in the *L5* gene. UXP locates with DBP in nucleoli and VLCs, is mostly uncharacterized but likely is associated with AdV DNA replication or RNA transcription (Tollefson et al., 2007; Ying et al., 2010). However, UPX does not seem to play supporting role in rAAV proliferation. Although expression was possible through the reintroduction of its promoter and first coding exon, along with the presence of its other exons, it did not increase rAAV titres. Given that the removal of *L5* eliminated contamination concerns and did not cause the decrease in virus genome titre, the 2.6 kb fragment including *L5* was excluded from all subsequent plasmids. The newly created 12 kb plasmid also lacked a ~0.6 kb sequence upstream of *E4* compared to pAdΔF6. As this sequence originated from the left ITR and adjacent bases of the AdV5 genome and was not likely essential for *E4* transcription, its exclusion was not considered detrimental. On the contrary, the right ITR was included as part of *E4*'s promoter sequence.

6.1.3 Assessment and Shortening of the *VA RNA* Fragment

The final part omitted in the 3.8 kb size reduction was a 0.5 kb fragment located between the *E2A* gene and the *VAI* sequence. This likely non-coding DNA originated from a fragment of the *L3* coded protease of AdV5. Additionally, pAdΔF6 only contained less than half of the VAII sequence, which was fully eliminated during the assembly of the new Helper plasmid. Since VAI is the main VA RNA species, executing all essential functions, and only a fraction of VAII was present in the parental plasmid, its full elimination was not considered a restrictive change for rAAV titres. Although the low abundance VAII is preferred by Dicer and RISC incorporation, it was shown to be redundant in AdV proliferation experiments and is likely a backup for the essential VAI (Vachon & Conn, 2016). Moreover, in experiments for the promotion of AAV related parvovirus B19 proliferation with AdV genes of Winter et al. (2012) demonstrated that the presence of VAI, but not VAII, leads to increased VP expression. VAII, on its own, even reduces capsid protein amounts. Therefore, the elimination of the partial *VAII* sequence was not deemed substantial and subsequent plasmids were created only with the *VAI* sequence. However, a reinstallation of the full *VAII*, similar to the experiment performed with the promoter less fiber gene, could confirm the lack of benefits or even drawbacks on rAAV production resulting from *VAII*.

Subsequently, an attempt was made to further minimize the *VA RNA* fragment utilized in the 12 kb plasmid. The fragment was reduced from its original size of 953 bp (full *VARNA*) to either 572 bp (short *VARNA*), excluding most of pAdΔF6 short *E2B* fragment, or to only *VAI* (160 bp). The 381 bp elimination was chosen due to the simplicity of a restriction digest approach using compatible restriction sites SalI and XhoI. This design resembled the approach used by Matsushita et al. (1998) and was consequently expected to be functional. However, the use of these restriction sites was only possible when the gene was cloned into a separate plasmid. Therefore, a 5-plasmid transfection system was employed, segregating all three AdV Helper genes onto different plasmids. Cloning of short *VA RNA* fragment into a regular Helper plasmid was performed subsequently, but not for this experiment. The *VAI*-only construction involved overhang-primer PCR amplification of *VAI* followed by restriction and ligation into the new restriction sites of the assembled plasmid (PspOMI+SalI). Functionality of the truncated *VA RNA* fragment was tested by transfecting HEK293 cells with the 5-plasmid approach (separating all AdV Helper genes on different plasmids) for the short *VA RNA* fragment and usual 3-plasmid transfection (Rep/Cap + Helper + ITR/GOI) for only *VAI*.

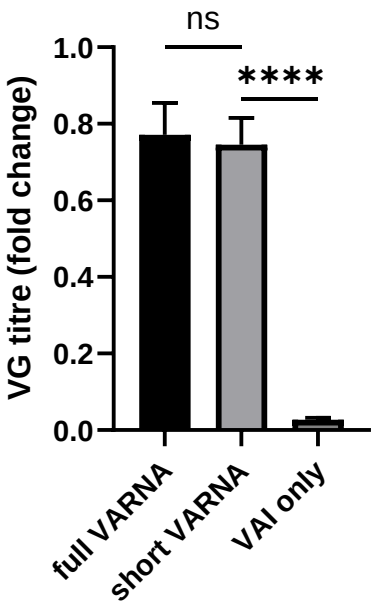


Figure 19: Evaluation of shortening the *VA RNA* fragment. VG fold change related to the control Helper plasmid 0.1. Transfections of 5 plasmids with a all three Helper genes split to different plasmids for the full *VA RNA* assembly fragment (953 bp, n=2) and the shorter version (572 bp, n=2). Regular 3 plasmid transfections (Rep/Cap + Helper + ITR/GOI) for the control and the short version of VAI (163 bp, n=4). Error bars show standard deviation, ordinary one-way ANOVA was performed for comparison of means.

Unfortunately, the exchange to only the VAI only caused an extreme drop in VG titre. Presumably, the inserted sequence did not promote sufficient VAI expression or even none at all. While RNA transcription should have been possible due to VA RNA pol III promoters within the transcribed sequence, the results demonstrated context specificity of the surrounding sequence for VAI transcription. The close proximity to the adjacent 3' end of *E2A*, halving the distance to 88 bp, might cause steric hindrance of the individual transcription machineries of the gene. Additionally, two bases of VAI were not annotated in the used gene annotation (GenBank: KX868466.2), resulting in these bases being absent in the VAI construct. Although general transcription from the conserved A and B Box promoter sequences should still have been possible, termination might have been an issue. Even though the primary terminator sequences T1A and T1B were present in the construct, the backup terminator T2 was not. The sharp decrease in rAAV titre could have been caused by the insufficient VAI amount or reduced DBP amounts. Both could have been caused by the loss of the VA RNA backup terminator T2, leading to increased known RNA pol III read-throughs and subsequently decreasing more frequent *VA RNA* reads and *E2A* transcription. As expected, the construct using the 572 bp *VA RNA* sequence performed as well as the larger fragment. The deletion of nearly 300 bp upstream of the first *VAI* bases was not expected to have any negative effects, as these bases were also absent in the plasmids that are based on the study of Matsushita et al. (1998). Unfortunately, the lack of the second terminator, potentially the reason for the titre decrease in the smaller version, was not detected at the time. Thus, it is possible that future Helper plasmids' sizes could be reduced by another 250 to 300 bp upstream of the VAI sequence.

6.1.4 Discovery of L4-33K/22K as a New AdV5 Helper Function for High Titre rAAV Production through Analysis of *E2A*'s 5'UTR

The largest remaining sequence without defined rAAV Helper function is located between the *E2A* gene and its 3 kb distant promoter. To investigate whether this 5'UTR of the *E2A* gene is required for high titre rAAV production or if it could be omitted to further reduce size, a plasmid was created without it. Instead, the E2E promoter was placed directly upstream of the *E2A* gene. This was done by inserting the promoter with and without the additional 3 kb 5'-UTR upstream of *E2A* via the previously established restriction sites NdeI and XbaI. The promoter variants were amplified from the original plasmid via PCR with overhang primers. The resulting plasmids are the previously discussed 12 kb plasmid (Helper 1.0, E2A) and a 9 kb large Helper plasmid (Helper 1.1, $E2A_{min2}$) with the standardly used, extended E2E promoter (486 bp) right in front of *E2A*'s CDS (Figure 20).

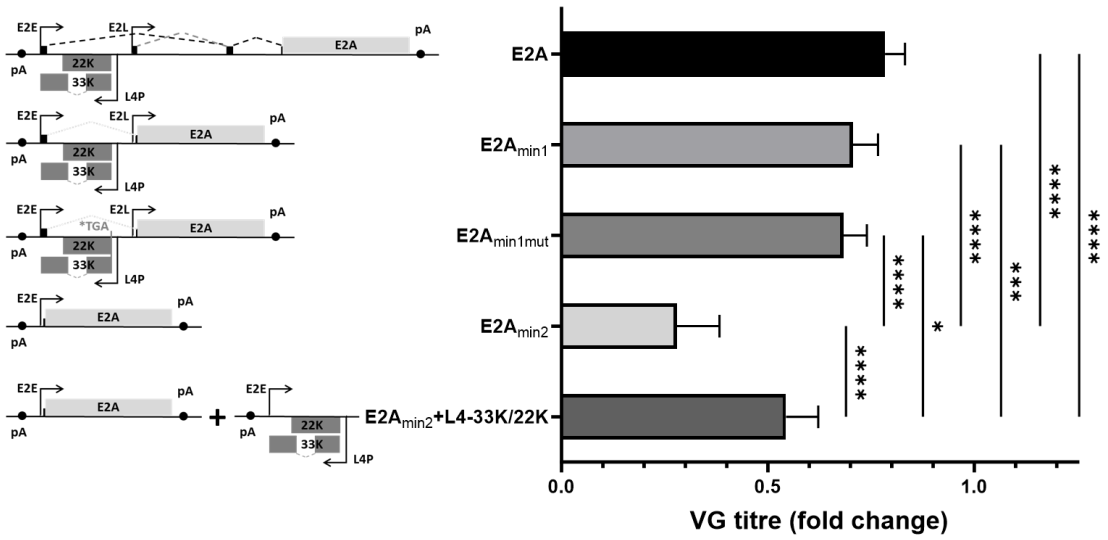


Figure 20: Functional evaluation of the *E2A* and *L4-33K/22K* Helper components for rAAV8 production. Schematic depiction of the *E2A* and *L4-33K/22K* open reading frames within Helper plasmid (*E4* and *VA RNA* not indicated). A L4 promoter-driven plasmid expressing only L4-33K/22K was constructed and co-transfected (n=9) in an equimolar ratio with the E2min2 version (9kb, Helper 1.1, n=11) of Helper plasmid 1.0 (E2A, n=11). The shortened *E2A* 5'-UTR version $E2A_{min1}$ including *L4-33K/22K* (n=11) was also created with mutated ATG→TGA start codon of *L4-33K/22K* ($E2A_{min1mut}$, n=5). Genes, 5' and 3'-UTRs, but not promoters (arrows) and polyA signals (dots) are drawn to approximate scale. HEK293 cells were triple transfected with each Helper plasmid, Rep/Cap 1.2 plasmid and transgene plasmid at 2:1:0.1 weight ratio. rAAV8 crude VG titres analysed 72 h post-transfection are displayed as fold changes to the control Helper plasmid 0.1. Data shown as means, error bars show standard deviation, statistical analysis was performed as an ordinary one-way ANOVA.

Transient transfection results showed dramatically reduced rAAV titres for the smaller (9 kb) plasmid without the *E2A*-5'-UTR ($E2A_{min2}$). VG titres of $E2A_{min2}$ were about 36% of the titres of the 12 kb large Helper plasmid or, in other words, the

regular *E2A* gene with its 5'-UTR (denoted E2A in Figure 20). The sharp decrease in VG titre demanded a deeper look into the excluded sequence and what it is encoding.

Firstly, *E2A* is transcribed by two promoters: the E2early promoter (E2E) located upstream (3 kb distance to start codon) and the E2late promoter (E2L) situated 1.8 kb away from start codon. As the names suggest, the early promoter is activated in the early phase of AdV infections, activated by E1A, whereas E2L is inhibited by E1A and activated in the intermediate and late phase (Hemström et al., 1991; Imperiale & Nevins, 1984; Takako et al., 1988). Due to the omission of the 3 kb space between E2E and the gene, E2L is absent in the 9 kb plasmid. However, several studies demonstrated that E2E is the much stronger promoter (Donovan-Banfield et al., 2020; Westergren Jakobsson et al., 2021; Zhao et al., 2014). Even though E2L transcription strength increases in the intermediate and late phases and E2E's usage frequency drops, E2A mRNAs mainly origin from E2E in all phases. Particularly in the present HEK293 based rAAV production system, it is likely that E2E exhibits far superior transcriptional activity throughout production compared to E2L. This assumption is based on the cells' stable E1A expression and the fact that only AdV early genes are present. Consequently, it is assumed that cellular conditions always resemble the early, but never intermediate or MLP gene product driven late phase of the AdV life cycle (Fessler & Young, 1998). Therefore, the drop in rAAV titre was likely not associated with reduced DBP amounts due to missing transcription of the late E2 promoter. Nevertheless, E2A mRNA amounts could have been affected for different reasons.

One potential reason is that E2A pre-mRNAs are known to contain introns and exons in the 5'-UTR. The 9 kb Helper plasmid could therefore exhibit altered mRNA amounts and reduced rAAV titres caused by the missing of the 5'-UTR. E2A mRNAs mainly undergo two splice events if originating from early or late promoter, although recent studies showed that there is a multitude of differently spliced mRNA variants, exhibiting more than or less than the usual two untranslated 5'-exons (Westergren Jakobsson et al., 2021). Generally, 5'-UTRs, in particular introns and their splicing, are known to affect protein expression to various extents (S.-Y. Kim et al., 2002; Petitclerc et al., 1995). Accordingly, it is possible that the exclusion of these introns and untranslated exons had an influence on mRNA expression and mRNA stability. The loss of these splice events and particularly the mRNA's untranslated 5'-UTR could have consequently result in reduced or increased E2A mRNA amounts. The lack of a Helper component, as well as a stoichiometric imbalance, can lead to drastic losses in rAAV titre due to an unideal production environment.

Moreover, E2E and its consecutive first exon contain a superimposed RNA pol III promoter sequence. Although potential functions of the very low abundance E2E

pol III RNAs are unknown, the opposing activity regulation through competition of the two different polymerases likely influences E2E's transcription strength for E2A mRNAs (Ellsworth et al., 2001). The 9 kb plasmid lacks E2E's original downstream sequence from +13 bp onwards. Since A Box and B Box of its RNA pol III promoter are located in +14 to +25 and +39 to +49 / +51 to +61, respectively, Helper 1.1 cannot express these non-coding RNAs, resulting in no transcription reduction due to polymerase competition. Notably, the study of Ellsworth et al. (2001) showed that downstream sequences are also of importance for E2E's pol II activity. Specifically, the regions 2 to +11 (present in all plasmids) and +14 to +21 (partly missing in $E2A_{min2}$) were crucial in for E2A transcripts in their experiments. Therefore, *E2A* transcription could have been impaired with this plasmid due to the absence of these sequences, as well as the splicing of E2A pre-mRNA introns and exons.

Lastly, there are several adenoviral coding sequences of the *L4* gene embedded in the opposite strand of the *E2A* 5'-UTR. These are the hexon assembly protein L4-100K, the L4-22K protein, and its splice variant, the L4-33K protein. All three are mainly transcribed from the MLP during the late phase of AdV infection (Donovan-Banfield et al., 2020; Shaw & Ziff, 1980). As the MLP is not present in the Helper plasmids, there is no transcription of the L4-100K protein. However, the L4-33K and 22K (33K/22K) mRNAs can also be transcribed through their own promoter, the L4 promoter (L4P). This for a longtime overseen promoter is required to express 33K/22K already in the early to intermediate phase since these two proteins, along with E4orf4, pIX and Iva2, are part of the AdV's mechanism to shift its expression patterns from early genes to late genes (Backström et al., 2010; Biasiotto & Akusjärvi, 2015; Farley et al., 2004; Lutz et al., 1997; Lutz & Kedinger, 1996; Somberg et al., 2009). In the shift from AdV early to late phase, L4-22K suppresses early gene expression and stimulate the MLP activity, while L4-33K functions as an alternative splicing factor for weak 3' splice sites and regulates the accumulation of AdV late gene mRNAs (Törmänen et al., 2006; Wu et al., 2012, 2013). For AdV infection, 33K is the more abundant of the two proteins, but its main transcription is driven by the MLP. Regarding L4P transcripts, 22K mRNAs are equally or even more abundant (Donovan-Banfield et al., 2020; Westergren Jakobsson et al., 2021). Since E1A transactivates L4P transcription, the expression of these two proteins could be expected (Morris & Leppard, 2009). To test the potential influence on rAAV production of these two proteins, two plasmids were created: one Helper 1.0 based plasmid with a truncated *E2A* 5'-UTR, only containing the sequence of *L4-33K/22K* ($E2A_{min1}$), and the same sequence on its own up until the 5' end of the E2E promoter extension to ensure an adequate polyadenylation of the L4 mRNAs on a separate

plasmid. The latter was used in a complementary co-transfection in an equimolar amount to the 9 kb Helper plasmid ($E2A_{min2}$).

The presence of 33K and 22K in the transfected cells was tried to be determined by LC-MS. As expected, analysis confirmed the presence of 33K in the 12 kb plasmid, but not in mock transfected cells. 22K was not detected in any samples, which is not surprising given the fact that both samples share large parts of their amino acid sequences and are therefore hard to distinguish with the used non-specialised MS method. Contradictory results were received for the 9 kb plasmid and the co-transfected L4-33K/22K plasmid, showing peptide hits for the 9 kb plasmid alone, which does not carry the *L4-33K/22K* gene, but not for the 9 kb Helper together with the L4-33K/22K plasmid. Due to very similar labelling and the absence of the encoding *L4* gene in the 9 kb plasmid it can be assumed that these samples were mixed up. The expression of 33K from the co-transfected plasmid can thereby be confirmed with reservations. L4-33K expression from the minimalized 5'-UTR could also be confirmed. Expression was even detected from a plasmid with *L4-33K/22K*'s start codon mutated to a stop codon (ATG to TGA). This indicates an alternative non-AUG translation start or usage of an AUG translation initiation downstream, resulting in a shortened variant, for example missing the first 34 amino acids.

Resulting titres of the transient transfections with the reintroduced *L4-33K/22K* in the same plasmid or co-transfected showed significant increases in rAAV production in both cases ($p<0.0001$ in both cases). The previous 64% reduction in VG titre due to the missing of the 5'-UTR was mostly recovered with $E2A_{min1}$ to 90% of the titre with the full *E2A* 5'-UTR. The co-transfection of the *L4* gene elevated the titre to 69% compared to the full *E2A* sequence. Titre increases with the minimal 5'-UTR could partly be explained by the reintroduced E2L promoter. Splicing could also be changed, although E2E's untranslated exon was not reinserted, making splicing changes less likely. The co-transfection of *L4-33K/22K* leads to the conclusion that L4-33K/22K proteins are involved in efficient rAAV production and are most likely the main contributor to enhanced rAAV production of *E2A*'s 5'-UTR.

The reason why the presence of one or both proteins is beneficial for rAAV production can only be speculated. L4-22K and 33K are known to be essential for AdV packaging; however, this mechanism does not fully translate to AAV. Although the exact mechanisms are still not fully resolved, L4-22K requires a TTTG motif in the adenoviral DNA close to the ITRs for its role in packaging. Apart from the differences in ITRs and general packaging of the dsDNA of AdV, this motif is only present in the ITR/GOI plasmid's backbone and not in between the ITRs. L4-33K interacts with the E2A DBP and might act similarly in packaging to L4-22K due to

the shared N-terminus, where they form a portal complex for the dsDNA (Ahi & Mittal, 2016; Ewing et al., 2007; Wu et al., 2013). Despite the differences, it is possibly that this complex and the interaction with DBP could also be utilized and useful for AAV genome packaging. Additionally, L4-33K is involved in AdV capsid assembly (Finnen et al., 2001). This function could also transfer to AAV. Further, both proteins are known for transcription modulation, required for their role as early to late phase switches. Interestingly, while some studies describe early gene silencing by L4-22K, others reported increased E1A and DBP expression, indicating more complex viral expression regulations by the two *L4* products (Biasiotto & Akusjärvi, 2015; Wu et al., 2012). Finally, L4-33K's and L4-22K's function as a splice factor might be beneficial for Helper mRNAs, but also rep and *cap* originated mRNAs, all of which are extensively spliced. Of the two, mainly 33K is known for its splicing enhancing function located in its unique C-terminal domain (Biasiotto & Akusjärvi, 2015). Farris and Pintel (2008) showed that splicing of *cap* derived mRNA is enhanced in the presence of Helper plasmid and further that increased splicing goes along with enhanced rAAV titres. On the one hand, increased Rep amounts due to the Helper presence, specifically DBP, are a reason for this enhanced splicing. L4-33K/22K could be indirectly involved through transcription cascades. On the other hand, Helper functions themselves, particularly the two *L4* products could be directly involved, too. That the mutated plasmid, likely producing an N-terminal truncated 33K was capable of enhancing rAAV production to a similar extent as the unmutated *L4* could highlight that splicing enhancement is a major Helper function of the L4P originated proteins. For the AdV based TESSA rAAV manufacturing system, *L4-33K/22K* was also found to be essential (Nony et al., 2001; Su, 2021; Su et al., 2022). However, it is uncertain if the suggested (*cis*-acting replication element) CARE-based Rep/Cap amplification suggested for the system can be transferred to the present Rep/Cap plasmid, which is not integrated and only possesses a minimal CARE in the p5 promoter downstream of *cap*. Nevertheless, the observed rAAV titre increases could be originated in other beneficial functions of L4-33K/22K and mechanisms that work similarly in both rAAV production systems, for example amplified splicing, packaging, or transcription enhancement.

The *L4* gene is not essential for rAAV production as seen in Figure 20. This was also shown by Matsushita et al. (1998), who tried to figure out if the *L4* region is required for rAAV production and exchanged it with a CMV promoter, resulting in similar rAAV titres. Nevertheless, our experiments demonstrate that the *E2A* intron and L4-33K/22K proteins are integral components for optimal, high-yield rAAV8 production. Consequently, the intron or the shortened version including only *L4* was kept as a standard feature in subsequent plasmid modifications.

6.1.5 Dissection and Analysis of *E4* and its Open Reading Frames Leading to Increased rAAV Titres through ORF Recombination

The Adenovirus Helper gene *E4*, which encodes seven proteins (E4orf1, E4orf2, E4orf3, E4orf3/4, E4orf4, E4orf6, and E4orf6/7), was also investigated for potential optimisation. Each of these proteins from the different *E4* open reading frames (ORFs) exhibits different functions, including promoting viral gene expression and replication, as well as modulation of TF activities, but only the 34 kDa E4orf6 protein facilitates the essential AAV Helper function as previously demonstrated by Huang and Hearing (1989). While this knowledge was available for some time, many Helper plasmids included the entire *E4* gene, with only a few carrying solely *E4orf6* (J. M. Allen et al., 2000; Grimm et al., 1998; Hildinger et al., 2001; Matsushita et al., 1998; Xiao et al., 1998). This led to the question of whether the other E4orf proteins might have beneficial functions for rAAV production. To determine the specific contribution of different *E4orfs*, the *E4* gene was dissected by constructing Helper plasmids containing different subsets of the ORFs. Sequential truncations from the 5' end of the *E4* gene were performed, progressively excluding ORF's from the 5' end. This way six plasmids containing different combinations of *E4orfs* were created and tested for their ability to mediate rAAV8 production. As always, the different helper plasmids were used in equimolar amounts in transient triple transfections with constant plasmid copy numbers, crude rAAV titres were then analysed 72 h post transfection (Figure 21).

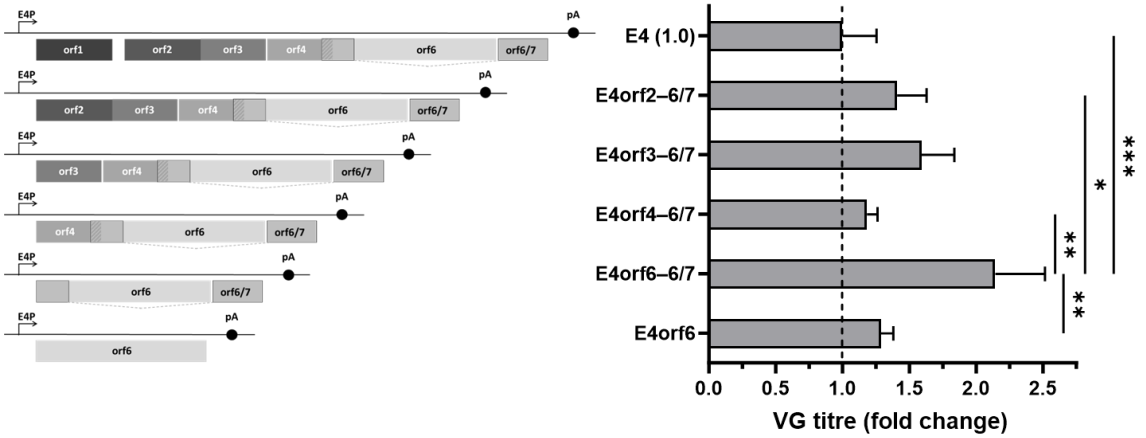


Figure 21: Sequential dissection of the *E4* ORFs for evaluation of their individual functions as components for rAAV8 production. Schematic depiction of the *E4* ORFs within the Helper plasmid (*E2A*, *L4-33K/22K* and *VA RNA* are not indicated). All components are drawn to approximate scale. rAAV8 VG titres are expressed as a fold change compared to the full-length *E4* gene (plasmid 1.0). Data shown as means with error bars indicating standard deviations of three biological replicates, statistical analysis was performed as an ordinary one-way ANOVA.

Confirming previously published results, *E4orf6* alone demonstrated the capability of rAAV production and provided the essential *E4* Helper functions (J. M. Allen et al., 2000; Ferrari et al., 1996; M. M. Huang & Hearing, 1989a; Samulski & Shenk, 1988). The virus genome titres with E4orf6 alone were equivalent to those observed with the full-length *E4* gene, representing a 1.3-fold increase compared to *E4*. Interestingly, the removal of *E4orf1* and *E4orf2* resulted in increases in rAAV titres by 22% and 37%, respectively, when compared to the E4 control. However, the removal of *E4orf3* reduced titres to the level of the *E4* control and did not follow the occurring trend of shorter *E4* sequences resulting in higher titres. In contrast, additional exclusion of *E4orf4* resulted in the highest measured titres, showing a 2.1-fold increase over the full-*E4* control. The combination of *E4orf6* and *E4orf6/7* performed much better than Orf6 alone, suggesting highly beneficial functions of E4orf6/7. At first, an influence of the 3'-UTR and polyadenylation was expected, because the plasmid carrying only *E4orf6* was constructed differently at first, resulting in the exclusion of 92 bp of *E4*'s 3'-UTR. However, a new *E4orf6* construct with the identical 3'-UTR to the other *E4* variants produced similar titres (0.85±0.11-fold of the full *E4*), confirming the beneficial influence of E4orf6/7.

This positive influence of E4orf6/7 is not unexpected, as it is known to enhance *E2A* transcription and stimulate other promoters that are also activated by the transcription factor E2F. The stimulation of cellular E2F targets facilitates the progression of the cell cycle into the S phase, supporting viral DNA replication (D. G. Johnson et al., 1993; Schaley et al., 2000). This function is similar to and probably cooperative with the induction of G1/S-Phase by Rep68 and followed S-phase cell cycle arrest by Rep78, as well as the interplay of E1A, E1B and E4orf6 supporting the same cell cycle control (Saudan 2000, Ben-Israel 2002). Increased expression of E4orf6/7 might therefore be beneficial because it increases DBP levels and can potentially substitute for lower levels of large Reps, E1A, E1B and E4orf6, facilitating S-Phase cell cycle entry and arrest. On the other hand, E4orf1's natural functions do not seem to be required or beneficial for rAAV production. The same could be the case for E4orf2, although its functionality might be generally miniscule, since there are still not characterized. On the other hand, the functions of E4orf1 and E4orf2 do not seem to be required or beneficial for rAAV production. Their exclusion may enhance rAAV titres due to the closer proximity of the E4 promoter to essential and supporting ORFs 3, 6 and 6/7, as well as the removal of other splice variants, resulting in increased mRNA amounts for these beneficial mRNAs.

The removal of *E4orf3*, the naturally most abundant ORF mRNA (Dix & Leppard, 1993; Westergren Jakobsson et al., 2021), reduced rAAV titres. E4orf3 possesses abilities of chromatin modulation and protein interactions, resulting in virus

promoting cell modulation and can thereby even substitute for E4orf6 in AdV replication. These abilities are likely beneficial for rAAV production, explaining the drop in rAAV titres upon its exclusion. Furthermore, E4orf3 is an activator of L4P (J. Wright & Leppard, 2013). Increased 33K/22K amounts likely enhance rAAV production, as it was previously demonstrated here that these proteins contribute to rAAV VG titre improvements. E4orf4 negatively regulates *E1A* and *E4* transcription, actively reducing rAAV-supporting Helper functions. Amongst others, it also impedes the L4 promoter, which is activated by E1A and E4orf3. The dephosphorylation of E1A mediated by E4orf4 counteracts the actions of E4orf6/7, inhibiting E2F-dependent promoter activation, including essential E2E promoter activation (Mannervik et al., 1999). Additionally, E4orf4's splicing function does not seem required or may even be counterproductive for rAAV production. In summary, it can be concluded that for rAAV production, only *E4orf6* is essential, *E4orf1* and *E4orf2* are dispensable or even counterproductive. *E4orf4* should be eliminated due to its negative influence on rAAV titres. Beneficial features for rAAV production were demonstrated by *E4orf3*, and mostly notably, by *E4orf6/7*. Therefore, the ORF subset fragment *E4orf6+6/7* was utilized in most of the subsequent plasmids.

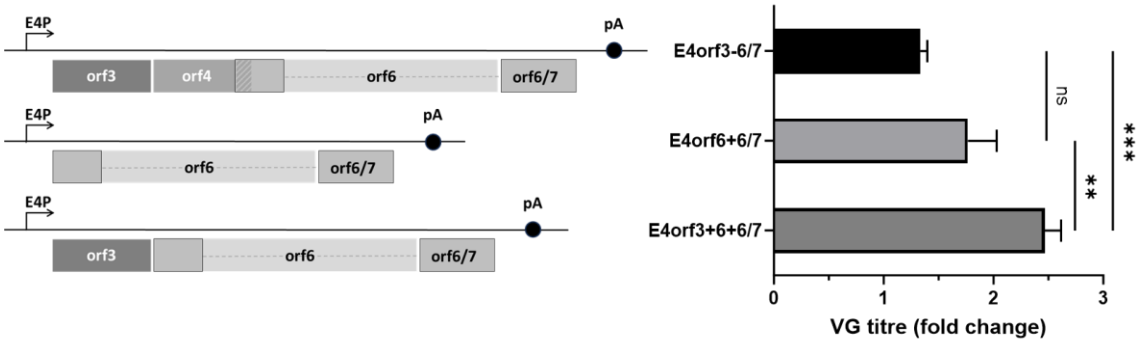


Figure 22: Combination of *E4* ORFs demonstrating improvement of *E4*'s capability as AAV8 Helper function. Schematic depiction of the *E4* ORFs within the Helper plasmid (*E2A*, *L4-33K/22K* and *VA RNA* are not indicated). All components are drawn to approximate scale. rAAV8 VG titres are expressed as a fold change compared to the control plasmid 0.1. Data shown as means with error bars indicating standard deviations of three biological replicates, statistical analysis was performed as an ordinary one-way ANOVA.

It was further hypothesised that a combination of ORFs, specifically excluding the repressing *E4orf4* and including the beneficial *E4orf3*, would maximise positive features of the *E4* gene for enhanced rAAV production. Consequently, a plasmid was created containing a combination of *E4orfs*: *3*, *6* and *6/7*. The achieved rAAV VG titres were 2.5-fold greater than those of the control plasmid Helper 0.1. While a direct comparison to the full *E4* of Helper 1.0 could not be made, an extrapolated value suggested an approximately 2.9-fold higher VG titre based on the observed 1.17-fold difference between Helper 0.1 and Helper 1.0 (Figure 17). The significant

increase in VG titre of *E4orf3+6+6/7* compared to the other *E4orf* subsets confirmed the hypothesis of an optimized composition of *E4orfs* improving rAAV production. The increased titres demonstrated that by simple deletion and recombination of individual ORFs, desired functions can be incorporated and possibly even amplified, while the elimination of undesired functions also has a positive impact on rAAV titres.

6.1.6 Change of the Helper Plasmid Backbone

The last part of the plasmid optimisation of the more than 20-year-old pADΔF6 Helper involved changing the plasmid backbone, which is responsible for its amplification in bacteria. The attempt to reduce the plasmid size by eliminating pADΔF6's 0.6 kb gap between *ori* and *KanR* and 0.3 kb gap between the antibiotic resistance and its promoter, using site-directed mutagenesis (SDM), resulted in drastically reduced rAAV titres. Despite unchanged amplification behaviour and full sequence confirmation by Sanger sequencing (data not shown), the plasmid seemed not able to support adequate rAAV production (Figure 23).

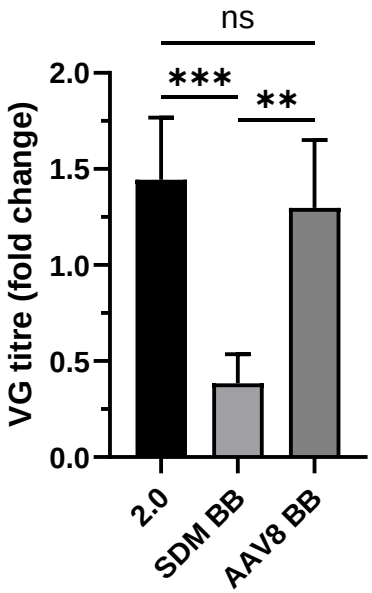


Figure 23: VG titres of E4orf6/7 (Helper plasmid 2.0) based Helper plasmids, differing in their plasmid back bones. VG titre fold change to the control plasmid 0.1. Means and error bars as standard deviation are shown of Helper 2.0 containing the pADΔF6 back bone (2625 bp, n=6), the SDM BB containing plasmid (1706 bp, n=4) and the AAV8 BB containing Helper plasmid (1897 bp, n=6), ordinary one-way ANOVA was performed for statistical analysis.

Different factors can influence the transfection and transcription efficiency of a plasmid, many of these potentially influenced by the plasmid's backbone. Most prominent are plasmid size and topology, as well as CpG content and CpG islands (Boye et al., 2022; Coban et al., 2005; Hornstein et al., 2016; Kreiss, 1999; Maucksch et al., 2013; Prösch et al., 1996). It is possible that the topology of the SDM-altered plasmid changed, or transcription efficiency, in particular of the E4 promoter was reduced due changed upstream sequence. For size reduction and elimination of the two unrequired backbone sequences in the Helper plasmid, the backbone was then exchanged to the one of the AAV8 Rep/Cap plasmid. As expected, plasmids with this backbone could be produced without any problems and rAAV production resulted in similar titres

to the original backbone of pADΔF6. The AAV8 plasmid backbone is in total 900 bp smaller than the previously used backbone. For future cloning purposes, an EcoRI restriction site was added downstream of the ori via SDM, which did not cause any titre variations. Furthermore, the restriction site FseI was inserted downstream of *E2A*, in front of its poly adenylation sites (pA) via site directed mutagenesis to expand the options of restriction enzyme-based exchanges of pAs and other elements at this position. For similar reasons, the mini multiple cloning site located downstream of *E4* and upstream of *E2A* was expanded in additional plasmid alterations. For future Helper, ITR/GOI and Rep/Cap plasmids, novel alternatives to antibiotic resistance-based backbones with much smaller sizes could be utilized. Minicircles offer the best size option, with about 100 bp, their low production yields do not make them feasible for a transient transfection cGMP process. However, the recently created nanoplasmid vectors are supposed to yield similar plasmid amounts with only 500 bp backbone constructs, while enhancing the level and longevity of expression, as well as mitigating cell-transfection-related toxicity and suppression of transgene caused by CpGs in conventional backbones (Williams & Paez, 2023).

6.1.7 Summary

- Non-essential DNA fragments were removed to reduce the size of the popular and commonly used, but large and for 20 years unchanged Helper plasmid pADΔF6 (15.8 kb)
    - Removal of non-essential and non-coding regions of the plasmid backbone, upstream and downstream of the *E4*, *E2A* and *VAI* genes
    - Removal of the *L5 (fiber)* gene did not affect virus production and eliminated a potential contamination which could be relevant for downstream processes
- Introduction of restriction sites for simplified molecular engineering, including exchange of genes and genetic elements like promoters
- Analysis of the *E2A* 5'-UTR revealed a so far unrecognized AAV Helper function that increases rAAV titres, L4-33K/22K
- Sequential truncation, analysis and recombination of the *E4* ORFs resulted in highly increased rAAV virus genome titres through the creation of Helper plasmids containing *E4orf* subsets excluding adverse and recombining advantageous *E4orfs*
    - E4orf4 was identified as restrictive for high titre rAAV production
    - E4orf3 and E4orf6/7 are supportive E4orfs for rAAV production
    - The novel subset of E4orf3+6+6/7 increased rAAV8 titres 2.5-fold

6.1.8 Conclusion

The initial phase of this work centred on the modernisation of the widely employed Helper plasmid pADΔF6. The successful reduction in plasmid size, from 15.8 kb to approximately 10 kb, has the potential to enhance transfection efficiencies and thereby reduce the quantity of costly cGMP plasmid material required. This downsizing also gives leeway to integrate or substitute new genetic components without risking a substantial drop in transfection efficiencies compared to the baseline plasmid. The integration of new restriction sites has streamlined the exchange of such genetic elements. Notably, the newly modular arrangement significantly simplifies future genetic engineering of these plasmids, constituting a cornerstone of the majority of this study's endeavours.

The comprehensive analysis of the AdV Helper genes *VA RNA*, *E2A*, and *E4* and their influence on rAAV VG titres underscores the dependence of rAAV vectors for gene therapy and their efficient production on (AdV) Helper functions. This reaffirms the Helper plasmid's significance as a promising engineering target for enhancing rAAV production. Moreover, the outcomes of these genetic engineering-based functionality analyses have unveiled an uncharted AdV helper function for improved rAAV proliferation—*L4-33K/22K*. Additionally, they have highlighted the potential of *E4* ORF subset recombination with an augmented capacity for enhancing rAAV production. These discoveries possess substantial commercial potential and could shape future plasmid designs beyond the confines of this study. As a consequence, the recombination of *E4* ORF subsets has already been incorporated into patent “Composition and Methods for Recombinant AAV Production” by REGENXBIO (P. Liu et al., 2023).

6.2 Investigation of Genetic Component Stoichiometries for Optimisation of rAAV Production

An overarching objective of this project was to elevate rAAV production by refining the stoichiometry of individual genetic components within the transient transfection system. The transition from virus replication to plasmid-based viral vector production introduces the possibility of an imbalance among AAV and AdV Helper components. Optimising viral gene expression and component ratios has the potential to elevate rAAV titres and improve product quality. Therefore, various experimental methodologies were employed to identify the requirements of increased or decreased expression of individual genes.

6.2.1 Plasmid Ratio Alterations

In a first approach it was aimed to alter stoichiometry at the plasmid level. Changing the amounts of the three individual plasmids can influence the abundance of their connected group of functions. While this does not allow for individual control, it can indicate if a specific group of genes, in form of the individual plasmids ITR/GOI, Rep/Cap and AdV Helper, requires adjustments for increased rAAV production. Other studies have shown that plasmid ratios and the ratio of plasmid DNA to PEI are important process parameters and optimisation can improve rAAV titres. As the general production process was predefined by REGENXBIO and the optimisation interest was focused on the gene abundancies, no experiments were performed to optimize the PEI to DNA ratio. Additionally, the ITR/GOI plasmid amount was speculated to be less critical due to its lower utilisation in the original process compared to the other two plasmids (molar plasmid ratio Helper : Rep/Cap : ITR/GOI = 0.92 : 1 : 0.12, mass ratio 2 : 1 : 0.1).

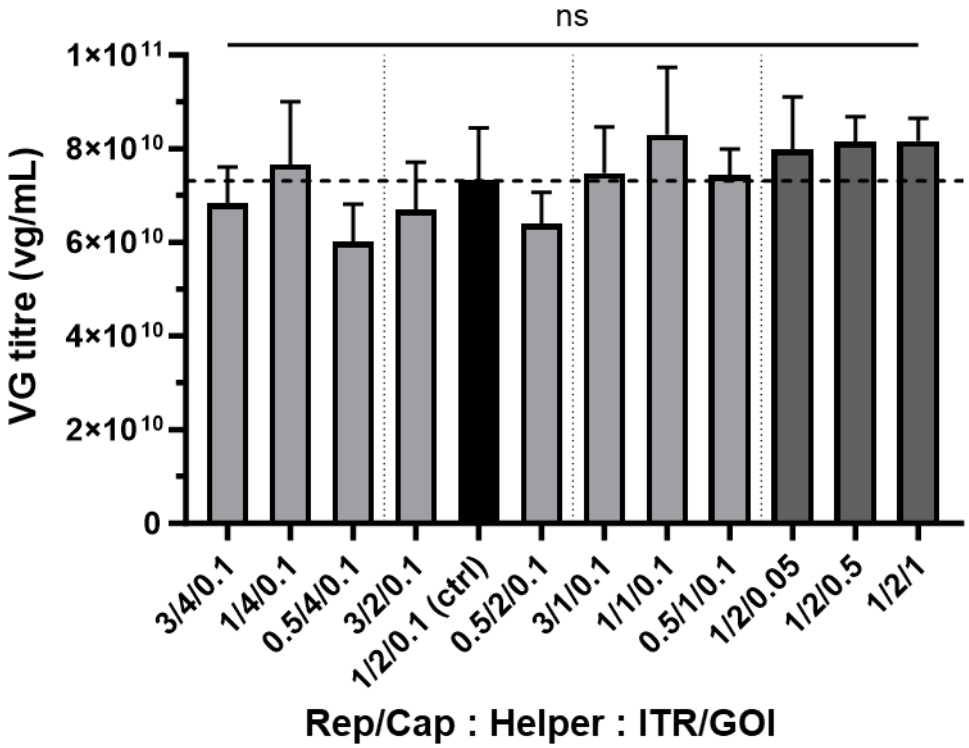


Figure 24: Plasmid ratio variation for rAAV production. VG titres per mL of transient triple transfections with the original plasmid system (Rep/Cap v1.2, pADΔF6 (0.1), ITR/GOI) with the originally used mass ratio 1 : 2 : 0.1 (equivalent to molar plasmid ratio 1.1 : 1 : 0.1). Indicated plasmid ratios describe mass ratios of Rep/Cap to Helper to ITR/GOI, the total plasmid amount was kept constant. Means with standard deviation as error bars are shown of two biological replicates for the DOE like approach (grey) compared the control (n=6, black). Variation of ITR/GOI plasmid amounts (dark grey) were transfected as biological duplicates in parallel. Ordinary one-way ANOVA was performed for statistical analysis.

As shown in Figure 24, increasing the ITR/GOI plasmid amount five- or even ten-fold did not result in increased rAAV titres. No statistically significant differences were observed between the different ITR/GOI plasmid amounts. It is possible that the AAV rolling hairpin-like genome replication process produces sufficient single-stranded genomes even when copy numbers of the ITR/GOI plasmid per cell are low. However, the increase of the ITR plasmid amount seemed to decrease data variation, indicated by a decrease of the data's standard deviation. A reason for this effect could be that the relatively small amount of ITR/GOI plasmid in the mix, in combination with the formation of PEI/DNA polyplexes, could lead to a less homogenous distribution of the plasmid across the polyplexes, resulting in more or less cells receiving sufficient amounts of the plasmid after uptake of polyplexes. With the very low number of only two replicates the observed effect of increased ITR/GOI plasmid amounts can only be assumed. Subsequently, in an attempt to increase transfection result reliability, a further investigation into this topic was performed to gather further evidence and insight (see 6.3.1).

Since changes of the ITR/GOI plasmid amount showed no effect, the DOE-like approach was conducted with varying amounts of only the Helper and Rep/Cap plasmids. Both factors were tested with three levels, with the middle one for Rep/Cap and the lower one for the Helper plasmid chosen based on the control ratio of 1:2:0.1. The upper and medium levels for the Helper plasmid were set at four and two times the control. As a result of the higher rAAV titres achieved with the new 12 kb Helper plasmid and higher Helper to Rep/Cap and ITR ratios of other groups, it was concluded that reducing the Helper plasmid would likely be counterproductive. The additional levels for the Rep/Cap plasmid were chosen as three times and half of the lower Helper amount. An increased Rep/Cap plasmid amount could potentially lead to a beneficial increase in capsid production, but too much large Rep could be cytotoxic. Therefore, the chosen levels included one higher and one lower than the control.

Although differences in rAAV8 titres were observed, they were not statistically significant, and not usable to create a model. Observing the three Helper plasmid levels individually (see dotted lines in Figure 24), revealed that neither increasing nor decreasing the Rep/Cap plasmid increased the titres. Further, this sorting indicated a slight increase in titre with the lowest Helper amounts. In accordance, the highest titre was achieved by the 1:1 mass ratio, which resulted in a roughly doubled copy number of Rep/Cap plasmids due to the Helper's larger size. However, the relatively small and statistically insignificant differences in titres were unexpected, given the success of plasmid ratio variations in other studies. For example, Grieger et al. (2016) showed an about 3.2-fold increase in virus genome titre with their optimized molar plasmid ratio of 1.5:2:1 (mass ratio of Rep/Cap : Helper : ITR/GOI). The systematic plasmid ratio DOE approach of Zhao et al. (2020) even yielded up to 6.1-fold increases in VG titre. Iterations of designs spaces lead to the plasmid mass ratio 5:1:0.31, decreasing the ITR/GOI plasmid similarly to the presently used ratio, but increasing proportion of Rep/Cap plasmid further. It was hypothesised that the expression of *cap* originated proteins is likely limiting and rAAV production can therefore be increased by increasing transcription of *cap* either with stronger promoters or by increasing the Rep/Cap plasmid amount as shown in previous studies (Vincent et al., 1997; Zhao et al., 2020). In their DOE experiments Zhao et al. (2020) varied plasmid ratios by up to 5-fold between different plasmids. Given the insignificant differences of the present experiment, the design space in would have needed to be extended much further. However, it is to note that the chosen differences between plasmid amounts were already considerably big. The conversion into molar ratio levels shows that the used mass ratio options 0.5/4/0.1 and 3/1/0.1 mean a difference of 24 times more Rep/Cap

copies compared than Helper copies. A further decrease of the Helper plasmid could possibly have resulted in at least slightly further increased rAAV titres, particularly with the results of Zhao et al. (2020) in mind. However, the results of other groups showing better rAAV production with more Helper plasmid than Rep/Cap plasmid hint that a much further decrease in Helper plasmid amounts would possibly decrease titres (Emmerling et al., 2016; Grieger et al., 2016; Nguyen et al., 2021; Xiao et al., 1998). In combination with the marginal titre differences of the present data, it can be assumed that the utilised plasmid ratio is already close to the optimum. A titre increase was not the main goal of the experiment, but more gaining more insight into ideal component stoichiometry. However, the marginal or no increases in titres suggest that the on the resolution level of plasmid ratios, the current stoichiometries are already near the optimum, supported by similar ratios used by other groups and the idea of lower ITR/GOI requirement confirmed by Zhao et al. (2020). The optimal molar ratio for Grieger et al. (2016) was 2:1:1.5, whereas the findings of Nguyen et al. (2021) showed with modelling and experimental data an optimal molar plasmid ratio of 1.5:2:1. However, the clustering of most components onto only two variables was a limitation for component stoichiometry resolution in this approach. Therefore, a different approach was conducted to gain more insight into the ideal component stoichiometries for rAAV production.

6.2.2 Deconstruction of Helper and Rep/Cap Plasmids

To achieve a better understanding of the individual genetic components' ideal abundance for rAAV production, both the Rep/Cap and Helper plasmids were split into multiple plasmids, separating individual components. The Rep/Cap plasmid was split into two plasmids containing the two genes *rep* and *cap*, respectively. The Helper plasmid's genes (*E2A*, *E4*, and *VA RNA*) were placed on different plasmids. This resulted in a 6-plasmid transfection system for rAAV production, with the ITR/GOI and all involved viral genes on separate plasmids, except for the genomically integrated *E1*.

Similar to the plasmid ratio experiment with the regular 3-plasmid system, a DOE like approach was planned for this 6-plasmid transfection system. Since the ITR/GOI amount was not of concern, the amount of this plasmid was regarded as constant. Further, a preliminary experiment revealed that increasing the amount of the *VA RNA* carrying plasmid up to three-fold had no effect on rAAV titres (see appendix Figure 49). Due to its internal RNA pol III promoter, variation of *VA RNA* amounts would have been difficult, making this observation a welcome result. A reason for the insignificance of increasing gene copy numbers of *VAI* is likely its general high transcription and resulting extraordinary abundance of $\sim 10^8$ cellular copies (Vachon & Conn, 2016). Thus, this plasmid was also regarded as a constant. For the

remaining four plasmids a higher and a lower level were utilised. As a control, the plasmids were transfected in an equimolar ratio. The equalised copy number was used as one of the levels. The other was chosen as follows. For pRep, the molar plasmid ratio level was halved for the lower level, for pCap and pE2A the upper ratio value was 2.5 and for pE4 it was doubled. Transient transfections were carried out in biological quadruplicates, the results are shown in Figure 25.

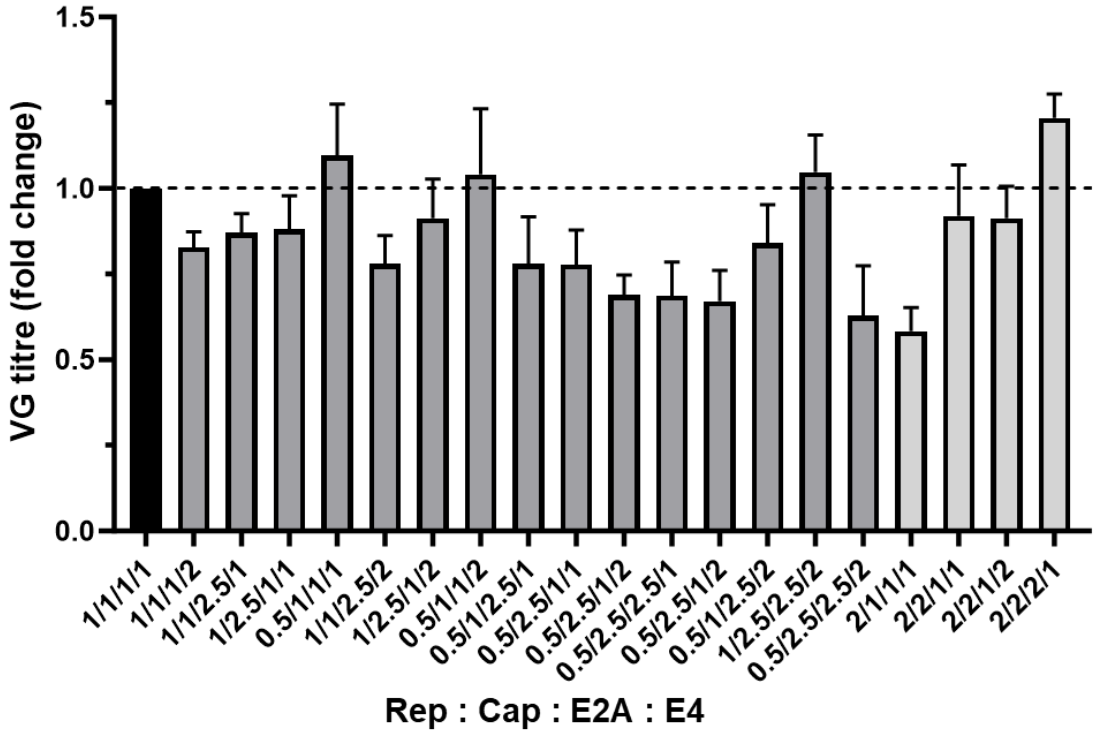


Figure 25: Different plasmid ratios of transient transfections with a 6-plasmid system for rAAV8 production, splitting the required genes of AAV and AdV5 onto different plasmids. Plasmids ratios are indicated as molar ratios in the order pRep:pCap:pE2A:pE4, the plasmid amounts of pVA and pITR/GOI were constant. Fold changes are respective to the control ration 1/1/1/1. Means with standard deviation as error bars are shown of the four biological replicates for the different ratios of the original design space (dark grey) and additional ratios (light grey) compared the control (black).

Although titre differences could be observed most changed molar plasmid ratios of the original design space achieved lower rAAV8 VG titres compared to the control. The exceptions were 0.5/1/1/1, 0.5/1/1/2 and 1/2.5/2.5/2. All three have a lowered rep amount compared to all other components in common. Also, the additional sample 2/1/1/1 follows this trend. The large Reps, particularly, Rep78, are known to be cytotoxic and therefore this finding is in line with previously published results stating to decrease large Rep production for increased rAAV titres (J. Li et al., 1997; Xiao et al., 1998). However, not all samples with decreased pRep resulted in increased titres. For example, the sample with the lowest *rep* content

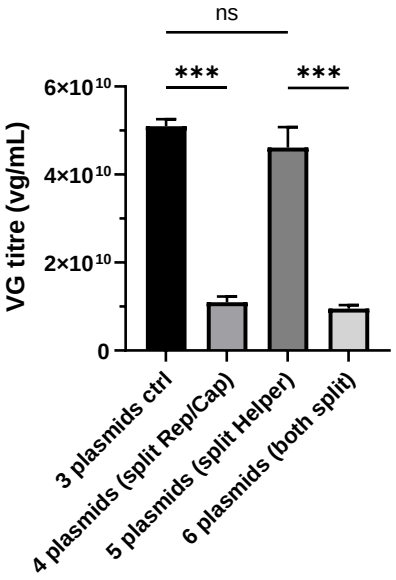


Figure 26: rAAV production with different transient transfection systems allocating AAV and AdV Helper genes to 3 plasmids (Rep/Cap v1.2, Helper 1.0, ITR/GOI), 4 plasmids (pRep, pCap, Helper 1.0, pITR/GOI), 5 plasmids (pRep/Cap v1.2, pE2A, pE4, pVA, pITR/GOI) or 6 plasmids (pRep, pCap, pE2A, pE4, pVA, pITR/GOI). Data show means and standard deviations as error bars of two independent biological replicates, ordinary one-way ANOVA was performed for statistical analysis.

(0.5/2.5/2.5/2) shows the lowest titre overall, indicating a too low amount of replicases for efficient rAAV production. The highest rAAV8 VG titre reached with the 6-plasmid system originated from the sample 2/2/2/1, which effectively displays a decreased *E4* amount. However, the lower level *E4* samples did not generally show higher titres. No obvious trends could be observed for pE2A and pCap.

Although the achieved results are not entirely unusable and meaningless, the experiment generally encountered three different errors, which did not allow the statistical evaluation required for a model and question the reliability these results. Firstly, one sample (0.5/2.5/1/2) was transfected with the low instead of the higher pRep level. Secondly, the differences between biological replicates were relatively high, an issue that reoccurred during the whole course of the project with different experiments, making it difficult to compare raw titres and therefore requiring normalisation to a control. The data was consequently presented as fold changes to each biological replicate of sample 1/1/1/1 in Figure 25, accounting for day-to-day biological and experimental variation. The dramatically decreased relative errors of all samples shows that inconsistencies mainly result from day-to-day variation. Thirdly, titres of all transfections with the 6-plasmid system were noticeably low. To determine whether this issue was due to the split system in general or a specific plasmid, transfections with intact Rep/Cap plasmid and split Helper (4-plasmid system), as well as intact Helper and split Rep/Cap plasmid (5-plasmid system), were conducted. The resulting crude lysate rAAV8 VG titres are shown in Figure 26.

Results of the transfection of the four different plasmid systems with differently split Helper and Ren/Cap plasmids clearly demonstrated a fault with the split Rep/Cap plasmid. Achieved titres with the intact Rep/Cap and split Helper plasmid could not be discerned from the original triple transfection system. In contrast, both systems using the split Rep/Cap (4- and 6-plasmid), showed significantly lower rAAV titres.

Consequently, either the pRep or pCap plasmid did not function sufficiently or splitting the two usually overlapping genes apart resulted in general misfunction. The separation of the two genes also turned out to be difficult in the same plasmid as described in chapter 6.3. Other studies, however, demonstrated that the separation of *rep* and *cap* can be done successfully and even increase rAAV titres (J. M. Allen et al., 2000; Emmerling et al., 2016; Whiteway et al., 2003). Detection of rAAV virus genomes showed that the created split Rep and Cap plasmid were able to produce rAAV8 to some extent, but the low titres do question the relevance of the results from the 6-plasmid system. One issue of the plasmids might have been the missing p5 promoter downstream of *cap*, which is likely required for p19 and p40 activation and turned out to be essential in the split Rep/Cap plasmid for the three-plasmid system (Pereira & Muzyczka, 1997a). Further analysis or optimisation of the six-plasmid system were dismissed, and another alternative approach was conceived to understand which components require an altered abundance for increased rAAV production. Since the underlying idea was to achieve optimised component stoichiometries with alternating promoters in the final system, it was conceptualised to work closer to this idea and utilize the well characterised and strong CMV promoter to increase Helper components. However, also the split Helper system (5-plasmid system) was used for further analysis of the utilized transient transfection system.

6.2.3 Evaluation of the Requirement of Individual Helper Genes and their Abundances with a Split Helper Plasmid System

As an alternative strategy, individual Helper genes were tested for their essentiality or benefit in the current system, expressed either from their endogenous promoter or from a strong CMV promoter. Different plasmids and combinations of plasmids with individual Helper genes were transfected with the regular Rep/Cap plasmid (v1.2) and the ITR/GOI plasmid. (5-plasmid system).

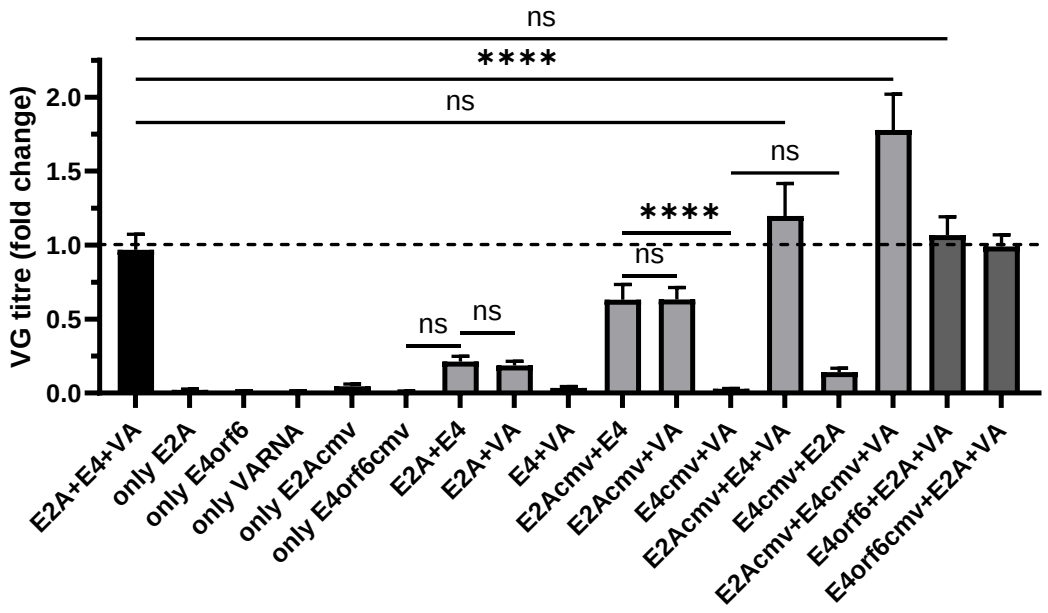


Figure 27: Transfections of individual AdV Helper genes placed on separate plasmids and combinations of these. Based on the 5-plasmid system (Rep/Cap v1.2, pE2A, pE4, pVA, pITR/GOI), transfected genes were expressed by either their endogenous promoter or a heterologous CMV promoter (grey). Bars of transfections containing *E4orf6* instead of the full *E4* gene are displayed in dark grey, the regular 5 plasmid system in black. VG titre fold changes to the 3-plasmid control (Helper 1.0) are displayed as means with standard deviation as error bars of three individual biological replicates. Data shown as means with error bars indicating standard deviations of three biological replicates, statistical analysis performed as an ordinary one-way ANOVA, pairwise comparisons displayed for selected data pairs.

The experiment was designed similar to the ones of Matsushita et al. (1998) and shows similar results. Neither of the three Helper genes that were described as essential for AAV proliferation, namely *E2A*, *E4*, and *VA RNA*, were capable of promoting rAAV production on their own when transcribed by their original promoter. However, the use of the CMV promoter for *E2A* showed a slight increase in VG titre. Apart from a potential, marginal increase in rAAV production this result could also be an artifact caused by DBP's natural function of binding single stranded DNA and thereby protecting AAV genomes from DNAseI digest prior to the ddPCR measurement, similar to the protection by the capsid. Therefore, it cannot be concluded that *E2A* alone is capable of promoting rAAV production to a minimal amount. Instead, a combination of at least one of the other two genes and *E2A* seems to be required. A high increase of these low rAAV titres (about 20% of the control) to 63% of the control's titre was achieved by using a CMV promoter for E2A expression in combination with pE4 or pVA. The addition of the third gene further boosted VG titres by 1.2-fold above control levels.

Considering *E4*'s various actions and interactions to promote AAV proliferation, it was not surprising that E2A+VA rAAV titres with *E4* are four times higher. The observed increase of only about two-fold when *E2A*'s transcription is controlled by

a CMV promoter might be associated with *E4orf6/7*. This ORF's primary function enhances transcription from *E2A*'s E2E promoter. Since CMV is a genuinely much stronger promoter and does not get co-activated by E4orf6/7, DBP amounts are likely sufficiently high regardless of E4orf6/7, and this protein might become redundant for rAAV production with a CMV promoter controlling *E2A*. To investigate this hypothesis, a transfection of the plasmids pE4orf6cmv, pE2Acmv, and pVA could be conducted. When *E4* or *E4orf6* were tested alone or in combination with *VA RNA*, neither was capable of elevating rAAV production in the absence of *E2A*. This finding established *E2A* as the only indispensable AdV5 Helper function for rAAV8 production, which is consistent with the results of Matsushita et al. (1998), stating that *E2A* contributes much more to AAV production than *E4* and *VA*. However, the cellular modulation by either E4 or VAI seems to be required, likely to shut down host cell processes and support AAV protein expression.

It is noteworthy that E4orf6cmv alone was not capable of providing all the required functions for sufficient rAAV production, contradicting the results of Allen et al. (2000), who postulated that E4orf6 is the only essential Helper function when expressed with a CMV promoter. The discrepancy could be attributed to the study's use of heterologous promoters for *rep* and *cap*, which seems to make a crucial difference. E4orf6 is known to be involved in the upregulation Rep and VP protein synthesis. Since p40 and p19 are only upregulated by Rep78/68 in presence of E1A these promoters should be functional in HEK293 cells. However, studies of Wang et al., (2018) showed that despite the present E1A in the genome of HEK293 cells, p19 and p40 transcripts were >5-times lower in the absence of a AdV Helper plasmid (substitution with a boca virus Helper). Both p5, and therefore in a cascading reaction, probably also p19 and p40 are activated by E1A but also E2A (DBP), likely making these Helper functions indispensable for rAAV production with the endogenous promoters p5, p19 and p40. The present results support that it is unlikely that this upregulation of p40 and p19 is facilitated by E4orf6. Consequently, heterologous promoters for *rep* and *cap* are required for efficient AAV protein expression when only *E4* is utilized as a Helper function to modulate the cell for virus production.

E4orf6's other functions, including the promotion of second-strand synthesis and MRN complex breakdown, seem to be beneficial for rAAV production but not essential. Similar to results of the 3-plasmid system, transfection of *E4orf6* instead of the full *E4* gene demonstrated to be as efficient, confirming previous studies identifying *Orf6* as the essential part of *E4* for AAV production. A dissection of the gene was performed to test if benefits and drawbacks of individual *E4orfs*, similar to the 3-plasmid system, were present (Figure 28).

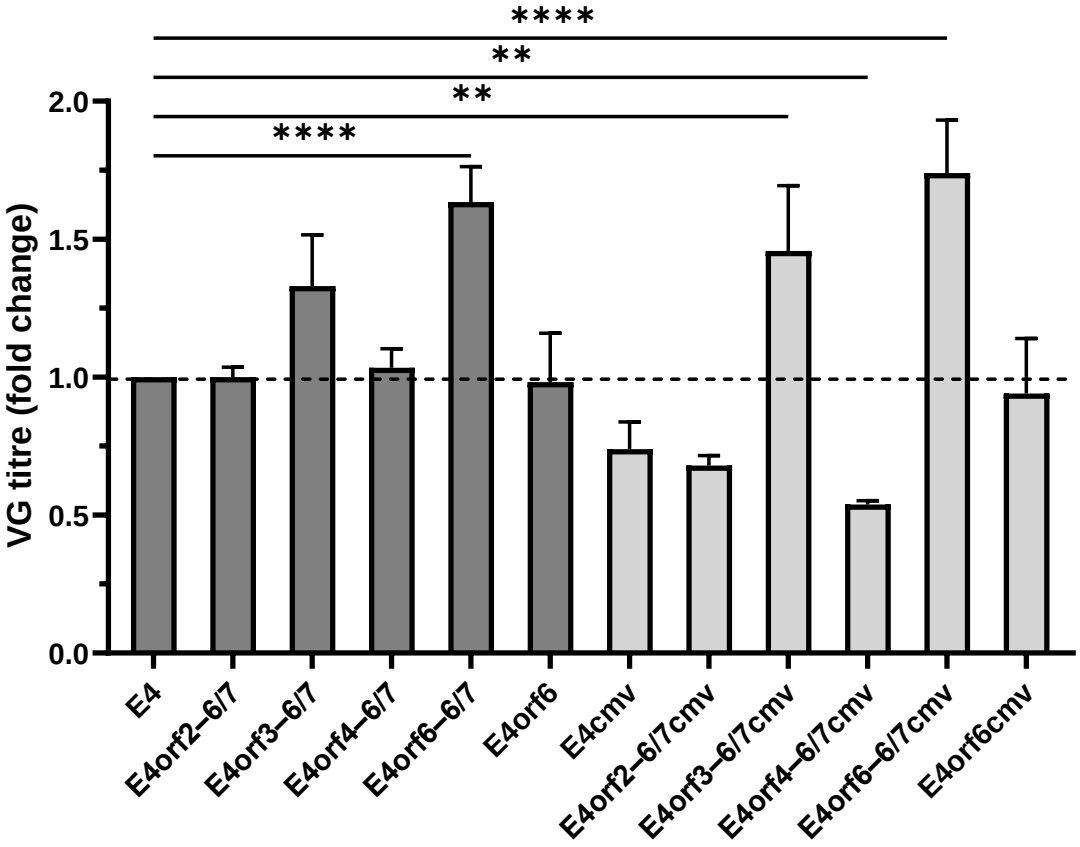


Figure 28: Sequential dissection of the *E4* ORFs in the 5-plasmid system with individual plasmids for each gene of the AdV5 helper. The *E4* gene was truncated from the 5'-end and the resulting *E4orf* subsets were expressed either with their endogenous E4P (darker grey) or a CMV promoter (lighter grey). rAAV8 VG titres are expressed as a fold change compared to the full-length *E4* gene transcribed from the E4 promoter. Data shown as means with error bars indicating standard deviations of three biological replicates. Statistical analysis was performed as an ordinary one-way ANOVA, pairwise comparisons displayed for selected data pairs.

Presumably increased *E4orf6* transcription with a CMV promoter did not change rAAV8 titres for the pE4orf6cmv transfections compared to the ones with pE4orf6. Usage of CMV for the whole *E4* gene instead decreased titres by more than a quarter. Generally, the dissection of the *E4* gene in the 5-plasmid system confirmed trends of the same experiment in in the 3-plasmid system (Figure 21). However, whereas the deletion of ORFs 1 and 2 did both have a positive impact in the 3-plasmid system, there was no effect for *E4orf1* deletions regardless of the promoter in the 5-plasmid system. Deletion of *E4orf1-2* increased titres, and further deletion of *E4orf3* decreased rAAV8 titres in all cases. This confirmed the previous findings of Orf3 providing beneficial AAV Helper functions and Orf4 being a repressor of rAAV production. The largest differences were obtained with the CMV controlled *E4orf*-subsets. The CMV promoter likely transcribes higher mRNA amounts. If these transcripts were mainly from the beneficial ORF's 3, 6 and 6/7, it would result in a positive effect, although titres did not exceed the E4P controlled equivalents.

However, if the CMV promoter was placed in front of a destructive ORF and this E4orf was presumably predominantly expressed, or other E4orfs produced to a lesser extent, it would result in decreased rAAV titres. However, for the abundance of an individual ORF, the splicing would need to change as well. To measure *E4orf* transcript abundances, mRNA deep sequencing could be deployed similar to experiments by Donovan-Banfield et al. (2020).

In contrast to *E2A*, the use of a CMV promoter for *E4* or *E4orf6* did not increase rAAV production, even when unfavourable *E4orfs* were removed (Figure 28). However, the combination of pE2Acmv and pE4cmv increased rAAV production significantly compared to the 5-plasmid control, as well as 1.8-fold compared to the 3-plasmid control (Helper 1.0) (Figure 27). Contrary to the results of Matsushita et al. (1998), the E2Acmv plasmid and the double CMV system worked much better than the controls, which all demonstrated similar AAV yields in their experiments. Balancing out the likely strongly increased *E2A* transcription with enhanced transcription of *E4* appeared to have a strongly positive effect on rAAV production in the utilised system. Encouraged by this result, CMV promoters were also deployed in the 3-plasmid system.

6.2.4 Effects of the CMV Promoter Driving Helper Gene Transcription and its Use with and without *E2A*'s 5'-UTR

The pE2Acmv plasmid in the 5-plasmid system used a CMV promoter immediately upstream of *E2A*'s translation start site. However, it is uncertain whether placing the CMV promoter in this position or in the original position of the E2E promoter (3 kb upstream) would be preferable for high titre rAAV production. On the one hand, the large distance between promoter and gene might reduce its transcriptional activity for the CDS, and effects of the intron and its splicing on mRNA abundance are unknown and uncertain variables when exchanging the promoter. On the other hand, the previously demonstrated benefits of *L4-33K/22K*, located in this sequence, were an argument in favour of using the 5'-UTR. Therefore, one of the primary objectives of this experiment was to determine whether the 5' UTR is favourable for rAAV production when E2A expression is controlled by a CMV promoter. Consequently, plasmids were created with CMV promoters based on Helper 1.0 (containing *E2A*'s 5'-UTR), 1.1 (without *E2A* 5'-UTR), and 2.0 with and without *E2A* 5' UTR. The obtained VG titres of transfection with these plasmids are shown in Figure 29.

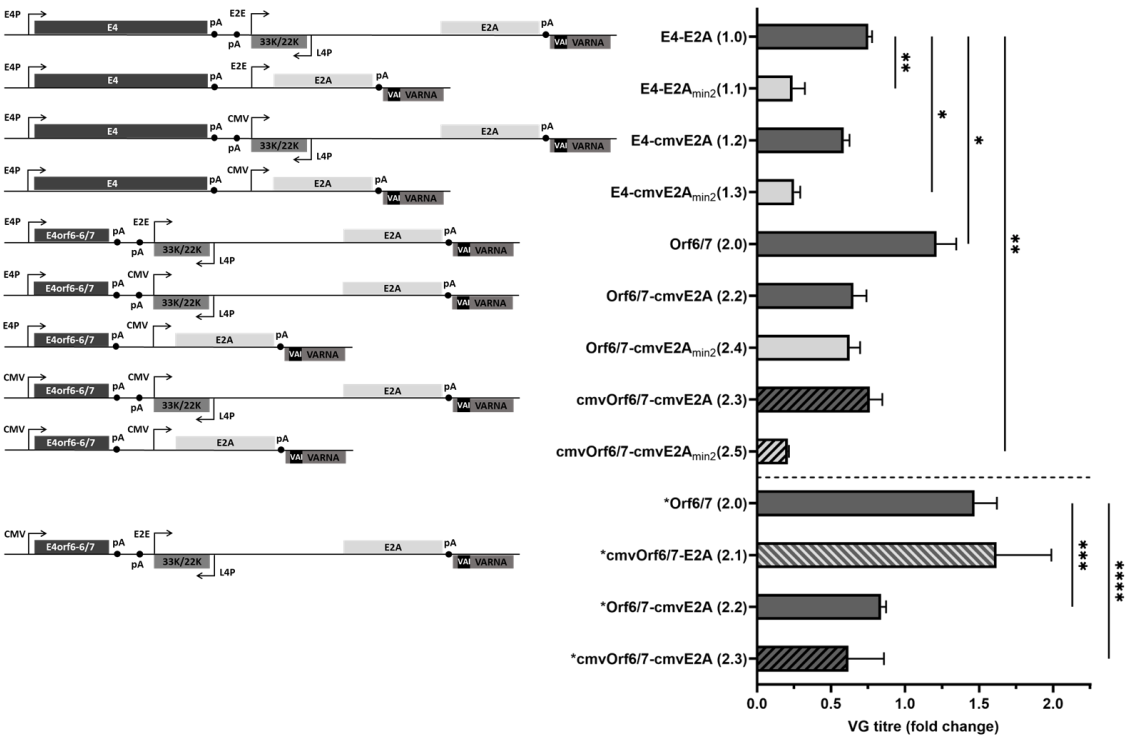


Figure 29: Helper plasmids containing CMV promoters to control *E2A* an *E4* transcription. Schematic depiction of the Helper plasmids next to associated VG titre fold changes compared to helper 0.1. Arrows denote promoters, black dots poly A sequences. Genes, 5' and 3'-UTRs, but not promoters and pAs are drawn to approximate scale. Data shown as means with error bars indicating standard deviations of three biological replicates. Constructs with *E2A* 5'-UTR as dark grey bars, without in light grey, constructs that contain a CMV promoter instead of E4P for *E4* expression control with hatched pattern. Data of bottom four bars (*) originating from separate experiment with 25% lowered total DNA concentration. Statistical analysis was performed as an ordinary one-way ANOVA, comparisons displayed to Helper 1.0 and Helper 2.0, respectively.

As previously observed, rAAV8 VG titres decreased with the exclusion of the *E2A* 5'-UTR (Helper 1.1) when the endogenous E2E promoters was used. Changing the promoter to a CMV promoter had no different effect. Additionally, titres were not increased with the CMV promoter driving DBP expression with and without the 5'-UTR (Helper 1.2 and 1.3), but remained similar to the E2E driven versions of the constructs. This result was not unexpected, as the slight 1.2-fold increase with the CMV promoter in the 5-plasmid transfection was not significant (one-way ANOVA p=0.2886). Nevertheless, it raises the question of whether further VG titre increases could have been achieved with the 5-plasmid system using the *L4-33K/22K* including *E2A* 5'-UTR. Combining CMV driven E2A transcription with the *E4orf6-6/7* containing Helper 2.0 containing the *E2A* 5'-UTR decreased titres to levels similar to the full *E4* containing Helper plasmids 1.0 and 1.1. However, no further reduction was observed when the *L4-33K/22K* containing 5'-UTR sequence of *E2A* was removed. It was assumed that based on E4orf6/7's E2E promoting function titres of Helper 2.2 being similar to the control could mean that the benefits of excluding the other *E4orfs* mainly increased *E4orf6/7* and consequently *E2A* transcript levels, but

not E4orf6 levels. The alternative explanation would be that if E4orf6 levels were increased too, increased E4orf6 levels did not influence rAAV titres. Evidence for the latter hypothesis are the minor differences in VG titre of Helper 2.1 in comparison to Helper 2.0. However, a later analysis of mRNA transcripts revealed that E4 transcript levels were similar between Helper 1.0 and 2.0, whereas E2A transcripts of Helper 2.0 were actually reduced at 72h post transfection (see Figure 40).

The addition of a second CMV promoter, controlling *E4orf6-6/7*, did not change VG titres significantly, supporting the hypothesis of increased E4or6 levels having a neglectable influence on rAAV titres. In another experiment with the same plasmids, but less total DNA transfected, rAAV titres with Helper 2.3 were even decreased compared to 2.2. The additional CMV might result in transcription competition for other components like the *cap* gene and *E2A*. The fact that the double CMV system was performing best in the 5-plasmid system but poorly in the 3-plasmid system indicates that a problem occurs when two of the strong CMV promoters are placed on the same plasmid. Less polymerase availability for *E2A* transcription is therefore likely the reason for these decreased titres. However, the same could be the case for the RNA pol III driven VAI, due to steric hindrance or other competition related effects reducing transcriptional activity.

The decrease without the *E2A* 5'-UTR was also observed with the double CMV construct (Helper 2.5), but not with Helper 2.4, containing E4P driven *E4orf6-6/7* and CMV driven *E2A* sequences. Helper 2.4 seems to be an exception that cannot be explained at this point. Nevertheless, all other comparisons between similar plasmids with and without the 5'-UTR demonstrated the already described beneficial effect, most plausibly generated by the expression of L4-33K/22K.

Highest titres were achieved with the Helper 2.1, utilising a CMV promoter for *E4orf6-6/7* transcription and the regular E2E promoter with its introns for *E2A* transcription. Like in the 5-plasmid approach titres of the *E4orf6-6/7* constructs with CMV and E4P promoter did not vary significantly (p=0.9981 / p=0.9849). Measurements of E4P promoter strength through eGFP expression demonstrated that the CMV promoter is about 1.3-times as strong as the E4 promoter, while it is 3-times as strong as the E2E in HEK293 cells (appendix 8.2.4.1). This might explain the little changes in VG titres with the CMV promoter controlling expression of E4 or E4orf6-6/7, compared to CMV-induced changes for E2A expression. Another explanation would be that increased E4 protein expression or *E4* products in general have less of an impact on AAV production than the DBP. The importance of *E2A* on rAAV production lies on one side in its stabilisation of ssDNA and the connected role in AAV genome replication and on the other side in its promoter transactivation, enhancing E1A 6-fold, p5 27-fold and its own E2E promoter 13-fold

(L. S. Chang & Shenk, 1990; Stracker et al., 2004; Ward et al., 1998). Subsequent experiments should, therefore, focus on ideal E2A transcription levels and guarantee a sufficient amount of *E2A* transcription, since both the *E4orf6/7* change (Helper 2.0) and the pE2Acmv indicate that increased *E2A* transcription might improve rAAV production. Although it seems logical that the *E4orf6-6/7* effect is mitigated by the exchange of E2E for a CMV promoter, the exhibited VG titre decrease was unexpected. This raises the question of what made the *E4orf6-6/7* E2E combination better than the potentially even higher expression driven by CMV. The intact regulation, similar as in wildtype AdV/AAV proliferation, could mean a better concentration and time dependent fine-tuning of expression levels. In contrast, the strong constitutive expression with the CMV promoter might produce too high DBP levels at the wrong time. Additionally, strong CMV promoter transcription in the opposite direction of *L4-33K/22K* could also cause a loss of function or strong reduction of this for rAAV production beneficial protein. However, the reduction in titre by the elimination of the 5'-UTR from Helper 2.3 to 2.5 and thereby complete loss of *L4-33K/22K* goes against this theory. Lastly, CMV driven DBP expression, which is likely 3-times higher than the previous E2E expression might also be too strong in general and cause cytotoxic stress on the cell (Johari et al., 2022; Klessig et al., 1984). Therefore, optimising the required promoter strength for *E2A*, but also *E4*, promises further increases in rAAV VG titres with the used triple plasmid system.

6.2.5 Summary

- Varying Helper and Rep/Cap plasmid ratios, as initial attempt of altering component stoichiometries of plasmids, only marginally effects rAAV titres
    - Increasing ITR amounts without influence on VG titres
    - Current plasmid ratio of 2:1:0.1 likely already near optimum
- Rep/Cap and Helper plasmids split into multiple plasmids to understand ideal genetic component abundance for rAAV production
    - Lowered *rep* amounts seem to be favourable, confirming results of other groups
    - Split Rep/Cap plasmid lowered titres significantly, challenging the relevance of 6-plasmid approach results
    - Split Helper plasmid is functional
- Testing individual Helper genes' essentiality or benefit using split Helper plasmid system
    - No AdV Helper gene promotes rAAV replication on its own, regardless of Helper gene promoters
    - *E2A* is found to be the only indispensable Helper gene; *E4* and *VA RNA* are required for optimal function

    - Use of CMV promoter for *E2A* and *E4* showed increased VG titres
        - CMV driving *E2A* only led to slight increase
        - CMV driving both genes led to strongly increased VG titres
    - 5-plasmid system results for sequential *E4* ORF dissection confirm 3-plasmid system results
        - Confirmation of *E4orf6* as essential part of *E4* for AAV production
        - Use of CMV promoter for *E4orf* subsets enhances positive and negative effects of ORFs, but does not increase rAAV production
- CMV driven expression of *E2A* and *E4*/*E4orf6-6/7* in 3-plasmid system were tested to determine requirement for transcription activity enhancement of both genes
    - Only marginally improved titre with CMV driving *E4* transcription
    - Exclusion of *E2A* 5' UTR led to decreased rAAV8 VG titres regardless of promoter change
    - No titre improvements with CMV driving *E2A*, regardless of position
    - CMV for *E2A* might make E4orf6/7 obsolete and reduces titre
    - Double CMV not boosting titre as in 5-plasmid system, but diminishing it, likely due to promoter interference and TF competition

6.2.6 Conclusion

To establish a general of favourable expression levels of AAV and AdV Helper genes within the system, with the aim of achieving enhanced rAAV titres, a comprehensive investigation was undertaken. Despite initial expectations, traditional DOE-like experiments involving the 3- and 6-plasmid systems did not provide the desired insights. Varying the Helper and Rep/Cap plasmid ratios yielded only marginal impacts on rAAV titres, implying that the current 2:1:0.1 plasmid ratio might already be near the optimum, precluding drastic improvements as well as valuable insights into component requirements. Furthermore, the 3-plasmid system only allowed titration of gene combinations, necessitating a more intricate approach for the observation of individual gene abundance requirements.

Dividing the Rep/Cap and Helper plasmids into discrete components was envisioned to yield a more nuanced understanding by allowing for individual gene variation on separate plasmids. Although, the 6-plasmid strategy suggested that reduced rep levels were advantageous, consistent with prior studies, the split Rep/Cap plasmid configuration substantially diminished titres, casting doubt on the applicability of the 6-plasmid system and the relevance of its results.

Transfection results with the split Helper plasmid and individual components of it, in combination with the intact Rep/Cap plasmid, aligned with Matsushita et al. (1998) findings on the essentiality of AdV Helper genes. Evidently, no single AdV Helper gene autonomously facilitated rAAV replication. *E2A* emerged as the only indispensable Helper gene, while *E4* and *VA RNA* contribute to optimal function. The results underline specific importance of the DBP for enhanced rAAV production. Its involvement in ITR/GOI sequence replication could be of great importance, but also *E2A*'s function as a transcription activator of the AAV genes might be pivotal for adequate transcription of the virus' genes. Consequently, the results of the 5-plasmid model suggested to increase DBP expression to improve rAAV production.

Employing CMV promoters for *E2A* and *E4* in the 5-plasmid system demonstrated mild VG titre increases with cmvE2A, while substantial improvements were achieved with combined CMV-driven E2A and E4 plasmids. This highlighted the potential of promoter substitutions within the Helper plasmid to elevate rAAV titres, alongside the possibility of required heightened expression of the two Helper genes. However, contrasting outcomes emerged from the 3-plasmid system, revealing diminished rAAV titres when two CMV promoters were employed, raising concerns about transcription factor competition and promoter interference. CMV-driven *E2A* expression seemed to compromise E4orf6/7's function, and CMV-regulated *E2A* transcription failed to elevate rAAV VG titres. The omission of *E2A* 5' UTR and the resultant *L4-33K/22K* loss led to decreased titres. Initial endeavours involving the introduction of heterologous promoters into the Helper plasmid showcased the general applicability of foreign promoters for the Helper genes. However, these attempts also brought to light that swapping promoters constitutes significant alterations to the system, necessitating methodical plasmid engineering to ensure compatibility and achieve enhanced outcomes.

In conclusion, the data regarding CMV-driven Helper genes suggests that introducing heterologous Helper gene promoters, such as CMV or alternative ones, could enhance rAAV production potential. However, substantial modifications to the plasmid architecture might be necessary for optimal functionality. Furthermore, adjustments of promoter strengths are imperative for optimal expression of the two mRNA-generating Helper genes. Overall, the results highlight the complexity of rAAV production with transient transfection systems, underscoring the intricate interplay between gene expression and plasmid design.

6.3 Combination of Engineered Rep/Cap and Helper Plasmids - The Design of Controllable Recombinant AAV Expression Systems for Enhanced Vector Production

Next to the extensive work on the Helper plasmid, it was also aimed to improve the Rep/Cap plasmid to create a new and controllable rAAV transient triple transfection system. At the start of the project the work was divided into a Helper-focused package, led by me, and a Rep/Cap package, with Dr. Yusuf Johari in charge. Most planning, experimental work and intellectual property regarding the Rep/Cap plasmid belong to Dr. Johari. However, until the respective departures of the two post-doctoral researchers, the project was in many regards a group effort by all three University of Sheffield-based members, namely Dr. Joseph Scarrott, Dr. Yusuf Johari, and me. Several experiments with the Rep/Cap plasmid were executed by me or in cooperation with Dr. Johari. Furthermore, after he left the project, the Rep/Cap work was continued by me. Finally, a joined draft of a manuscript for publication was compiled by the two of us as equivalent coauthors, combining Rep/Cap and Helper plasmid work. Given these circumstances, a brief description of the performed Rep/Cap work that was not or only partially performed by me will be given. The full extend of the work can be viewed in 8.2.4, exhibiting the coauthored publication.

Molecular Design of Controllable Recombinant AAV Expression Systems for Enhanced Vector Production

- Yusuf B. Johari*, Thilo H. Pohle*, Joseph M. Scarrott, Ping Liu, Ayda Mayer, David C. James

The Rep/Cap work aimed, similar to the Helper work, to develop a controllable Rep/Cap plasmid to efficiently produce rAAVs for the use in gene therapy. Previous research showed that unregulated overexpression of Rep78/68 inhibited rAAV production, while reduced levels of Rep78/68 improved rAAV titres. To address this issue and create a controllable Rep/Cap plasmid with potentially optimised component stoichiometries, instrumental modifications to the Rep/Cap plasmid were made. The most crucial of these changes are listed in the following.

1. Truncation of the p5 promoter and introduction of p5 downstream of cap: The p5 promoter was truncated to attenuate the expression of large Rep proteins, addressing the issue of unregulated overexpression of Rep78/68, which inhibits rAAV production. Moreover, a full p5 downstream of cap was required, potentially to retain the expression from p19 and p40. These changes were similarly reported by Xiao et al. (1998).

2. Separation of *rep* and *cap* genes: The *rep* and *cap* genes were split by cloning p40 and *cap* downstream of the *rep* gene's stop codons. The separation required mutations of the p40 promoter, TATA box and start codon of remaining partial *cap* 5' sequence overlapping *rep*'s 3' sequence to prevent expression of truncated viral gene products.
3. Introduction of restriction sites up- and downstream of *rep* and *cap* genes to enable simplification of further changes, including promoter replacements, to unlock improved controllability.
4. The 5' *cap* intron turned out to be crucial for efficient rAAV production and needed to be reintroduced.
5. Replacement of the p40 promoter with the heterologous CMV promoter driving *cap* expression, was needed to enable rAAV titres similar to the parental Rep/Cap plasmid (v1.2)
6. A poly adenylation signal sequence allowed VG titres 1.2-fold increased to the control (Rep/Cap plasmid v3.7).
7. Despite reports of Rep78 being dispensable, it was shown to be essential in present system.
8. An additional Rep52/40 copy with an optimised Kozak sequence was introduced to increase Rep52/40 amounts and reach a higher small to large Rep ratio for enhanced genome packaging without increased cytotoxicity.

These modifications collectively contributed to the development of a controllable Rep/Cap plasmid system, enabling improved rAAV production for potential gene therapy applications. Key constructs of the work and resulting rAAV8 VG titres of transfections with these are shown in Figure 30.

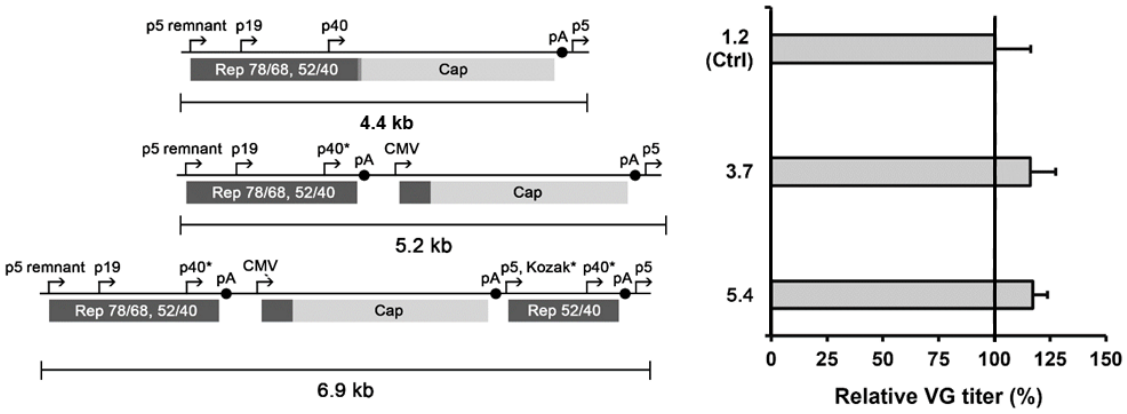


Figure 30: Rep/Cap plasmid constructs most relevant for this work. Schematic depiction of Rep/Cap plasmid constructs show all components are drawn to approximate scale. Replication (Rep) and capsid (Cap) open reading frames are indicated, arrows denote promoters, asterisks denote sequence mutations, black circles poly As. HEK293 cells were triple transfected with each Rep/Cap plasmid, Helper 0.1 (pADΔF6) plasmid and the ITR/GOI plasmid at 1:2:0.1 weight ratio. rAAV8 crude viral genome (VG) titres were analysed 72 h post-transfection, expressed as a relative percentage compared to Rep/Cap plasmid 1.2. Data shown are the mean with standard deviation of three independent biological replicates.

The paper draft also includes the modifications and findings regarding the *E2A* and *E4* genes of the Helper plasmid, presented in chapter 6.1. Additionally, the new Rep/Cap plasmids and Helper plasmids were combined to test their compatibility and compile a first version of a triple plasmid system for rAAV production that has the potential for adjustment and controllability of specific components. The results are fully described and analysed in the paper draft. As these experiments were also a larger part of my own work, a description and graphical display of the results will also be included here.

Of all constructed Rep/Cap plasmids, highest titres were achieved with the Rep/Cap plasmids v3.7 and v5.4, which resemble one another apart from the added Rep52/40 copy downstream of *cap* in v5.4. Helper plasmid 2.0 (E4orf6-6/7) performed best of all the Helper plasmids tested in the work of the paper draft and enhanced the titre to 185% of the E4 control titre. Therefore, the 2.0 Helper plasmid was to be combined with the two best Rep/Cap plasmids. Unfortunately, both Rep/Cap versions were incompatible with Helper 2.0 as seen by VG titres drops to 68% (v3.7) and 74% (v5.4) compared to Helper 2.0 with the control Rep/Cap plasmid v1.2. It was postulated that the use of the strong and highly complex CMV promoter in Rep/Cap v3.7 and v5.4 titrated away the limited pool of available TF molecules from the Helper promoters E4P and E2E. Consequently, the two Rep/Cap plasmids were also tested with versions of the E4orf6-6/7 Helper with CMV

promoters in exchange of the endogenous promoters to enhance transcription of the *E4* and *E2A* Helper components.

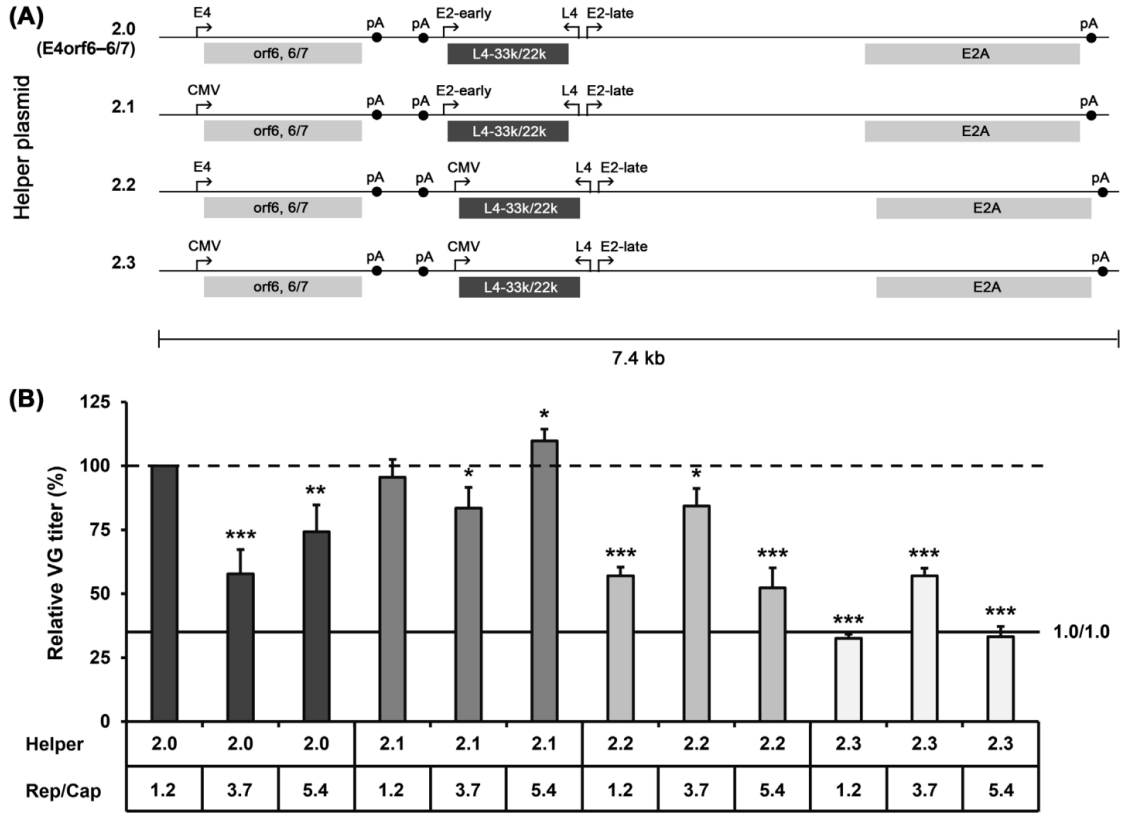


Figure 31: Evaluation of engineered Helper and Rep/Cap plasmid combinations for rAAV8 production. (A) Schematic depiction of the Helper plasmid constructs (*VA RNA* not indicated). Arrows denote promoters, black circles poly A signal sequences. All components are drawn to approximate scale. (B) rAAV8 crude viral genome VG analysed 72 h post-transfection, expressed as a relative percentage compared to the Helper 2.0 and Rep/Cap 1.2 plasmid combination. Data shown as mean with error bars displaying standard deviation. Data were analysed using unpaired Student's t-test with respect to the Helper 2.0 and Rep/Cap 1.2 plasmid combination.

Employing the CMV promoter for *E4orf6-6/7* transcription (Helper 2.1) restored rAAV titres comparable to those achieved with Helper 2.0 and Rep/Cap v1.2 (83-110%). However, substituting the E2E promoter with the CMV promoter (Helper 2.2) led to reduced titres, particularly when combined with Rep/Cap v1.2 and v5.4. Further decreases in rAAV VG titres were observed with Helper 2.3, utilizing the CMV promoter to drive both E4orf6–6/7 and E2A expression. These findings were consistent with our idea of competing TF binding sites, employing a system with a total of four CMV promoters, all competing for the same pool of TFs. The data also showed that Rep/Cap v5.4 with Helper 2.0 or 2.1 yielded higher titres compared to Rep/Cap 3.7. To explore this observation, a sub-panel of different Rep/Cap and Helper combinations was selected for quantification of fully assembled, intact capsids that can be compared to measured VG titres of the same samples to determine the ratio of full to empty particles (Figure 32).

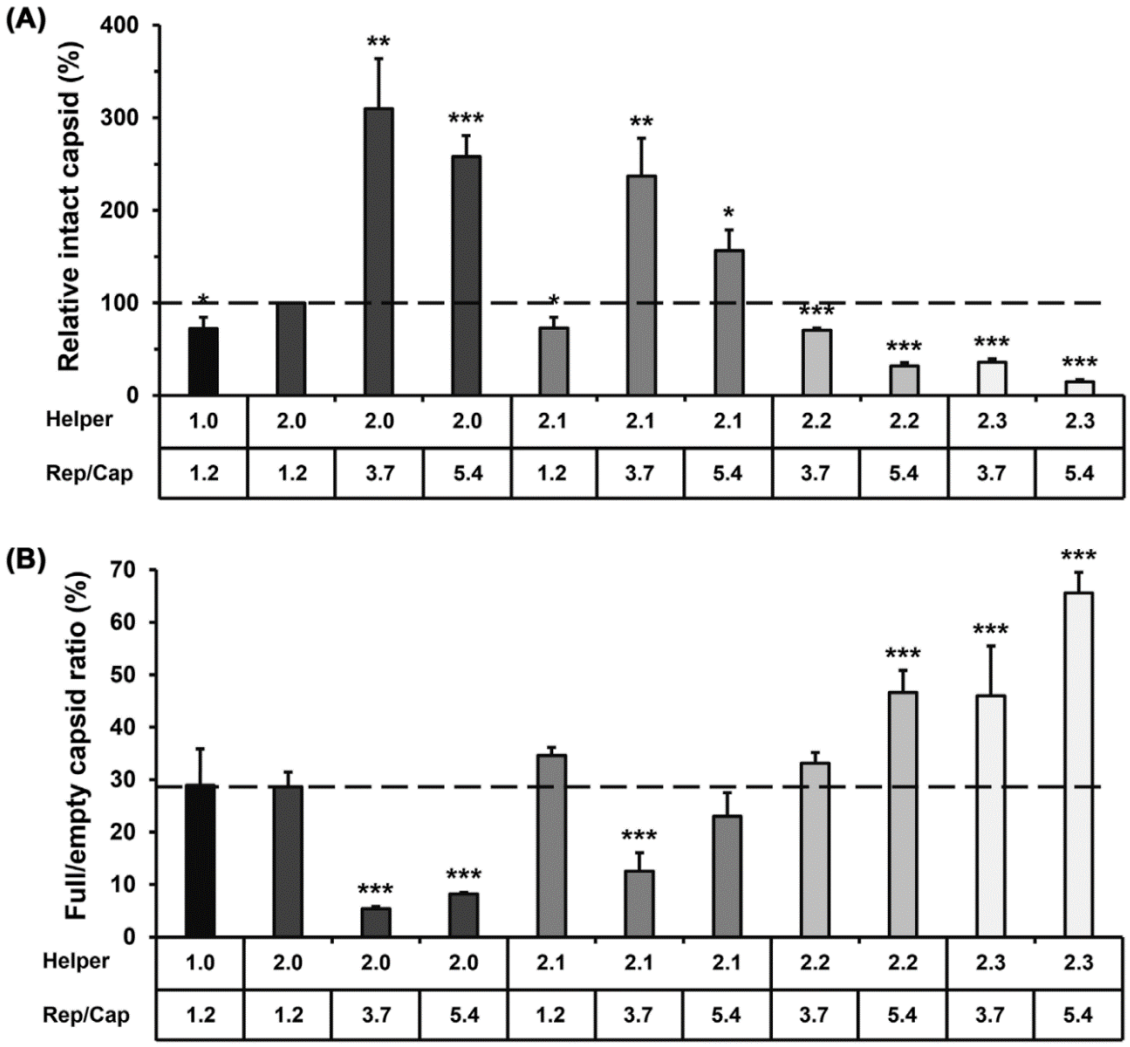


Figure 32: Determination of product quality of the engineered Helper and Rep/Cap plasmids for rAAV8 production. (A) Intact capsids were quantified at 72 h post-transfection using rAAV8-specific capsid ELISA and expressed as a relative percentage compared to the Helper 0.1 and Rep/Cap 1.2 plasmid combination. (B) The full/empty capsid ratio was calculated from the measured intact capsids in A and its VG titre. Data shown as means with standard deviation as error bars of three independent biological replicates. Data were analysed using unpaired Student's t-test with respect to the Helper 2.0 and Rep/Cap 1.2 plasmid combination.

The analysis of total capsid concentrations revealed that Rep/Cap plasmids v3.7 and 5.4 boosted capsid production 3.1 to 3.5-fold with the Helpers 0.1, 2.0 and 2.1, respectively, compared to the same helpers in combination with Rep/Cap v1.2. However, since VG titres did not increase proportionally, this meant that full/empty ratios decreased to as low as 5.3% (v3.7) and 8.0% (v5.4). Consequently, the CMV promoter in both new Rep/Cap plasmids lead to a large increase in empty capsids with these three Helper plasmids. Interestingly, Helper 2.1 showed less capsid production than 2.0. In alignment with the promoter and TF competition theory, it is possible that the CMV promoter employed for E4orf6-6/7 expression tuned down VP production, resulting in higher full/empty ratios. The fact that also rAAV VG titre were increased with this helper construct and the otherwise mostly empty capsid

producing Rep/Cap plasmids v3.7 and v5.4, highlights the importance of adjusted component stoichiometries. Further, it might demonstrate that also a time component is important, as shown in the modelling experiments by Nguyen et al. (2021). Although Helper 2.0 facilitated more capsids with v3.7 and v5.4, these could not be filled with genomes in time. Additionally, capsid production with the CMV promoter was presumably consistently high until the harvest 72 h post transfection. But according Nguyen et al. (2021) genome replication is not and strongly declines after about 24 h, which would mean that for the next 48 h mostly empty capsids were produced. Measurement of full and empty capsids at different time points with the different promoters would be great to analyse these events and kinetics.

Likely, a significant factor affecting recombinant production of multiple component expression systems with strong promoters is cellular resource allocation. The highly abundant VPs expressed from the CMV promoter were potentially efficiently assembled, as assembly is a fast, non-limiting step in AAV proliferation, but due to the extensive allocation of resources for these processes, specifically the VP expression, rAAV genome replication, the expression of helper functions, or Rep might have been compromised. The use of a CMV promoter for *E4orf6-6/7* potentially alleviated this effect. The decrease in total capsids observed in transfections with the Helper plasmids 2.2 and 2.3 (the E2Acmv or double CMV, respectively) support this hypothesis. However, it remains unclear why the use of a CMV promoter for *E2A* decreased capsid amounts, whereas the cmvE4orf6-6/7 construct did not exhibit the same effect. Lower VG titres and total capsid titres were achieved by the combinations of Rep/Cap v3.7 and v5.4 with the Helpers 2.2 and 2.3. Part of the titre decrease could have been caused by the potentially cytostatic or cytotoxic effects of high DBP levels resulting from replacement of the self-regulating loop mechanisms of *E4orf6/7* and *E2A* with high and uncontrolled CMV-driven *E2A* transcription (Klessig et al., 1984). However, cell concentration/viability data showed no differences between the endogenous E2E and CMV promoter-driven *E2A* constructs (data not shown). It is speculated that the detrimental effects of high DBP level in HEK293 cells originated in negative regulation of capsid expression rather than direct exertion of cytotoxicity on the host cells. Another contributing factor to the highly increased full to empty capsid ratios (up to 66%) could be attributed to the increase in DBP, potentially leading to increased Rep levels and ITR/GOI replication. DBP itself, and increased large Rep levels due to DBP transactivation, could support large Rep-ITR mediated rAAV genome replication. Through transcription activation cascades elevated small Rep levels could improve ITR genome packaging. Comparing full/empty ratios of Rep/Cap plasmids v3.7 and v5.4, which are consistently higher with plasmid v5.4 containing a second Rep52/40 copy, also

demonstrates a higher packaging rate of the single-stranded DNA contained between the ITRs due to an increase in small Rep amounts.

The high empty capsid titres achieved with the two new Rep/Cap plasmids are currently not ideal, particularly from a downstream purification point of view. Furthermore, questions should be raised regarding the VP composition. Studies in the baculovirus system suffered from unfavourable VP stoichiometries, likely caused by the CMV promoter used for *cap* transcription, resulting in poor infectivity of the manufactured rAAVs (Mader et al., 2013; Mietzsch et al., 2015; Rumachik et al., 2020). Western blot analysis for VP ratios and transduction experiments are suggested for future quality controls and experiments with the engineered Rep/Cap plasmids.

Taken together virus genome titre increases of the new Rep/Cap (v3.7/v5.4) and Helper (2.0/2.1) were not additive, but the observed VG titre decrease of the combination of 2.0 and v3.7/v5.4 could be reverted by the use of a CMV promoter for *E4orf6-6/7*. The new systems' VG titres were 1.8-fold higher than the ones of the original system, but >3-fold higher capsid titres suggest an even greater potential with an optimised balance of component amounts. The present experiments, which involved the addition of a second copy of Rep52/40, showcased that packaging rates can be increased with molecular biological tools. Future fine-tuning of component amounts was enabled with the new Rep/Cap and Helper plasmids that were created during this work. Rather than using the constitutive, strong, and uncontrollable transcription from the diverse CMV promoter, synthetic promoters could prove to be well-suited tools for achieving precise control over the system. Notably, the achieved full/empty particle ratios of the control and generally with the regular Rep/Cap plasmid v1.2 are similar to reports from other groups, such as the study Mietzsch et al. (2021), which reported 20% full virions with the AAV8 *cap* and AAV2 *rep* combination that was also used in the present experiments. Their research also presents an alternative strategy to increase packaging by using alternative Rep proteins, yielding up to 43% full particles in rAAV8. Therefore, the combination of better-suited Rep proteins, respectively to the used serotype, and an optimised expression profile of Rep and VPs with synthetic promoters could possibly further boost packaging and VG titres.

6.3.1 Transfection Process Adaptation for Increased Consistency

In the course of the research presented so far, experimental day-to-day variation of transient transfections was an accompanying issue. After the Rep/Cap and Helper combination experiments, which were thought to stay cohesive with the rest of the work for the manuscript draft, it was decided to optimise the experimental setup for

higher consistency between biological replicates. Part of this work will be presented briefly in the following.

In a first experiment, the total plasmid amount per million cells was lowered, as a saturation by the regularly used 0.63 µg/10^6 cells was feared. This could lead to greater inhomogeneity and variation between transfections. Furthermore, reducing plasmid DNA usage is beneficial in rAAV production due to its immense contribution to upstream costs, particularly in cGMP processes (Cameau et al., 2019; Merten, 2016). Transfections with the initial plasmid system (Helper 0.1+Rep/Cap v1.2) were carried out as usual in 24-well plates with constant DNA to PEI ratio of 1:1.75 but varying total DNA concentrations.

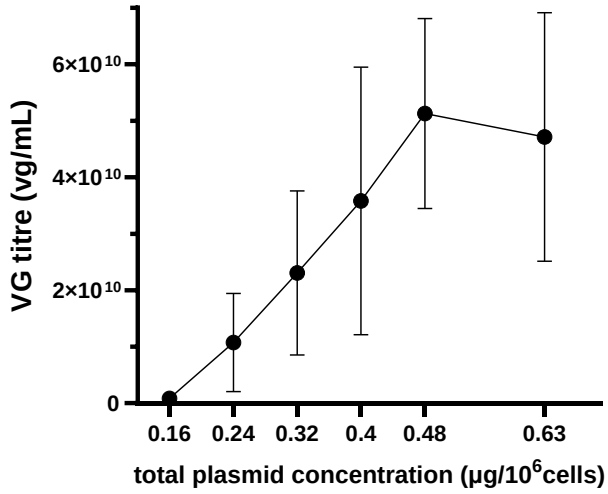


Figure 33: rAAV VG titre as function of increasing total plasmid concentration. Data points measured in independent biological triplicates (black dots) display means, error bars standard deviations.

The resulting VG titres, as shown in Figure 33 were highest with transfections of 0.48 µg DNA per 10^6 cells. Notably, the lowest DNA concentration still yielded about 9×10^8 vg/mL, which would be better visible on a logarithmic scaled y-axis. The chosen decimal scalation, however, is a better display of the linear increase of VG titre with the total DNA load up to its maximum at 0.48 µg DNA per 10^6 cells. The previously used DNA amount seemed to be too much DNA for the cells. Measurement of viable cell densities could have demonstrated if the lower VG titre was correlating with a higher cell mortality. Differences between biological replicates, transfected on different days, were high in all cases, with relative errors being slightly smaller at 0.48 µg compared to 0.4 and 0.63. Although, a sample size of n=5 is too small to give any meaning to this, the three-quarter lower DNA concentration did at least not seem to increase variability. Since it also displayed a marginally higher titre, meaning potential cost reduction in industrial production

processes, it turned out to be the preferred total DNA concentration for future transfections.

Further concerns referring to the transfection method regarded the polyplex formation of DNA and PEIpro, the low ITR/GOI plasmid amount relative to the other two plasmids and inhomogeneity of transfections with cells in different sub cultivation passages. Experiments with different polyplexing times and methods revealed that the usual method with a 10 to 15 minute polyplex formation period was still the preferable choice (see appendix Figure 50). The other two points were addressed with an experimental setup transfecting different ratios of ITR/GOI plasmid (1-fold, 5-fold and 10-fold) with the three helper plasmids 0.1, 2.0 and 2.2 as technical triplicates of each of the biological triplicates in 24-well plates. The transient transfections were carried out at 0.48 µg/10^6 cells in 24-well plates. A sub passaged culture of HEK293 cells from a single cryovial represented each of the three biological replicates and was transfected at passages three, five and seven.

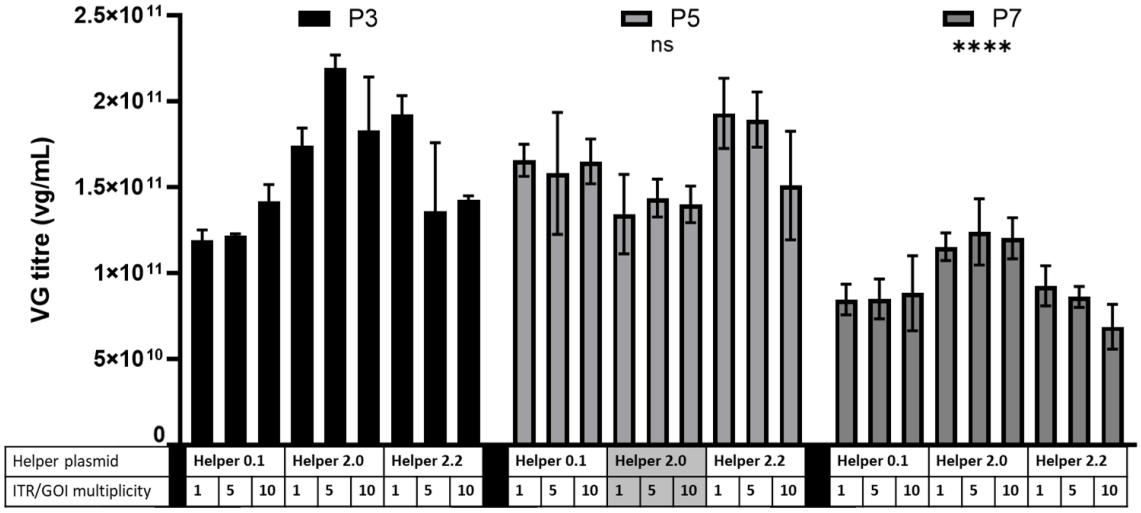


Figure 34: Consistency of transfections and resulting VG titres with different helper plasmids (0.1, 2.0 and 2.2) and different ITR/GOI plasmid amounts (1-, 5-, and 10-fold of the regular amount of the plasmid ratio 1:2:0.1) transfected at different sub cultivation passages (P3, P5, P7, passaging was performed in a 3-day rhythm). Bars represent three biological replicates, each transfected and analysed in technical triplicates of which the determined VG tires were subsequently pooled. Error bars display standard deviation of the biological replicates. Statistical analysis was performed as an ordinary one-way ANOVA on the means of each cultivation passage, pairwise comparison displayed to the mean of P3.

The bars in Figure 34 represent the average VG titres of biological replicates and the error bars represent the standard deviation of these, whereas VG titre measurements of the technical replicates (transfections of the same cell pool in different wells) were averaged beforehand. Therefore, the results in Figure 34 show clear differences between the biological replicates regardless of helper plasmid and ITR amount used. Interestingly, the variations were larger between different passage numbers and transfection days of the same biological replicate than

between the biological replicates on the same day. Surprisingly, no specific trend was observed regarding variation with different ITR amounts. However, the at that time best and most used Helper plasmid, Helper 2.0, showed slightly higher titres at 5x ITR/GOI plasmid amounts. This increase in ITR/GOI amount changed the concentrations of the other two plasmids only marginally but prevented a possible inhomogeneous distribution of the ITR/GOI plasmid in polyplexes. Such inhomogeneities could mean that some cells that take up DNA have no or only small amounts of the ITR/GOI plasmid, insufficient for specific high titre rAAV production. Considering that likely not more than 5% of the initially applied plasmid is taken up inside the cells with PEI and less than 1% gets into the nucleus, it is surprising that the very low ITR/GOI amount that was previously chosen was not limiting and didn't cause more inconsistency (Nguyen et al., 2021; Pollard et al., 1998). Furthermore, these low triple transfection numbers are also confirmed by a study of Dash et al. (2022), which demonstrated that less than 10% of all cells of transfections for rAAV production do produce virus. This is a different aspect of rAAV production optimisation which needs to be improved, but it was not supposed to be addressed in this work, as the focus of this work layed on the molecular engineering of the plasmid system.

The increased ITR amount was chosen for upcoming transfections. It was additionally decided to carry out transfection at cell sub cultivation passage number three, whenever possible, since the results seemed to show a trend of declining VG titres with higher passage numbers. This trend of declining VG titres with higher passage numbers was also observed by others, although Grieger et al. (2016) suggested a much longer usage of cells up to passage 30-40. Notably, the volumetric rAAV virus genome titres and VG titre trends of plasmid constructs were consistent between different passages, except for Helper 2.0 in passage 5 (labelling marked in grey). This inconsistency was attributed to a different plasmid preparation stock used for this specific transfection, even though no apparent issues were detected with the plasmid purification process. Consequently, from that point onwards, attention was paid to perform transfections of the same experiment but different replicates always with the same plasmid stocks.

Lastly, the transfection in 24-well plates was investigated, because it was hypothesised that transfections with larger culture volumes in 50 mL TubeSpin reactors or Erlenmeyer shake flaks would be less prone to variation and errors of different origins. General handling of larger volumes should be less variable and the transfection of more cells at a time should lead to greater uniformity. Moreover, it was speculated that larger vessels might provide better cultivation conditions, leading to higher cell densities during the three days post transfection and

consequently higher VG titres. This speculation was particularly relevant for the new HEK293 cell line introduced to the project at the time, which showed slightly higher clumping. Transfections were carried out with three different helper plasmids (0.1, 2.0 and 2.2) in 50 mL TubeSpin reactors with 5 mL culture volume and 24-well plates with their usual 0.7 mL culture volume.

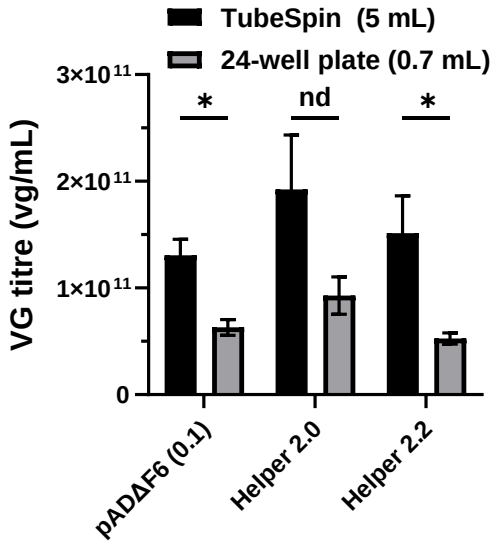


Figure 35: Comparison of rAAV8 VG titres from transfections of different plasmids in either 50 mL TubeSpin reactors (5 mL culture volume, black) or 24-well plates (0.7 mL culture volume, grey). Data shown as means with error bars as standard deviation of three individual biological replicates.

The results of the investigation showed that the measured rAAV8 VG titres were consistently higher with transfections in the 50 mL TubeSpin reactors. Calculated t-tests demonstrated significantly higher titres for Helper 0.1 and 2.2, primarily due to relatively large differences among the three biological replicates with Helper 2.0. Although higher VG titres could be achieved in 50 mL TubeSpin reactors, the relative errors were similar or even higher compared to 24-well plates (between 11.5% and 26.5% for 50 mL TubeSpin reactors and 11.5% and 18.8% in 24-well plates). To investigate the reason for the higher titres with the larger culture volume, additional transfections were performed with Helper 0.1 as technical and biological triplicates in both culture vessels. For these transfections VCDs were measured at the harvest (three days post transfection) and transfection efficiencies measured through the expressed GFP from the ITR/GOI plasmid were determined 48 h post transfection (Figure 36). In this experiment the 50 mL TubeSpin reactor transfections again yielded higher VG titres on average, but no significance could be attested. In this case, a much high relative error of the 24-well plate titres was determined (48.8%). Whereas the 50 mL TubeSpin reactor transfections only showed a relative error of 17.1%. Comparing these results to the previously described experiment (Figure 35), the general higher variability of the plate transfections becomes obvious.

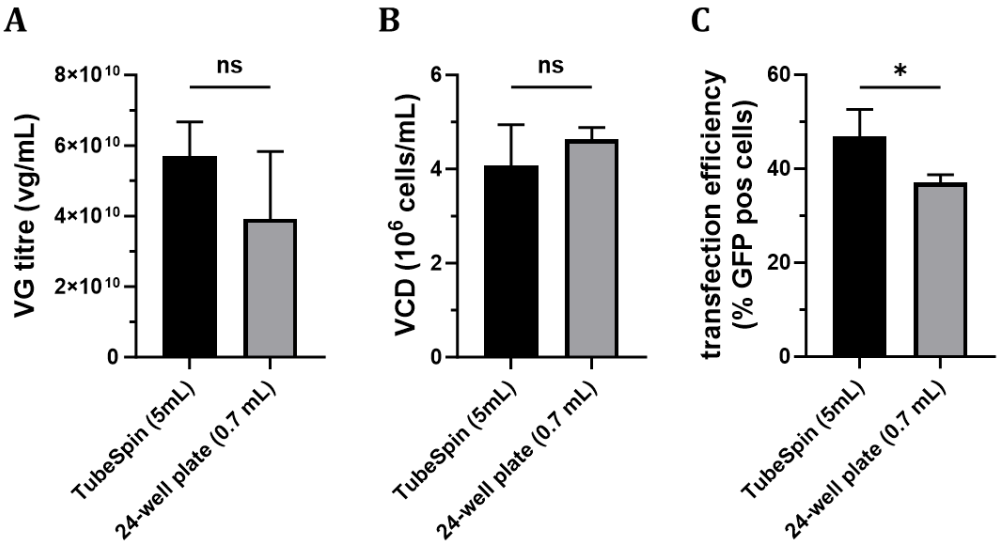


Figure 36: Comparison of transfections in 50 mL TubeSpin reactors (5 mL culture volume, black) and 24-well plates (0.7 mL culture volume, grey). rAAV8 VG titres (A) and viable cell densities (B) were measured at the time of cell harvest 72 h post transfection, transfection efficiency as GFP positive cells 48 h post transfection. Data shown as means with error bars as standard deviation of three individual biological replicates, with pooled data of three technical triplicate transections for each biological replicate. Unpaired t-tests were performed for statistical analysis.

The results indicated that the differences in rAAV titres were likely a result of improved transfection efficiency, which was significantly increased from 37±5.4% to 46±7.4% in the 50 mL TubeSpin reactor transfections. Differences in viable cell densities at the time of cell harvest could not be detected (Figure 36 B). It is hypothesised that better mixing in the larger vessel during transfection itself and cultivation, particularly during the first hours when PEI polyplexes are taken up by the cells, are responsible for the higher transfection efficiency. Since the cells used for both 50 mL TubeSpin reactor and 24-well plate transfections came from the same culture of origin, other factors that influence transfection efficiency such as passage number, cell health, cell cycle phases, and DNA quantity and quality were guaranteed to be consistent.

As a consequence of all the conducted experiments concerning transfection quality and consistency, the method was adapted as follows: whenever possible, HEK293 cells were transfected at passage number three; using the same plasmid DNA stocks of high Midi or Maxi prep grade; at a total DNA concentration of 0.48 µg per 10^6 cells; with a mass plasmid ratio of 1:2:0.5; in 5 mL culture volume in 50 mL TubeSpin reactors.

6.3.2 Summary

- Rep/Cap plasmid with separated and individually controllable *largeReps*, *cap*, and *smallReps* were created

    - Modifications restored rAAV titres and slightly above (Rep/Cap v3.7 and v5.4 ~1.2-fold of v1.2)
    - Rep78 essential in this system
    - p5 downstream of *cap* or additional *Rep52/40* beneficial
    - Intron of *cap* and short CMV promoter (80% activity of CMV) instead of p40 required for high rAAV titres
    - p5 is essential downstream of *cap* or additional *Rep52/40*
- Compatibility of new Rep/Cap plasmids with *E4orf6-6/7* based Helper (2.0) required CMV controlled *E4* transcription (Helper 2.1)
    - VG titre 1.8-fold higher compared to the original system
- CMV controlled *cap* in Rep/Cap plasmids v3.7 and v5.4 boosted capsid production up to 3.5fold, but with mainly empty capsids
- TF competition by multiple CMV promoters
    - Reduced CMV controlled capsid production when CMV promoters also control Helper genes, especially with two CMV promoters in the helper plasmid
- Packaging increased with second *Rep52/40* copy
- Lowering total plasmid amount per million cells to 0.48 µg improved VG titres, potentially reducing variability and avoiding potential saturation issues and high DNA usage
- Investigation of ITR/GOI plasmid ratios and passage numbers revealed minimal impact on VG titres but showed higher variability among different passage numbers
- Transfections in larger culture volumes using 50 mL TubeSpin reactors consistently yielded higher VG titres due to enhanced transfection efficiency likely because of improved mixing and cultivation conditions
- Refined method parameters were established, involving transfection at passage three, coherent plasmid DNA stocks, 0.48 µg total DNA per 10^6 cells, and 5 mL culture volume in 50 mL TubeSpin reactors, aiming to achieve consistent and efficient rAAV production.

6.3.3 Conclusion

In conclusion, Rep/Cap and Helper plasmids engineered for individual control over AAV and AdV Helper genes elevated rAAV VG titres on their own and could be used together. However, the new Rep/Cap plasmids v3.7 and v5.4 with their CMV driven *cap*, showed incompatibility with the Helper 2.0. The use of a CMV promoter for *E4orf6-6/7* was required to restore rAAV VG titres. The addition of a second small Rep copy proved its intended purpose, improved genome packaging. The shortened CMV promoter used for *cap* lead to uncontrolled VP expression, and consequently very high amounts of capsid, which were predominantly empty. CMV controlled *E2A* expression decreased capsid numbers and increased full/empty ratios. The competition between CMV promoters and Helper promoters raised questions about transcription factor availability and cellular resource allocation, with CMV-driven VP expression potentially compromising other crucial processes. This work highlights the complex interplay between various components, suggesting the importance of precise stoichiometry control, presumably addressable with the created system and synthetic promoters with defined transcription strength. While yielding valuable insights, further research is required to optimize the full/empty and harness the full potential of the engineered Rep/Cap and Helper plasmid system to elevate vector production efficiency for gene therapy applications.

Furthermore, the consistency of transient transfections was aimed to be improved. Challenges in day-to-day variation were addressed by optimizing the experimental setup. Lowering the total plasmid amount per million cells led to increased VG titres, highlighting the impact of DNA concentration on production. Biological variability and passage number effects were observed, emphasizing the need for stringent experimental controls. Transfections in larger culture volumes using TubeSpin reactors yielded higher VG titres, attributed to enhanced transfection efficiency. Consequently, a refined protocol was established: transfecting cells at passage three with standardized plasmid DNA stocks, using DNA concentrations0.48 µg total DNA per 10^6 cells with an increased ITR/GOI plasmid ratio in 50 mL TubeSpin reactors. These adaptations strive to mitigate variability and promote consistent and efficient recombinant AAV production.

6.4 Assessment of Impacts on rAAV Production, *E2A* and *E4* Transcription by Helper Gene Rearrangements and Recombinant Plasmid Engineering

Changes to the Helper plasmid demonstrated great potential for rAAV production optimisation in this work. The creation of simplified and size-reduced helper plasmids (6.1and the use these with CMV promoters for the transcription of *E4orf* subsets and *E2A* (6.2) confirmed the hypothesis that different transcriptional activities can influence and improve rAAV VG titres. The positioning, orientation, and context of genes can impact transcriptional accessibility, expression strength, and component stoichiometries in multigene expression vectors (Eszterhas et al., 2002; Patel et al., 2021). As the Helper plasmid is a multigene expression vector, the order of genes and their orientation were changed with Gibson Assembly approaches and so to speak reshuffled. These eleven reshuffled Helper plasmids (3.0-3.10) were based on the design of helper plasmid 2.0 and therefore contained the genes *E4orf6-6/7*, *E2A* with its *L4-33K/22K* containing 5'-UTR and *VAI*, in the about 1 kb large *VA RNA* fragment, in various orders. The Helper plasmid 3.0 maintained the gene order and orientation of Helper 2.0 but differed in the plasmid backbone, which was used from the Rep/Cap AAV8 plasmid v1.2 and contained additional restriction sites downstream of *E2A* and *E4*. Although these minor changes should not have made any difference, slightly higher titres were measured (data not shown). These could either be caused by the marginally size reduction of 0.5 kb to 10.1 kb, or by changes in mRNA contents. The latter might be caused by the minor sequence changes to the genes' 3'-UTR and E2E environment due to the new restriction sites. The schematic configuration of all eleven plasmids and the titres resulting from transfections with them are shown in Figure 37.

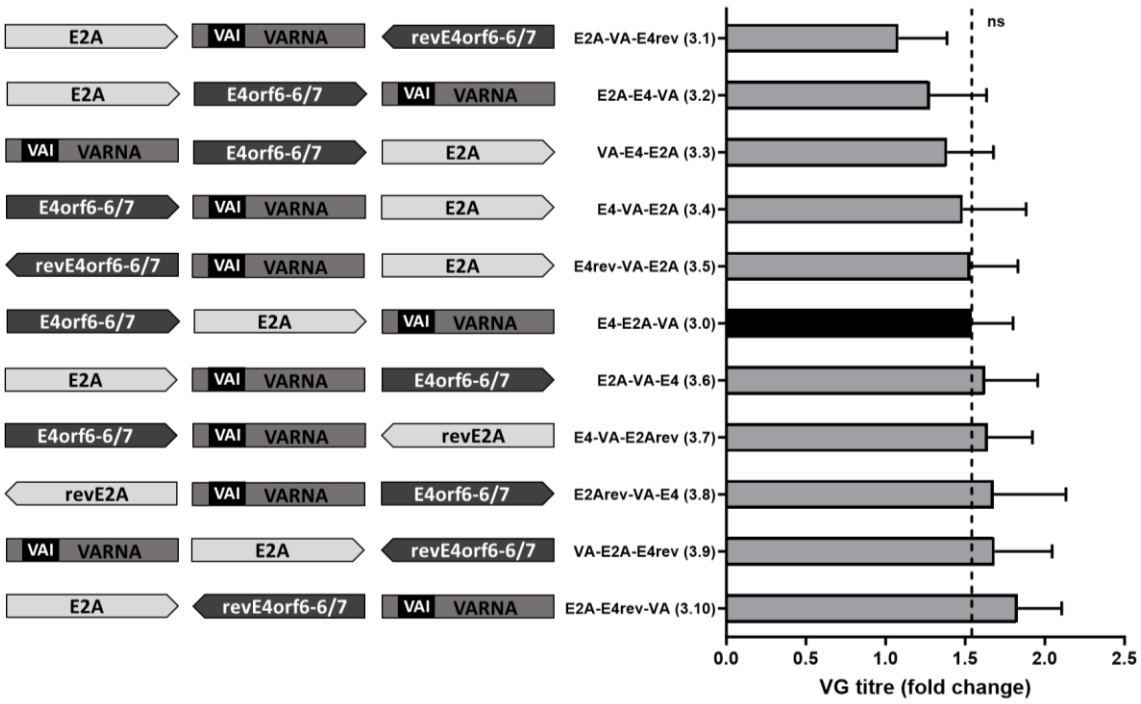


Figure 37: Reshuffled Helper plasmids based on Helper 2.0. Schematics of the gene order and orientation are not drawn to the actual size of the respective gene sequences. rAAV VG tires are fold changes compared to the control helper plasmid 0.1 and are displayed as means with error bars as standard deviation of three individual biological replicates, statistical analysis was performed as ordinary one-way ANOVA.

The created Helper plasmids with reshuffled gene orders (3.0 to 3.10) exhibited a range of rAAV production capacities, with differences in VG titre fold changes from 1.08 to 1.83-fold compared to the control (0.1). While the differences between the individual constructs were not statistically significant, they still confirmed the impact of gene order and orientation on the transcriptional activity of individual transcriptional units. Transcriptional run-through is one of the factors that changed between the constructs. In AdV, transcripts of the E4 promoter can include E2A translatable mRNAs (Donovan-Banfield et al., 2020). The same might also apply to the helper plasmid in any kind of tandem gene arrangements (same transcription direction, *E4→E2A* or *E2A→E4*), but is probably less likely with the presence of *VA RNA* or BB sequences between the genes. Polymerase read throughs and also steric hinderances can be important factors for transcriptional accessibility in multigene vectors. The highly active transcription of VAI might also influence the transcription initiation or termination of the other two genes in close proximity. However, the positioning of VAI did not follow a definite pattern of positive or negative impact on VG titres. Two of the three best performing vectors have the VAI sequence on the same stand right in front of their promoters of *E2A* and *E4*, respectively.

Some trends were discernible regarding gene arrangements, but no definite rules could be observed regarding gene preferred arrangements of the genes. Constructs with consecutive *E4* and *E2A* or *E2A* and *E4* gene sequences were in the bottom half achieved of VG titres, with the Helper 2.0 like configuration of Helper 3.0 being the best of these and exactly in the middle of all configurations. Transcriptional interference caused by the run-through of the RNA II polymerase transcribing the upstream gene might be a reason for the lowered titres (Shearwin et al., 2005). The tandem arrangement favours expression of the upstream gene vectors (Eszterhas et al., 2002; Patel et al., 2021), leading to the hypothesis that the two Helper genes are likely required to be transcribed similar amounts, as none of the two configurations *E4*→*E2A* or *E2A*→*E4* was generally much better than the other.

The two best constructs resembled each other's configuration with convergent *E2A* and *E4* genes in a polyA-to-polyA configuration, potentially preventing polymerase run-through and enable more efficient termination. Divergent gene arrangements often exhibit even better transcriptional activity (Eszterhas et al., 2002). This arrangement is exhibited by the third best reshuffled Helper (3.8), with a separation by the *VA RNA* fragment. No divergent constructs were created with head-to-head promoter configuration because interference of the promoters was feared, impeding results when these promoters would be exchanged. Apart from Helper 3.1, all configurations separating *E2A* and *E4* with the *VA RNA* fragment and the plasmid backbone showed similar titres, ranging in the midfield of all the configurations. The best one of these, Helper 3.8, was taken forward together with the best performing configuration 3.10 to test a further minimisation. Helper 3.8 was chosen instead of 3.9 because of 3.9's similarity to 3.10 and the possible benefit entirely different transcription directions of 3.8 and 3.10 with a separation of the divergent promoters by the *VA RNA* fragment in Helper 3.8. Additionally, the 2.0 configuration resembling Helper 3.0 (*E4*→*E2A*→*VA*) and Helper 3.2, which has put *E2A* upstream of *E4* (*E2A*→*E4*→*VA*), were used to explore synthetic molecular biology techniques for multicistronic gene expression in rAAV production.

6.4.1 Minimizing of Reshuffled Helper Plasmids

In chapter 6.1.4, it was observed that reducing the *E2A* 5' UTR to its essential part, the *L4-33K/22K* gene, resulted in a slight but insignificant decrease in rAAV VG titres. The 1.8 kb reduction in plasmid size made it reasonable to use the smaller 5'-UTR ($E2A_{min1}$) since similar titres were achieved. Moreover, a plasmid size of Helper 3.0 still >10 kb could be potentially improved in terms of transfection efficiencies (Hornstein et al., 2016; Kreiss, 1999), as dramatically increased efficiencies were observed when the size was below 10 kb in experiments published by Kreiss (1999). Additionally, the reduced size might then allow the use of other

genes or genetic components on the same plasmid without impairing transfection efficiencies too much. Consequently, the E2A$_{min1}$ 5‘-UTR was constructed in the Helper plasmids 3.0, 3.2, 3.8 and 3.10.

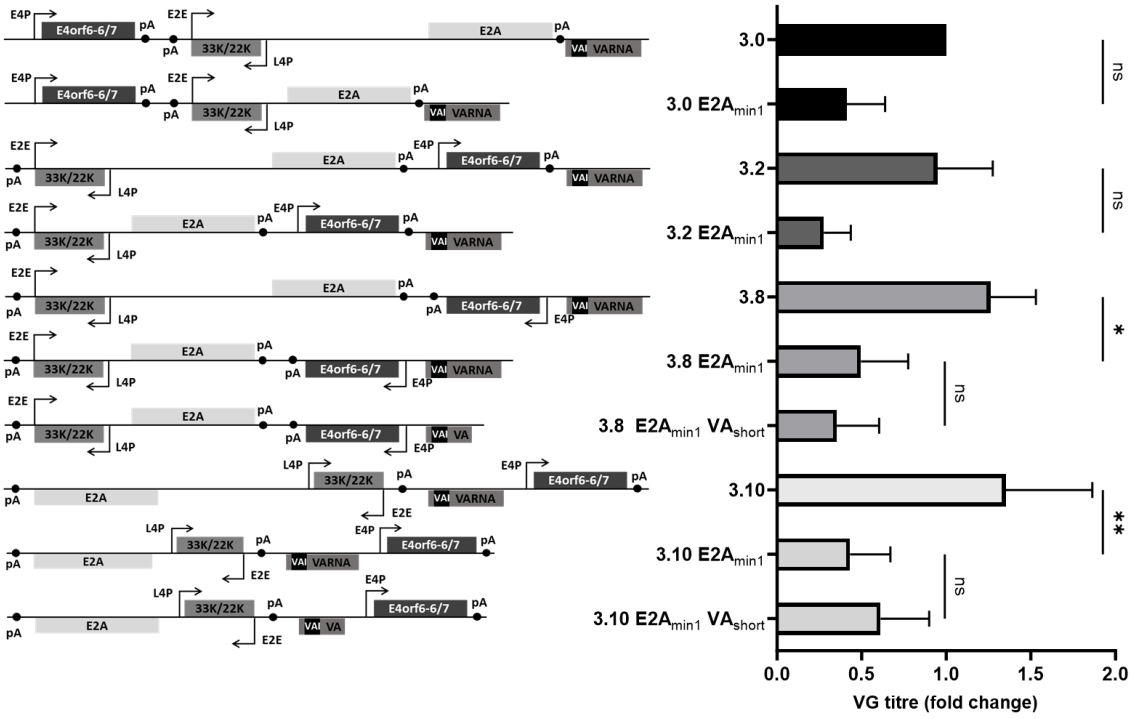


Figure 38: Size reduction of the *E2A* 5’-intron of the Helper plasmids 3.0, 3.2, 3.8 and 3.10. Schematic depiction of the Helper plasmids next to associated VG titre fold changes compared to Helper 3.0. Arrows denote promoters, black dots poly A sequences. Genes, 5’ and 3’-UTRs, but not promoters and polyA signals are drawn to approximate scale. Data shown as means with error bars indicating standard deviations of four biological replicates. Ordinary one-way ANOVA was performed for statistical analysis.

All four Helper plasmids exhibited strong drops in VG titre with the smaller *E2A* 5’-UTR. Although not all changes were statistically significant, rAAV8 titres dropped for Helper 3.0 and 3.8 by nearly 60% and for Helper 3.2 and 3.10 by about 70% on average when the *E2A* 5’-UTR was reduced to the *L4-33K/22K* gene (E2A$_{min1}$). It remains unclear why E2A$_{min1}$ 5’-UTR restored VG titres compared to E2A$_{min2}$ and showed similar titres to the full *E2A* 5’-UTR when used in the Helper 1.0 plasmid, but not in these plasmid versions with the optimised *E4orf6-6/7* gene.

The intact E2E promoter should have provided sufficient expression of DBP, and E4orf6/7's transactivation effect on E2E should have remained intact. Interestingly, the reduction in titre was observed regardless of the gene order to comparable extents. Additionally, no issues were expected with the expression of L4-33K and 22K, as confirmed by mass spectrometry with the Helper 1.0 E2A$_{min1}$ plasmid, and no changes were made to the gene's close environment.

The second change that was tested with the $E2A_{min1}$ versions of plasmids 3.8 and 3.10 involved exchanging of the *VA RNA* fragment for the smaller version, *VA* RNA_{short}, which had already been successfully used in the 5-plasmid system. Here, no significant differences were observed, confirming the results of the 5-plasmid system. Therefore, it was decided to use the 400 bp smaller *VA RNA* sequence in upcoming plasmid versions. However, further investigation was deemed necessary for the *E2A* 5' UTR, or it should be left intact. For promoter exchanges, this is not ideal, as the 3 kb distance between promoter and translation start site induces uncertainty regarding the translatability of previously tested promoter strengths.

6.4.2 Synthetic, Bicistronic Helper Plasmids and Repositioning of *L4-33K/22K*

Rather than discarding the full *E2A* 5'-UTR, including *L4-33K/22K*, or keeping it intact and having a 3 kb gap between promoter and gene, or using shortened 5'-UTR, a fourth option was explored. Versions of Helper plasmids 3.8 and 3.10 were designed to move the promoter right in front of the *E2A* gene while retaining the beneficial *L4-33K/22K* gene upstream of the promoter used for *E2A* transcription control. To ensure termination of L4P transcripts, a rabbit beta-globin poly adenylation signal was placed downstream of the gene (Lanoix & Acheson, 1988). The divergent gene configuration of *E2A*, L4-33K/22K, could even be preferable, because it separates the two genes and stops their existing transcription competition when transcribing the same sequence in opposite directions. *E2A* transcription was controlled by a CMV or SV40 early promoter (SV40) and also the terminator of *E2A* was exchanged to the commonly in recombinant mammalian processes used SV40 poly adenylation site (SV40pA).

Furthermore, Helper plasmids were designed containing an internal ribosome entry site (IRES) from encephalomyocarditis virus (EMCV) or the 2A self-cleaving peptide P2A for a bicistronic expression approach of *E2A* and *E4* helper genes for rAAV production (Bochkov & Palmenberg, 2006; Douin et al., 2004; Z. Liu et al., 2017). To achieve efficient transcription termination and possibly enhanced mRNA stability, the Woodchuck Hepatitis Virus (WHV) Posttranscriptional Regulatory Element (WPRE) was utilized with a bovine growth hormone polyA (bGH pA) downstream of the bicistronic gene cassette (Higashimoto et al., 2007; Sun et al., 2009).

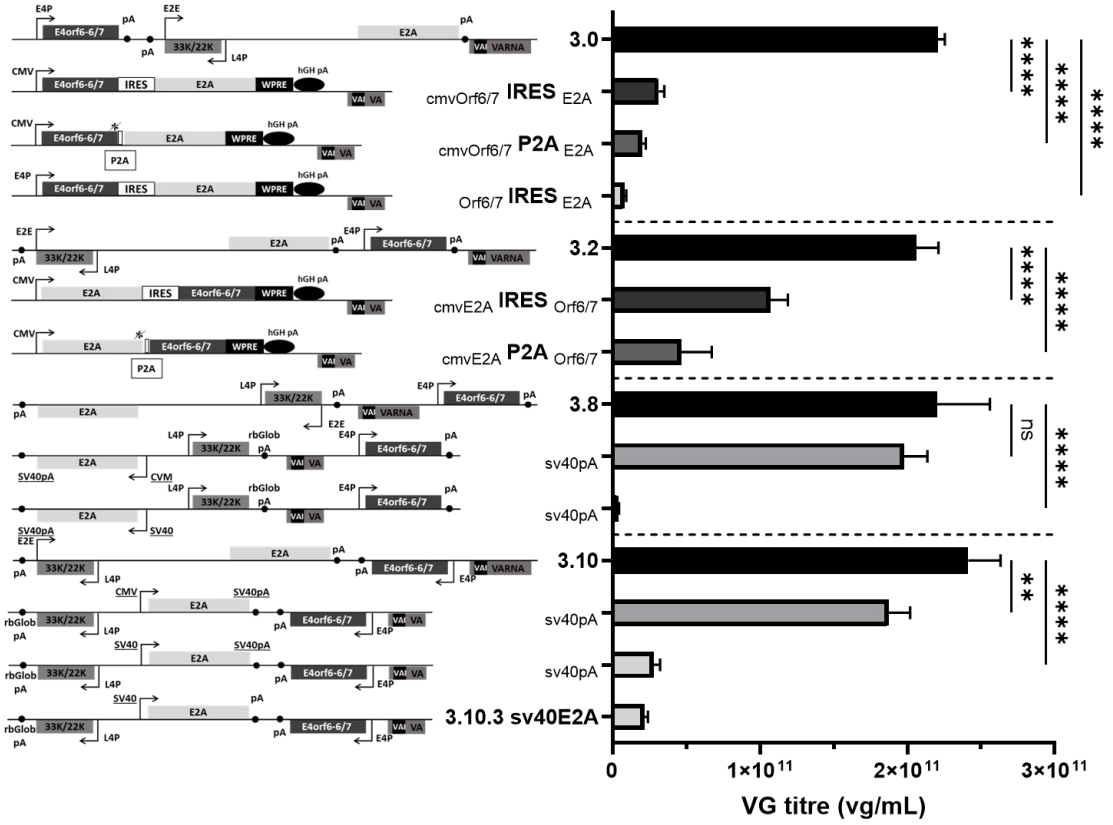


Figure 39: Synthetic Helper plasmids designs containing heterologous promoters, EMCV IRES or P2A sites for a bicistronic *E2A*/*E4* gene expression, and Helper plasmids with relocated L4-33K/22K gene. Schematic depiction of the Helper plasmids next to associated VG titres. Arrows denote promoters, black dots poly A sequences, crossed out asterixis demonstrate the missing of the usual stop codon. Genes, 5' and 3'-UTRs, IRES and P2A, but not promoters and polyA signals are drawn to approximate scale. Data shown as means with error bars indicating standard deviations of three biological replicates. Ordinary one-way ANOVA was performed for statistical analysis comparisons displayed to parental Helper 3.0, 3.8 or 3.10.

Synthetic biology methods for bicistronic protein expression are not new to AAV production systems. Many GOI combination constructs deploy an IRES, and also Rep expression was previously controlled with an IRES in the TESSA-DS-Cap-EMCV-Rep construct (Su, 2021). The present work represents the first use of this synthetic biology technique for the AdV Helper transient triple plasmid method used for rAAV production, although not for AdV Helper genes in general. While this work was conducted, Gu et al. (2020) worked on stably integrated, bicistronically expressed Helper genes in HEK cell lines and patented these. Transfections of the Helper plasmids 3.01 to 3.2.2, presented in Figure 39, demonstrates that rAAV can be produced with bicistronic helper gene expression, with the lowest produced average rAAV8 titre being 7.8×10^9 vg/mL (3.0.3), which is still far above the detection limit of the ddPCR for the utilized 4000-times dilution of the crude sample. However, VG titres were much lower than with the monocistronic parental Helper plasmids. For Helper 3.0 the rAAV8 VG titres decreased to about 15% with the

cmvIRES construct and to about 9% when a P2A site was used instead. Helper 3.2, which used the CMV directly upstream of *E2A* and utilized the IRES or P2A for the E4orf6-6/7 expression, showed only a decrease from 2.1×10^{11} vg/mL to 1.1×10^{11} vg/mL (52%) for the IRES construct and a reduction to 23% of titre of Helper 3.2 with the P2A construct. Differences between the consecutive gene orders follow a trend that was generally seen in previous experiments of this work, the greater importance of adequate E2A than E4 expression. Both IRES and P2A worked better with the gene order *E2A*→*E4*. It is known for both bicistronic expression mechanisms, despite their different mechanistic foundations, that the expression of the second gene is usually lower. For EMCV IRES constructs, the second, IRES-dependent protein is usually expressed 0.2-0.5-times the level of the first protein (Mizuguchi et al., 2000). For P2A, the second protein is typically expressed at a similarly decreased level of about 0.3-fold of the first protein (Z. Liu et al., 2017). Consequently, the results suggest that the *E2A* gene seems to require higher transcriptional activity or is inherently more important for high titre rAAV production than the *E4* gene. This is confirming the findings of the 5-plasmid system. However, the *E4orf6-6/7* based CMV containing systems showed slightly positive or no titre changes for cmvE4 and titre decreases with cmvE2A. This discrepancy seems contradictive as long as it is assumed that CMV driven E2A expression much stronger than E2E driven expression. That this is not the case is shown in 6.4.3. Furthermore, the patented system of Gu et al. (2020) also utilises the gene order *E2A*→*E4* with an IRES for their stable cell lines, confirming the here presented results.

Generally, usage of the P2A peptide for bicistronic expression achieved lower VG titres than the use of the EMCV IRES when a CMV promoter was used. The use of the regular E4 promoter (3.0.3), which is weaker than CMV (~75% of CMV), also resulted in very low VG titres when the IRES was used. Likely, *E2A* transcription is far too low with this promoter and plasmid configuration, leading to low rAAV production due to insufficient DBP levels. The generally higher VG titres achieved with the consecutive monocistronic gene arrangement instead of the bicistronic approach could also be due to the downstream promoter increasing expression levels. This effect was shown by Sakaguchi et al. (2014) with different promoters but was strongest with the triple promoter setup downstream of the 3' polyAsignal of hTERT, SV40 and CMV promoters, making it uncertain if such effects are present with E4P and E2E promoters, especially since the uninterrupted *E2A*→*E4* and *E4*→*E2A* gene sequence orders performed worse than other gene orders in the reshuffling of the Helper plasmid (Figure 37).

rAAV production with Helper 3.2.1 might also be improvable with the reinstallation of the missing *L4-33K/22K* gene, potentially reinstating virus genome titre yields to levels of parental plasmid 3.2. The positioning of the gene could be similar to the Helper plasmids 3.8.1 to 3.10.3. Here the *L4-33K/22K* gene was placed in front of the *E2A* driving CMV promoter. rAAV VG titres drop slightly for both plasmid configuration (3.8 and 3.10), although not statistically significant in the case of 3.8 to 3.8.1. In contrast, when an SV40 promoter was used instead to control DBP expression, virus genome titres crashed to 11% or less of the parental plasmids' titres. This clearly indicated that the SV40 promoter was too weak for *E2A* transcription. Although our measurements with production of empty rAAVs showed that the E2E promoter is only about ~32% as strong as CMV (appendix 8.2.4.1) and the SV40 promoter is usually described as strong as ~37% of CMV in 293T cells (Qin et al., 2010), it seems to be much weaker in the current system and does not support high titre rAAV production. Two factors potentially influence the strength of the SV40 promoter in comparison to CMV in HEK293 cells that could diminish its strength in the present transfection system. Firstly, the large T antigen present in 293T cells acts as a transactivator for different promoters, but specifically for the SV40 promoter (Gruda et al., 1993; Gruda & Alwine, 1991). Therefore, transcription strength of the promoter is likely much lower in the used HEK293 cells compared to the numbers of the 293T cells (Qin et al., 2010; Toktay et al., 2022). Most promoter strength experiments containing SV40 were carried out with HEK cell containing the large T antigen, but data of Green et al. (2023) suggest a strength in Expi293F cells of about 30% of CMV, whereas data of Raup et al. (2016) indicates a much weaker eGPF expression with SV40 in HEK293 than with CMV. Secondly, the CMV promoter is known to be activated by E1A, which is always present in HEK293 cells, whereas the SV40 promoter is repressed by E1A (Cockett et al., 1991). Therefore, far greater differences in transcription strength of SV40 and CMV promoter and consequently in DBP concentrations could be prevalent in the present transections of HEK293 cells, explaining the extreme VG titre differences between 3.8.1 and 3.8.2 and 3.10.1 and 3.10.2, respectively. The use of *E2A*'s regular 3'UTR and poly adenylation site lowered the already low VG titre with the SV40 promoter only slightly. Both polyA sites seem to be function adequately for *E2A*. Taken together, the results of Figure 39 suggest a strong promoter like CMV for the *E2A* Helper gene for high titre rAAV production with heterologous promoters in a plasmid configuration like Helpers 3.8.1 and 3.10.1.

6.4.3 Examination of *E2A* and *E4orf6+6/7* mRNA Levels of Different Helper Plasmid

To validate previously assumed actions and interactions of components and gain further insight into why some vectors performed better than others in terms of plasmid and promoter configurations, quantification of mRNA levels of the two helper genes *E2A* and *E4* was attempted. RNA extractions were performed and normalised RNA samples were measured using two different primer-probe sets in the same OneStep-RT ddPCR reaction. The results of these measurements for different Helper plasmid constructs and their corresponding titres from the same transfections are presented in Figure 40.

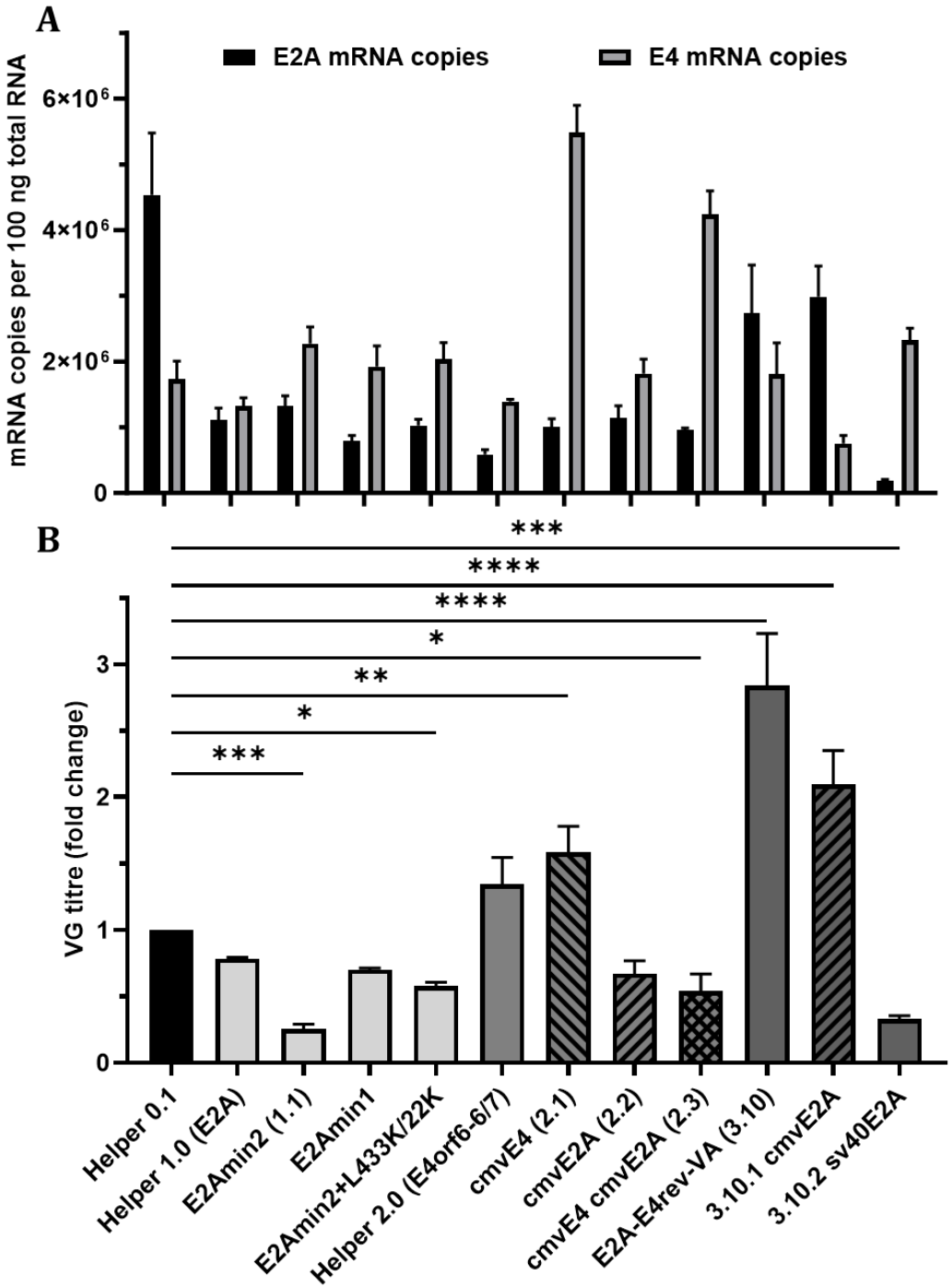


Figure 40: Determination of mRNA levels of the AdV5 Helper genes *E2A* and *E4orf6+6/7* in transient triple transfections for rAAV production with different Helper plasmids. (A) E2A (black) and E4orf6+6/7 (grey) mRNA copies of different Helper plasmids constructs measured with different primer-probe sets in the same ddPCR reaction. Data presented as mean and standard deviation of three biological replicates. (B) rAAV8 VG titre fold changes respectively to control plasmid Helper 0.1 (black) of transfection samples presented in A. Samples of the same plasmid generation in same grey shade, CMV promoter-controlled plasmids marked with hatched pattern. Data presented as mean and standard deviation of three biological replicates. Ordinary one-way ANOVA was performed for statistical analysis, pairwise comparisons displayed to parental Helper 3.0, 3.8 or 3.10.

The previously reported small titre difference between Helper 0.1 and the 12 kb large Helper 1.0 (about 14%) could not be explained by the missing of sequences including the *L5 (fiber)* gene. The mRNA analysis of E2A and E4 transcripts does indicate that lowered DBP concentrations could be a reason. This explanation is supported by findings of the Helpers 3.10.1 and 3.10.2, which clearly demonstrate that a sharp decrease in E2A transcripts is associated with plummeting rAAV VG titres. The previous assumption that the SV40 promoter does not provide sufficient activity for adequate E2A mRNA amounts could be confirmed. Differences in E2A transcripts between Helper 0.1 and 1.0, as well as all other created Helper plasmid in this work, as they are descendants of Helper 1.0, could originate in the used *E4* 3'-UTR sequence and the included polyA site. As this sequence does not fully resemble the one of Helper 0.1, polymerase run-through might be a bigger issue. The distance between *E4* and *E2A* might also play a role, suggesting that the excluded *L5* gene could have an influence on E2A transcripts, even if it was not resembled in rAAV VG titres. The assumption of decreased of E2A mRNAs with the deletion or part deletion of the *E2A* 5'-UTR could not be unconditionally confirmed. It is possible that the overall increase in transcription with the smaller plasmid could be due to its smaller size and potential higher transfection efficiency (Kreiss, 1999). Additionally, in $E2A_{min2}$, even though *E4orf6+6/7* levels are increased, E2A levels remained similar to those of Helper 1.0, indicating that increased transcription interference or the partially missing E2E downstream sequence (+14 to +21), which is crucial according to Ellsworth et al. (2001) could still be reasons for reduced *E2A* transcription in this plasmid. However, $E2A_{min1}$, which features this E2E downstream sequence, has only slightly decreased E2A mRNA levels, refuting the hypothesis of this element being the cause for decreased VG titres. The fact that mRNA levels between $E2A_{min2}$, $E2A_{min1}$ and the co-transfection with *L4-33K/22K* showed similar mRNA amounts for both genes is further evidence that the differences in VG tire caused by the *E2A* 5'-UTR are in fact associated with L4-33K/22K.

Contrary to expectations, the CMV promoter did not increase mRNA levels of E2A compared to the E2E promoter. This finding is a surprising as in-house promoter activity measurements showed that the E2E promoter had only an activity of 32% of CMV when expressing the reporter protein GFP, refuting the hypothesis of decreased VG titres due to DBP overexpression and cytotoxicity. The influence of context specificity, promoter and gene environment become evident when the E2A mRNA levels of the Helper plasmid 3.10 with its reshuffled gene order and orientation are regarded. The examination of Helper plasmid 3.10 with its reshuffled gene order and orientation revealed that VG titre increases were accompanied by a

general increase in Helper gene transcription, likely enabled by the new order and orientation of the genes. The most significant enhancement in mRNA level was observed for E2A, with a 4.7-fold increase compared to Helper 2.0. This suggests that the change in gene orientation and order had a more substantial impact than the use of the CMV promoter. This observation supports the idea that the CMV promoter is likely not much stronger in this context than E2E, if at all. As creating promoters much stronger than CMV is challenging in HEK cells, but high E2A transcript levels seem beneficial, alternative strategies such as gene orientation may be required to enhance gene transcript levels, which can have an impact on VG titre (Johari et al., 2022). Some experiments in HEK293 cells suggest that other heterologous promoters like EF-1α or CAG may be stronger than CMV, but this is not always the case (Johari et al., 2022; Qin et al., 2010; Xu et al., 2019). Therefore, another alternative could be the modification of the endogenous E2E promoter to increase E2A transcript levels. Both the E2L and E2E promoters lack an ideal TATA box. Optimising E2L's TATA box by a single-point mutation increased its activity by 10-fold (D. H. Huang et al., 1988). A similar modification could potentially increase E2E's activity, allowing for increased DBP levels without disrupting the natural regulatory mechanisms of the viral transcription units. However, such mutations might not titrate transcriptional strength as finely as a suite of synthetic promoters can.

In contrast to the E2E promoter, the exchange of the E4 promoter for a CMV promoter drastically changed E4 transcript levels. However, as described earlier, the increase in E4orf6-6/7 mRNA levels resulting from this change had no or only a marginal improving impact on rAAV titres. Interestingly, transcript levels of E2A were not impaired when the CMV promoter was used for *E4orf6-6/7* transcription. On the contrary, E4 mRNA transcripts were affected when E2A expression was driven by a CMV promoter. This result perfectly confirms our TF competition hypothesis, as the overlap of TFREs of CMV is larger with the E4 promoter than with E2E (appendix 8.2.4.1). When two CMV promoters were utilized, E4 mRNA levels were increased 3-fold compared to Helper 2.0, but E2A mRNA levels only 1.6-fold. The consecutive gene order *E4→E2A* is like the reason for this, as other studies have shown the decrease of expression of the downstream gene in tandem gene arrangements (Eszterhas et al., 2002; Patel et al., 2021). The use of an insulator gene element between the genes could minimize this effect (Hasegawa & Nakatsuji, 2002).

Many of the results of the mRNA determination align with the made changes and outcomes; however, amounts of the E2A and E4 transcripts, as well as their ratios, do not seem to tell the whole story and are no absolute indicator whether high rAAV

titres can be achieved with a certain Helper plasmid. Helper 2.0 shows greatly reduced mRNA levels of E2A, and the E4 mRNA amount is slightly reduced compared to Helper 0.1; however, VG titres are about 1.4-times as high. Similarly, while the E2A level of Helper 2.0 is further decreased compared to Helper 1.0, VG titres are increased, contradicting the previously made hypothesis of high E2A levels as a requirement for high rAAV titres. Additionally, the overall transcriptional increase of Helper 1.1 compared to 1.0 and 2.0 is not representative of the achieved VG titres. Although a trend suggests that less overall Helper gene transcription might be beneficial, potentially due to enhanced capacities for AAV gene expression, the much higher titres of Helper 3.10 contradict this hypothesis entirely.

In summary, the E2A mRNA levels were similar with the E2E and CMV promoters, indicating similar strength in this context. Notably, the 2.x cmvE2A constructs possessed the 3 kb 5'-UTR separating gene and promoter, whereas construct 3.10.1 employed the promoter directly upstream of the gene. Therefore, CMV could not enhance *E2A* transcription despite its position. However, E2A mRNA levels could be increased by changing gene order and orientation. As this change was associated with a significant increase in VG titre, it is hypothesised that the promoter strength for *E2A* transcription should be increased for rAAV production improvement. E4*orf6-6/7* transcription enhancement had only marginal impact on rAAV titres. The use of two strong consecutive promoters was impacting the downstream promoter as expected. Therefore, an alternative gene arrangement should preferably be used when promoter exchanges are supposed to be tested. For a better result promoter exchanges and general plasmid development of a multi-plasmid system for rAAV production should be performed by testing combinations of different Helper and Rep/Cap plasmids, as only transcription levels are not indicative of superior performance and interactions are likely to change different parts of the system.

The experiment's results confirmed many previous assumptions, providing more insight into the transcriptional landscape and are thus relevant for promoter and other transcriptional design decisions for the Helper plasmid. However, the impact of this data has limitations and needs to be put into perspective. The main constraint of the experiment is that only the mRNA landscape at the time of harvest is displayed. As mRNA levels are fluctuant during cultivation and production, specifically with the intricate viral gene regulations, the drawn picture might look very different 6, 10, 16, 24 or 48 hours post transfection. Despite the costs and the labour intensity of this mRNA-extraction-based method, testing it at multiple time points during cultivation and investigating all significant plasmid modifications would provide more insight into Helper gene transcript level developments over

time, component interactions, as well as when and how different transcript levels directly influence rAAV VG titres. The investigation of Helper 2.0 and the impact of E4orf6/7 on E2A levels over the production period would be of particular interest. Previously, it was assumed that the reason this Helper performed better than 1.0 and others was its potentially upregulated *E2A* transcription, caused by increased E4orf6/7 levels. This hypothesis might still hold true, but only at earlier time points and not 72 hours post transfection when it E2A mRNA levels were even below those of Helper 1.0. Furthermore, observing mRNA levels of AAV genes simultaneously would be more informative. Although impossible with the used RT-Onestep ddPCR approach, mRNA sequencing techniques, microarray-based system or the novel nano-string technology could be employed (Heller, 2002; Stutika et al., 2016; Tsang et al., 2017).

6.4.4 Summary

- Plasmids with reshuffled Helper gene arrangements created based on the design of Helper plasmid 2.0
    - Helper plasmid gene order and orientation influenced VG titre from 1.08 to 1.83-fold compared to the control
    - Two of the best designs, Helper 3.8 and 3.10, were taken forward for further genetic engineering
- Minimalization of *E2A* 5'-UTR not compatible with *E4orf6/7*
    - Contrary to Helper 1.0 $E2A_{min1}$ diminished VG titres compared to the full *E2A* 5'-UTR
    - Alternative incorporation of *L4-33K/22K* upstream of *E2A*'s promoter is functional for high -titre rAAV production (Helper 3.8.1 / 3.10.1)
- Synthetic biology techniques for bicistronic gene expression of Helper genes in transient triple transfection system were explored
    - IRES superior to P2A in present Helper plasmid
    - Bicistronic Helper gene expression requires *E2A* upstream of the *E4* and a strong promoter, confirming published results for stable cell lines
        - Indication of the greater importance of *E2A* than *E4* for enhanced rAAV production
- E2A and E4 mRNA levels were examined for different helper plasmid constructs
    - High importance of *E2A* gene transcription for efficient rAAV production was confirmed
    - Changes in gene order and orientation had substantial impacts on transcription levels and VG titres

- CMV promoter did not increase *E2A* transcription in comparison with E2E, regardless of position or intron presence
- CMV promoter did increase *E4* transcription drastically, but VG titres only marginally
- Reduced mRNA levels did not consequently lead to lower VG titres
  - Complex interactions and intricacy of the system require individual titration and multi-time point analysis of further components additional to *E2A* and *E4*

6.4.5 Conclusion

The reconfiguration of the Helper plasmid yielded significant enhancements in VG titres and holds promise for further enhancements as it enables further improvements by genetic engineering. Surprisingly, the attempt to streamline *E2A*'s 5'-UTR to its essential element, *L4-22K/33K*, proved unsuccessful within the reshuffled Helper plasmids. However, an alternative strategy involving the insertion of the gene upstream of the *E2A* regulating promoter resulted in the production of robust rAAV genome titres. This modification also established a more dependable platform for future promoter substitutions. Notably, the successful integration of $E2A_{min1}$ into Helper 1.0 underscored that certain alterations to the plasmid system may not be universally applicable and could necessitate additional optimisation of other components to achieve full functionality. The novel Helper plasmids, designated as 3.8.1 and 3.10.1, represent the most refined iterations thus far, enabling systematic promoter replacements and, potentially, the manipulation of polyA sites to finely tune transcriptional control and achieve optimal Helper component ratios.

The application of bicistronic expression to E2A and E4orf6-6/7 demonstrated the feasibility of leveraging these synthetic biology tools within the Helper plasmid framework for transient rAAV production. Given further optimisation, these could prove valuable as they reduce the number of required promoters, at least if similarly strong and simultaneous expressions of both Helper genes turned out to be ideal. The bicistronic Helper plasmids also highlighted the pivotal role of adequate *E2A* transcription and emphasize its significance in comparison to *E4*.

Although the mRNA analysis indicated a similar trend in Helper gene expression, no experiments testing weak promoters or conditions that might diminish E4 transcript levels have been conducted. Conceivably, both genes demand robust transcription, with a potential bias towards higher E2A transcript levels. Since the CMV promoter did not alter E2A transcript levels, the observation that gene order and orientation exert a substantial influence on mRNA levels underscores the

viability of alternative strategies, beyond relying solely on distinct promoters, to finely adjust component stoichiometries.

While the conducted analysis of mRNA levels offers valuable insights into the requirements of Helper transcript levels and the effects of made plasmid modifications, a more comprehensive understanding of component requirements demands a more intricate examination. This would entail employing a diverse array of analytical techniques to investigate all pertinent components at various stages during production, offering deeper and more dependable insights.

6.5 Molecular Design for a Controllable Recombinant AAV Expression System Encompassing Novel Helper and Rep/Cap Plasmids

The ultimate objective of this project was to achieve precise control over AAV and AdV helper genes by utilizing promoters of defined strength, thereby optimizing the expression levels of each component to maximize rAAV titres. To accomplish this goal, two distinct molecular engineering tasks were undertaken. First, the focus was on modernizing the rAAV plasmid system to enable independent regulation of all four mRNA-producing genes and the multitude of their resulting proteins with promoters as unimpeded and individual as possible. Secondly, a suite of promoters with varying transcriptional activities was developed to yield different amounts of individual components and achieve efficient system control. For the second task, the CMV promoter, renowned for its strength and ubiquitous utility, along with its TFRE composition and its *cis*-regulatory modules (CRMs) underwent in-depth analysis. A set of deletion variants of the CMV promoter was then constructed to exhibit distinct transcriptional strengths in HEK293 cells. While I contributed to some experimental work, scientific discussions, idea conceptualisations, and review and editing processes, it is important to note that the primary authorship and intellectual property belong to Dr. Yusuf Johari. Therefore, only the abstract of the co-authored article will be quoted and presented here, with the full article is presented in 8.2.2.

Engineering of the CMV Promoter for Controlled Expression of Recombinant Genes in HEK293 Cells

- Yusuf B. Johari, Joseph M. Scarrott, Thilo H. Pohle, Ping Liu, Ayda Mayer, Adam J. Brown, David C. James

“Expression of recombinant genes in HEK293 cells is frequently utilized for production of recombinant proteins and viral vectors. These systems frequently employ the cytomegalovirus (CMV) promoter to drive recombinant gene transcription. However, the mechanistic basis of CMV-mediated transcriptional activation in HEK293 cells is unknown and consequently there are no strategies to engineer CMV for controlled expression of recombinant genes. Extensive bioinformatic analyses of transcription factor regulatory elements (TFREs) within the human CMV sequence and transcription factor mRNAs within the HEK293 transcriptome revealed 80 possible regulatory interactions. Through in vitro functional testing using reporter constructs harbouring discrete TFREs or CMV deletion variants we identified key TFRE components and clusters of TFREs (*cis*-regulatory modules) within the CMV sequence. Our data reveal that CMV activity in HEK293 cells is a function of the promoters various constituent TFREs including

AhR:ARNT, CREB, E4F, Sp1, ZBED1, JunB, c-Rel, and NF-κB. We also identified critical Sp1-dependent upstream activator elements near the transcriptional start site that were required for efficient transcription and YY1 and RBP-Jκ binding sites that mediate transrepression. Our study shows for the first time that novel, compact CMV-derived promoters can be engineered that exhibit up to 50% higher transcriptional efficiency (activity per unit DNA sequence) or 14% increase in total activity compared to the wild-type counterpart."

Based on this work, particularly the identification of active and repressive TFREs within the CMV promoter in HEK293 cells, a suite of synthetic promoters was created following Dr. Johari's departure from the project. The CMV study used the exact TFRE sequences from the CMV promoter for the TFRE analysis with homotypic promoters (enhancer regions comprised of seven copies of the same TFRE upstream of a CMV core promoter). Additionally, the study tested the consensus sequence of some of the TFREs (MYBL1, Oct, E2F), all of which exhibited higher transcriptional activity than the versions present in CMV (measured as relative fluorescence levels of GFP compared to GFP expressed with a CMV promoter). Consequently, consensus sequences for the utilised TFREs, obtained from the JASPAR open-access database of transcription factor binding profiles, were employed instead of the original CMV-derived sequences (Castro-Mondragon et al., 2022). The selected TFREs for the synthetic promoter assemblies were the 12 most active ones from the paper, along with YY1 as a repressor. The following design principles were applied for the synthetic enhancer elements: (i) no spacers between TFREs, (ii) length <500 bp, (iii) maximal four times the same TFRE, (iv) random mixing of used TFREs, but (v) maximal two consecutive copies of the same TFRE. Five different designs were employed, as depicted in Figure 41 A. The promoter activities of the created synthetic promoters, measured as relative fluorescence levels compared to the full CMV promoter, are presented in Figure 41 B.

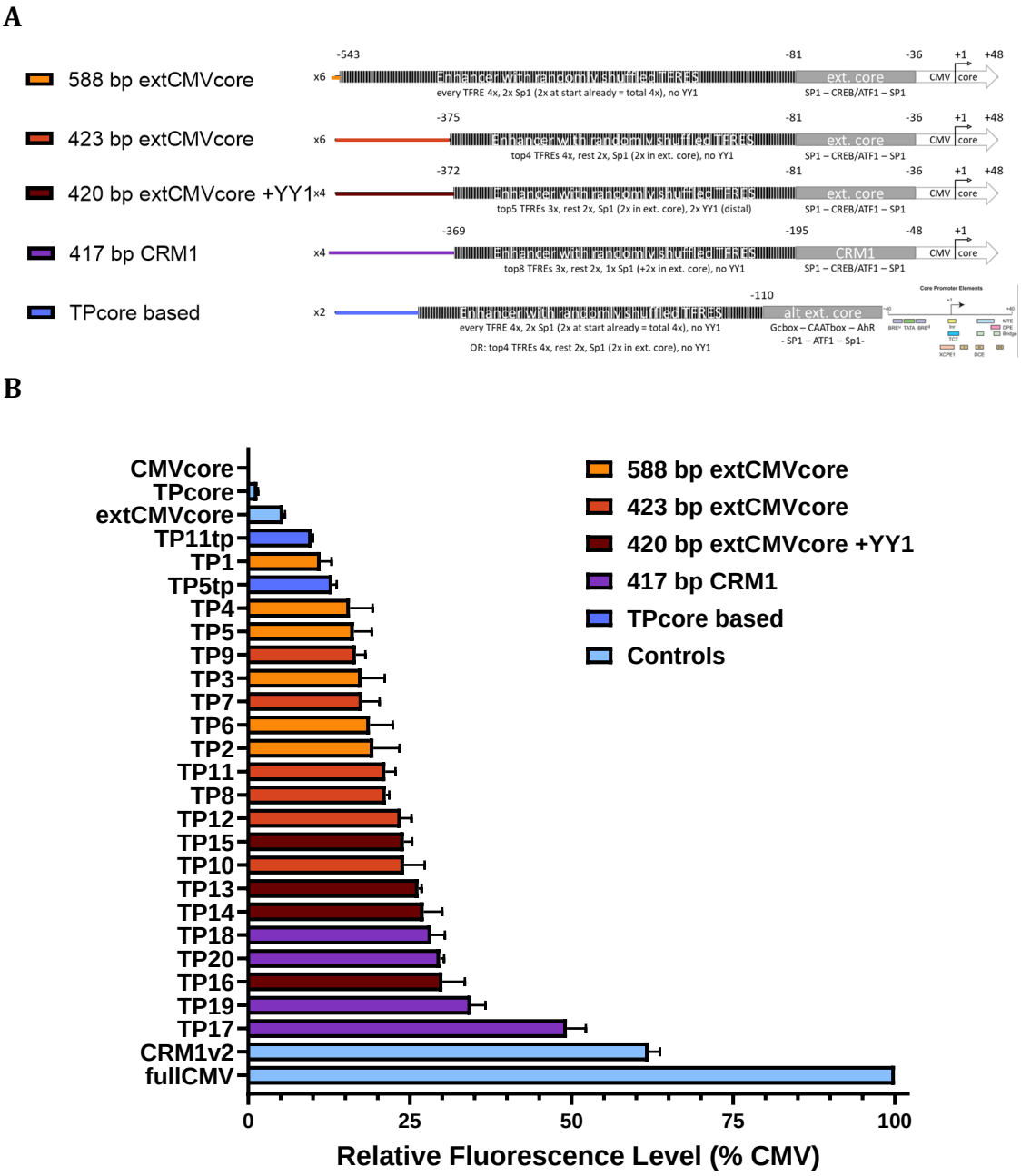


Figure 41: (A) Synthetic promoter designs conceptualised in this work. Indicated are sequence length in base pairs, relative to the transcription start site (TSS, +1). Important TFREs of the proximal promoter region here called extended core (-36 to -81 of CMV, orange, red, brown), CRM1 (-48 to -195 of CMV, purple) and the alternative extended core (-1 to -110 of promoter consensus sequences, blue) are listed. These three mentioned proximal and core promoter regions were equipped with synthetic enhancer sequences of five different TFRE composition designs. Through random shuffling of TFREs a number of synthetic promoters (indicated in the depiction) of the five basic TFRE composition designs were created (by DNA synthesis). All designs utilized a CMV core promoter, except for the TPcore based ones (blue) which used a synthetically designed promoter core based on consensus sequences of mammalian core promoter elements. Top x TFREs refer to homotypic screening in Johari et al. (2022), ranked sequences listed in appendix Figure 51. (B) Relative transcriptional activity exhibited by synthetic promoters compared to the full CMV promoter. Synthetic promoter controlled GFP expressing reporter plasmids were transfected into HEK293 cells using PEIpro and cultured in 50 mL TubeSpin reactors (5 mL culture volume) at 37°C. GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the CMV promoter. Data shown are the mean value and standard deviation of three independent biological replicates.

The previous study had demonstrated that the CMV CRM1, encompassing the CMV core promoter (48 to -36 bp relative to TSS (+1)) and the CMV proximal enhancer sequence up to -195 bp, exhibited 67% of the transcriptional activity of the full CMV promoter (Johari et al., 2022). In an effort to increase promoter strength beyond CMV levels, four synthetic promoters were created with transcriptional highly active TFREs as extensions of the already relatively strong CRM1 enhancer. Unfortunately, the four constructs did not meet the expectations, and the effect of highly active TFREs (top 8 from the homotypic screening) was repressive on the CRM1 promoter (Figure 41, purple). Nevertheless, due to the usage of the already relatively strong CRM1, these four promoters (TP17 to TP20) were among the five strongest synthetic promoters in this library, with TP17 achieving the highest transcriptional activity of all synthetic promoters at 49.3% of CMV.

Generally, the individual promoters within each design exhibited similar behaviour, with the 417 bp CRM1 designs being the strongest, followed by the 420 bp extCMVcore designs containing YY1, and then the designs without YY1 and the longest enhancer sequences, the 588 bp extCMV core designs. The extended CMV core used in the constructs contain the important TFREs Sp1-CREB/AF1-SP1. The Sp1 sites are also known as GC box and the CREB/ATF1 element is similar to the CAAT box, both of which are essential elements in the proximal promoter for transcription initiation. The two SP1 sites were crucial for the activity of the CMV promoter, which is why the -81 sequence of the CMV promoter was included in the constructs as an extension to the CMV core. It was assumed that this way, functionality and a base level of transcription would be guaranteed. As expected, all synthetic promoters that were designed with the extended CMV core showed transcriptional activity with at least 11.2% activity of CMV (TP1). However, the extended core itself was already transcriptionally active (5.4% of CMV), meaning that the used enhancer elements only increased its activity by 2.5 to 6.8-fold, with TP16 performing best at 30% of CMV.

In the CMV study, it was clearly shown that YY1 acts as a repressor in the CMV promoter. Exclusion of a YY1 containing segment of CMV (-241 to -373, CRM1+2+7, named CMV4.01) resulted in a 1.15-fold stronger CMV promoter. Surprisingly, almost all synthetic promoters containing YY1 outperformed the YY1-less constructs that were also designed based on the extended CMV core. TP16 even performed better than two of the CRM1-based constructs. The reason for this behaviour is uncertain. While YY1 is known, for example, from *rep*'s p5 promoter, as a repressor that turns into an activator in the presence of helper virus, this function is mediated by E1A, which is always present in HEK293 cells and would, therefore,

have had the same effect in the CMV study (Lee et al., 1995; Pereira et al., 1997; Shi et al., 1991).

The newly, fully synthetically designed TPcore and its proximal region, composed of mammalian consensus sequences of the specific regions and elements, did not perform as well as the CMVcore and its upstream extension used in all other constructs. The enhancer regions of the two TPcore constructs were the same as the ones of TP5 and TP11, respectively. Both versions with the synthetic core were less active (20% for TP5, 54% TP11) than their extCMVcore-based counterparts. It would be interesting to test these synthetic promoter constructs side by side in different cell lines to determine if this phenomenon is specific to HEK cells, as CMV is known to be particularly strong in these cells, or if the new TPcore and its installed proximal promoter region are generally weaker than the first 129 bp of CMV. The low expression with this promoter core was unexpected due to its similarity to the synthetic so-called "super core promoter 1", which demonstrated superior performance to the CMV core in HeLa cells (Juven-Gershon et al., 2006).

In summary, transcriptional activity of the synthetic promoter library, based on the CMV analysis study, was lower than expected. Nevertheless, the library provided a range of different promoter strengths, which was one of the primary goals. In combination with the CRM-CMV promoter library, a wide spectrum of different promoter activities was created, offering promoters with discrete and defined transcriptional activity with activities as low as 1% and up to 115% of CMV. These promoters enabled altering rAAV component amounts with predefined transcription levels. For optimised transferability of these promoter activities to the conditions during rAAV production, selected promoters could also have been tested as co-transfections during rAAV transient triple transfections. This experimental setup should be considered for future promoter developments for rAAV plasmid systems, as Rep and Helper functions are known to modulate the transcriptional landscape to drastic measures (Babiss, 1989; L. S. Chang & Shenk, 1990; Di Pasquale & Stacey, 1998; Farley et al., 2004; Marton et al., 1990; Shi et al., 1991; Song et al., 1995; Timpe et al., 2006; Velcich & Ziff, 1985; Weger et al., 1999).

6.5.1 Assembly of a Helper Plasmid Based on Previous Plasmid Optimisations for Individualised and Unimpeded Helper Gene Expression

The task of creating Rep/Cap and Helper plasmids in which these promoters can be easily inserted, exchanged, and utilized with as few restrictions as possible, was broadly addressed in the previous chapters. For the Helper plasmid, the *L4-33K/22K-E2A* gene order and plasmid design was most suited for this task. Among these Helper plasmid designs, Helper 3.10.1 performed best and was selected as the basis for creating a Helper plasmid that combines the knowledge and results from previous experiments to achieve an optimized gene composition for the exchange of promoters and achieve an optimized gene expression strength of helper components. Additionally, the aim was to use only heterologous promoters from this point onward, if possible, to eliminate the uncontrollable regulation and feedback loops of the viral promoters and their products. Thus, based on Helper plasmid 3.10.1, a design was used with CMV promoters driving expression of both *E2A* and *E4orf6-6/7* (Helper 4.0). Considering the promoter competition theory and the decreased VG titres observed with the double CMV Helper plasmid 2.3, an alternative version was also constructed with the heterologous E4 promoter for E4orf6-6/7 transcription (4.0_{E4P}).

The SV40pA signal was deemed slightly advantageous for *E2A* (see 3.10.2 vs 3.10.3) and is commonly used for recombinant gene expression with strong viral promoters in mammalian cells. Furthermore, it is known that promoter and terminator build a functional unit that determines strength and efficiency of transcription and should therefore be attuned to each other (Al-Husini et al., 2020). With multi gene containing plasmids like the Helper plasmid, transcriptional interference is of specific importance when genes are close together as it is the case in plasmids (S. West & Proudfoot, 2009). The SV40pA signal is one of the strongest for mammalian recombinant protein expression and therefore often used in combination with CMV promoters. Consequently, it seemed feasible to use it for both genes. However, there are two versions of the SV40pA, the SV40 early (SV40EpA, here commonly referred to as SV40pA) and SV40 late (SV40LpA) version, whereas one is simply the reverse complement of the other. The stronger terminator is SV40EpA (Chao et al., 1999). To give both genes the same termination sequence, the same SV40EpA sequences were used downstream of each gene. The resulting structure was a double terminator of SV40EpA-SV40LpA due to the reverse complimentary nature of the two polyA sites. The constructed Helper plasmids are shown schematically alongside the corresponding VG titres in Figure 42.

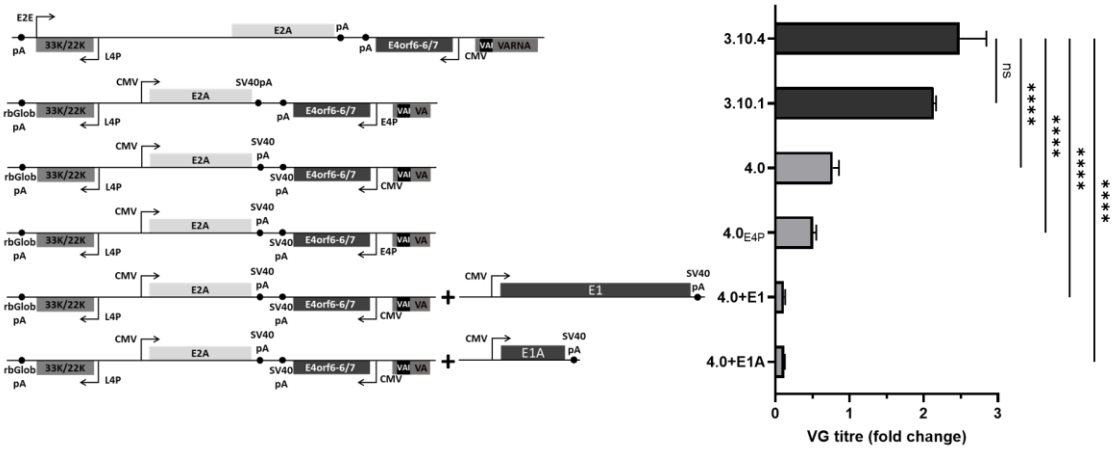


Figure 42: Helper plasmids based on Helper 3.10.1 with heterologous promoter and polyA sequences, as well as co-expression of AdV5 E1 and E1A with Helper plasmid 4.0 for rAAV8 production. Schematic depiction of the Helper plasmids next to associated VG titres fold changes compared to control Helper plasmid 0.1. Arrows denote promoters, black dots poly A sequences. Genes, 5' and 3'-UTRs, but not promoters and polyA signals are drawn to approximate scale. Data shown as means with error bars indicating standard deviations of three biological replicates. Ordinary one-way ANOVA was performed for statistical analysis, pairwise comparisons shown with respect to Helper 3.10.4.

The newly constructed Helper plasmid 4.0 did not perform as expected. VG titres were 33% lower than the original Helper plasmid 0.1 (pADΔF6). In contrast, the parental plasmids 3.10.1 and 3.10.4 achieved much higher rAAV, reaching 2.5-fold and 2.1-fold of Helper 0.1, respectively. With Helper 3.10.4 reaching titres similar to 3.10 (data not shown), it was confirmed again that for *E4orf6-6/7* transcription in the CMV promoter worked equally to the E4 promoter to achieve high rAAV titres. Helper 4.0_{E4P} also confirmed this hypothesis, showing only slightly, statistically not significant, decreased VG titres compared to 4.0 with its CMV promoter (one-way ANOVA p=0.3249). This also meant, in contrast to Helper 2.3, that the double CMV composition of this Helper plasmid was not performing worse and transcription factor competition seemed to be less of an issue. The Rep/Cap plasmid used (v1.2) without any additional CMV promoters probably had an influence, and the changed gene orientation might be preferable for similar promoters.

However, the idea of two SV40 terminators in succession, intended as a backup for enhanced efficiency of transcription termination, was likely flawed and responsible for the substantial reduction in rAAV VG titres. It is possible that the SV40EpA is a weaker terminator than the endogenous AdV5 polyA sites, and the convergent transcription directions lead to polymerase collision (Hobson et al., 2012). Improvements could be achieved with a more spaced-out design or potentially with stronger polyA sites or a combination of both. Nevertheless, the decrease in rAAV titre and the non-ideal gene orientation for the installation of different polyA signals was regarded as problematic, particularly with the cooperation of polyA and

promoter concept in mind. Further, convergent gene orientations can cause downregulation of gene expression through RNA interference pathways (Proudfoot, 2016).

The Helper 4.0 was also co-transfected with the AdV5 E1A gene or the full E1 gene. E1A is an important helper function in AAV proliferation and rAAV production. Its transactivation of the early AdV gene promoters and AAV promoters. Particularly p5 transactivation of E1A is vital for efficient rAAV production. E1A enhances the activity of the E4 promoter 10 to 50-fold, E2E up to 21-fold and p5 up to 40-fold (L. S. Chang & Shenk, 1990; Gilardi & Perricaudet, 1986). E1B is essential for rAAV production in its cooperative role with E4orf6, but also to counteract E1A to prevent premature apoptosis (Meier et al., 2020). Since the E1 gene is not transfected with the other helper genes, there is no control over it yet. Potentially varying copy numbers of E1 in HEK293 clones open the possibility of reduced rAAV production due to a lack of this gene. Additionally, the variation in genomic E1 copy numbers could explain the potential for VG titre boosts by cell line development and cellular cloning. To compensate for potential lack of E1 products and boost titres in general, controllable E1 or E1A transcripts could be added. Previous studies have shown positive effects of overexpressing the E1A gene on rAAV production (Gu & Wang, 2022; Tratschin et al., 1984). However, transfections with Helper 4.0 and E1A or E1, respectively, resulted in a further significant decrease in rAAV VG titres. The incompatibility of E1 or E1A overexpression with the current system could have several reasons. Firstly, the published and patented system might benefit from additional E1A expression due to a lower basal production from their HEK293 cell line, which may not be the case for the present cell line. With sufficient E1 product levels in the system, overexpression might not have a benefit and could instead imbalance component stoichiometries. Secondly, E1A causes apoptosis, likely explaining the low rAAV titres with strong co-expression driven by a CMV promoter, either with E1A or the full E1 gene overexpressed (Rao et al., 1992; White, 2001b). Measurement of VCD during the production period and at the cell harvest could have confirmed this theory. Additionally, using CMV promoters instead of E2E and E4 promoters reduces the potential positive effects of E1A transcription activation in this specific helper setup. Lastly, the presence of the CMV promoter on the equimolar co-transfected plasmid may cause transcription factor competition or depletion, as it was the third CMV promoter in the system. TFRE similarities of the AdV and AAV promoters to CMV, and particularly the other CMV promoters in the system suggest that these promoters are likely affected by the use of multiple copies of the strong CMV promoter titrating away TFs from each other and therefore weakening their own protein expression.

In summary, the overexpression of the *E1* gene or only its first product, E1A, resulted in decreased rAAV VG titres with Helper 4.0 instead of the expected increase in titres based on other studies. Furthermore, Helper 4.0 itself did not perform as expected, with the convergent gene orientation with a double SV40 polyadenylation setup being a major issue. Furthermore, the use of a multitude of CMV promoters is likely not ideal. Therefore, the suite of CMV derivatives and synthetic promoters of different transcriptional strengths is a valuable functional tool to address different genes in the system for distinct transcriptional activities.

Besides regulating the genes already present in the system, overexpression of other potentially beneficial genes could be tested and fine-tuned using these synthetic promoters. Although the overexpression of E1A was not successful in this case, other groups have shown positive results (Gu & Wang, 2022). The assembly activating protein (AAP) is a considerable target for overexpression, but in experiments of Grosse et al. (2017) it did not increase rAAV titres. Changes to the MAAP showed promising titre improvements, but its overexpression does not appear to be generally beneficial, as it only slightly enhanced VG titres but increased packaging of contaminant sequences (Galibert et al., 2021). Lastly, the combination of AdV5 Helper genes with helper gene functions of other AAV helper viruses, such as genes from human bocavirus 1 (NS2, NP1, and BocaSR) or common helper HSV 1, which modulate the cell and transcriptional landscape for improved virus production, could provide benefits when co-expressed (C. Li & Samulski, 2020; Meier et al., 2020; Z. Wang et al., 2017, 2018).

Given the multiple drawbacks observed with Helper 4.0, particularly the substantial drop in VG titre with the use of heterologous promoters and polyA sites, a new design was required to optimise the system's component stoichiometries using heterologous promoters with defined transcriptional activities. Consequently, a new design for the optimisation of the systems' component stoichiometries controlled by heterologous promoters with defined transcriptional activity was required. Instead of attempting a new design based on Helper 3.10.1, it was decided to return to the design of Helper 3.8.1, which included the shortened *VA RNA* fragment separating *E2A* and *E4* genes on different strands. For the *E4* gene, the optimized ORF subset *E4orf3+6+6/7* was utilized. A set of Helper plasmids was created using different promoters from the synthetic promoter library and the CRM-CMV library. The production results of rAAV8 with these plasmids are shown in Figure 43.

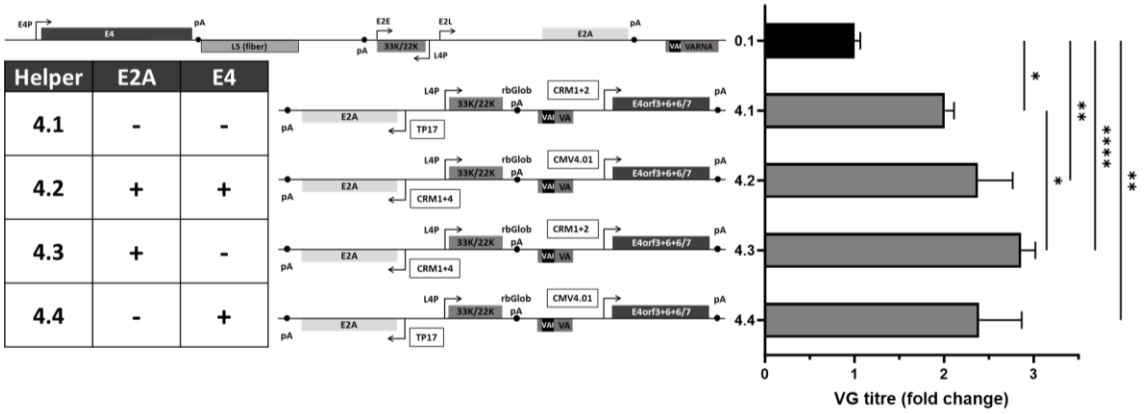


Figure 43: Helper plasmid designs 4.1 to 4.4, based on the design of Helper 3.8.1, deploying different promoters from the synthetic promoter library and the CRM-CMV library. Schematic depiction of the Helper plasmids next to associated VG titres fold changes compared to control Helper plasmid 0.1. Arrows denote promoters, black dots poly A sequences. Genes, 5' and 3'-UTRs, but not promoters and polyA signals are drawn to approximate scale. The table indicates the promoter choices for *E2A* (- =TP17 / + =CRM1+4) and E4orf3+6+6/7 (- =CRM1+2 / + =CMV4.01). VG titre data shown as means with error bars indicating standard deviations of three biological replicates. Ordinary one-way ANOVA was performed for statistical analysis.

Unlike Helper 4.0, Helpers 4.1 to 4.4 did not use heterologous polyA signals. Instead, they employed heterologous promoters with much greater success, achieving rAAV8 titres beyond 2.5×10^{11} vg/mL with all four plasmid variations. The highest titres overall were achieved with Helper 4.3, yielding 3.7×10^{11} vg/mL or 2.86-fold of the parental base plasmid Helper 0.1 (pADΔF6). The four Helper plasmids (4.1-4.4) only differed in their promoter composition, using stronger and weaker promoters for both controllable genes. For *E2A*, TP17 with a relative strength of 49.3% compared to CMV and CRM1+4 with 85% of CMV transcription activity were employed as stronger and weaker promoters, respectively. For E4orf3+6+6/7, CRM1+2 with 75% of CMV's transcriptional activity was used as the weaker promoter, while the stronger option was the strongest promoter of the CRM-CMV library, 4.01, which has an activity of 115% of the full CMV promoter (Johari et al., 2022). The promoter choices for the two genes were based on their endogenous promoters, E2E with 32% and E4P with 76% of the transcriptional activity of CMV. Consequently, the promoters for *E4* were chosen to be stronger than those for *E2A*. However, the best performing combination, Helper 4.3, demonstrated that the stronger promoter option for *E2A* and the weaker one for *E4orf3+6+6/7* yielded the best results. Even though all four helpers showed high titre rAAV production, the Helpers with the weaker TP17 promoter for *E2A* transcription performed worse. These results are consistent with the previous findings, showing that a too low or lower *E2A* transcription limits rAAV VG titres. However, unlike the SV40 promoter, the stronger CRM1+4 promoter did not diminish titres when used for *E2A* transcription, indicating sufficient expression of the DBP with promoters of

strengths at least 50% of CMV and higher. Most likely, unlike the SV40 promoter, there was no repression by E1A with the synthetic promoter, which is a favourable quality. Again, for the *E4* gene, a difference in transcription strength did not seem to matter as much. Further experiments should investigate lowered promoter strengths for *E4orf3+6+6/7*. Additionally, it would be interesting to explore whether *E4orf6/7* could be omitted when *E2A* transcription is regulated by a synthetic promoter that is not co-transactivated by E4orf6/7, instead of the E2E promoter. Future experiments should test higher promoter strengths for the *E2A* controlling promoter to find an optimal fine-tuning of promoter strengths for the two mRNA-producing Helper genes. Ideally, such experiments should also include the AAV genes.

6.5.2 Attempted Rearrangement of the Rep/Cap Plasmid for Individualised and Unimpeded Gene Expression

To change the promoters of the AAV genes, new plasmid compositions were created based on Rep/Cap plasmids v3.7 and v5.4. These new plasmids were designed with altered gene order and orientation, and the promoters of the genes were exchanged to remove endogenous promoters, except for the p19 promoter located in the 5' part of *rep*. To control the small Reps, the second copy, placed downstream of *cap*, in the v5.4 derived plasmids, was equipped with the TP14 promoter. TP14 has a transcription activity of 27% relative to CMV, which is similar to the one of p5, the promoter that was previously used to control the transcription of the second *smallRep* copy in v5.4. Although the CMV version used in the Rep/Cap plasmids was not a full CMV, but a truncated version with a relative strength of 80% to the full CMV, it was possibly still too strong, as it was leading to the production of mainly empty capsids. On the other hand, the far weaker p40 promoter (50% CMV without AAV production) was too weak, achieving titres no higher than 30% of the control (Rep/Cap v1.2, see 8.2.4). To find a suitable *cap* driving promoter, the CRM-CMV promoter CRM1 was chosen with a transcription activity of 67%, closer to but lower than the truncated 80% CMV promoter. The weak, remnant p5 promoter upstream of the *rep* gene had only marginally higher transcription activity than the CMV core and was therefore nearly non-transcriptionally active. It was replaced with the TPcore promoter, which has a transcriptional activity of 1% relative to CMV. Similar to the reshuffled Helper plasmids, the gene order of *rep* and *cap* was changed for the Rep/Cap v5.4 derived plasmids v7.2 and v7.3. The Rep/Cap v3.7 derived plasmid v6.2 also had the orientation of the *cap* gene changed to a convergent gene arrangement.

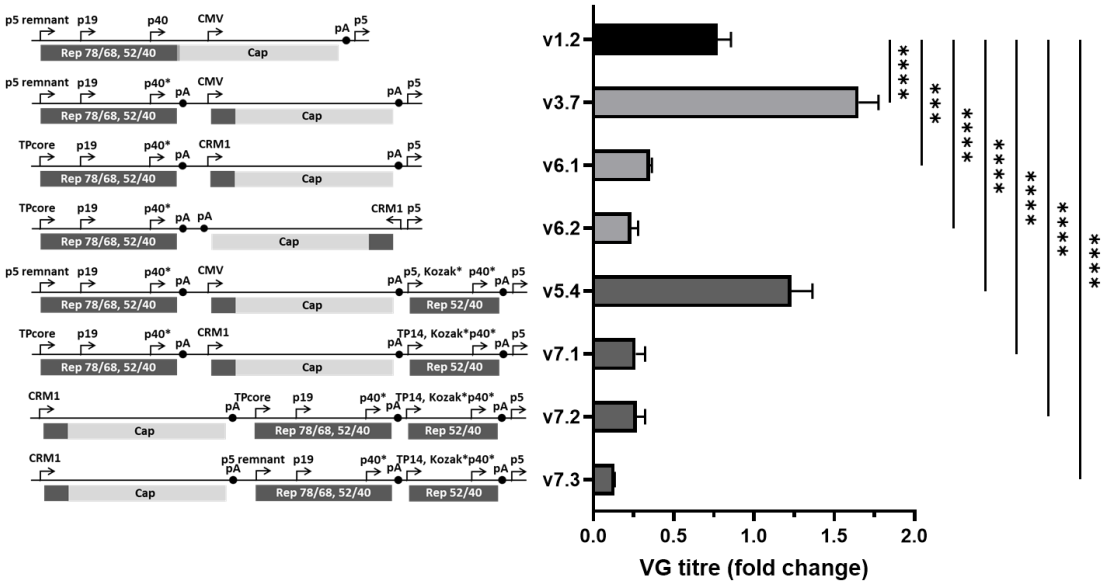


Figure 44: Rep/Cap plasmid designs with changed gene order and orientation, based on either Rep/Cap plasmid v3.7 (light grey bars) or v5.4 (dark grey bars). Transfections for rAAV production of respective Rep/Cap plasmids performed with Helper plasmid 4.0 and the regular ITR/GOI plasmid. rAAV VG titres fold changes compared to control Helper plasmid 0.1 are presented next to schematic depiction of Rep/Cap plasmid designs. Arrows denote promoters, black dots poly A sequences. Genes, 5' and 3'-UTRs, but not promoters and polyA signals are drawn to approximate scale. Data shown as means with error bars indicating standard deviations of three biological replicates. For statistical analysis an ordinary one-way ANOVA was performed, pairwise comparisons shown with respect to Rep/Cap plasmid v1.2.

All in Figure 44 shown transfections were performed with Helper 4.0, which performed worse than Helper 0.1 with Rep/Cap v1.2. Surprisingly, the combination of Helper 4.0 and Rep/Cap v3.7 and v5.4 improved the performance of both Helper and Rep/Cap plasmids. rAAV VG titres of v3.7 and v5.4 were improved by about 1.2-fold compared to Rep/Cap v1.2 when used in combination with Helper 0.1. The combination of Helper 4.0 with v3.7 or v5.4, increased rAAV VG titres by 1.65 and 1.23-fold, respectively, compared to Rep/Cap v1.2 combined with Helper 0.1. The synergistic effect could be attributed to Helper 4.0's two CMV promoters combined with the use of a CMV promoter for transcriptional control of *cap* in v3.7 and v5.4. Even though Helper 2.3 that also used two CMV promoters did not increase VG titres, it improved full/empty ratios and significantly reduced the very high amount of empty capsids produced by the two Rep/Cap plasmids v3.7 and v5.4. The decrease in empty capsids and TF competition could be the trigger for other mechanistic effects leading to the synergistic VG titre improvement seen with the combination of these plasmids. This example highlights the importance of an optimized combination of Rep/Cap and Helper plasmids and their specific promoters.

All five of the newly created Rep/Cap plasmids (6.1-7.3) generated much lower rAAV VG titres than their parental plasmids, achieving only 6-21% of the parental

VG titre. The intention behind these Rep/Cap plasmid changes was to increase VG titres through favourable *cap* positioning and promoter choice. As downstream genes in consecutive gene arrangements are known to be expressed at a reduced extent (Eszterhas et al., 2002), it was anticipated that the convergent gene orientation of Rep/Cap v6.2 would outperform Rep/Cap v6.1. Surprisingly, despite no statistical significance (ANOVA p=0.6116), Rep/Cap v6.2 exhibited a lower VG titre than v6.1. The presence of two adjacent polyA sites, similar to Helper 4.0, might have resulted in polymerase crashes or distorted termination, which could be a contributing factor. Additionally, VP overproduction could still occur with the CRM1 promoter, but it might have been lowered due to the consecutive gene arrangement, potentially influencing VG titres. Changes in other gene orders and orientations did not seem to impact VG titres. Since different gene arrangements and promoter changes were tested together, the exchange of promoters appeared to be the common factor leading to the decreased VG titres. It became necessary to identify which specific change was responsible and needed to be reverted.

When analysing the results in combination with the results of the v5.4 derived Rep/Cap plasmids v7.1 and v7.2, two features were consistently present: the CRM1 promoter and the TPcore promoter, which replaced the 80% CMV and remnant p5 promoter, respectively. The use of CRM1 in all plasmids for *cap* transcription activation might have caused a drop in VP production, which was initially favoured. However, the expression with the promoter exhibiting only 67% of CMV might have been too weak in the split Rep/Cap plasmid system, similar to p40 with its 50% activity respective to CMV. Further experiments should be conducted to explore different gene orientations, including various promoter choices for the essential expression of *cap*-originated proteins. Additionally, the routine measurement of total capsids would be favourable. Considering costs, low throughput, and labour intensity of these analysis with the PROGEN ELISA kits, available HPLC-based methods should be established first (Khatwani et al., 2021).

The reinstallation of the p5 remnant promoter in plasmid v7.3 resulted in even lower VG titres, possibly because it is even weaker than the TPcore promoter, leading to too low Rep78/68 production. Although it is commonly accepted to reduce the quantity of large Rep proteins to improve rAAV titres due to their toxicity, they are still essential and, in some cases, limiting for AAV replication when expressed too low quantities, leading to no or low virus production (Sha et al., 2021). For instance, the results of the manuscript draft clearly indicated that the loss of Rep78 significantly reduced rAAV production. Therefore, it is conceivable that all the newly created Rep/Cap plasmids (v6.1-7.3) generated insufficient Rep78/68 levels. It was hypothesised that Rep78/68 amounts are higher with the previous

plasmids (v3.7 and v5.4) because of presence of the p5 promoter downstream of *cap* or second *Rep52/40* copy, and still drives large Rep expression. This full p5 promoter, only separated from its regular gene product by the plasmid backbone, might still serves its natural function – the transcription of *Rep78/68*. The Rep/Cap research presented in 6.3 demonstrates the downstream p5's beneficial effect on rAAV production that is associated with its function as a p5 remnant competitor and its transactivation features for p19 (Lackner & Muzyczka, 2002). It was speculated that its significance in common rAAV production might be extended to its function as a distant promoter for its gene products Rep78/68. This hypothesis does not interfere with the lowered expression of the large Rep proteins which were reported previously (Xiao et al., 1998), but are rather caused by the increased distance between gene and promoter than the loss of the promoter. However, similarly low rAAV titres with plasmids containing other promoters, but still the full downstream p5 promoter and the remnant p5 in front of *rep* refuted this theory (6.5.3). It is to note that this function would only have been possible in plasmid transfections due to the circular nature. An experiment using linearized DNA could test if there is a benefit of the full p5 as a distant promoter for the expression of the large Rep protein.

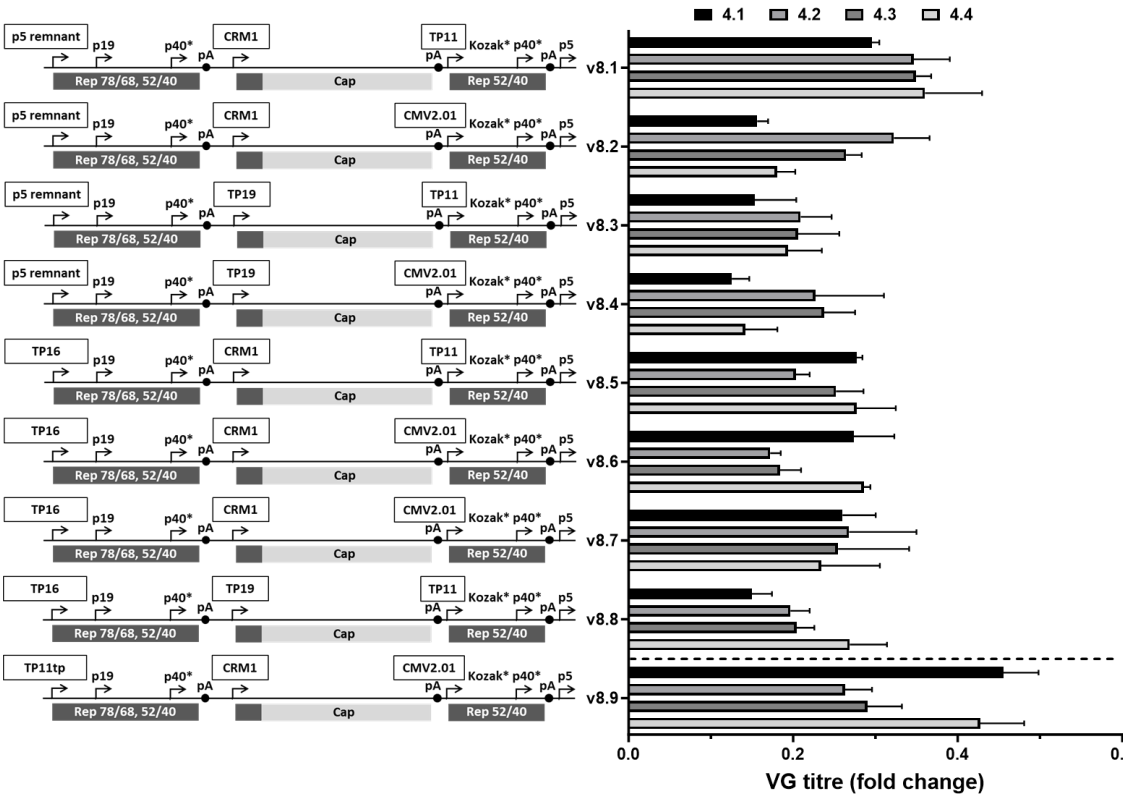
As it was hypothesised at the time that the lowered rAAV VG titres observed with Rep/Cap plasmids v6.1-7.3 might be attributed to the disruption of the potential p5 distant promoter function and the absence of the complete p5-BB-p5remn-rep constellation, further promoter exchanges for the *rep* and *cap* genes were planned with Rep/Cap v5.4. Furthermore, the p5 sequence was considered potentially valuable due to its contained CARE sequence. This *cis*-Acting Replication Element (CARE), located in the p5 promoter downstream of *cap*, is capable of plasmid replication as shown by Nony et al. (2001), potentially reducing the required ITR/GOI plasmid quantity.

Lastly, no definitive statement can be made about the exchange of p5 to TP14 for the control of the second *Rep52/40* copy. There might not be a substantial difference due to the similar promoter strengths, except for the elimination of E1A activation required for p5's activity and other control mechanisms like feedback loops of Reps repressing p5. Consequently, a heterologous, preferably synthetic promoter should be considered as it could be as good or even advantageous. It is recommended for future improvements with this plasmid configuration to exchange the promoters for *cap* and both *rep* gene varieties, ideally in combination with the promoters of all other genes in the system, including the Helper plasmid genes.

6.5.3 Improving Promoter Exchanges for rAAV Production: A First Experimental Attempt

The foundation for broad-ranging promoter exchanges of all viral genes to enhance rAAV production was established, enabling the addressability of individual Rep/Cap genes and the exchange of promoters. In addition, suites of synthetic promoters tailored for this purpose were developed. A first experimental attempt towards achieving this goal is presented in the following. A set of Rep/Cap plasmids was created, each containing options for a stronger and weaker promoter for the three individually controllable genes of v5.4. The design of these Rep/Cap plasmids and their respective VG titres produced in a design of experiments (DOE) approach can be observed in Figure 45 A. Transfections were carried out using Helper plasmids 4.1 to 4.4, which also featured two levels of promoter strength for each of the two mRNA-producing Helper genes, *E2A* and *E4orf3+6+6/7*. The promoter setups of the plasmids and the respective promoters for each of the individually controllable genes, along with their strengths, are listed in tables of Figure 45 B.

A



B

Helper	E2A	E4
4.1	-	-
4.2	+	+
4.3	+	-
4.4	-	+

Rep/Cap	Rep	Cap	smallRep
v8.1	-	+	-
v8.2	-	+	+
v8.3	-	-	-
v8.4	-	-	+
v8.5	+	+	-
v8.6	+	+	+
v8.7	+	-	-
v8.8	+	-	+

Gene	-	+
1. Rep	p5rem (0%)	TP16 (30%)
2. Cap	TP19 (34%)	CRM1 (67%)
3. smallRep	TP11 (21%)	CMV2.01 (87%)
4. E4(orf3+6+6/7)	CRM1+2 (77%)	CMV4.01 (114%)
5. E2A	TP17 (49%)	CRM1+4 (86%)

Figure 45: DOE approach for Rep/Cap and Helper plasmids with two levels of promoter strength for each individually controllable gene. (A) Schematic depiction of Rep/Cap plasmids next to associated VG titres fold changes compared to control Helper plasmid 0.1 combined with Rep/Cap v1.2. Arrows denote promoters, black dots poly A sequences. Genes, 5' and 3'-UTRs, but not promoters and polyA signals are drawn to approximate scale. VG titre data shown as means with error bars indicating standard deviations of three biological replicates. Rep/Cap plasmids v8.1-8.8 were transfected with the Helpers 4.1-4.4 in a DOE approach. Rep/Cap 8.9 (separated by dotted line) was transfected in parallel with the same Helper plasmids, its design and promoter use are depicted, too (TP11tp =10% of CMV). (B) The tables indicate the promoter setups of the individual Helper and Rep/Cap plasmids. Additionally, the two promoter levels (high + / low -) of each gene are translated to the respective promoters that were utilised and their relative strength compared to CMV.

The combination of the Rep/Cap plasmids of the v8.x designs with the Helper plasmids of the 4.x designs revealed that these plasmid combinations were incompatible for high titre rAAV production. While all four Helper plasmids (4.1-4.4) achieved very high VG titres with Rep/Cap plasmid v1.2 (Figure 43), ranging between 2.0 and 2.86-fold of the control (Helper 0.1+Rep/Cap v1.2), the maximum VG titre fold change achieved with the best Rep/Cap 8.x design (v8.9) was 0.46-fold

respective to the control. Unfortunately, no clear trends were observed, and no particular setup outperformed the others. The data was subjected to analysis using the DOE platform tool Desice.io, and the main effects of the experiment can be summarized as follows: (i) very marginal VG titre increases result from weaker promoter option for *Rep78/68*, (ii) *cap* transcription activity should probably be increased further (>67% CMV), (iii) the weaker promoter of the second *Rep52/40* copy was the better choice or 87% CMV were too strong, (iv) the stronger promoter was marginally better for *E2A*, (v) E4 promoter changes had no influence. It is to note that all differences are very slim and only the factors Cap and small Rep had a significant influence. Nevertheless, the obtained results of the analysis confirm previous trends.

(i) A weaker expression of large Reps aligns with previously published reports. TP16, with its similarly strength to p5, was found to be too strong, consistent with previous studies and our own results that increased rAAV titres by weakening the p5 promoter (Emmerling et al., 2016; J. Li et al., 1997; Qiao et al., 2002; Xiao et al., 1998). The positive results of Rep/Cap v8.9, employing a medium-strength promoter option with TP11tp, in between the two DOE level options, shows likewise to the insignificance of most parameters that the model results of this experiment should only be considered as possible trends, calling for further improvement and optimisation experiments with more promoter choices per gene, which would be much more complex, but presumably more insightful.

(ii) Increasing the transcription of VPs and other *cap* products appears logical for enhanced rAAV production. Other studies also recognized capsid production as a limiting step and exchanged the p40 promoter accordingly (Farris & Pintel, 2008; Vincent et al., 1997). It also makes sense that the weaker promoter was worse for *cap*, since the even less transcriptionally active p40 promoter was already inefficient for high rAAV titre production with the split Rep/Cap system. Given the absolute importance of balanced *cap* transcription for efficient rAAV production, conducting further testing of various promoters dedicated to this gene to optimise VG titres and full/empty capsid ratios is advised for future experiments.

(iii) Increased transcription of *Rep52/40* and consequently higher abundance of small Reps was expected to boost full/empty ratios even more than the addition of the second *Rep52/40* already did. However, the CMV2.01 promoter with its 87% activity of CMV seemed to be too strong. Fine-tuning of the promoter strength for the second *smallRep* copy, between p5's 31% and CMV2.01, is suggested in future experiments.

(iv) Similar to the swaps of the endogenous E4 promoter for a CMV, no substantial changes in VG titre occurred with the two CMV-derived promoter variants. For *E4orf3+6+6/7* a sufficiently high expression seems to be enough. Therefore, reducing the promoter strength for *E4orf3+6+6/7* in future experiments may be of interest to see how weak a promoter for this gene can be until an effect on VG titre can be observed.

(v) Sufficiently strong transcription of *E2A* is crucial for high rAAV titre production, as already reported in previous chapters. The promoter chosen to produce DBP should preferably be at least as strong as the utilised upper-level option CRM1+4 (86% of CMV activity). Additional experiments for promoter fine-tuning are advised to use promoters with strengths only slightly below 80% of CMV and up to the strongest options available.

Generally, the choice of employing very strong promoters, and multiple such promoters for various genes in the plasmid system, may not be ideal. While strong, often viral, consecutive promoters are effective for pharmaceutical productions that involve overexpressing one or two genes, a complex multi gene expression system like rAAV necessitates at least four or five promoters, in addition to the one for the GOI. Inevitably, competition for transcription factors (TFs) arises, potentially leading to deregulation and depletion of TFs required by crucial cellular components, such as pol δ, the MCM complex, and components of the transcriptional machinery, all essential for AAV replication (Karreth et al., 2014; Munteanu et al., 2010; Samulski & Muzyczka, 2014; Zabet & Adryan, 2013). Especially when these promoters are composed of similar TFREs and utilise the same pool of TFs, competition can disrupt the functionality of the system. This experiment likely encountered both issues, too strong promoters and TF competition. The observation that stronger promoter choices often outperformed others for specific genes may be attributed to competition among the promoters. Moreover, the CRM-CMV library and the synthetic TP promoter library share similar TFRE compositions. Although these are diversly positioned, the TF pool the promoters address is very much the same. It is therefore hypothesised that a more diverse promoter library could be of greater potential for a multi gene expression system of this complexity. Additionally, experiments with such a library might benefit from the utilisation of weak to medium-strength promoters, allowing more transcriptional activity for host cellular components, which are also essential for rAAV proliferation.

In addition to a more diverse TFRE variety and a larger, more diverse synthetic promoter library, alternative promoter functionalities and mechanisms should be considered. The promoters used for the expression of *cap*-derived proteins serve as an example of why the decoupling of promoter and product interactions, as done in

this study with synthetic promoters, may not always be preferable. The truncated CMV promoter, with 80% of the full CMV's activity, mainly produced empty capsids (Figure 32). On the other hand, the slightly weaker promoter CRM1, with 67% CMV activity, possibly did not provide sufficient strength for rAAV production. With a promoter constitutively producing VP and capsids from about 6 hours post transfection until harvest, the issue of empty capsids may persist, considering the AAV production model of Nguyen et al. (2021). The model study suggests a decoupling of AAV replication and capsid production, which can be done with the split Rep/Cap plasmid system developed by our group. Time-shifted transfection of a *cap* only plasmid could be used, but retransfection of cells might turn out problematic. Alternatively, the use of inducible promoters or the creation of promoter circuits should be considered to design a system where components are produced in favourable stoichiometries and optimised quantities as needed, while being shut off when not required or beneficial (Das et al., 2016; Kitada et al., 2018; Lienert et al., 2014). The CMV promoter analysis and the created promoter libraries serve as valuable foundations for such endeavours (Johari et al., 2022). Further development involving diversification and combination with concepts of promoter engineering and synthetic biology could significantly enhance the control and regulation of gene expression in this highly complex system.

Furthermore, the high number of empty capsids produced with strong promoters for *cap* transcription does not necessarily indicate a negative outcome. It rather demonstrates that the cells possess a greater production capacity than what the currently achieved rAAV VG titres reflect. Enhancing packaging efficiency is crucial to unlock this potential. Strategies to achieve this goal may include modifying heterologous or synthetic promoters to exert more process control, improving intron splicing, or optimizing Rep proteins through protein engineering. The optimisation of individual component concentrations at the molecular genetic level presents a promising concept to improve rAAV production, as demonstrated by the presented experiments. The advancement in synthetic promoter development could play a key role in executing and refining this concept.

6.5.4 Use of Small Molecule Chemical Additives for Augmentation of rAAV Production

Another important aspect of rAAV production that can significantly influence the expression levels of proteins and should therefore ideally be considered and developed in parallel is process development, specifically cultivation conditions and media optimisation. This aspect of rAAV production was also addressed in this project. However, this work was performed by Dr. Joseph Scarrott and only minor experimental work of this package was carried out by me. A journal article titled

"Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives" resulted out of this work. As with the published CMV promoter study, the main intellectual property of this process development focused publication does not belong to me and my part regarding this published work was limited to some experimental work, scientific discourse about results and interpretations, as well as reviewing and editing of the manuscript. Additionally, a follow-up experiment based on results of this publication was performed afterwards. Therefore, the abstract of the article is being presented here to give an overview of the relevant work. The full article is attached in 8.2.3.

Increased Recombinant Adeno-Associated Virus Production by HEK293 Cells Using Small Molecule Chemical Additives

- Joseph M. Scarrott, Yusuf B. Johari, Thilo H. Pohle, Ping Liu, Ayda Mayer, David C. James

"Recombinant adeno-associated virus (rAAV) has established itself as a highly efficacious gene delivery vector with a well characterised safety profile allowing broad clinical application. Recent successes in rAAV-mediated gene therapy clinical trials will continue to drive demand for improved rAAV production processes to reduce costs. Here, we demonstrate that small molecule bioactive chemical additives can significantly increase recombinant AAV vector production by human embryonic kidney (HEK) cells up to three-fold. Nocodazole (an anti-mitotic agent) and M344 (a selective histone deacetylase inhibitor) were identified as positive regulators of rAAV8 genome titre in a microplate screening assay. Addition of nocodazole to triple-transfected HEK293 suspension cells producing rAAV arrested cells in G2/M phase, increased average cell volume and reduced viable cell density relative to untreated rAAV producing cells at harvest. Final crude genome vector titre from nocodazole treated cultures was >2-fold higher compared to non-treated cultures. Further investigation showed nocodazole addition to cultures to be time critical. Genome titre improvement was found to be scalable and serotype independent across two distinct rAAV serotypes, rAAV8 and rAAV9. Furthermore, a combination of M344 and nocodazole produced a positive additive effect on rAAV8 genome titre, resulting in a three-fold increase in genome titre compared to untreated cells."

Building on the promising outcomes, yielding nearly three-fold increased rAAV VG titres by adding nocodazole and M344 to the commonly used Helper 0.1 + Rep/Cap v1.2 transfection system, the two small molecule chemicals were also tested with the best Rep/Cap and Helper plasmids created during the presented work. The results of the chemical addition 4 hours post transfection with

the highest rAAV producing Rep/Cap and Helper plasmids combinations v1.2+4.3, v3.7+4.3 and v5.4+4.3 are presented in Figure 46.

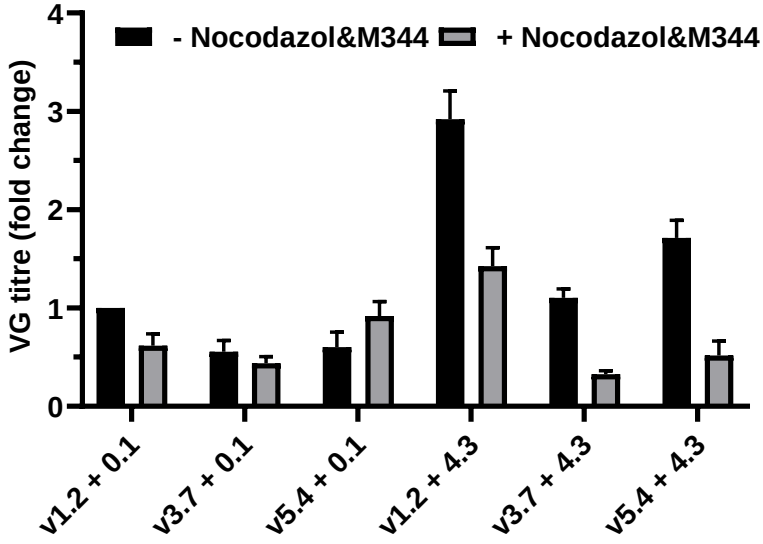


Figure 46: Addition of small molecule chemical additives nocodazole (4 µM) and M344 (2.5 µM) 4 h post transfection to HEK293 cultures transfected with different combinations of Rep/Cap (v1.2/ 3.7/v5.4) and Helper (0.1/4.3) plasmids for rAAV8 production. Results of rAAV VG titre analysis 72 h post transfection presented as mean and standard deviation of three individual biological replicates of cultures with (grey) and without chemical addition (black).

Surprisingly, the resulting rAAV VG titres with the addition of nocodazole (4 µM) and M344 (2.5 µM) to different Rep/Cap and Helper plasmid combinations did not align with expectations. While the crude VG titres of the combination of Rep/Cap v1.2 and Helper 0.1 previously increased nearly three-fold due to the addition of nocodazole and M344, they decreased to 0.6-fold of the same plasmid combination without chemical additives. The only observed increase in VG titre was with the combination of Rep/Cap v5.4 and Helper 0.1. This result was unexpected, considering the consistently positive effects observed by the addition of these two chemicals. Additionally, the lowered titres of the combinations of Helper 0.1 and Rep/Cap v3.7/5.4 showed lower titres without chemical addition instead if an about 1.2-fold increase that was previously observed. Otherwise, titres of transfections with Helper 4.3 were within the expected ranges, with the combination of Rep/Cap v1.2 and Helper 4.3 reaching the highest titres of about 3×10^{11} vg/mL and a fold change to the control of 2.9, respectively. The combination of the best Helper plasmid 4.3 with the best Rep/Cap plasmids v3.7 and v5.4 demonstrated, once again, incompatibility of the high titre producing Helper plasmids with these Rep/Cap plasmids, resulting in VG titres decreased by 62% and 41% compared to Rep/Cap v1.2. The stronger, heterologous promoters used for *E2A* and *E4orf3+6+6/7* in the Helper plasmid, in combination with the CMV promoter controlling cap

transcription, seemed to compete in a way that negatively affected rAAV output. Virus genome titres with chemical addition followed a similar trend as without.

The observed reduction in titre could have two underlying reasons. Firstly, the use of a different cell line may be a contributing factor. Although derived from the same parental cell pool, the HEK293 cells used from the reshuffled Helper gene plasmid experiments onward were monoclonal cells provided by the industrial partner REGENXBIO at that time. As it is known that small molecule addition effects are cultivation conditions and media, but also cell line dependent, it could simply be that the effects of nocodazole and M344 are not as pronounced in the monoclonal cell line. However, their general mechanistic effects would not be reverted. Nocodazole's improvement of transfection efficiency that was also shown in CHO cells, and its general cell cycle modulation, should genuinely be beneficial for rAAV production (Tait et al., 2004). With regard to the latter, viral components modulate the cell cycle similar to nocodazole, which is a main function of helper and Rep proteins. Studies have shown increased rAAV production, due to the cell modulation into G2/M phase (Barnes et al., 2021; Berk, 2005; Berthet et al., 2005; Franzoso et al., 2017; Meyer et al., 2017; Raj et al., 2001; Saudan et al., 2000; White, 2001b). This so to speak helper function for rAAV production should be universal, but the proliferation reduction of the molecule points out that adequate dosing is required (Meyer et al., 2017). Thus, observed VG titre decreases can be attributed to a potential reduction in VCD, likely resulting from the addition of nocodazole and possibly M344. Unfortunately, due to issues with the cell counter at the time, VCD measurements were not possible. It is essential to include VCD evaluation in future experiments of this nature. Secondly, the chemical stock used might have contributed to the outcome. The same stock of chemicals, dissolved in DMSO and stored at -20°C, was used as in previous experiments. The time span between the experiments could have meant a possible breakdown over time. Although there was no time left to order new chemical stocks and repeat the experiment, it is evident that the beneficial effect of these small molecules on rAAV production is transferable to different cells and plasmid systems. However, a change in cell line and plasmid system warrants a re-evaluation of chemical concentrations and application timing, as the efficacy and application of these chemicals depend on cell line, process, and product parameters (M. J. Allen et al., 2008; Johari et al., 2015; W. C. Yang et al., 2014).

Nocodazole exerts extensive modulation of the cellular environment, cell cycle and consequently the transcriptional landscape. As plethora of small molecules possess similar cell-modulating capabilities to nocodazole, incorporating chemical screening and using specific chemicals in experiments for transcriptional control is a logical approach. The development of synthetic promoters, new plasmid

variations, and the use of small molecules in joint experiments, coupled with detailed analysis, present a highly complex task but hold great potential for informed and directed engineering of the intricate multi-gene expression system that rAAV production entails.

6.5.5 Summary

- Based on the findings of the co-authored paper "Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells" (Johari et al., 2022), a library of synthetic promoters was created
    - TFRE extensions of CMV-CRM1 were not as strong as expected and showed repressive effects on CRM1
        - TP17 achieved the highest activity at 49.3% of CMV, which is only 0.74-fold the activity of CRM1 (67% activity of CMV)
    - The synthetic promoters covered a wide range of activity, which is useful for transcription titration of the rAAV plasmid system
    - YY1, which repressed CMV, improved synthetic promoters' performance, outperforming YY1-less constructs
- A Helper plasmid was designed to accommodate the exchange of promoters with minimal restrictions on promoter activity and efficient gene expression
    - A first approach used Helper 3.10.1 as the basis for optimising gene composition and expression strength using heterologous promoters (Helper 4.0)
    - Helper 4.0, using convergent gene orientation and double SV40 terminators, showed lower rAAV VG titres than expected
    - Overexpression of E1A or E1 genes resulted in decreased rAAV VG titres, possibly due to imbalanced component stoichiometries or apoptosis induction
- Revised Helper plasmid designs based on Helper 3.8.1 created with synthetic promoters achieved improved rAAV titres (Helpers 4.1 to 4.4)
    - Helper 4.3 performing the best, with the combination of the stronger *E2A* promoter and the weaker promoter option for *E4orf3+6+6/7*
        - Differences between Helper plasmids emphasise the importance of promoter selection for different genes
- Attempted rearrangement of the Rep/Cap plasmid gene orders and orientations based on Rep/Cap plasmids v3.7 and v5.4 diminished rAAV VG titres
    - Utilisation of CRM1 promoter for *cap* and TPcore promoter for *largeReps* potentially with negative impact on VG titres

- The combination of the highly functional Helper plasmids 4.1-4.4 with synthetic promoter containing Rep/Cap plasmids generated similarly low rAAV VG titres as the combination with the poorly performing Helper 4.0
- Despite low titres and many statistically non-significant results, the DOE approach with two promoter strength levels confirmed many previously seen trends
    - First experimental attempt to improve rAAV production by simultaneously exchanging promoters for Rep/Cap and Helper genes using a selection of promoters from the CMV-CRM and synthetic TP promoter libraries
- The use of small molecule chemical additives to enhance rAAV production was explored in the co-authored article "Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives" (Scarrott et al., 2023)
- Results identifying nocodazole and M344 as positive regulators of rAAV8 production with the original plasmid system in HEK293 cells (up to three-fold increase) could not be reproduced in a slightly different cell line with engineered Rep/Cap and Helper plasmids

6.5.6 Conclusion

The synthetic promoter library based on CMV analysis generated a range of promoter strengths, contributing to a wide spectrum of transcriptional activities for precise gene expression regulation. While the synthetic TP promoter library encompasses the lower range of activities, the upper range, extending up to 115% of CMV's activity, is covered by the CMV-derived promoters developed by Dr. Johari. These promoters hold the potential for finely tuning rAAV component quantities, given their subtle strength differences. Further investigations could explore their applicability in rAAV plasmid systems across varying conditions. Nonetheless, a greater TFRE pool should be considered in an expansion of these synthetic promoter libraries to avoid transcription factor competition and depletion. Furthermore, innovative promoter modifications addressing time and concentration-dependent regulation needs of the transient rAAV production system were deliberated and should be considered for future promoter exchanges in Helper and Rep/Cap plasmids.

The Helper plasmid 4.3 embodies the culmination of pivotal findings from this thesis, converging into a design that attains the highest overall rAAV8 VG titre of 3.7×10^{11} vg/mL. This achievement signifies a remarkable 2.86-fold VG titre elevation compared to the initial plasmid, Helper 0.1 (pADΔF6). The incorporation of novel features transforms it into a modernized plasmid, facilitating streamlined

genetic manipulation and enhanced controllability. Although further improvements to the design of the plasmid and its transcriptional control might be possible, the current iteration already delivers a substantial advancement in industrial rAAV production.

In contrast, the Rep/Cap plasmid designs and promoter exchanges made in this part of the work did not yield any improvements. Substantial alterations in sequence design and promoter development are likely necessary for this plasmid. Nevertheless, the separated gene configuration holds great promise for future process optimisation, which could result in notable enhancements in VG titre and product quality.

The general benefits of small molecule additives for rAAV production, as highlighted by remarkable VG titre enhancements resulting from the media supplementation with nocodazole and M344, are highly promising. Future endeavours should involve comprehensive screening for cell-modulating chemicals, providing a cost-effective avenue for boosting production yields. Chemical compounds can be viewed as supplementary Helper functions, which demonstrated to be most efficient in increasing rAAV titres in the current work. Consequently, the combination of likely synergetic small molecule chemicals and genetic engineering should be considered, as it holds the potential to amplify VG titres and enhance process control on various levels.

7 Conclusion and Outlook

This PhD thesis represents a comprehensive investigation into the optimisation of recombinant adeno-associated virus (rAAV) production through systematic genetic engineering and synthetic biology methodologies. The primary research objective was to enhance the production of rAAV through the development of an engineerable and controllable plasmid system, with particular emphasis on the Helper plasmid. By engineering this system, it was aimed to optimise component quantities to increase rAAV virus genome (VG) titres.

While there is potential for further enhancements and promoter fine-tuning is still pending, the thought after plasmid system was successfully created. This novel system enables control of individual component expression, and its modular design simplifies genetic engineering. This new Rep/Cap and Helper plasmid system already exhibited a remarkable 1.8-fold increase in rAAV8 VG titres compared to the initial plasmid system. The combination of the most advanced Helper plasmid with the initial Rep/Cap plasmid resulted in even higher VG titres with a nearly threefold increase compared to the original configuration.

The initial phase of this research involved the modernisation and streamlining of the widely used Helper plasmid pADΔF6. These foundational modifications paved the way for subsequent experiments. Dissection and analysing the Helper plasmid unveiled *L4-33K/22K* in *E2A*'s 5' UTR as a component of the Helper plasmid that serves AAV helper functions and enhances rAAV yields. Furthermore, an optimised subset of *E4* open reading frames was engineered, excluding adverse ORFs and including beneficial set of *E4orf3+6+6/7* for high-titre rAAV production. This detailed analysis of the Helper plasmid genes provided novel insights into their influence on rAAV production, resulting in streamlined and optimized gene variants that are novel in the realm of rAAV production systems. Even though, the combination of these improved Helper genes required additional genetic modifications, they resulted in a highly advantageous Helper plasmid configuration. These structural improvements to the Helper plasmid and its genes represent important knowledge about this plasmid and its genes for rAAV production, as well as valuable improvements. The commercial value of these enhancements is underscored by the inclusion of the *E4* subset findings in the REGNEXBIO owned patent (P. Liu et al., 2023).

To make well-informed design decisions, it would be ideal to know the systems' exact demands of specific component quantities. Implementing these necessary quantities through a top-down engineering strategy would be the ideal course of action. Consequently, experiments involving transfections of virus genes on

different plasmids and variation of their ratios were conducted. Unfortunately, problems with this experiment made this approach of deciphering the system unsuccessful. Furthermore, it must be acknowledged in hindsight that an attempt of unravelling the complex interactions and component demands of rAAV production would necessitate significantly more comprehensive and detailed analytics. As a result, the strategy shifted towards a continued focus on changing and optimizing individual components of the system in an iterative manner before integrating them into a complete system. This bottom-up engineering approach proved to be more successful, just as it was for the endogenous components of the Helper plasmid, as well as for the Rep/Cap plasmid and their modularisations.

The separation of the Helper genes onto different plasmids remained valuable, pointing out their essential roles in rAAV production and foundational insights into component stoichiometry requirements. Notably, *E2A* emerged as the only indispensable Helper gene, highlighting its pivotal role in the current plasmid system. The decomposition of the Helper plasmid, bicistronic helper gene expression, and transcript analysis emphasized the importance of sufficient helper gene expression, specifically of *E2A*, and gave an idea about transcription requirements. The need for strong *E2A* transcription was postulated as an attribute of high-titre rAAV expression. However, mRNA analysis revealed that overall mRNA transcript levels and Helper transcript ratios cannot solely indicate rAAV production capabilities, illustrating the intricate nature of this multi-gene expression system.

The utilisation of stronger promoters for *E4orf* subsets yielded limited benefits yet became crucial when combined with the new Rep/Cap plasmids that employ a truncated CMV promoter for *cap* transcription regulation. Issues related to increased empty capsid production due to the short CMV promoter for VP expression were revealed with the new Rep/Cap plasmids. Titres were restored by using the CMV promoter to regulate *E4* transcription. The effect of TF competition became evident when a CMV promoter was also used for *E2A* transcription, as E4P and multiple CMV promoters all require the same transcription factors. The use of a weaker promoter for *cap* to reduce empty capsid amounts decreased titres similar as the endogenous p40 promoter. If a configuration allowing high-titre rAAV production with lowered *cap* expression cannot be found, the surplus capsid production needs to be regarded as an opportunity to be capitalized upon. Although the strategy of overexpressing MAAP and AAP didn't yield successful outcomes for enhancing rAAV titres in other groups' more common systems, this high VP production system, equipped with the capability for component adjustment, could potentially benefit from the overexpression of these genes. Furthermore, it's

plausible that the strong promoter for *cap* is not solely required for VP production but also for ensuring sufficient AAP or MAAP production. If these genes were expressed independently, a weaker promoter for the VPs might be applicable.

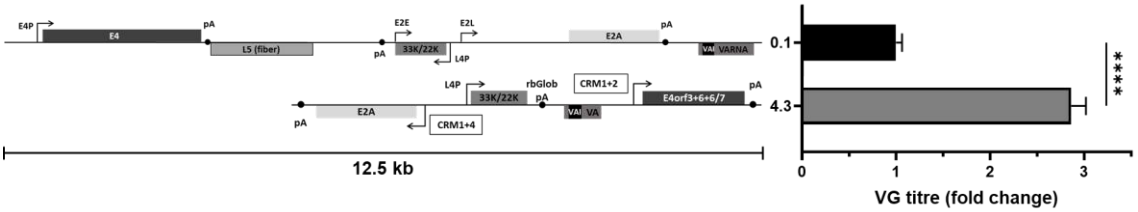


Figure 47: Industry standard and starting point Helper plasmid 0.1 in comparison to the best performing engineered Helper plasmid 4.3, displaying a 45% size reduction (15.8 kb to 8.6kb including the plasmid backbone) and a 2.9-fold VG titre increase. Data displayed as means of three biological replicates with error bars representing their standard deviation. Statistical analysis was performed with an ordinary one-way ANOVA.

The new Helper plasmid design utilised for Helper 4.3 represents the culmination of this work, combining all previous advantageous Helper plasmid concepts and findings, while also incorporating the newly created CMV-based synthetic promoters. Through iterative modifications and necessary genetic engineering adaptations, the amalgamation of the individual Helper findings proved effective, yielding the highest rAAV8 VG titres up to this point with 3.7×10^{11} vg/, which represents a 2.86-fold increase over the parental base plasmid Helper 0.1 (pADΔF6). Additionally, this plasmid design stands out for its adaptability in genetic part exchange, distinguishing it from other Helper plasmids. Its configuration ensures controlled gene expression and transcription controllability using new promoters and potentially other regulatory genetic elements. Future enhancements might involve refining these promoters and polyA signals. As the full *E2A* 5'-UTR was often beneficial for rAAV VG titres, it should be tested to reinstall it, and modify the promoter regardless of its distance. Also, Kozak and codon-optimisations could be considered. Apart from the further minimalization of the *VA RNA* fragment, the study's thorough analysis and subsequently optimisation of the Helper genes imply that no further improvements in this regard are anticipated. Should the whole system utilise only heterologous promoters, it could be tried to eliminate the short *E4orf7* sequence, potentially increasing E4orf6 amounts by abolishing *E4orf6/7* transcripts. Integrating Helper genes from other AAV helper viruses could potentially amplify host cell modulation. The significantly reduced size of the new Helper plasmids may allow such integration without exceeding the dimensions of the original plasmid. The reduction in size, from pADΔF6's 15.8 kb to 8.6 kb (Helper 4.3), could also facilitate production using a different *E. coli* strain, distinct from the currently utilized NEBstable strain. The NEBstable strain was necessary due to plasmid replication issues in DH5α. If these issues were primarily linked to the large

size of the plasmids rather than the viral sequences, the new plasmids, in conjunction with a more efficient bacterial strain, could potentially elevate plasmid yields and consequently mitigate costs. The implementation of the engineered Helper plasmids in industrial transfection processes could diminish upstream costs even more, due to their notable improvements in titres. Helper plasmid 4.3's VG titre increase translates to an approximate cost reduction of $3000 per dose, assuming manufacturing costs of $25000 per dose of 1×10^{14} VG/kg and a UPS cost amount of 35% (Cameau et al., 2019; Lyle et al., 2023).

To enhance VG titres even further by plasmid engineering, currently the greatest potential lies within the Rep/Cap plasmids. The promoter changes made to Rep/Cap plasmid v5.4 resulted in greatly diminished rAAV8 VG titres, similar to the v7.x Rep/Cap plasmids. A re-evaluation of the plasmid design should be considered, as the tandem sequence *rep*→*cap* is possibly not ideal for increased *cap* transcription. Comprehending the reasons behind the difficulties encountered in separating *rep* and *cap* genes within this system, in contrast to the experiences of other groups with their plasmid systems, should stand as a central objective for forthcoming inquiries. Additionally, tweaking the design to enhance performance and reach usability with different promoters is imperative. High capsid production capabilities hold promise, but packaging and expression control need refinement, potentially involving staggered *cap* transcription induction and very finely regulated expression of both Rep species and the different *cap* originated proteins.

The exploration of synthetic promoters and chemical additives showcased promising avenues for boosting rAAV production. The synthetic promoter suites exhibited a spectrum of strengths, enabling precise transcriptional control, while small molecule additives demonstrated potential for enhancing VG titres. The created promoter libraries are valuable resources for future modifications to rAAV production systems, be it transient or stable manufacturing. Particularly, insights from the CMV promoter study provide novel understandings of this renowned transcription regulatory element, with potential applications in recombinant protein production in HEK293 cells for rAAV and other products. Similar to recombinant protein production, the use of small molecules could amplify product titres for rAAV manufacturing, representing a non-viral approach to cell modulation. Utilisation of this cost efficient and effective way of cell and process manipulation should be further explored, not only in parallel but potentially in combination with genetic engineering.

The overall conclusion emphasizes the complexity of the interrelated factors governing rAAV production and the potential for advancements through a combination of genetic engineering, synthetic biology techniques, and chemical

interventions. The work demonstrates that rAAV production is a highly complex biotechnological system, where slight changes on molecular genetic level can exert profound impacts. Achieving balance within this system is critical, as evidenced by the many modifications that did not enhanced VG titres but led to unexpected reductions. The hypothesis of an imbalanced initial system was confirmed by the increased VG titres that were achieved through the alterations of genes, their positioning and promoters. Further fine-tuning of component quantities holds great potential to significantly enhance production capacities and qualities. To gain full control over the system, more fundamental work is imperative. AAV replication is a highly regulated process. While decades of research have contributed to our comprehension of this system, employing novel analytical and bioinformatic tools is crucial for deeper insights into these regulations and pathways. Such knowledge can then be leveraged to create optimised synthetic versions, in the form of controllable circuits, contributing to the development of a truly synthetically engineered rAAV production system.

This project has established solid foundations for more advanced research and has already led to significant improvements in both the plasmid system itself and rAAV production yields.

8 Appendix

8.1 COVID-19: Research Project and Thesis Impact

The COVID-19 pandemic had an impact on all our lives, unfortunately also on the research of this doctoral thesis. Caused by lockdowns in the UK and restrictions from the University, our laboratory was closed from March 2020 until August 2020. Even afterwards, the accessibility of the laboratories was partly restricted and material shortages did their fair share too, delaying research progress. However, in contrast to most other PhD students I was in the lucky position not to be restricted to work only from home during the first six months of the pandemic. Instead, I could work in the lab for nearly the whole period of time.

At the beginning of the pandemic a call for help from the Department of Immunology of the Sheffield Teaching Hospitals NHS Foundation Trust reached the David James Lab to produce the SARS-COV-2 spike protein and optimize the production of this difficult to express trimeric protein with 22 N-glycosylation sites. The hospital's immunology department required the material to invent an ELISA based assay for antibodies against the virus. Basically, this was a first, plate-based, antibody COVID test, when there were no quick lateral flow tests for antigen or antibody detection developed yet. The tests and the biological material for it were required to test hospital staff and keep the hospital running as safe as possible at the beginning of the pandemic. The spike protein is an important antigen of the virus that represents one of the best possibilities for an accurate serological assay, which can declare if the tested person carries antibodies in their blood due to a SARS-COV-2 infection. Preliminary tests of the hospital's immunology department showed greater specificity for the spike protein than just the receptor binding domain (RBD), so this was the protein of choice, although it was more difficult to manufacture.

To produce enough material for testing, previously performed methods did not yield sufficient amounts of spike protein. Also, these methods are very costly and labour intensive. Consequently, a group of mainly six members of the David James Lab, including me, tried to optimize the existing HEK293-transfectionbased system. In a next step, the production was transferred in CHO cells, transfected with plasmids carrying different promoter constructs regulating spike expression, similar to the HEK approach. CHO cells were then used to generate a cell pool with stably integrated, consecutive expressed spike gene copies. The creation of these cells and a process optimisation with them was the primary area of responsibility of Dr. Stephen Jaffé and me, yielding a maximum of 53 mg/L equalling an up to 25 to 50-fold increase compared to results reported by other groups using the original

HEK transfection system (Esposito et al., 2020; Johari, Jaffé, et al., 2021). The produced spike protein was afterwards purified by Dr. Tuck Seng Wong's lab and used for the development of a serological assay by Dr. Ravishankar Sargur and his group. Being part of this project was an amazing and fulfilling experience, albeit it also meant progression in my PhD thesis was put on hold due to this extraordinary situation. The results of the work are published in the journal Biotechnology and Bioengineering. Additionally, a paper about the purification of the manufactured material was composed and uploaded to the biorxiv by the Wong group (Lan Tee et al., 2020). Since, this work during the pandemic is likely one of my most significant contributions to the scientific community with an immediate real-life impact on my life and also represents timewise a significant part of my PhD, the resulting research article is attached below 8.2.1.

8.2 Co-Authored Papers

8.2.1 Production of trimeric SARS-CoV-2 spike protein by CHO cells for serological COVID-19 testing

Received: 5 August 2020 | Revised: 14 October 2020 | Accepted: 27 October 2020

DOI: 10.1002/bit.27615

COMMUNICATION TO THE EDITOR

BIOTECHNOLOGY BIOENGINEERING WILEY

Production of trimeric SARS-CoV-2 spike protein by CHO cells for serological COVID-19 testing

Yusuf B. Johari[1] | Stephen R. P. Jaffé[1] | Joseph M. Scarrott[1] | Abayomi O. Johnson[1] | Théo Mozzanino[1] | Thilo H. Pohle[1] | Sheetal Maisuria[2] | Amina Bhayat-Cammack[2] | Giulia Lambiase[1] | Adam J. Brown[1] | Kang Lan Tee[1] | Philip J. Jackson[1] | Tuck Seng Wong[1] | Mark J. Dickman[1] | Ravishankar B. Sargur[2] | David C. James[1]

[1]Department of Chemical and Biological Engineering, University of Sheffield, Mappin St., Sheffield, UK

[2]Department of Immunology, Sheffield Teaching Hospitals NHS Foundation Trust, Sheffield, UK

Correspondence
David C. James, Department of Chemical and Biological Engineering, University of Sheffield, Mappin St., Sheffield S1 3JD, UK.
Email: d.c.james@sheffield.ac.uk

Funding information
Sheffield Teaching Hospitals NHS Foundation Trust; University of Sheffield

Abstract

We describe scalable and cost-efficient production of full length, His-tagged severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) spike glycoprotein trimer by Chinese hamster ovary (CHO) cells that can be used to detect SARS-CoV-2 antibodies in patient sera at high specificity and sensitivity. Transient production of spike in both human embryonic kidney (HEK) and CHO cells mediated by polyethyleneimine was increased significantly (up to 10.9-fold) by a reduction in culture temperature to 32°C to permit extended duration cultures. Based on these data GS-CHO pools stably producing spike trimer under the control of a strong synthetic promoter were cultured in hypothermic conditions with combinations of bioactive small molecules to increase yield of purified spike product 4.9-fold to 53 mg/L. Purification of recombinant spike by Ni-chelate affinity chromatography initially yielded a variety of co-eluting protein impurities identified as host cell derived by mass spectrometry, which were separated from spike trimer using a modified imidazole gradient elution. Purified CHO spike trimer antigen was used in enzyme-linked immunosorbent assay format to detect immunoglobulin G antibodies against SARS-CoV-2 in sera from patient cohorts previously tested for viral infection by polymerase chain reaction, including those who had displayed coronavirus disease 2019 (COVID-19) symptoms. The antibody assay, validated to ISO 15189 Medical Laboratories standards, exhibited a specificity of 100% and sensitivity of 92.3%. Our data show that CHO cells are a suitable host for the production of larger quantities of recombinant SARS-CoV-2 trimer which can be used as antigen for mass serological testing.

KEYWORDS
bioproduction, CHO cells, COVID-19, SARS-CoV-2, serological assay, spike trimer

Yusuf B. Johari and Stephen R. P. Jaffé contributed equally to this study.

1 | INTRODUCTION

Immune response represents the first line of defense against severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2) infection that has caused the coronavirus disease 2019 (COVID-19) pandemic. The spike glycoprotein that protrudes from the surface of the virus is highly immunogenic with the receptor-binding domain (RBD) being the target of many neutralizing antibodies (Yuan et al., 2020). Utilizing a stabilized version of the full-length SARS-CoV-2 spike protein, a very robust and accurate serological enzyme-linked immunosorbent assay (ELISA) for antibodies in patient sera has recently been developed (Amanat et al., 2020) and approved for use by the US FDA. However, very low production titers (1–2 mg/L) of the spike trimer were reported using the human embryonic kidney (HEK) Expi293 Expression system (Esposito et al., 2020; see Figure S1), therefore effectively limiting its widespread utilization as a preferred antigen in serological assays for COVID-19. The low production titer is not surprising considering that the SARS-CoV-2 spike is a large homotrimer (~670 kDa) with 22 *N*-linked glycosylation sites per monomer (Watanabe et al., 2020). In this study, using improved vector engineering and production process strategies we describe the development of a stable recombinant spike manufacturing platform utilizing Chinese hamster ovary (CHO) cells as a preferred production host. Transient expression was initially employed to

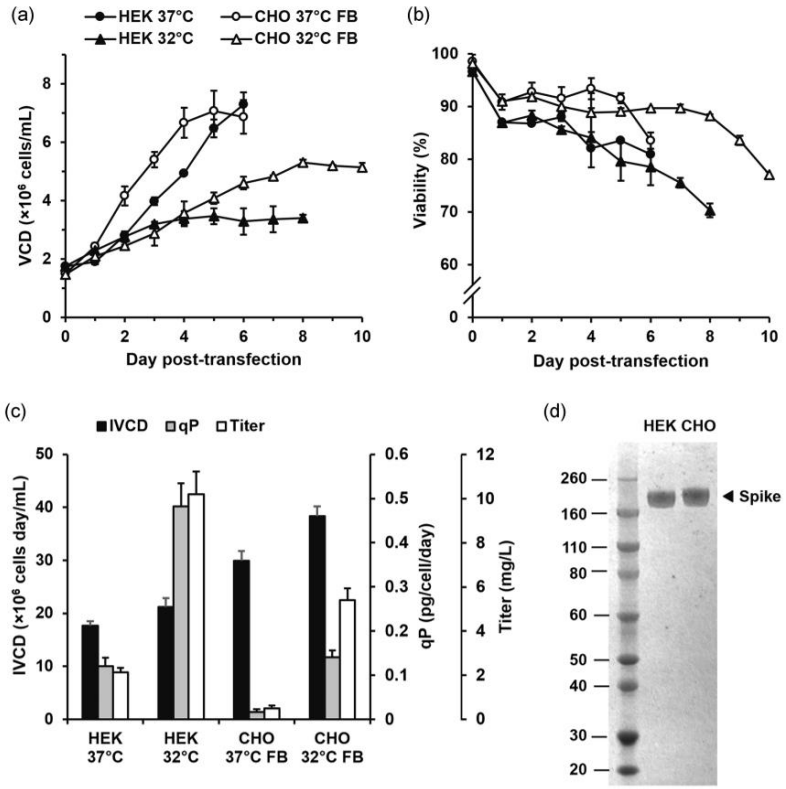


FIGURE 1 Transient production of recombinant spike in HEK and CHO cells. HEK Expi293F cells and CHO-S cells were transfected with plasmids encoding spike gene using PEI under optimized conditions and cultured at 37°C or 32°C. CHO cultures were fed every 2 days with 5% v/v EfficientFeed B. (a) Viable cell density and (b) cell viability posttransfection (>70%). (c) IVCD, recombinant qP and purified spike titer, quantified using Bradford assay. We note that quantification using the A280 method produced ~25% higher titers. (d) Coomassie-stained reducing SDS-PAGE gel of purified HEK and CHO cell-derived spike (~200 kDa) based on the improved imidazole gradient elution method. CHO, Chinese hamster ovary; HEK, human embryonic kidney; IVCD, integral of viable cell density; PEI, polyethyleneimine; SDS-PAGE, sodium dodecyl sulfate polyacrylamide gel electrophoresis

fast-track production of recombinant protein to enable biophysical analyses and early clinical evaluation in serological assays, as well as to evaluate product manufacturability and refine production process and purification conditions.

We have previously shown that for difficult-to-express (DTE) proteins, transient production processes need to be tailored to negate the protein-specific negative effects of recombinant gene overexpression in host cells (e.g., unfolded protein response [UPR] induction, limited cell-specific productivity [qP]; Johari et al., 2015). Using the plasmid construct from the Krammer Laboratory (see Section 2 for details), HEK Expi293F and CHO-S cells were transiently transfected with the CAG-driven expression plasmid using polyethyleneimine (PEI) at an optimal gene dosage for spike production in both cases (data not shown). Further, we utilized a mild hypothermic condition, an effective process engineering intervention for production of DTE proteins (e.g., Estes et al., 2015; Johari et al., 2015) and to extend culture longevity (Figure 1a,b). As shown in Figure 1c, the qP of HEK cells increased 2.4-fold from 0.20 to 0.48 pg/cell/day when the culture temperature was lowered from 37°C to 32°C. Additionally, the prolonged batch culture duration at 32°C enabled a 4.1-fold increase in titer, yielding 10.2 mg/L of purified spike. Greater enhancement was observed with CHO cells where mild hypothermia resulted in an 8.5-fold higher qP than that at 37°C, and a further increase in titer (10.9-fold, 5.4 mg/L) was obtained via increased cell accumulation. We anticipate that improved CHO systems (e.g., ExpiCHO-S cell line and ExpiCHO medium) as well as co-expression of genetic effectors and chemical chaperone addition (Cartwright et al., 2020; Johari et al., 2015) would significantly increase spike transient production in CHO cells. The CHO-derived spike exhibited a monomeric molecular mass of ~200 kDa by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) (Figure 1d) and a trimeric mass of ~670 kDa was measured using analytical size exclusion chromatography (Figure S2). The material was further validated using peptide mapping in conjunction with mass spectrometry analysis (Figure S3). Critically, the preliminary COVID-19 antibody serological test demonstrated its suitability for the ELISA (data not shown) thus permitting the development of CHO stable production platform.

For DTE proteins, very low yielding transient expression systems can be an early indication of reduced stable production (Mason et al., 2012), where particular engineering strategies may be required to obtain stable cells with desirable production characteristics. To generate CHO cells stably expressing recombinant spike trimer, we tested two in-house CHO synthetic promoters, namely 40 RPU (~90% CMV activity) and 100 RPU (~220% CMV activity) promoters (see Brown et al., 2017; Johari et al., 2019). Although the use of extremely strong promoters may be counterintuitive for DTE proteins, we reasoned that only those transfectants harboring a sub-UPR threshold productivity, and thus capable of proliferation would survive. Thus, if cell proliferation attenuation and apoptosis occur as ER functional capacity is exceeded, this is a condition that would directly deselect poorly performing stable transfectants. The promoters and spike gene were inserted into a vector construct encoding glutamine synthetase (driven by an SV40 promoter) and the electroporated CHO-S host cells were subjected to a single round of selection at 25 or 50 μM methionine sulfoximine (MSX), using suspension culture. After 19 days, recovered CHO cell populations were screened for the ability to produce spike in 3-day batch culture (Figure 2a). These data showed that transfectant pools derived from genetic constructs harboring the strong 100 RPU promoter expressed recombinant spike whereas those using the weaker 40 RPU promoter did not. More stringent selection conditions (50 μM MSX) yielded transfectant pools exhibiting lower productivity. Accordingly, CHO cell pools harboring the 100 RPU promoter under 25 μM MSX were taken forward for the manufacturing process.

To rapidly produce recombinant spike, stable transfectant pools (rather than clonally derived populations) were employed. Based on

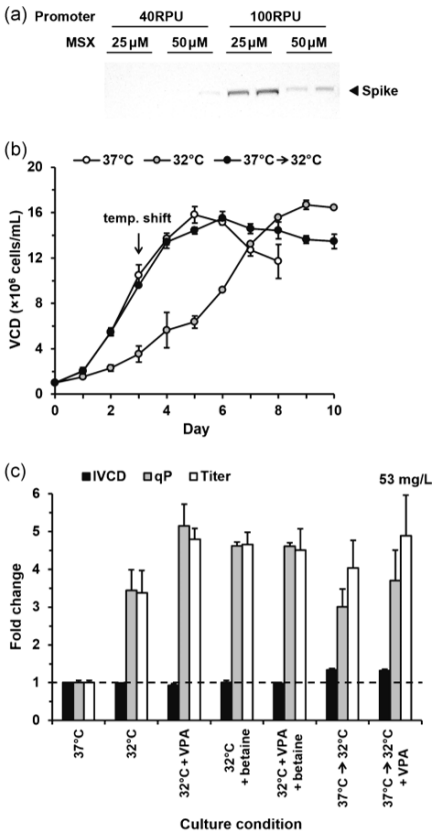


FIGURE 2 (See caption on next page)

CHO transient process data (Figure 1), we tested the hypothesis that an optimal 10-day fed-batch stable production process could be executed at 32°C and further enhanced by chemical chaperone additives (e.g., Johari et al., 2015). We compared this strategy to an alternative approach utilizing culture temperature shift to achieve maximal cell density (biphasic), as well as constant 37°C as control. These data are shown in Figure 2b,c, while the screening data for eight small molecule chemical additives are shown in Figure S4. Compared to the 37°C control culture, hypothermia resulted in a clear (33%) initial reduction in cell-specific proliferation rate over the first 5 days of culture (Figure 2b). However, the qP of the latter was 3.4-fold higher over control and addition of valproic acid (VPA) at Day 6 further enhanced qP 5.1-fold, yielding 51 mg/L of spike after purification by immobilized metal affinity chromatography (IMAC; Figure 2c). Similar enhancement was observed with betaine although there was no synergistic effect when the two molecules were utilized together. Reduction in culture temperature after 3 days culture improved the integral of viable cell density (IVCD) 1.4-fold and when combined with VPA addition at Day 4, 53 mg/L of spike was attained after IMAC purification. These data demonstrate that the optimal process engineering intervention for recombinant spike production identified for rapid transient gene expression was generally translatable to the stable production process. Further, as low-level, sub-UPR threshold expression is likely required to permit adequate cell growth (Figure S1), we reasonably expect that application of mammalian inducible expression technology (e.g., cumate; Poulain et al., 2017) to switch on spike production using an intensified biphasic culture system would be particularly useful to maximize stable production.

FIGURE 2 Development of a stable production platform for SARS-CoV-2 spike in CHO cells. (a) Generation and analysis of CHO stable transfectant pools expressing recombinant spike under the control of synthetic promoters. CHO-S cells were electroporated in duplicate with plasmids containing a GS gene driven by an SV40 promoter and a spike gene driven by either a 40 or 100 RPU synthetic promoter, followed by selection in glutamine-free media containing 25 or 50 μM MSX under suspension condition. Recovered cell pools were assessed for their ability to express spike in 3-day batch culture by Western blot analysis. Figure shown is a representative Western blot analysis of two technical replicates. (b) Cells from the best performing pools in A were inoculated and cultured at 37°C, 32°C, or 37°C with a shift to 32°C at Day 3. Cultures were fed every 2 days with 5% v/v EfficientFeed B. (c) Comparison of the fed-batch culture production performance without or with a chemical addition (chemical screening data are shown in Figure S4).1 mM VPA and/or 12.5 mM betaine were added at Day 4 for the biphasic cultures or at Day 6 for the 32°C cultures. Purified spike titer was quantified using Bradford assay. Data are normalized with respect to culture at 37°C without any chemical addition. Data shown are the mean ± *SD* of two independently generated stable pools each performed in duplicate. CHO, Chinese hamster ovary; MSX, methionine sulfoximine; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

To purify spike protein from culture supernatant, IMAC was initially performed using a step-elution of 250 mM imidazole according to Stadlbauer et al. (2020). Figure 3a shows SDS-PAGE of eluted proteins, and reveals the presence of protein impurities not derived from recombinant spike (Figure 3b), which were identified using tandem mass spectrometry as CHO host cell-derived proteins (HCPs). While all of the identified extracellular HCPs have previously been shown to be present in CHO cell culture supernatant (Park et al., 2017), HSPG in particular has been reported to occur in CHO cells at relatively higher level than HEK cells (Goey et al., 2018; Lee et al., 2016), illustrating the need for an improved purification method especially at high IVCDs. To increase recombinant spike purity, a revised gradient elution profile up to 250 mM imidazole was implemented. As shown in Figure 3c, HCPs were eluted at a lower imidazole concentration than recombinant spike, permitting recovery of high purity (>95%) product for use in serological assays (Figure 3d).

COVID-19 antibody tests would help reveal the true scale of the pandemic in a population and the persistence of immunity, whether vaccines (many of which are based on the production of neutralizing antibodies against spike protein) designed to protect from infection are effective, as well as identify highly reactive human donors for convalescent plasma therapy. The CHO-spike anti-SARS-CoV-2 ELISA was developed based on the Krammer Laboratory's assay, and validated to ISO 15189 Medical Laboratories standards. Initially, we tested a panel of 234 negative samples taken pre-COVID-19 outbreak (June–August 2019) and 26 positive samples taken during the COVID-19 outbreak (≥15 days post-positive polymerase chain reaction [PCR] test). ELISAs were performed by 1/20 dilution of the individual serum samples and the cut-off index of 1.4 was determined using the cut-off OD value (ROC curve with 100% specificity) and the negative control. In this particular evaluation, the assay had an overall specificity of 100% and sensitivity of 92.3% as illustrated in Figure 4a. To establish the reproducibility of the ELISA, positive samples were tested on 5 separate assays over 2 days at three different dilutions to determine the inter-assay variations. The data (Figure 4b) show that the assay performed within the standard range for precision with inter-assay %CV of ≤5%. To be able to interpret serosurveys correctly, the ELISA was evaluated for potential cross-reactivity from individuals with other medical conditions where zero positives were observed in all cases (Table S1).

Overall, our work serves as an exemplar for a development process of characteristically difficult-to-express spike manufacturing platform utilizing CHO cells. This itself is a significant and useful finding, as many DTE recombinant proteins cannot be produced using this industry standard production vehicle—for example, a recent study reported that for over 2200 human genes encoding secreted proteins expressed in CHO cells, almost 50% did not yield target protein (Uhlen et al., 2018). On the other hand, the spike production in HEK cells was highly dependent on the very expensive Expi293 medium (we note that spike production using FreeStyle 293 medium resulted in an even lower titer [<40%]; data not shown). While it is potentially easier to produce a "monomer," the native

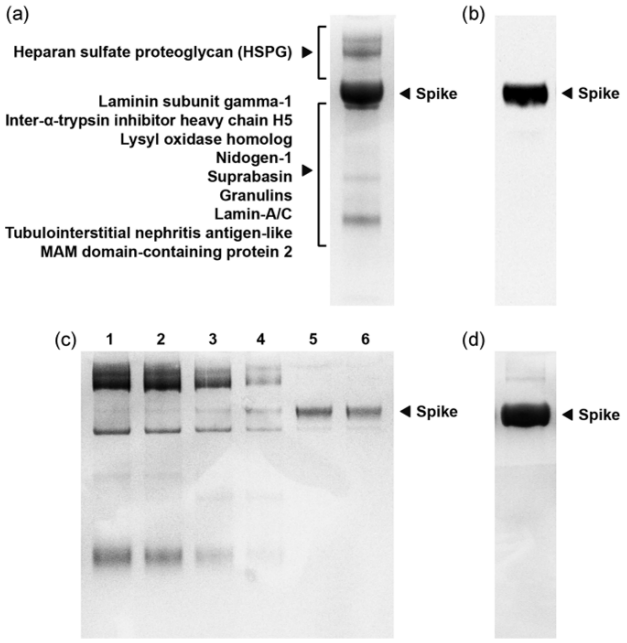


FIGURE 3 Optimization of affinity chromatography purification strategies for spike protein using HisTrap columns. (a) Coomassie-stained gel of the initial purification strategy of spike utilizing the method from Stadlbauer et al. (2020) with associated impurities identified using tandem mass spectrometry. (b) Assessment of spike sample shown in A by Western blot analysis. (c) Gradient elution of spike protein starting from 10 mM imidazole up to a final concentration of 250 mM imidazole. Lanes 1–6: 115, 125, 135, 145, 180, and 190 mM imidazole, respectively. (d) Purified spike from optimized step elution affinity chromatography

trimeric structure of the spike protein on the surface of SARS-CoV-2 is rationally more desirable for optimal serological assay and research. Indeed, further modifications on trimeric spike's molecular architecture to improve its stability and production titers have recently been reported (Hsieh et al., 2020; Stuible et al., 2020; Xiong et al., 2020). Our data demonstrate how process and vector development could complement protein engineering efforts particularly where rapid product generation is required.

With 1 mg of spike providing serological assays for approximately 3500 patient samples, the rapid, scalable transient platform

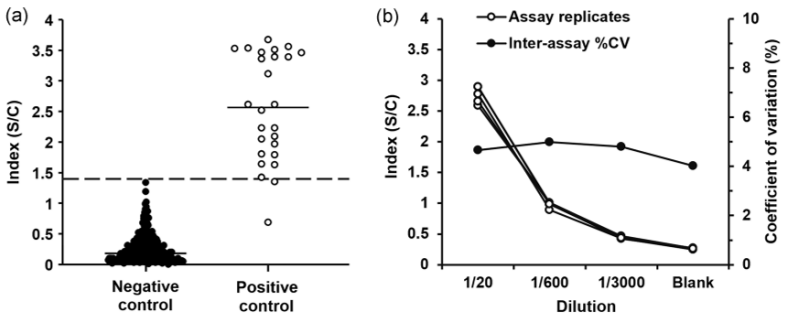


FIGURE 4 Evaluation of CHO-spike anti-SARS-CoV-2 ELISA. (a) 234 negative serum samples (taken pre-COVID-19 outbreak) and 26 positive serum samples (taken ≥15 days post-positive PCR test) were used to evaluate the assay performance, yielding an overall sensitivity of 92.3% anti-SARS-CoV-2 antibodies. (b) To determine the assay precision, one serum sample was assayed in triplicates at five separate times over 2 days (n= 15). COVID-19, coronavirus disease 2019; ELISA, enzyme-linked immunosorbent assay; PCR, polymerase chain reaction; SARS-CoV-2, severe acute respiratory syndrome coronavirus 2

was adequate for local population antibody tests and research studies. To enable large, constant clinical supply of spike, we showed that it was possible to generate CHO stable transfectants expressing the very complex glycoprotein, while high titers could be achieved via a combination of process engineering approaches designed for both high qP and cell biomass accumulation. The refinement of the IMAC affinity purification process permitted greatly enhanced purity of the CHO spike product following an extended 10-day culture, ensuring suitability for use in serological immunity testing. The assay has been implemented at local hospitals with ~7200 staff tested (as of July 31, 2020) which resulted in ~16% positive COVID-19 antibody detection, thus supporting the global effort to limit and mitigate the impact of SARS-CoV-2. Furthermore, it is highly likely that the cell and process engineering interventions designed for SARS-CoV-2 spike production is generically applicable to spike from different coronavirus strains.

2 | MATERIALS AND METHODS

2.1 | Cell cultures and chemical chaperones

Expi293F cells were cultured in Expi293 Expression medium (Thermo Fisher Scientific) in Erlenmeyer flasks maintained at 37°C, 125 rpm under 8% CO_2, 85% humidity. CHO-S clonal isolate cells (C1-80; Fernandez-Martell et al., 2018) were cultured in CD CHO medium (Thermo Fisher Scientific) supplemented with 8 mM L-glutamine maintained at 37°C, 140 rpm under 5% CO_2, 85% humidity. Cells were seeded at 2×10^5 viable cells/ml and were subcultured every 3–4 days. Cell viability and VCD were measured using the Vi-CELL XR (Beckman Coulter). The IVCD was calculated as follows:

$$\mathrm{IVCD} = \left(\frac{\mathrm{VCD}_{t-1} + \mathrm{VCD}_t}{2} \times \Delta t\right) + \mathrm{IVCD}_{t-1}, \quad (1)$$

where t is the time point (day). VPA, NaBu, dimethyl sulfoxide, glycerol, betaine, trimethylamine *N*-oxide, and proline were obtained from Sigma while tauroursodeoxycholic acid was obtained from Merck.

2.2 | Transient production in HEK and CHO cells

pCAGGS plasmids encoding the stabilized full-length SARS-CoV-2 spike protein trimer (polybasic furin cleavage site removed alongside K986P and V987P substitutions, P1213 addition of thrombin cleavage site, T4 trimerization domain and 6×His-tag; Amanat et al., 2020) or RBD were provided by the Krammer Laboratory (Icahn School of Medicine at Mount Sinai). The plasmids were amplified and purified using QIAGEN Plasmid Plus kit (Qiagen). For the optimized Expi293F transfection, cells were grown to 1.75×10^6 cells/ml, centrifuged and resuspended at a density of 3.5×10^6 cells/ml, followed by sequential addition of 0.85 μg of DNA and 2.55 μl of PEI MAX (each pre-diluted in 10 μl of 150 mM NaCl) per million cells. At 24 h posttransfection, the cells were diluted 2× by adding fresh medium, and where applicable culture was shifted to 32°C. For CHO transfection, cells were seeded one day before transfection and grown to 1.5×10^6 cells/ml. For every 1.5×10^6 cells, 1.3 μg of DNA and 4.55 μl of PEI MAX (each pre-diluted in 15 μl of 150 mM NaCl) were combined and incubated at room temperature (RT) for 2 min before being added into culture. Where applicable, culture was shifted to 32°C at 4 h posttransfection. For fed-batch production, 5% v/v CHO CD EfficientFeed B was added at Days 2, 4, 6, and 8.

2.3 | Generation of stable CHO pools and fed-batch production

A stable vector containing an SV40 promoter-driven GS gene was provided by AstraZeneca. The spike gene was cloned by PCR, inserted into the vector downstream of 40 or 100 RPU synthetic promoter (Brown et al., 2017) and the plasmid constructs were confirmed by DNA sequencing. 10×10^6 cells per cuvette were electroporated with 7 μg linearized DNA using Cell Line Nucleofector Kit V system (Lonza) and transferred to a TubeSpin containing 10 ml glutamine-free culture medium with the addition of 25 or 50 μM MSX after 48 h. The cells were left to recover under suspension conditions and recovered pools were cryopreserved when the cell viability reached >90%. For fed-batch production, 5% v/v CHO CD EfficientFeed B was added at Days 2, 4, 6, and 8.

2.4 | Western blot analysis

Proteins in culture supernatant were precipitated by TCA/DOC, resuspended in LDS loading buffer with BME and heated to 70°C. SDS-PAGE was performed using 4–12% NuPAGE Bis-Tris gels and resolved proteins were transferred to nitrocellulose membranes by iBlot system (Thermo Fisher Scientific). Membranes were blocked in 5% milk/Tris-buffered saline with Tween-20 (TBS-T) before being incubated with horseradish peroxidase (HRP)-conjugated anti-HisTag antibody (Bio-Rad) and visualized by enhanced chemiluminescence (Thermo Fisher Scientific).

2.5 | Recombinant protein purification and quantification

Spike protein was harvested by centrifugation at 3000*g* for 20 min at 4°C and supernatant was filtered through a 0.22 μm filter. Protein was purified using the ÄKTA Pure system (Cytiva) and a 5-ml HisTrap HP column (Cytiva). The column was washed with 5 column volumes (CVs) of buffer B (50 mM sodium phosphate, 300 mM NaCl, 250 mM imidazole, pH 8.0), and equilibrated with 5 CVs of buffer A (50 mM sodium phosphate, 300 mM NaCl, 10 mM imidazole, pH 8.0). To reduce nonspecific binding, the supernatant was adjusted to 20 mM imidazole

using buffer B before sample loading. After sample loading, the column was washed in three steps using 5 CVs of buffer A, 5 CVs of 4.5% v/v buffer B, and 10 CVs of 9% v/v buffer B. Protein was eluted using 100% v/v buffer B. Eluted protein fractions were pooled and buffer exchanged into storage buffer (20 mM Tris, 200 mM NaCl, 10% v/v glycerol, pH 8.0) using a PD-10 desalting column (Cytiva). Protein was quantified using the Pierce Coomassie Plus (Bradford) Assay kit and bovine serum albumin for the calibration curve (Thermo Fisher Scientific) and analyzed by reducing SDS-PAGE. An orthogonal quantification method was performed using an A280 measurement (NanoDrop One^{C}; Thermo Fisher Scientific) with spike extinction coefficient of 428 255 M^{-1} cm^{-1} and M_W of 412.516 kDa. A complementary quantification of spike in culture supernatant was performed using CR3022 antibody ELISA (see below).

2.6 | Protein identification by mass spectrometry

All materials were supplied by Thermo Fisher Scientific unless otherwise stated. Briefly, protein samples in 50 mM ammonium bicarbonate (ABC), 5 mM tris(2-carboxyethyl)phosphine–HCl were reduced by incubation at 37°C for 30 min. *S*-alkylation was performed by the addition of 1 µl 100 mM methyl methanethiosulfonate in isopropanol. For proteolytic digestion, 1.5 µl 0.2% ProteaseMax surfactant in 50 mM ABC and 2 µl 0.2 g/L trypsin/endoproteinase Lys-C mixture (Promega) were added followed by incubation at 37°C for 16 h. Proteolysis was stopped and the surfactant hydrolyzed by the addition of 0.5% trifluoroacetic acid (TFA). The samples were desalted using HyperSep Hypercarb solid-phase extraction tips and dried by vacuum centrifugation. For RPLC-MS, samples in 0.5% TFA, 3% acetonitrile (ACN) were injected. Peptides were separated using an RSLCnano system with a PepSwift PS-DVB monolithic column using a gradient from 97% solvent A (0.1% formic acid) to 35% solvent B (0.1% formic acid, 80% ACN). Mass spectra were acquired on a Q Exactive HF quadrupole-Orbitrap instrument, with automated data-dependent switching between full-MS and tandem MS/MS scans. Proteins were identified by converting the MS data into Mascot Generic Format (MGF) files and analyzed against human and Chinese hamster reference proteome databases with the spike glycoprotein construct sequence inserted (www.uniprot.org) using Mascot Daemon v.2.5.1 with Mascot server v.2.5 (Matrix Science).

2.7 | CR3022 antibody ELISA for spike quantification

96-wells were coated overnight with 100 µl of anti-SARS-CoV spike CR3022 antibody (absolute antibody; 5 µg/ml in PBS) at 4°C. The coating solution was removed and washed twice (with 0.1% TBS-T). 100 µl of blocking solution (5% nonfat milk in 0.1% TBS-T) was added for 1 h and washed twice. 100 µl of the sample was added and incubated for 2 h at RT. A standard curve (Figure S5) was generated from serially diluted, purified and quantified CHO spike (Figure 1d) using CD CHO medium as diluent. Plate was washed twice before incubation with 100 µl HRP-conjugated anti-HisTag antibody (Bio-Rad) diluted 1:500 with 5% milk/TBS-T for 2 h at RT. Plate was washed three times and 100 µl of SigmaFast OPD solution (Sigma) was added to each well. The reaction was allowed to proceed for 10 min at RT before being stopped by the addition of 50 µl of 3 M HCl. Plate was read at 492 nm using a SpectraMax iD5 plate reader (Molecular Devices).

2.8 | Spike ELISA for serological testing

The ELISA protocol was adapted from Stadlbauer et al. (2020) using spike protein with >95% purity. Microtiter plates (96-well) were coated overnight with 50 µl of spike per well (2 µg/ml in PBS pH 7.4) at 4°C. The coating solution was removed and 300 µl of blocking solution (3% nonfat milk in 0.1% PBS-T) was added for 1 h and washed three times (with 0.1% PBS-T). Samples were added at 1/20 dilution and incubated for 2 h at RT. Plate was washed three times and 100 µl of anti-human immunoglobulin G conjugate was added to the wells and incubated for 1 h at RT. Plate was washed three times and 100 µl of substrate was added and incubated in the dark for 45 min. The reaction was stopped by the addition of 50 µl of 3 M HCl and the plate was read at 490 nm using the Agility ELISA system (Dynex Technologies). The index value was calculated as follows:

$$\text{Index value} = \frac{\text{Sample OD}}{\text{Mean of negative controls} + 3\text{SDs}}. \quad (2)$$

The cut-off value was calculated with 100% specificity using ROC curves of calculated index values.

ACKNOWLEDGMENTS

This study was supported by Sheffield Teaching Hospitals NHS Foundation Trust and the University of Sheffield, UK. The authors are grateful to Prof. Florian Krammer (Icahn School of Medicine at Mount Sinai, New York) for providing the spike plasmid via Dr. Thushan de Silva (University of Sheffield), Dr. Martin Nicklin (University of Sheffield) for providing the Expi293F cells, Molly Smith (University of Sheffield) for a preliminary test of HEK transfection procedures, Prof. William Egner (Sheffield Teaching Hospitals) for support and help with ELISA development work and validation. Prof. Mark Dickman acknowledges support from the Biotechnology and Biological Sciences Research Council UK (BBSRC) (BB/M012166/1).

CONFLICT OF INTERESTS

The authors declare that there are no conflict of interests.

AUTHOR CONTRIBUTIONS

Yusuf B. Johari: Conceptualization, methodology, investigation, formal analysis, validation, data curation, visualization, writing—original draft, Writing—review & editing, Project administration. Stephen R.P. Jaffé: Conceptualization, methodology, investigation, formal analysis, validation, data curation, writing—review & editing,

project administration. Joseph M. Scarrott: Conceptualization, methodology, investigation, formal analysis, validation. Abayomi O. Johnson: Investigation, formal analysis, data curation. Théo Mozzanino: Investigation, formal analysis, data curation. Thilo H. Pohle: Investigation, formal analysis, data curation. Sheetal Maisuria: Investigation, formal analysis, validation, data curation. Amina Bhayat-Cammack: Investigation, formal analysis, validation, data curation. Giulia Lambiase: Investigation, formal analysis, data curation. Adam J. Brown: Conceptualization, methodology. Kang Lan Tee: Conceptualization, methodology, investigation, formal analysis, validation, data curation. Philip J. Jackson: Conceptualization, methodology, investigation, formal analysis, validation, data curation. Tuck Seng Wong: Conceptualization, methodology, investigation, formal analysis, validation, data curation. Mark J. Dickman: Supervision, conceptualization, methodology, formal analysis, validation, data curation, writing—review & editing. Ravishankar B. Sargur: Supervision, conceptualization, methodology, funding acquisition, project administration. David C. James: Supervision, conceptualization, writing—review & editing, funding acquisition, project administration.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the Supporting Information of this study.

ORCID

Yusuf B. Johari http://orcid.org/0000-0001-9933-5764
Stephen R. P. Jaffé http://orcid.org/0000-0001-7138-9699
Joseph M. Scarrott https://orcid.org/0000-0002-6046-7687
Adam J. Brown http://orcid.org/0000-0002-3290-4560
Kang Lan Tee http://orcid.org/0000-0001-9458-733X
Philip J. Jackson https://orcid.org/0000-0001-9671-2472
Tuck Seng Wong https://orcid.org/0000-0001-7689-9057
Mark J. Dickman https://orcid.org/0000-0002-9236-0788
Ravishankar B. Sargur https://orcid.org/0000-0002-8535-630X
David C. James http://orcid.org/0000-0002-1697-151X

REFERENCES

Amanat, F., Stadlbauer, D., Strohmeier, S., Nguyen, T. H. O., Chromikova, V., McMahon, M., Jiang, K., Arunkumar, G. A., Jurczyszak, D., Polanco, J., Bermudez-Gonzalez, M., Kleiner, G., Aydillo, T., Miorin, L., Fierer, D. S., Lugo, L. A., Kojic, E. M., Stoever, J., Liu, S. T. H., … Krammer, F. (2020). A serological assay to detect SARS-CoV-2 seroconversion in humans. *Nature Medicine*, *26*, 1033–1036. https://doi.org/10.1038/s41591-020-0913-5

Brown, A. J., Gibson, S. J., Hatton, D., & James, D. C. (2017). *In silico* design of context-responsive mammalian promoters with user-defined functionality. *Nucleic Acids Research*, *45*, 10906–10919. https://doi.org/10.1093/nar/gkx768

Cartwright, J. F., Arnall, C. L., Patel, Y. D., Barber, N. O. W., Lovelady, C. S., Rosignoli, G., Harris, C. L., Dunn, S., Field, R. P., Dean, G., Daramola, O., Gibson, S. J., Peden, A. A., Brown, A. J., Hatton, D., & James, D. C. (2020). A platform for context-specific genetic engineering of recombinant protein production by CHO cells. *Journal of Biotechnology*, *312*, 11–22. https://doi.org/10.1016/j.jbiotec.2020.02.012

Esposito, D., Mehalko, J., Drew, M., Snead, K., Wall, V., Taylor, T., Frank, P., Denson, J.-P., Hong, M., Gulten, G., Sadtler, K., Messing, S., & Gillette, W. (2020). Optimizing high-yield production of SARS-CoV-2 soluble spike trimers for serology assays. *Protein Expression and Purification*, *174*, 105686. https://doi.org/10.1016/j.pep.2020.105686

Estes, B., Hsu, Y. R., Tam, L. T., Sheng, J., Stevens, J., & Haldankar, R. (2015). Uncovering methods for the prevention of protein aggregation and improvement of product quality in a transient expression system. *Biotechnology Progress*, *31*, 258–267. https://doi.org/10.1002/btpr.2021

Fernandez-Martell, A., Johari, Y. B., & James, D. C. (2018). Metabolic phenotyping of CHO cells varying in cellular biomass accumulation and maintenance during fed-batch culture. *Biotechnology and Bioengineering*, *115*, 645–660. https://doi.org/10.1002/bit.26485

Goey, H. C., Alhuthali, S., & Kontoravdi, C. (2018). Host cell protein removal from biopharmaceutical preparations: Towards the implementation of quality by design. *Biotechnology Advances*, *36*, 1223–1237. https://doi.org/10.1016/j.biotechadv.2018.03.021

Hsieh, C.-L., Goldsmith, J. A., Schaub, J. M., DiVenere, A. M., Kuo, H.-C., Javanmardi, K., Le, K. C., Wrapp, D., Lee, A. G., Liu, Y., Chou, C.-W., Byrne, P. O., Hjorth, C. K., Johnson, N. V., Ludes-Meyers, J., Nguyen, A. W., Park, J., Wang, N., Amengor, D., … McLellan, J. S. (2020). Structure-based design of prefusion-stabilized SARS-CoV-2 spikes. *Science*, *369*, 1501–1505. https://doi.org/10.1126/science.abd0826

Johari, Y. B., Brown, A. J., Alves, C. S., Zhou, Y., Wright, C. M., Estes, S. D., Kshirsagar, R., & James, D. C. (2019). CHO genome mining for synthetic promoter design. *Journal of Biotechnology*, *294*, 1–13. https://doi.org/10.1016/j.jbiotec.2019.01.015

Johari, Y. B., Estes, S. D., Alves, C. S., Sinacore, M. S., & James, D. C. (2015). Integrated cell and process engineering for improved transient production of a "difficult-to-express" fusion protein by CHO cells. *Biotechnology and Bioengineering*, *112*, 2527–2542. https://doi.org/10.1002/bit.25687

Lee, S., Kim, M. G., Kim, N., Heo, W. D., & Lee, G. M. (2016). Heparan sulfate proteoglycan synthesis in CHO DG44 and HEK293 cells. *Biotechnology and Bioprocess Engineering*, *21*, 439–445. https://doi.org/10.1007/s12257-015-0688-6

Mason, M., Sweeney, B., Cain, K., Stephens, P., & Sharfstein, S. T. (2012). Identifying bottlenecks in transient and stable production of recombinant monoclonal-antibody sequence variants in Chinese hamster ovary cells. *Biotechnology Progress*, *28*, 846–855. https://doi.org/10.1002/btpr.1542

Park, J. H., Jin, J. H., Ji, I. J., An, H. J., Kim, J. W., & Lee, G. M. (2017). Proteomic analysis of host cell protein dynamics in the supernatant of Fc-fusion protein-producing CHO DG44 and DUKX-B11 cell lines in batch and fed-batch cultures. *Biotechnology and Bioengineering*, *114*, 2267–2278. https://doi.org/10.1002/bit.26360

Poulain, A., Perret, S., Malenfant, F., Mullick, A., Massie, B., & Durocher, Y. (2017). Rapid protein production from stable CHO cell pools using plasmid vector and the cumate gene-switch. *Journal of Biotechnology*, *255*, 16–27. https://doi.org/10.1016/j.jbiotec.2017.06.009

Stadlbauer, D., Amanat, F., Chromikova, V., Jiang, K., Strohmeier, S., Arunkumar, G. A., Tan, J., Bhavsar, D., Capuano, C., Kirkpatrick, E., Meade, P., Brito, R. N., Teo, C., McMahon, M., Simon, V., & Krammer, F. (2020). SARS-CoV-2 seroconversion in humans: A detailed protocol for a serological assay, antigen production, and test setup. *Current Protocols in Microbiology*, *57*, e100. https://doi.org/10.1002/cpmc.100

Stuible, M., Gervais, C., Lord-Dufour, S., Perret, S., L'Abbe, D., Schrag, J., St-Laurent, G., & Durocher, Y. (2020). Rapid, high-yield production of full-length SARS-CoV-2 spike ectodomain by transient gene expression in CHO cells. *bioRxiv*. https://doi.org/10.1101/2020.09.08.286732

Uhlen, M., Tegel, H., Sivertsson, Å., Kuo, C., Gutierrez, J. M., Lewis, N. E., Forsström, B., Dannemeyer, M., Fagerberg, L., Malm, M., Vunk, H.,

Edfors, F., Hober, A., Sjöstedt, E., Kotol, D., Mulder, J., Mardinoglu, A., Schwenk, J. M., Nilsson, P., … Hober, S. (2018). The human secretome—The proteins secreted from human cells. *bioRxiv*. https://doi.org/10.1101/465815

Watanabe, Y., Allen, J. D., Wrapp, D., McLellan, J. S., & Crispin, M. (2020). Site-specific glycan analysis of the SARS-CoV-2 spike. *Science*, *369*, 330–333. https://doi.org/10.1126/science.abb9983

Xiong, X., Qu, K., Ciazynska, K. A., Hosmillo, M., Carter, A. P., Ebrahimi, S., Ke, Z., Scheres, S. H. W., Bergamaschi, L., Grice, G. L., Zhang, Y., CITIID-NIHR COVID-19 BioResource CollaborationNathan, J. A., Baker, S., James, L. C., Baxendale, H. E., Goodfellow, I., Doffinger, R., & Briggs, J. A. G. (2020). A thermostable, closed SARS-CoV-2 spike protein trimer. *Nature Structural & Molecular Biology*, *27*, 934–941. https://doi.org/10.1038/s41594-020-0478-5

Yuan, M., Liu, H., Wu, N. C., Lee, C. D., Zhu, X., Zhao, F., Huang, D., Yu, W., Hua, Y., Tien, H., Rogers, T. F., Landais, E., Sok, D., Jardine, J. G., Burton, D. R., & Wilson, I. A. (2020). Structural basis of a shared antibody response to SARS-CoV-2. *Science*, *369*, 1119–1123. https://doi.org/10.1126/science.abd2321

SUPPORTING INFORMATION

Additional Supporting Information may be found online in the supporting information tab for this article.

How to cite this article: Johari YB, Jaffé SRP, Scarrott JM, et al. Production of Trimeric SARS-CoV-2 Spike Protein by CHO Cells for Serological COVID-19 Testing. *Biotechnology and Bioengineering*. 2021;118:1013–1021. https://doi.org/10.1002/bit.27615

8.2.2 Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells

Received: 31 January 2022 | Revised: 7 April 2022 | Accepted: 23 April 2022

DOI: 10.1002/biot.202200062

Biotechnology Journal

RESEARCH ARTICLE

Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells

Yusuf B. Johari[1] | Joseph M. Scarrott[1] | Thilo H. Pohle[1] | Ping Liu[2] | Ayda Mayer[2] | Adam J. Brown[1,3] | David C. James[1,3]

[1]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK

[2]Cell Line Development, REGENXBIO Inc., Rockville, Maryland, USA

[3]Syngensys Ltd., Sheffield, UK

Correspondence
David C. James, Department of Chemical and Biological Engineering, University of Sheffield, Mappin St., Sheffield S1 3JD, UK.
Email: d.c.james@sheffield.ac.uk

Present address: Yusuf B. Johari, Lonza Biologics, Cambridge, UK

Abstract

Expression of recombinant genes in HEK293 cells is frequently utilized for production of recombinant proteins and viral vectors. These systems frequently employ the cytomegalovirus (CMV) promoter to drive recombinant gene transcription. However, the mechanistic basis of CMV-mediated transcriptional activation in HEK293 cells is unknown and consequently there are no strategies to engineer CMV for controlled expression of recombinant genes. Extensive bioinformatic analyses of transcription factor regulatory elements (TFREs) within the human CMV sequence and transcription factor mRNAs within the HEK293 transcriptome revealed 80 possible regulatory interactions. Through in vitro functional testing using reporter constructs harboring discrete TFREs or CMV deletion variants we identified key TFRE components and clusters of TFREs (cis-regulatory modules) within the CMV sequence. Our data reveal that CMV activity in HEK293 cells is a function of the promoters various constituent TFREs including AhR:ARNT, CREB, E4F, Sp1, ZBED1, JunB, c-Rel, and NF-κB. We also identified critical Sp1-dependent upstream activator elements near the transcriptional start site that were required for efficient transcription and YY1 and RBP-Jκ binding sites that mediate transrepression. Our study shows for the first time that novel, compact CMV-derived promoters can be engineered that exhibit up to 50% higher transcriptional efficiency (activity per unit DNA sequence) or 14% increase in total activity compared to the wild-type counterpart.

KEYWORDS
CMV promoter, HEK293 cells, transcription factor, transcriptional regulation, transient gene expression

Abbreviations: AAV, adeno-associated virus; CHO, Chinese hamster ovary; CMV, cytomegalovirus; CRM, *cis*-regulatory modules; GFP, green fluorescent protein; HEK, human embryonic kidney; TPM, transcripts per million; ORF, open reading frame; RNA-Seq, next-generation RNA sequencing; TF, transcription factor; TFRE, transcription factor regulatory element; TGE, transient gene expression; TSS, transcriptional start site.

1 | INTRODUCTION

The human embryonic kidney 293 (HEK293) cell line is the most commonly utilized human cell line for manufacturing of therapeutic proteins and viral vectors. The cell line serves as a fitting expression host for proteins with particular requirement for human post-translational modifications,[1] and possesses integrated E1 genes required for adeno-associated virus (AAV) production.[2] Moreover, facile cell culture and transfection enable transient gene expression (TGE) methods to be effectively employed for rapid production of potential drug candidates and expression of toxic viral genes.[3,4] Accordingly, various cell and process development have been carried out to improve TGE in HEK293 cells. Examples include optimization of transfection methods,[5] and co-expression of the genes encoding cell cycle-regulators p18 and p21[6] or knock-out of the pro-apoptotic genes Bax and Bak.[7] Further efforts to boost TGE levels are exemplified via process engineering strategies by implementing mild hypothermia[8,9] as well as the addition of chemical inducers such as valproic acid[6,10] and sodium butyrate.[11] Despite these improvements, at the core of HEK293-based production systems, recombinant gene transcription is most often directed by the human cytomegalovirus immediate early (hCMV-IE) promoter,[6,8–10,12,13] although there is little mechanistic understanding of how CMV transcriptional activity can be controlled and enhanced. Engineering fundamental synthetic processes in mammalian cell factories such as HEK293 therefore remains a highly desirable objective.

hCMV-IE promoter (henceforth referred to as the CMV promoter) is a highly complex element comprising binding sites (transcription factor regulatory elements [TFREs]) for numerous ubiquitously expressed transcription factors (TFs).[14] This is not surprising considering that the promoter has evolved to function in a broad cell tropism.[15,16] However, promoter activity in any given host is regulated by a system-specific combination of interactions between the promoter's constituent TFREs and the cells repertoire of endogenous TFs.[17] Therefore, transcriptional activity of the CMV promoter is highly context-specific and cell type-dependent expression has been observed both in vivo[18,19] and in vitro[20,21]. With respect to the latter, we have for example, demonstrated that CMV-driven TGE in Chinese hamster ovary (CHO) cells was largely a function of transactivation mediated through just two discrete TFREs (NF-κB and CREB).[22] Further, the CMV promoter comprises binding sites of several transcriptional repressors such as YY1 — conferring on cytomegalovirus the ability to establish latent infection.[16,22–24] Accordingly, it is likely that the CMV promoter is fundamentally sub-optimal for use in unnatural, specific processes such as recombinant gene expression in HEK293 cells. Despite this, no previous studies have examined the specific CMV promoter interactions within the HEK293 transcriptional landscape that drive recombinant gene transcription, and the lack of this mechanistic information currently renders optimization of CMV-driven TGE in HEK293 cells intractable.

In this study, we identify the regulators of CMV-mediated TGE in HEK293 cells through mechanistic dissection of the CMV promoter. We performed an extensive bioinformatic analysis on the promoter's TFRE composition, coupled with a detailed in vitro comparative analysis of the relative influence of CMV component parts on gene expression to identify functional elements (TFRE sequences and *cis*-regulatory modules [CRMs]) that critically control promoter activity in HEK293 cells. We demonstrate, for the first time, that the wild-type CMV promoter can be re-engineered specifically for HEK293 cells to derive highly compact and transcriptionally efficient novel promoters with increased transcriptional activity.

2 | MATERIALS AND METHODS

2.1 | HEK and CHO cell cultures

Suspension-adapted HEK293 cells were provided by REGENXBIO and cultured in Dynamis medium (Thermo Fisher Scientific) supplemented with L-glutamine (Thermo Fisher Scientific). Expi293F cells (Thermo Fisher Scientific) were cultured in Expi293 Expression medium (Thermo Fisher Scientific). CHO-S cells (Thermo Fisher Scientific) were cultured in CD CHO medium (Thermo Fisher Scientific) supplemented with 8 mM L-glutamine. Cells were maintained in Erlenmeyer flasks (Corning) at 37°C, 140 rpm under 5% CO_2, 85% humidity and were sub-cultured every 3 to 4 days by seeding at 3×10^5 viable cells per mL. Cell viability and viable cell density were measured using a Vi-CELL XR (Beckman Coulter).

2.2 | Vector construction

pmaxGFP vector (Lonza) was utilized as a backbone. The CMV promoter and chimeric intron of pmaxGFP were deleted by digestion with BsrGI and KpnI, and replaced with a short DNA fragment containing EcoRI and HindIII cloning sites. A full-length hCMV-IE promoter (–550 to +48 relative to the TSS) was synthesized (Eurofins Genomics) and inserted directly upstream of the green fluorescent protein (GFP) open reading frame (ORF) of the promoter less vector backbone. A minimal CMV core promoter (–36 to +48 relative to the TSS) was also synthesized and inserted directly upstream of the GFP ORF. To create TFRE reporter plasmids, synthetic oligonucleotides containing 7 × repeat copies of the TFRE sequences in Table S1 were synthesized, PCR amplified (Q5 high-fidelity 2 × master mix; NEB), and purified (QIAquick PCR Purification kit; Qiagen). The PCR products were then digested, gel extracted (QIAquick Gel Extraction kit; Qiagen) and inserted into the cloning sites upstream of the CMV core promoter. Discrete regions of the CMV promoter sequence were PCR amplified and inserted upstream of the CMV core promoter. Mutated promoter constructs were synthesized and inserted upstream of the CMV core promoter. The CBh promoter was excised from pSpCas9(BB)-2A-GFP plasmid (Addgene) by digestion with KpnI and AgeI and inserted directly upstream of the GFP ORF. Clonally derived plasmids were purified using a QIAGEN Plasmid Plus kit (Qiagen). The sequence of all plasmid constructs was confirmed by restriction enzyme analysis and DNA sequencing (Eurofins Genomics).

2.3 | PEI-mediated transient transfection

One day before transfection, cells were sub-cultured in an Erlenmeyer flask, grown to 1×10^6 cells per mL and aliquots of 10 mL were added to each TubeSpin bioreactor tube (TPP). 8 μg of DNA and 24 μL of PEI MAX (1 mg mL^{-1}; Polysciences) were each pre-diluted in 150 μL of NaCl (150 mM; Polyplus-transfection), combined and incubated at room temperature for 4 min before being added into culture. Transfected cells were cultured for 48 h at 37°C, 230 rpm under 5% CO_2, 85% humidity.

2.4 | Measurement of recombinant GFP expression in vitro

GFP expression was quantified using a SpectraMax iD5 microplate reader (Molecular Devices) 48 h post-transfection. Prior to fluorescence read (excitation: 485 nm, emission: 535 nm), culture medium was removed by centrifugation at $200 \times g$ for 5 min. 1.5×10^6 viable cells were resuspended in 750 μL Dulbecco's phosphate-buffered saline (DPBS; Sigma) and then transferred to a 96-well microplate at 3×10^5 cells (150 μL) per well. To measure transfection efficiency, cells were analyzed using Attune Acoustic Focusing Cytometer (Thermo Fisher Scientific). Background fluorescence/absorbance was determined in cells transfected with a promoterless vector.

2.5 | In silico analysis of transcription factor regulatory elements

Genomatix Gene Regulation software (MatInspector Release 8.4 and MatBase Version 11.2; Precigen Bioinformatics Germany) was used to analyze the CMV promoter to find putative human TFREs. To capture all possible binding sites of different TF subtypes within the promoter, analysis was performed using the Individual Matrix function (rather than Matrix Families), with the Core Similarity set to 0.80 and the Matrix Similarity set to 'Optimized'. Cognate TF of each TFRE matrix (totaling 116; Table S2) was obtained from previously published studies as listed in MatBase. Selective mutation of a specific TFRE was performed by mutating at least of one the four highly conserved nucleotides (core sequence) defined in MatBase. Mutated sequence was subjected to the same analysis using the software to ensure that the specific TF binding site was removed, and that neither any overlapping TFRE was perturbed nor new TFRE was introduced.

2.6 | Analysis of HEK293 transcription factor expression

HEK293 cells were seeded at 1×10^6 viable cells per mL and cultured as described above. From day 3, cells were fed daily (1% v/v) with feed medium containing 130 g L^{-1} glucose, 29.23 g L^{-1} L-glutamine, 25 g L^{-1} arginine and 20 g L^{-1} serine. Total RNA was extracted from duplicate cultures during exponential ($\sim 5 \times 10^6$ cells per mL) and stationary phases ($\sim 1.6 \times 10^7$ cells per mL) of growth. For each sample, 3×10^6 viable cells were collected by centrifugation at $200 \times g$ for 5 min. Cell pellets were immediately resuspended in 300 μL of RNAprotect Cell Reagent and stored at −80°C. RNA-seq libraries were prepared and sequenced by GENEWIZ using an Illumina NovaSeq (Illumina). Galaxy (usegalaxy.org) and R software were used to analyze the RNA-seq data using Salmon alignment tool and human GRCh38 GTF and FASTA files from www.ensembl.org. A curated database of ~1600 human TFs was obtained from Lambert et al.[26] (see Table S3).

2.7 | Statistical analysis

Microsoft Excel 2016 (Microsoft) was used to analyze the difference between the means (GFP expression) of two promoter constructs. For samples comprising only HEK293 cells, analysis was performed using Student's *t*-test with p-value < 0.05 was considered significant. For samples comprising HEK293 and Expi293F cells, analysis was performed using Two-Factor ANOVA with Replication with p-value < 0.05 was considered significant.

3 | RESULTS

3.1 | In silico and in vitro identification of regulators of CMV promoter transcriptional activity

In order to identify potential regulatory elements in CMV capable of recombinant gene transactivation in HEK293 cells, we performed bioinformatic survey of (i) putative TFREs (binding sites) in the promoter, and (ii) the TF repertoire of HEK293 cells based on RNA-seq datasets. With regard to the latter, although gene expression analysis does not permit precise quantification of active TF levels, it provides useful information on general TF expression profile where genes with more than two transcripts per million (TPM) can be considered active.[27] Using the Genomatix search tool, 108 discrete TFREs from 74 TF families were identified in the CMV promoter at copy numbers ranging from one to six. However, the gene expression analysis indicated that 22% (24/108) of the TFREs' cognate TFs were not expressed in HEK293 cells (exponential phase $\log_2$ TPM < 1; Figure 1A). Further, as we wanted to identify key regulatory elements, we focused our search on TFs that exhibit gene expression activities in both exponential and stationary phases of culture (i.e., "context-specific" expression can extend beyond cell-type), thus eliminating an additional four TFs that were not expressed in the latter phase of culture — yielding 80 potential TFREs. We note that two TFREs may have identical or overlapping sequences within the CMV promoter (e.g., NF-κB and NFAT5). Table S2 lists the identified TFREs and their cognate TFs.

To minimize the TFRE pool for functional testing, we filtered out TFREs with substantially overlapping binding sites and selected two TFREs from each TF family — yielding a subset of 25 TFREs. Figure 1B shows the map of select TFREs in the CMV promoter. To measure the

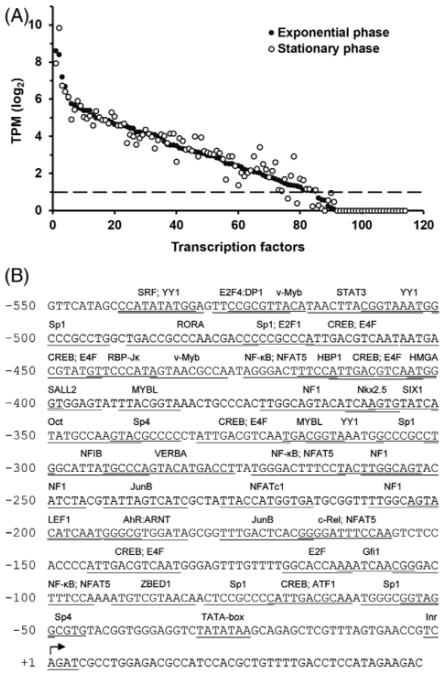


FIGURE 1 In silico identification of potential transcriptional regulators of CMV promoter activity. CMV promoter (−550 to +48 relative to the transcription start site; TSS) was surveyed for the presence of putative transcription factor regulatory elements (TFREs) using Genomatix software. One hundred and eight discrete TFREs identified in CMV promoter were subsequently analyzed for the presence of their cognate transcription factors (TFs) in HEK293 cells. (A) RNA-seq analysis of HEK293 cell transcriptome determined the relative gene expression level of TFs. Points represent the expression level (transcripts per million; TPM) of each TF sampled at exponential and stationary phases of culture. Genes with more than two transcripts per million ($\log_2$ TPM > 1) was considered as actively transcribed genes. (B) CMV promoter sequence with 25 selected TFREs for in vitro analysis. The TSS is indicated with an arrow

relative ability of TFREs to activate transcription of recombinant genes in HEK293 cells, we created a set of GFP reporter constructs that contained seven repeat copies of a specific TFRE in series, upstream of a minimal CMV core promoter (−36 to +48 relative to the TSS, containing a TATA box and an Inr motif) as previously described.[25,28] We note that the transcriptional output from a single TF binding site is often insufficient to drive detectable levels of recombinant gene expression. Optimized PEI-mediated transient transfection of plasmid DNA into suspension HEK293 cells yielded a transfection efficiency of ~ 94% with a cell viability of ~ 90% at 48 h post-transfection (measured using a vector harboring a CMV promoter). Additionally, preliminary experiments confirmed that GFP fluorescence intensity in HEK293 cell host is directly proportional to GFP mRNA levels post-transfection (Figure S1).[29] Measurement of GFP expression after transient transfection of HEK293 cells with each TFRE reporter plasmid is shown in Figure 2A. This analysis revealed eight TFREs with significantly increased expression (> 10-fold, $p < 0.01$) over basal expression from the minimal core promoter, that is, AhR:ARNT, CREB/ATF1, CREB/E4F, Sp1, ZBED1, JunB, c-Rel, and NF-κB. We note that in some instances, TFRE sequences with competing (overlapping) binding sites may be resolved by utilizing their consensus sequence (e.g., CREB and E4F; Figure 2A). Other TFRE reporter constructs displayed no obvious increase in GFP above core control level, suggesting alternative mechanisms of TF-mediated transcriptional activation or suboptimal TF binding sequences. To elucidate the latter, we tested the consensus sequence of MYBL1, Oct and E2F (selected based on a posteriori knowledge). This analysis revealed that the consensus sequences exhibited between 10 and 53-fold increase in expression over the core promoter (Figure 2A), indicating that the TFREs were essentially able to mediate activation of recombinant gene transcription in HEK293 cells using available TF activity.

In order to both confirm and further demonstrate the distinctive transcriptional landscape of HEK293 cells that influence CMV-mediated TGE, we tested the NF-κB and CREB sequences in CHO cells, as well as NF-κB p65 subunit sequence that is not present in the CMV promoter (Table S1). This analysis (Figure 2B) indicated that NF-κB and CREB were highly active in CHO cells (133% and 62% of CMV activity respectively), in line with our previous study that identified these two elements as key positive regulators of CHO cell-specific CMV promoter activity.[22] While NF-κB p65 subunit was five times more active than NF-κB in HEK293 cells, the activity was only one-fifth of that observed in CHO cells (we note that relative TFRE activities were not proportional to cognate TF expression levels). We therefore deduced that HEK293 cell-specific regulation of CMV promoter activity, in contrast to CHO, was a function of cooperative interactions amongst a broader range of TFREs.

3.2 | CMV promoter-mediated gene expression in HEK cells is regulated by proximal elements

In its natural context the CMV promoter can be divided into two modular components, the proximal and distal enhancers (Figure 3A).[14] Furthermore, TFREs often occur together in clusters as *cis*-regulatory modules (CRMs) where some elements may require interactions with adjacent or nearby TFRE partners in order to drive transcription.[30] To identify DNA sequence regions that are required for regulating gene expression in HEK293 cells, we inserted seven ~150-bp CRMs upstream of the CMV core in GFP reporter vectors (CRMs 1–7; Figure 3A). Figure 3B shows transient GFP reporter production from each CRM. CRMs from within the proximal enhancer sequence were generally more active than those from the distal, with CRM 1 alone yielding 67% of CMV's transcriptional activity. Analysis of the TFRE composition indicated that all positive regulators identified in the

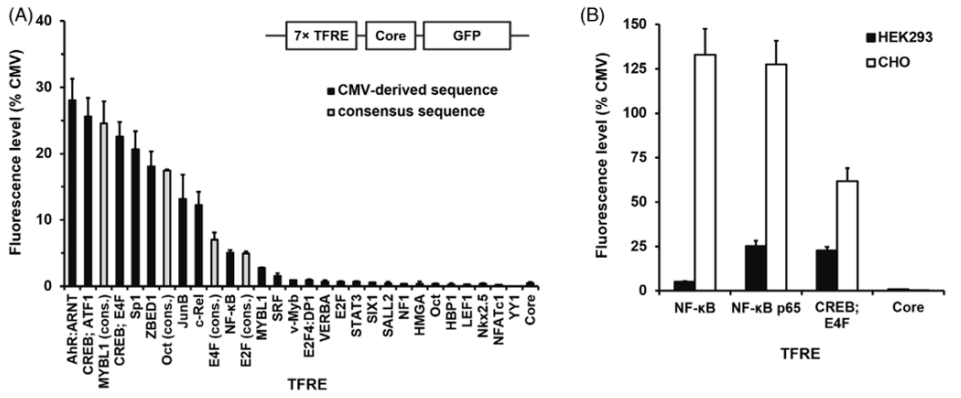


FIGURE 2 Identification of active transcription factor regulatory elements (TFREs). (A) TFRE sequence derived from the CMV promoter (black bars) or its consensus sequence (gray bars) was cloned in series (7 × copies) upstream of a minimal CMV core promoter in GFP-reporter vectors. HEK293 cells were transfected with each homotypic TFRE-reporter using polyethylenimine (PEI) and cultured in tube-spin bioreactors at 37°C. GFP expression was quantified 48 h post-transfection. (B) NF-κB p65 consensus sequence was cloned in series (7 × copies) upstream of a minimal CMV core promoter in GFP-reporter vectors and transfected into HEK293 and CHO-S cells alongside NF-κB and CREB/E4F constructs from A. Cells were cultured in tube-spin bioreactors at 37°C and GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the CMV promoter. Data shown are the mean value ± SD of two independent experiments each performed in duplicate

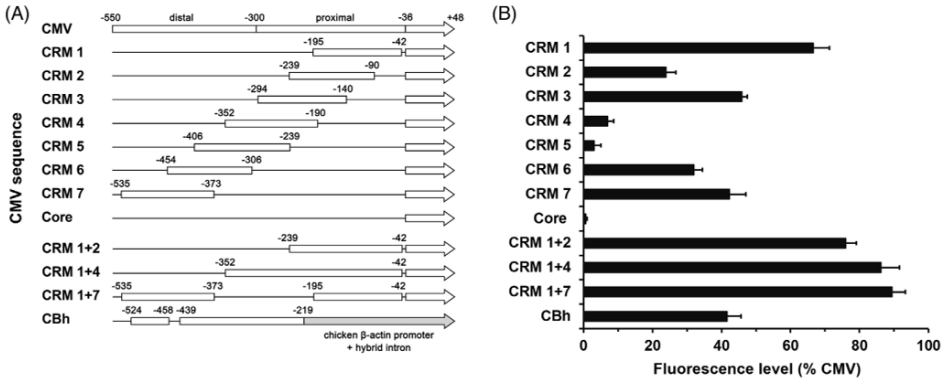


FIGURE 3 Relative transcriptional activity exhibited by CMV promoter structural elements. (A) The CMV promoter contains the proximal and distal enhancers and clusters of TFREs (*cis*-regulatory modules; CRMs). Each element was cloned upstream of a minimal CMV core promoter in GFP reporter plasmids while the CBh promoter (793 bp) was inserted directly upstream of the GFP open reading frame. (B) Reporter plasmids were transfected into HEK293 cells using PEI and cultured in tube-spin bioreactors at 37°C. GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the CMV promoter. Data shown are the mean value ± SD of two independent experiments each performed in duplicate

functional screen (Figure 2A) occurred in CRM 1, with one copy each of AhR:ARNT, CREB/ATF1, CREB/EF4, ZBED1, JunB, c-Rel, and NF-κB and two copies of Sp1. Moreover, multiple copies of CREB/E4F and Sp1 were present in CRMs 6 and 7, yielding 32% to 42% of CMV's activity. Conversely, CRMs from the middle of the CMV promoter (i.e., CRMs 4 and 5) did not display observable activity (≤ 7% of CMV). This was not unexpected considering that the constituent TFREs of these CRMs were mostly inactive in the functional screen.

Assembly of CRMs and comparison of their relative activity provided further analysis of individual CRM functions (Figure 3B). Combining CRM 1 and CRM 7 (67% and 42% of CMV activity respectively) yielded a promoter with only 90% CMV activity (CRM 1+7), suggesting a partially redundant function of the distal enhancer and/or spatial effects. On the other hand, adding inactive CRM 4 (7% CMV) onto CRM 1 significantly enhanced the transcriptional activity to 86% CMV (CRM 1+4; $p = 0.005$). This data implies a synergistic interaction of specific TFREs within the proximal enhancer. To expound this observation, we constructed an extended CRM 1 reporter vector (CRM 1+2) incorporating the NFATc1, NF1 and LEF1 binding sites. Even though these TFRE sequences were not active on their own (Figure 2A) the extended promoter displayed a 15% increase in activity ($p = 0.007$; Figure 3B), possibly via NFATc1–c-Rel interaction (−117 and −57 relative to the TSS respectively).[31] Critically, the data in Figure 3B reveal that CRM 2 exhibited 64% lower activity than CRM 1 despite a significant sequence overlap (Figure 3A), suggesting that either additional TFREs within the 5′ region of CRM 2 functioned to negatively regulate transcription, or essential regulators of CMV-mediated TGE in HEK293 were located in the 3′ region of CRM 1. With regard to the former, the apparent increase in GFP activity of CRM 1+2 compared to CRM 1 (see above) discounted the possibility of a specific transrepression effect of CRM 2. To substantiate the latter, we constructed a reporter vector utilizing a CMV enhancer/chicken β-actin hybrid (CBh) promoter (Figure 3A).[32] The promoter, comprising a practically complete CMV enhancer apart from CRM 1, exhibited only 41% of CMV's activity. Combining all observations made above, we inferred that (i) TFREs within the proximal enhancer functioned synergistically to drive transcription, and (ii) critical regulators of CMV promoter activity in HEK293 were located in the 3′ region of the proximal enhancer sequence (i.e., approximately −90 to −42 relative to the TSS).

3.3 | Sp1 binding sites near the TATA-box are essential for efficient CMV promoter-mediated gene expression in HEK293 cells

In order to specifically determine the key regulators of CMV-mediated gene expression in HEK293 cells we created CMV promoter variants with specific TFREs within −107 to −45 relative to the TSS "knocked-out". Proximal CMV (−300 to +48 relative to the TSS, ~ 84% CMV activity) rather than full-length CMV promoter was utilized for maximal impact of a single TFRE knock-out (i.e., minimal potential "noise" by other elements). Selective mutation was performed on the core sequence of a specific TFRE in order to disrupt the binding site without perturbing overlapping or introducing new TFREs (Figure 4A). Further, given the complexity of CMV promoter, we hypothesized that different HEK293 hosts (in different medium formulations) may potentially vary in their TF repertoires that could significantly influence CMV promoter regulation. In order to evaluate this, we determined the activity of the synthetic proximal CMV constructs in our standard HEK293 cell line as well as the commercially available Expi293F cell line, cultured in Dynamis medium and Expi293 Expression medium respectively. Measurement of GFP production after transient transfection of HEK293 and Expi293F cells with the knocked-out proximal CMV promoters is shown in Figure 4B. We observed that relative promoter activities were very similar in both cell lines, invalidating our hypothesis.

Our data (Figure 4B) also show that removal of NF-κB and ZBED1 binding sites, either individually or simultaneously, did not reduce GFP expression. This result is in line with the above finding (Figure 2A) that NF-κB had a very minimal activity in HEK293 cells but was not fully anticipated for the relatively active ZBED1. Utilizing the TFRE identification tool at a lower stringency, the in silico analysis identified a weak ZBED1 binding site at the mutated sequence (matrix similarity 0.734, optimal matrix threshold 0.76) suggesting that the ZBED1 mutation did not fully knock-out the TFRE. Removal of Sp1, CREB/ATF1, and Sp4 binding sites individually reduced promoter activity to ~ 62%, ~ 74%, and ~ 44% ($p < 0.008$) of that deriving from wild-type proximal CMV, respectively. Additionally, removal of the Sp1 site with CREB/ATF1 or Sp4 (mutations 3+4 and 3+6) led to further decrease in promoter activities. Critically, when the two Sp1 sites were simultaneously removed (mutations 3+5) GFP expression was reduced to the lowest level, that is, ~ 25% compared to the wild-type proximal CMV. No further reduction in promoter activity was observed when the Sp4 was mutated in conjunction with the two Sp1 sites — indicative of the Sp1's vital regulatory function. However, considering the relatively weak activity of the Sp1 homotypic promoter in Figure 2A, our data do not support the conclusion that Sp1 blocks (or any other TFRE) could support high transcriptional activity alone. We therefore deduced that the two Sp1 sites act as an upstream activator element[33] for CMV promoter-mediated transcription in HEK293 cells.

3.4 | Knock-out of repressor elements results in increased gene expression in CMV promoter variants

The above in silico analysis of regulation of CMV promoter activity by sequence elements (Figure 1) also identified two TFRE components that have previously been shown to negatively regulate transcription from the murine CMV-IE promoter in cytomegalovirus-infected mouse kidneys, YY1, and RBP-Jκ,[23] as well as Gfi1 where its overexpression has been shown to repress hCMV-IE promoter activity in mouse fibroblast cells.[34] We hypothesized that CMV promoter could be optimized for TGE by disrupting transrepression mediated by these TFREs. To evaluate the functional activity of these TFREs as regulators of CMV-mediated TGE in HEK293 cells, we synthesized CMV(-derived) promoters with repressor elements knocked-out and inserted them into GFP reporter vectors (Figure 5A,B). We note that the YY1 (three binding sites) and RBP-Jκ (one binding site) are located in the distal enhancer while the Gfi1 (one binding site) is located in the proximal enhancer. Additionally, previous studies suggested a fourth YY1 binding site at −343 to −353 relative to the TSS[14] which was identified as a weak binding sequence in this study (matrix similarity 0.889, optimal matrix threshold 0.94). GFP expression levels in HEK293 and Expi293F cells were measured 48 h post-transfection.

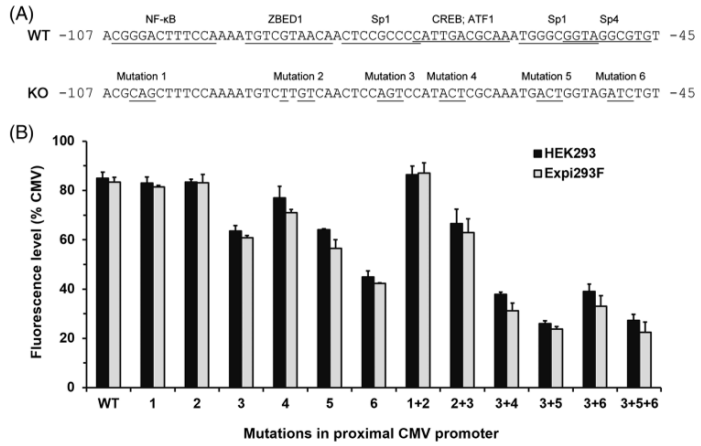


FIGURE 4 A proximal CMV promoter devoid of two Sp1 sites near the TATA box is unable to drive transcription in HEK293 cells. (A) Wild-type (WT) and mutated proximal CMV promoters (−300 to +48 relative to the TSS) with specific TFREs knocked-out (KO) were synthesized and cloned into GFP reporter vectors. Selective mutation was performed on a specific TFRE to disrupt the binding site without perturbing overlapping or introducing new TFREs. (B) The relative activity of each proximal CMV promoter construct was determined in HEK293 and Expi293F cells. Reporter plasmids were transfected into HEK293 cells using PEI and cultured in tube-spin bioreactors at 37°C. GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the full-length CMV promoter. Data shown are the mean value ± SD of two independent experiments each performed in duplicate

Removal of the repressor elements in full length CMV (promoters 1.01 and 1.02; Figure 5C) did not result in increased GFP expression compared to the wild-type control ($p > 0.603$), suggesting that the TFREs were not critical regulators affecting transcriptional activity under the conditions employed. This is possibly due to positive TF–TFRE interactions within the proximal enhancer decreasing the influence of distal enhancer-mediated processes (see above, Figure 3A,B). To further investigate the impact of proximal enhancer, we truncated the 5′ region of the enhancer (promoter 2.01) which resulted in a ∼ 13% decrease in transcriptional activity compared to the wild-type CMV (see Figure S2). In this promoter construct, removal of YY1 and RBP-Jκ binding sites increased the promoter activity by 12% (promoter 2.02; $p = 0.023$), indicating that the TFREs can act as negative regulators of CMV promoter in HEK293 cells. No additional increase was observed with further removal of Gfi1 (promoter 2.03) — this was not entirely unexpected considering that Gfi1 gene was lowly expressed ($\log_2$ TPM = 2.23) whereas YY1 and RBP-Jκ genes exhibited expression levels above the 90th percentile ($\log_2$ TPM = 5.51 and 5.53 respectively; Figure 1A).

To confirm the ability of YY1 and RBP-Jκ to mediate transrepression of recombinant gene transcription in HEK293 cells, we constructed CMV promoter variants with minimal proximal and distal enhancers containing only one site each of YY1, RBP-Jκ, and Gfi1 (promoters 3.01–3.03; see CRM 1+7 above). As anticipated, removal of the YY1 and RBP-Jκ binding sites increased the promoter activity by 11% ($p = 0.031$) compared to its non-mutated counterpart while no significant change was observed with further removal of Gfi1 (we note that removal of the YY1 binding alone increased the promoter activity by ∼ 7% [$p = 0.053$]; data not shown). Moreover, these shorter promoter sequences displayed similar activities to promoters 2.01–2.03. In this regard, we assumed that the deletion of distal enhancer's 3′ region effectively removed transrepression mediated by the fourth YY1 repressor motif (excluded in our in silico survey of CMV-constituent TFREs). Based on the above observations, we constructed promoter 4.01 that was devoid of repressor elements while retaining the active regions. This engineered CMV promoter displayed a ∼ 14% increase in expression ($p = 0.005$; Figure S2) while being 25% smaller in size compared to the wild-type CMV. To illustrate the enhanced capability of the promoters in driving transcription, we calculated a "transcriptional efficiency" for each promoter as a function of transcriptional output per promoter length. This analysis indicates that promoters 3.03 and 4.01 were ∼ 50% more transcriptionally efficient compared to the wild-type CMV promoter (Table S4). We conclude that the CMV promoter can be engineered for improved TGE in HEK293 via disruption of transrepression mediated by YY1 and RBP-Jκ and removal of redundant sequences.

4 | DISCUSSION

The vast majority of current HEK293 cell TGE systems utilize the CMV promoter for high-yield production of therapeutic proteins[6,8] and improved lentiviral and AAV expression vectors.[13,35,36] Our comparative transient expression analyses revealed that the CMV promoter

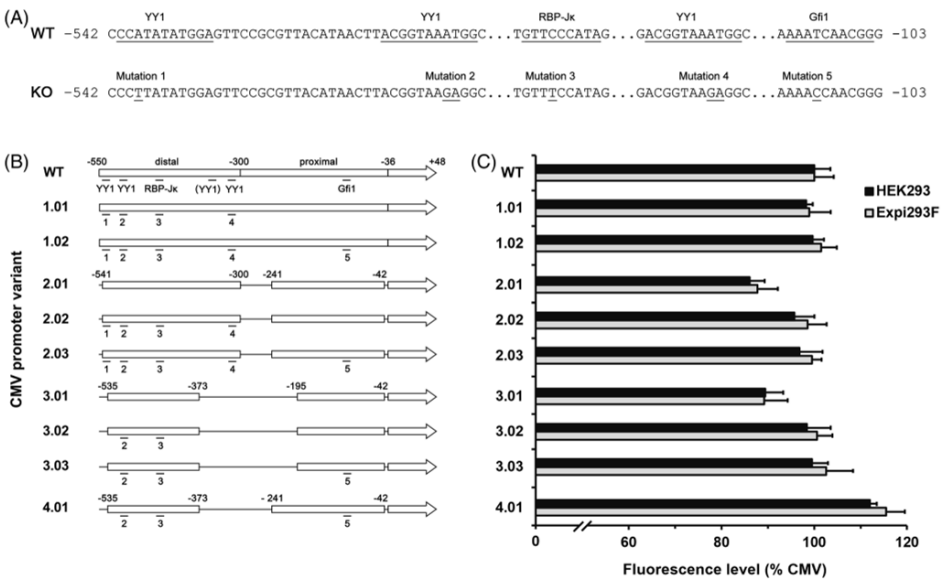


FIGURE 5 Removal of transrepression mediated by YY1 and RBP-Jκ and redundant sequences enhances CMV activity. (A) Selective mutation was performed on a specific TFRE to disrupt the binding site without perturbing overlapping or introducing new TFREs. (B) Wild-type (WT) CMV promoters (−550 to +48 relative to the TSS) and mutated CMV variants with specific TFREs knocked-out (KO) were synthesized and cloned into GFP reporter vectors. The locations of the repressor elements in CMV promoter are underlined. Numbers denote the corresponding TFRE knock-out in A. A fourth putative YYI binding site excluded by the TFRE analysis in this study is shown in bracket. (C) The relative activity of each promoter construct was determined in HEK293 and Expi293F cells. Reporter plasmids were transfected into HEK293 cells using PEI and cultured in tube-spin bioreactors at 37°C. GFP expression was quantified 48 h post-transfection. Data are expressed as a percentage with respect to the GFP expression of a vector containing the wild-type CMV promoter. Data shown are the mean value ± SD of two independent experiments each performed in duplicate

activity in HEK293 cells was a function of the promoter's various constituent TFREs including AhR:ARNT, CREB, E4F, Sp1, ZBED1, JunB, c-Rel, and NF-κB. This is a very significant and useful finding, as they form the basis of promoter engineering containing enhanced binding sites,[37] or can be directly utilized as modular building blocks to construct synthetic promoters de novo.[25,28] We further identified several sub-optimal TF binding sequences (MYBL1, Oct, E2F) which suggests an immense opportunity for maximizing CMV promoter's transcriptional output. Hundreds of TFRE motif sequence variants can be characterized simultaneously via in vitro use of high-throughput parallel screening methods, allowing determination of their optimal binding affinity. The major challenge with such functional tests is the difficulty in identifying TFREs underpinning the more complex regulation governing synergistic transactivation[30,31] which would require intricate screens of TFRE motif pairs with position-sensitive function. However, this limitation can be circumvented by using the TF decoy technology developed in this laboratory[38] to inhibit specific TFRE(s) within the CMV promoter architecture, obviating the need to characterize spatial effects between two TFRE motif pairs. Furthermore, this work, in effect, generated a novel library of promoter sequences (Figure 2) to control gene expression over a wide range. These highly compact promoters could be utilized in multigene vectors to give predictable stoichiometries (e.g., optimization of monoclonal antibody heavy chain to light chain ratio),[39] especially with non-overlapping sequences to avoid homologous recombination-mediated silencing (see below).

Bioinformatic analysis on the CMV promoter sequence indicated that Sp1 family is predominant in the promoter, in line with the notion that the element is essential for prevention of de novo methylation of CpG islands.[40] Importantly, our results show that each of the two Spl binding sites near the TATA box contributes to full activation of the CMV promoter in HEK293 cells — resembling the previous report in which mutation of these Sp1 binding sites caused inefficient CMV promoter transcription and cytomegalovirus replication in human fibroblast cells.[41] Similar transcription activation mechanism had been reported for the simian virus 40 (SV40) promoter in which Sp1 binding to its cognate sequences upstream of the TATA box enhanced the activity of RNA polymerase II.[42,43] Indeed, analyses of synthetic core promoters indicated that Sp1 binding sites, when placed

upstream of an Inr and/or TATA box, acted as an upstream activator element for efficient transcription initiation in vitro[33] and in HEK293 cells.[44] Nevertheless, our previous studies[22,25,38] as well as data in Figure 2B showed that CHO cells were able to drive efficient recombinant gene transcription in the absence of such upstream activator elements, illustrating that engineering strategies to improve CMV promoter activity have to be cell-type specific for maximum efficacy. We further conjecture that a Spl-dependent upstream activator element is a design prerequisite for construction of strong synthetic/hybrid promoters for HEK293 cells. This is in contrast to modulation at the translational level (e.g., engineering of 5′UTR elements) that is generally not cell-type dependent.[45]

Another important outcome of this study is the identification of negatively acting cellular TFs, and that a substantial proportion of the CMV sequence may be functionally redundant for recombinant gene expression in HEK293 cells. Specifically, our results showed that YY1 and RBP-Jκ-mediated transrepression of the CMV promoter could be removed by designing engineered CMV constructs with inactive cognate binding sites. It is worth noting previous studies have also shown that ERF (Ets-2 Repressor Factor) was able to repress the CMV promoter by binding to the 21 bp repeat motifs overlapping YY1 and Sp1 within the distal enhancer (see Figure S3),[24] and that the ERF gene was highly expressed in HEK293 cells ($\log_2$ TPM = 4.91; Table S2). Therefore, we postulate that the deletion of 3′ region of the distal enhancer (promoters 3.03 and 4.01) effectively removed the YY1 as well as an ERF binding site, permitting a more defined, improved regulation of recombinant transcriptional activity and with relatively small promoter size. The engineered promoters may further confer additional advantages in dynamic bioprocess conditions in respond to changes in cellular transcriptional landscape. For example, differential gene expression analysis on the HEK293 transcriptomic data showed that RBP-Jκ was (slightly) upregulated from the mid-exponential to early-stationary phase ($\log_2$ fold-change = 0.362, p-adj = 0.0012), suggesting that the positive impact of RBP-Jκ knock-out would be more pronounced in long-term, fed-batch production processes for example. It also interesting to note that our bioinformatic analysis indicated that the CMV promoter did not contain the binding sites of TFs associated with the unfolded protein response (ATF4, ATF6, eIF2α, and XBP1), suggesting that CMV promoter activity is not affected by cellular stress that may result from recombinant gene overexpression.

Lastly, the data presented in this study may offer benefits to systems beyond TGE. For instance, long-term stable expression can be compromised by the occurrences of sequence features such as repeat elements (homologous recombination-mediated silencing)[46] and CpG islands (methylation-mediated silencing).[47] With regard to the former, the CMV promoter contains two copies of 21 bp repeat motif in the distal enhancer as mentioned above. Promoters 3.03 and 4.01 indirectly removed one of these repeat elements, therefore avoiding potential genetic homologous recombination events associated with gene deletion. With regard to the latter, it may be possible to reduce the number of CpG dinucleotides within the CMV promoter by mutating TFREs with no/low activities, thus minimizing the formation of methylation-mediated epigenetic silencing linked to production instability. Minimal CpG dinucleotides is also a desirable feature for gene therapy vectors in which CpG motifs have immunostimulatory effects (e.g., promoter 3.03 contains 20% less CpG dinucleotides compared to the wild-type CMV promoter; Table S4).[48] Accordingly, it should now be possible to rationally design synthetic CMV promoter variants in order to equip HEK293 cells with new transcriptional machinery optimally suited for a specific intended purpose. We anticipate that similar approaches can be used to deconstruct and reconstruct other promoters for optimal functionalities in particular cell types.

ACKNOWLEDGMENTS

This study was supported by REGENXBIO Inc., U.S.A. The authors thank Adrian Bourke (University of Sheffield) for technical assistance in RNA-seq analysis, Selase Enuameh (REGENXBIO) and Jie Li (REGENXBIO) for useful discussions.

CONFLICT OF INTERESTS

The authors have a patent application filed based on the work in this paper.

DATA AVAILABILITY STATEMENT

The data that supports the findings of this study are available in the supplementary material of this article.

ORCID

Yusuf B. Johari https://orcid.org/0000-0001-9933-5764
Joseph M. Scarrott https://orcid.org/0000-0002-6046-7687
Thilo H. Pohle https://orcid.org/0000-0003-4437-3231
Adam J. Brown https://orcid.org/0000-0002-3290-4560
David C. James https://orcid.org/0000-0002-1697-151X

REFERENCES

1. Dumont, J., Euwart, D., Mei, B., Estes, S., & Kshirsagar, R. (2016). Human cell lines for biopharmaceutical manufacturing: History, status, and future perspectives. *Critical Reviews in Biotechnology*, *36*, 1110–1122.
2. Robert, M. A., Chahal, P. S., Audy, A., Kamen, A., Gilbert, R., & Gaillet, B. (2017). Manufacturing of recombinant adeno-associated viruses using mammalian expression platforms. *Biotechnology Journal*, *12*, 1600193.
3. Gutiérrez-Granados, S., Cervera, L., Kamen, A. A., & Gòdia, F. (2018). Advancements in mammalian cell transient gene expression (TGE) technology for accelerated production of biologics. *Critical Reviews in Biotechnology*, *38*, 918–940.
4. Abaandou, L., Quan, D., & Shiloach, J. (2021). Affecting HEK293 cell growth and production performance by modifying the expression of specific genes. *Cells*, *10*, 1667.
5. Backliwal, G., Hildinger, M., Hasija, V., & Wurm, F. M. (2008). High-density transfection with HEK-293 cells allows doubling of transient titers and removes need for a priori DNA complex formation with PEI. *Biotechnology and Bioengineering*, *99*, 721–727.
6. Backliwal, G., Hildinger, M., Chenuet, S., Wulhfard, S., De Jesus, M., & Wurm, F. M. (2008). Rational vector design and multi-pathway modulation of HEK 293E cells yield recombinant antibody titers exceeding 1 g/l by transient transfection under serum-free conditions. *Nucleic Acids Research*, *36*, e96.
7. Arena, T. A., Chou, B., Harms, P. D., & Wong, A. W. (2019). An anti-apoptotic HEK293 cell line provides a robust and high titer platform for transient protein expression in bioreactors. *MAbs*, *11*, 977–986.

8. Swiech, K., Kamen, A., Ansorge, S., Durocher, Y., Picanço-Castro, V., Russo-Carbolante, E. M., Neto, M. S., & Covas, D. T. (2011). Transient transfection of serum-free suspension HEK 293 cell culture for efficient production of human rFVIII. *Bmc Biotechnology [Electronic Resource]*, *11*, 114.
9. Lin, C. Y., Huang, Z., Wen, W., Wu, A., Wang, C., & Niu, L. (2015). Enhancing protein expression in HEK-293 cells by lowering culture temperature. *Plos One*, *10*, e0123562.
10. Kiszel, P., Fiesel, S., Voit, S., Waechtler, B., Meier, T., Oelschlaegel, T., Schraeml, M., & Engel, A. M. (2019). Transient gene expression using valproic acid in combination with co-transfection of SV40 large T antigen and human p21 CIP /p27 KIP. *Biotechnology progress*, *35*, e2786.
11. Zhao, H., Lee, K. J., Daris, M., Lin, Y., Wolfe, T., Sheng, J., Plewa, C., Wang, S., & Meisen, W. H. (2020). Creation of a high-yield aav vector production platform in suspension cells using a design-of-experiment approach. *Molecular Therapy – Methods & Clinical Development*, *18*, 312–320.
12. Román, R., Miret, J., Scalia, F., Casablancas, A., Lecina, M., & Cairó, J. J. (2016). Enhancing heterologous protein expression and secretion in HEK293 cells by means of combination of CMV promoter and IFNÎ±2 signal peptide. *Journal of Biotechnology*, *239*, 57–60.
13. Allen, J. M., Halbert, C. L., & Miller, A. D. (2000). Improved adeno-associated virus vector production with transfection of a single helper adenovirus gene, E4orf6. *Molecular Therapy*, *1*, 88–95.
14. Stinski, M. F., & Isomura, H. (2008). Role of the cytomegalovirus major immediate early enhancer in acute infection and reactivation from latency. *Medical Microbiology and Immunology*, *197*, 223–231.
15. Sinzger, C., Digel, M., & Jahn, G. (2008). Cytomegalovirus cell tropism. *Current Topics in Microbiology and Immunology*, *325*, 63–83.
16. Forte, E., Zhang, Z., Thorp, E. B., & Hummel, M. (2020). Cytomegalovirus latency and reactivation: An intricate interplay with the host immune response. *Frontiers in Cellular and Infection Microbiology*, *10*, 130.
17. Coulon, A., Chow, C. C., Singer, R. H., & Larson, D. R. (2013). Eukaryotic transcriptional dynamics: From single molecules to cell populations. *Nature Reviews Genetics*, *14*, 572–584.
18. Mella-Alvarado, V., Gautier, A., Le Gac, F., & Lareyre, J. - J. (2013). Tissue and cell-specific transcriptional activity of the human cytomegalovirus immediate early gene promoter (UL123) in zebrafish. *Gene Expression Patterns*, *13*, 91–103.
19. Vasey, D., Lillico, S., Sang, H., King, T., & Whitelaw, C. (2009). CMV enhancer-promoter is preferentially active in exocrine cells in vivo. *Transgenic Research*, *18*, 309–314.
20. Qin, J. Y., Zhang, L., Clift, K. L., Hulur, I., Xiang, A. P., Ren, B. Z., & Lahn, B. T. (2010). Systematic comparison of constitutive promoters and the doxycycline-inducible promoter. *Plos One*, *5*, e10611.
21. Xia, W., Bringmann, P., McClary, J., Jones, P. P., Manzana, W., Zhu, Y., Wang, S., Liu, Y., Harvey, S., Madlansacay, M. R., McLean, K., Rosser, M. P., MacRobbie, J., Olsen, C. L., & Cobb, R. R. (2006). High levels of protein expression using different mammalian CMV promoters in several cell lines. *Protein Expression and Purification*, *45*, 115–124.
22. Brown, A. J., Sweeney, B., Mainwaring, D. O., & James, D. C. (2015). NF-Î°B, CRE and YY1 elements are key functional regulators of CMV promoter-driven transient gene expression in CHO cells. *Biotechnology Journal*, *10*, 1019–1028.
23. Liu, X. F., Yan, S., Abecassis, M., & Hummel, M. (2008). Establishment of murine cytomegalovirus latency in vivo is associated with changes in histone modifications and recruitment of transcriptional repressors to the major immediate-early promoter. *Journal of Virology*, *82*, 10922–10931.
24. Bain, M., Mendelson, M., & Sinclair, J. (2003). Ets-2 Repressor Factor (ERF) mediates repression of the human cytomegalovirus major immediate-early promoter in undifferentiated non-permissive cells. *Journal of General Virology*, *84*, 41–49.
25. Johari, Y. B., Brown, A. J., Alves, C. S., Zhou, Y., Wright, C. M., Estes, S. D., Kshirsagar, R., & James, D. C. (2019). CHO genome mining for synthetic promoter design. *Journal of Biotechnology*, *294*, 1–13.
26. Lambert, S. A., Jolma, A., Campitelli, L. F., Das, P. K., Yin, Y., Albu, M., Chen, X., Taipale, J., Hughes, T. R., & Weirauch, M. T. (2018). The human transcription factors. *Cell*, *172*, 650–665.
27. Wagner, G. P., Kin, K., & Lynch, V. J. (2013). A model based criterion for gene expression calls using RNA-seq data. *Theory in Biosciences*, *132*, 159–164.
28. Johari, Y. B., Mercer, A. C., Liu, Y., Brown, A. J., & James, D. C. (2021). Design of synthetic promoters for controlled expression of therapeutic genes in retinal pigment epithelial cells. *Biotechnology and Bioengineering*, *118*, 2001–2015.
29. Ferreira, J. P., Peacock, R. W., Lawhorn, I. E., & Wang, C. L. (2011). Modulating ectopic gene expression levels by using retroviral vectors equipped with synthetic promoters. *Systems and Synthetic Biology*, *5*, 131.
30. Hardison, R., & Taylor, J. (2012). Genomic approaches towards finding cis-regulatory modules in animals. *Nature Reviews Genetics*, *13*, 469–483.
31. Pham, L. V., Tamayo, A. T., Yoshimura, L. C., Lin-Lee, Y. C., & Ford, R. J. (2005). Constitutive NF-Î°B and NFAT activation in aggressive B-cell lymphomas synergistically activates the CD154 gene and maintains lymphoma cell survival. *Blood*, *106*, 3940–3947.
32. Gray, S. J., Foti, S. B., Schwartz, J. W., Bachaboina, L., Taylor-Blake, B., Coleman, J., Ehlers, M. D., Zylka, M. J., McCown, T. J., & Samulski, R. J. (2011). Optimizing promoters for recombinant adeno-associated virus-mediated gene expression in the peripheral and central nervous system using self-complementary vectors. *Human Gene Therapy*, *22*, 1143–1153.
33. Smale, S. T., Schmidt, M. C., Berk, A. J., & Baltimore, D. (1990). Transcriptional activation by Sp1 as directed through TATA or initiator: Specific requirement for mammalian transcription factor IID. *Proceedings of the National Academy of Sciences of the United States of America*, *87*, 4509–4513.
34. Zweidler-McKay, P. A., Grimes, H. L., Flubacher, M. M., & Tsichlis, P. N. (1996). Gfi-1 encodes a nuclear zinc finger protein that binds DNA and functions as a transcriptional repressor. *Molecular and Cellular Biology*, *16*, 4024–4034.
35. Vink, C. A., Counsell, J. R., Perocheau, D. P., Karda, R., Buckley, S. M. K., Brugman, M. H., Galla, M., Schambach, A., McKay, T. R., Waddington, S. N., & Howe, S. J. (2017). Eliminating HIV-1 packaging sequences from lentiviral vector proviruses enhances safety and expedites gene transfer for gene therapy. *Molecular Therapy*, *25*, 1790–1804.
36. Wang, Z., Cheng, F., Engelhardt, J. F., Yan, Z., & Qiu, J. (2018). Development of a novel recombinant adeno-associated virus production system using human bocavirus 1 helper genes. *Molecular Therapy—Methods & Clinical Development*, *11*, 40–41.
37. Nong, L., Zhang, Y., Duan, Y., Hu, S., Lin, Y., & Liang, S. (2020). Engineering the regulatory site of the catalase promoter for improved heterologous protein production in Pichia pastoris. *Biotechnology Letters*, *42*, 2703–2709.
38. Brown, A. J., Mainwaring, D. O., Sweeney, B., & James, D. C. (2013). Block decoys: Transcription-factor decoys designed for in vitro gene regulation studies. *Analytical Biochemistry*, *443*, 205–210.
39. Patel, Y. D., Brown, A. J., Zhu, J., Rosignoli, G., Gibson, S. J., Hatton, D., & James, D. C. (2021). Control of multigene expression stoichiometry in mammalian cells using synthetic promoters. *ACS Synthetic Biology*, *10*, 1155–1165.
40. Brandeis, M., Frank, D., Keshet, I., Siegfried, Z., Mendelsohn, M., Names, A., Temper, V., Razin, A., & Cedar, H. (1994). Spl elements protect a CpG island from de novo methylation. *Nature*, *371*, 435–438.
41. Isomura, H., Stinski, M. F., Kudoh, A., Daikoku, T., Shirata, N., & Tsurumi, T. (2005). Two Sp1/Sp3 binding sites in the major immediate-early

proximal enhancer of human cytomegalovirus have a significant role in viral replication. *Journal of Virology*, *79*, 9597–9607.
42. Dynan, W. S., & Tjian, R. (1983). The promoter-specific transcription factor Sp1 binds to upstream sequences in the SV40 early promoter. *Cell*, *35*, 79–87.
43. Vigneron, M., Barrera-Saldana, H. A., Baty, D., Everett, R. E., & Chambon, P. (1984). Effect of the 21-bp repeat upstream element on in vitro transcription from the early and late SV40 promoters. *Embo Journal*, *3*, 2373–2382.
44. Emami, K. H., Navarre, W. W., & Smale, S. T. (1995). Core promoter specificities of the Sp1 and VP16 transcriptional activation domains. *Molecular and Cellular Biology*, *15*, 5906–5916.
45. Eisenhut, P., Mebrahtu, A., Barzadd, M. M., Thalén, N., Klanert, G., Weinguny, M., Sandegren, A., Su, C., Hatton, D., Borth, N., & Rockberg, J. (2020). Systematic use of synthetic 5′-UTR RNA structures to tune protein translation improves yield and quality of complex proteins in mammalian cell factories. *Nucleic Acids Research*, *48*, e119.
46. Jasin, M., & Rothstein, R. (2013). Repair of strand breaks by homologous recombination. *Cold Spring Harbor perspectives in biology*, *5*, a012740.
47. Kim, M., O'Callaghan, P. M., Droms, K. A., & James, D. C. (2011). A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. *Biotechnology and Bioengineering*, *108*, 2434–2446.
48. Bessis, N., GarciaCozar, F. J., & Boissier, M. C. (2004). Immune responses to gene therapy vectors: Influence on vector function and effector mechanisms. *Gene Therapy*, *11*, S10–S17.

SUPPORTING INFORMATION

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

How to cite this article: Johari, Y. B., Scarrott, J. M., Pohle, T. H., Liu, P., Mayer, A., Brown, A. J., & James, D. C. (2022). Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells. *Biotechnology Journal*, *17*, e2200062. https://doi.org/10.1002/biot.202200062

8.2.3 Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives

Received: 31 August 2022 | Revised: 4 December 2022 | Accepted: 5 December 2022

DOI: 10.1002/biot.202200450

Biotechnology Journal

RESEARCH ARTICLE

Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives

Joseph M. Scarrott[1] | Yusuf B. Johari[1] | Thilo H. Pohle[1] | Ping Liu[2] | Ayda Mayer[2] | David C. James[1]

[1]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK

[2]Cell Line Development, REGENXBIO Inc., Rockville, Maryland, USA

Correspondence
David C. James, Department of Chemical and Biological Engineering, University of Sheffield, Mappin St., Sheffield S1 3JD, UK.
Email: d.c.james@sheffield.ac.uk

Present address
Yusuf B. Johari, Lonza Biologics, Cambridge, UK

Abstract

Recombinant adeno-associated virus (rAAV) has established itself as a highly efficacious gene delivery vector with a well characterised safety profile allowing broad clinical application. Recent successes in rAAV-mediated gene therapy clinical trials will continue to drive demand for improved rAAV production processes to reduce costs. Here, we demonstrate that small molecule bioactive chemical additives can significantly increase recombinant AAV vector production by human embryonic kidney (HEK) cells up to three-fold. Nocodazole (an anti-mitotic agent) and M344 (a selective histone deacetylase inhibitor) were identified as positive regulators of rAAV8 genome titre in a microplate screening assay. Addition of nocodazole to triple-transfected HEK293 suspension cells producing rAAV arrested cells in G2/M phase, increased average cell volume and reduced viable cell density relative to untreated rAAV producing cells at harvest. Final crude genome vector titre from nocodazole treated cultures was >2-fold higher compared to non-treated cultures. Further investigation showed nocodazole addition to cultures to be time critical. Genome titre improvement was found to be scalable and serotype independent across two distinct rAAV serotypes, rAAV8 and rAAV9. Furthermore, a combination of M344 and nocodazole produced a positive additive effect on rAAV8 genome titre, resulting in a three-fold increase in genome titre compared to untreated cells.

KEYWORDS
bioprocess engineering, CHO cells, gene expression, mammalian cells, protein expression, recombinant proteins, synthetic biology, systems biology

Abbreviations: CHO, Chinese hamster ovary; ddPCR, digital droplet PCR; eGFP, enhanced green fluorescent protein; GMP, Good Manufacturing Practice; HEK, human embryonic kidney; HPT, hours post-transfection; LiCl, lithium chloride; NaBu, sodium butyrate; PEI, polyethyleneimine; PI, propidium iodide; rAAV, recombinant adeno-associated virus; TMAO, trimethylamine *N*-oxide; TUDCA, tauroursodeoxycholic acid; VPA, valproic acid.

1 | INTRODUCTION

Recent regulatory approvals for recombinant adeno-associated virus (rAAV) mediated gene therapy products (Luxturna in 2017 and Zolgensma in 2019) – together with a significant increase in the number of rAAV gene therapies in clinical trials[1,2] – has highlighted the need for improved production process technology with respect to titre, product quality and cost. Indeed, the high cost of manufacturing rAAV gene therapies to support the typically high therapeutic dosages used (e.g., 6×10^{13}–2×10^{14} vector genomes/kg[3]) has limited the economic viability of AAV gene therapies despite evidence of promising clinical efficacy. Therefore, there is an urgent need for robust, intensified process technology to support production of Good Manufacturing Practice (GMP) quality material. As nearly all current rAAV production processes rely on co-transfection of three plasmids encoding the necessary AAV and helper genes into HEK293 cell hosts for transient production of rAAV within 3–5 days,[4] efficient use of plasmid DNA (itself produced using a costly GMP-approved process) is also of paramount importance.

rAAV is a complex macromolecular product requiring a diverse network of cellular processes and molecular interactions to enable coordinated cellular synthesis – involving host-cell proteins, transiently expressed viral helper, capsid and replicase genes, as well as the single stranded therapeutic viral DNA payload itself.[5–7] Small molecule enhancers of recombinant protein production have demonstrated efficacy in a wide variety of mammalian cell lines[3,8,9] and the addition of chemicals as diverse as sodium chloride, sodium butyrate (NaBu) and soy peptones have been shown in previous studies to improve rAAV production yields.[9–11] Targeting of discrete pathways involved in the replication, packaging and trafficking of viral particles by bioactive small molecule cell culture additives offers a simple and cost-effective way of increasing viral titre and reducing overall production costs.

Here, we show that microplate plate-based screening of chemical additives can be used to rapidly identify positive effectors of rAAV synthesis in human embryonic kidney (HEK) cells and that significant improvements in viral genome titre can be obtained by subsequent optimisation of dosing regimen. Furthermore, we also demonstrate that a combination of two functionally distinct small molecule enhancers may act additively, resulting in a three-fold increase in viral vector genome titre and thus a substantial improvement in the efficiency of plasmid DNA usage.

2 | EXPERIMENTAL SECTION

2.1 | Cell culture

Proprietary suspension adapted HEK293 cells were provided by REGENXBIO (Rockville, MD). Cells were cultured in a serum free medium supplemented with L-glutamine (Thermo Scientific, Waltham, MA). Cells were maintained in an orbital shaking incubator (Infors, Bottmingen, Switzerland) at 30 ml culture volumes in 125 ml Erlenmeyer flasks (Corning) at 37°C, 5% CO_2 and 85% humidity, with agitation at 140 rpm. Smaller scale cultures were grown in shallow-well 24-well plates (0.7 ml culture volume/well, 240 rpm shaking) (Corning, Acton, MA) using the Deutz system.[12] Routine cell density and cell viability measurements were performed on a ViCell automated cell counter (Beckman-Coulter, Brea, CA). Measurements of density and viability and determination of mean cell volume during rAAV production runs were performed using the Norma HT system (Iprasense) utilising 3 μl total cell culture and 20 μm slide chambers. Mean cell volume (V) was calculated as $V = \frac{4}{3}\pi r^3$ where r = measured cell diameter/2.

2.2 | rAAV vector production in suspension adapted HEK293 cells

An HEK293-derived cell line adapted to suspension in Dynamis media, SKMB (REGENXBIO Inc.), was subjected to a triple-transfection protocol as previously described,[13] with the following modifications: (1) cis plasmid containing an enhanced green fluorescent protein (eGFP) expression cassette flanked by AAV2 inverted terminal repeats (ITRs); (2) trans plasmids pAAV2/8 and pAAV2/9 containing the AAV2 rep gene and capsid protein genes from AAV8 and 9, respectively and (3) adenovirus helper plasmid pAdDeltaF6 were transfected using polyethyleneimine (PEI) (Polyplus Transfection, Illkirch, France) according to the manufacturer's instructions. Plasmids and PEI were diluted in media and incubated at room temperature, then added to cells. For analysis, 450 μl of total cell culture was added to 50 μl 10x cell lysis buffer containing 1x cOmplete, EDTA-free Protease Inhibitor Cocktail (Roche) and incubated for 1 h at 37°C with gentle agitation. Samples were briefly centrifuged at 12,000 RPM to remove cell debris and the resulting supernatant used to determine viral genome and capsid titre.

2.3 | Measurement of rAAV genome titre

Genome titre was quantified by digital droplet PCR (ddPCR). Aliquots of post-lysis supernatant (5 μl) from total cell culture were treated with DNAse I (Roche, Basel, Switzerland) to remove residual plasmid DNA. DNAse I-treated samples were diluted 1000- or 10,000-fold in PCR dilution buffer (GeneAmp PCR Buffer I [Thermo Scientific], 0.02% UltraPure Salmon Sperm DNA Solution [Invitrogen, Waltham, MA], 0.1% Pluronic F-68 non-ionic surfactant). Droplet formation and subsequent post-PCR droplet analysis were performed using the QX200 system (Bio-Rad, Hercules, CA), with absolute quantification of AAV genome copies/μl determined using the Quantasoft analysis software (Bio-Rad). Genome detection was achieved using primers and an FAM-labelled probe targeting the PolyA sequence of the pAAV-CAG-GFP plasmid. Capsid titre quantification was performed using the AAV8 titration ELISA (Progen, Heidelberg, Germany) from total cell lysis supernatant diluted in 1x ASSB assay buffer (Progen).

2.4 | Chemical additive screening in microplates

Chemical additives for screening were diluted in either sterile water or dimethyl sulphoxide (DMSO, 100% v/v) where appropriate. Low, medium and high concentrations for each chemical were based on available literature describing their observed effects on recombinant protein expression in vitro (listed in Table S1). Cells were initially grown to a density of 4×10^6 viable cells/ml in Erlenmeyer flasks and mid-exponential phase cells were transfected with AAV-encoding plasmids prior to transferring to 24-shallow-well microplates (Corning) at 2.8×10^6 cells per well, 20 min post-transfection. Chemical additives were added to cells 24 hours post-transfection (HPT). Cell cultures were harvested at 72 HPT for viral genome titre analysis by ddPCR.

2.5 | HEK cell DNA content analysis by flow cytometry

Cells (1×10^6) were harvested 72 HPT and fixed in 70% v/v EtOH at 4°C for 30 min, with gentle vortexing to prevent clumping. EtOH was removed and the cells washed twice in 1x PBS. Fixed cells were treated with 100 μl RNAse A (100 μg/ml in PBS; Qiagen) for 5 min at RT and the DNA stained with the subsequent addition of 400 μl propidium iodide (PI) (50 μg/ml in PBS; Thermo Scientific) at room temperature for a minimum of 30 min. PI-stained cells were analysed by flow cytometry using an LSRII instrument (BD Biosciences). Gated single cell populations were detected based on PI signal and the resulting histograms analysed using FlowJo software to determine the relative distribution of cells within the cell cycle (G1/S/G2-M) phases.

2.6 | Immunocytochemistry

rAAV producing cells were harvested 24 h after nocodazole addition. 1×10^6 cells were spun down at 250 g for 5 min, culture media was removed and cells resuspended in 1 ml PBS. Resuspended cells were incubated with no agitation for 30 min at RT and allowed to adhere by gravity sedimentation to lysine-coated glass coverslips in 24-well plates. After 30 min, PBS and unadhered cells were removed by gentle aspiration. Adhered cells were fixed by 10 min incubation at RT with 0.5 ml 4% v/v paraformaldehyde in PBS. Fixed cells were washed with PBS before permeabilisation with 0.5 ml 0.5% v/v Triton-X100 for 10 min. Permeabilised cells were blocked by 10% normal goat serum (Life Technologies, Carlsbad, CA; #016201) in PBS for 30 min. Cells were then incubated with anti-Fibrillarin antibody (Abcam, Cambridge, UK; ab5821, 0.1 μg/ml in blocking buffer) for 1 h at RT before incubation with goat anti-Rabbit Alexa-594 conjugated secondary antibody (Abcam; ab150080, 1:1000 dilution in blocking buffer) for 45 min at room temperature. Coverslips were transferred to a glass slide and mounted with Fluoroshield mounting media containing DAPI (Sigma–Aldrich; F6507) for visualisation of nuclei. Slides were imaged using a Nikon Eclipse Ti fluorescent microscope and images adjusted for brightness and contrast using ImageJ software.

2.7 | Request for materials

All requests for materials used in this article should be directed to the primary author at: j.scarrott@sheffield.ac.uk.

3 | RESULTS

3.1 | Initial screening of chemical additives to increase AAV genome titre

To identify novel small molecule enhancers of rAAV production, a panel of 15 small molecule culture additives was chosen based on either their reported propensity to increase recombinant protein expression in mammalian cell systems or their reported positive effects on rAAV transduction efficiency. Due to the diverse mechanisms involved in AAV vector production, the panel featured chemicals that exert their influence on recombinant protein expression via a broad range of functional mechanisms. Specifically, we tested chemicals reported to be "chemical chaperones" (tauroursodeoxycholic acid [TUDCA], trimethylamine *N*-oxide [TMAO], betaine),[14–16] cell cycle modulators (nocodazole, BI-2536),[17,18] caspase inhibitors (z-VAD-fmk),[19] histone deacetylase inhibitors (NaBu, valproic acid [VPA], M344, apicidin),[8,20–22] insulin-mimetics (lithium chloride [LiCl] and zinc sulphate)[23] and proteasome inhibitors (ONX0912, MLN9708, MG132).[24–26]

To enable multi-parallel screening of small molecule additives, we employed pre-optimised small-scale cultures (700 μl culture volume) in 24-well microplates that exhibited a similar rAAV production profile to shake flask cultures (data not shown). Triple transfection was performed using PEI in a serum-free medium, and chemical additives were introduced to rAAV8-producing suspension HEK293 cells at 24 HPT. Each additive was used at three different concentrations (low, medium and high – values based on previous literature, see Table S1). Total cell culture was harvested at 72 HPT, cells were chemically lysed and the rAAV8 genome titre in the crude supernatant was measured by ddPCR. Crude genome titre in cells containing chemical additives was compared to that of untreated rAAV8 producing cells to determine changes to titre mediated by the small molecules. Several chemicals displayed a clear concentration-dependent impact on genome titre, both positively (z-VAD-fmk) and negatively (LiCl, ONX0912) (Figure 1A). Several compounds with similar mechanistic properties showed a marked reduction in rAAV titre at the tested concentrations – most notably inhibitors of proteasomal function (ONX0912, MLN9708, MG132) and the reported chemical chaperones (TUDCA, TMAO, betaine). Notably, the screening showed that high relative dose nocodazole (4 μM) added to culture 24 HPT resulted in a 1.34-fold increase in rAAV8 titre compared to untreated control cells ($p = 0.62$, $n = 2$), while M344 appeared to be effective at low dose (2.5 μM) with a 1.36-fold increase in rAAV8 titre ($p = 0.64$, $n = 2$). Of note, nocodazole treatment increased mean cell volume at all three concentrations, an attribute that appears to correlate positively with measured genome titre ($r^2 = 0.79$, $p < 0.0001$) (Figure 1B) and has been shown in previous reports to be a major

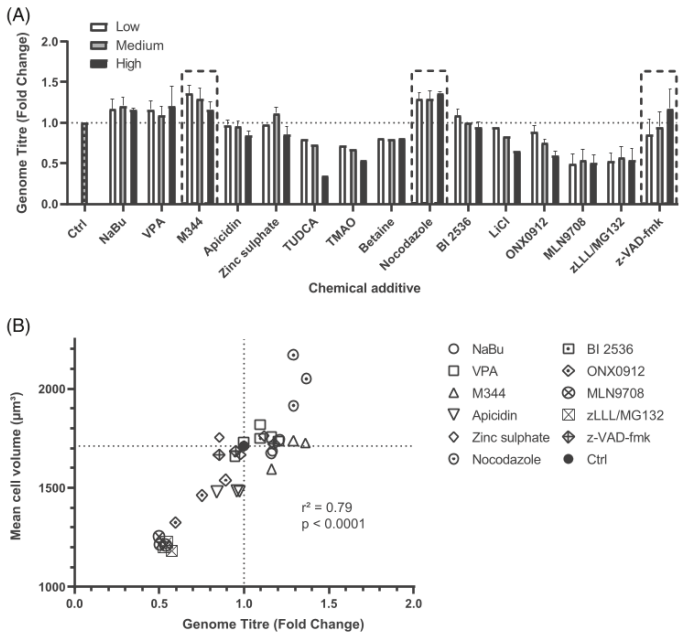


FIGURE 1 Initial screen of small molecule culture additives identifies nocodazole, M344 and z-VAD-fmk as novel putative enhancers of rAAV genome titre. Cells were analysed 72 HPT for rAAV8 genome titre, shown here as the fold change relative to rAAV8 producing cells containing no small molecule enhancers (Ctrl). Nocodazole, M344 and z-VAD-fmk are highlighted (dashed boxes) (A). Mean cell volume versus genome titre fold change is shown in (B), dashed lines indicate control values. R^2 and *p*-value calculated from Pearson correlation coefficient of mean cell volume versus genome titre. Data shown in (A) are mean ± SEM for two independent experiments carried out in technical duplicate, with the exception of TUDCA, TMAO, betaine and LiCl which are mean only data from one independent experiment carried out in technical duplicate (data from these experiments is omitted in (B)). HPT, hours post-transfection; LiCl, lithium chloride; NaBu, sodium butyrate; rAAV, recombinant adeno-associated virus; SEM, standard error of mean; TMAO, trimethylamine *N*-oxide; TUDCA, tauroursodeoxycholic acid; VPA, valproic acid

cellular determinant of recombinant protein productivity in Chinese hamster ovary (CHO) cells.[27–29] Nocodazole treated cells were found to be up to 27% larger than control cells. Nevertheless, this phenotypic variation was not observed with M344, in which the increase in titre was not accompanied by an increase in cell volume. It may therefore be inferred that these molecules modulate rAAV expression via distinct mechanisms independent of each other and that increased control of rAAV production may be achieved by the application of specific combinations of small molecule effectors.

3.2 | Early addition of nocodazole enhances rAAV8 production in HEK293 cells

A recent study showed that *cap* gene expression, translation and assembly of Cap proteins into empty AAV particles is a kinetically rapid process.[30] Approximately 80% of capsids are assembled within the first 24 HPT whereas Rep mediated replication of rAAV genome from the *cis*-ITR plasmid is slower, peaking after capsid protein levels have plateaued due to the inhibitory action of Rep binding to the packaging plasmid and inhibiting transcription or by translational repression of *cap* mRNA. As capsid secretion from the nucleus into the cytoplasm is independent of genome loading, pre-assembled empty capsids can be secreted from the site of rAAV genome loading, thus depleting the pool of empty capsids available to newly replicated rAAV genome.[32,31] Therefore, we hypothesised that the effect of a given chemical effector may be positive (or further enhanced), neutral or negative with respect to the timing of its deployment.

To further investigate the impact of nocodazole in enhancing rAAV production, we deployed high dose nocodazole (4 μM) between 0 and 48 HPT (Figure 2). Addition of 4 μM nocodazole immediately (0 HPT) and 4 HPT resulted in a considerable reduction in VCD at 72 HPT of 74% and 64% respectively, compared to rAAV8 producing cells without nocodazole (Ctrl) (Figure 2A). Also apparent at early addition timepoints was a slight reduction in cell viability ($p > 0.05$; Figure 2B) and a significantly higher mean cell volume compared to cultures not treated

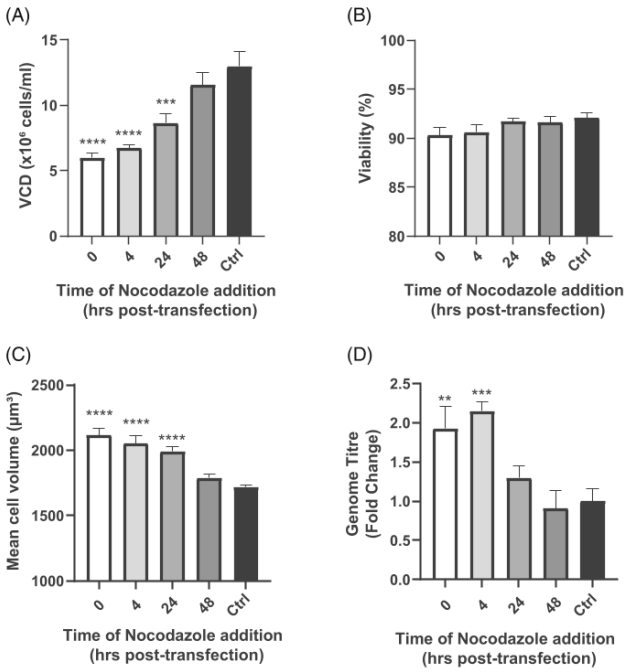


FIGURE 2 Addition of nocodazole to rAAV8 producing cells to increase genome titre is time-dependent. rAAV8 producing cells were cultured in 24-well microplates and 4 μM nocodazole was added at 0-, 4-, 24-, or 48-HPT or left untreated (Ctrl). At 72 HPT cells were harvested and VCD (A), viability (B) and mean cell volume (C) was measured. Final crude genome titre expressed as a fold change relative to Ctrl is shown in (D). Data shown for VCD, viability and mean cell diameter are the mean ± SEM. $n > 3$ independent biological replicates. Data shown for genome titre fold change are the mean ± SEM. $n \geq 3$ independent biological replicates. Data were analysed by one-way ANOVA followed by post-hoc Holm-Šídák multiple comparisons test with respect to Ctrl. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$. ANOVA, analysis of variance; HPT, hours post-transfection; SEM, standard error of mean; VCD, viable cell density

with nocodazole (up to 23%; Figure 2C), consistent with previous work by Tait et al. (2004) who observed increased cell size and decreased viability in nocodazole treated CHO cells producing a recombinant monoclonal antibody. Most strikingly, addition of nocodazole at 4 HPT resulted in a 2.2-fold increase in rAAV8 genome titre at harvest compared to control cells (Figure 2D). Addition at the earlier timepoint of 0 HPT was sub-optimal, with a 10% reduction in rAAV8 titre compared to the titre of cells treated at 4 HPT (Figure 2D). We therefore conclude that early addition (~4 HPT) of nocodazole to rAAV producing cells is an effective positive mediator of rAAV production in a small-scale, transient expression and suspension HEK cell system.

3.3 | Nocodazole improves rAAV genome titre in two different serotypes

To investigate the effect of nocodazole addition in a system more representative of large scale rAAV production, culture volume was scaled up from the initial microplate culture volume of 0.7–30 ml culture volume in shake flasks. To rule out a serotype-dependent effect of nocodazole addition, plasmids for both rAAV8 and rAAV9 serotypes were separately transfected in HEK293 cells and cultured with and without the addition of 4 μM nocodazole. rAAV genome titre was measured from 24 to 72 HPT. All measured parameters of nocodazole induced cell cycle arrest (reduced VCD, decreased viability and increased mean cell volume) and increased rAAV genome titre were consistent with those seen at smaller culture volumes (Figure 3A,B). Comparable to microplate-based cultures, mean viral genome titre was increased by up to 2.5-fold in nocodazole-treated cultures compared to untreated cultures (Figure 3C). Importantly there were no observed significant differences in the above-described measurements between cultures producing rAAV8 or rAAV9, suggesting a broad applicability for nocodazole addition in rAAV manufacturing.

3.4 | rAAV producing cultures treated with nocodazole have an increased proportion of cells in G2/M phase

Nocodazole is an anti-mitotic agent, used as both a chemotherapeutic and as a common agent of cell cycle synchronisation.[17,32–34] Nocodazole exerts its effect by reversibly inhibiting the polymerisation of β-tubulin, destabilising microtubules and preventing the formation of mitotic spindles, thus arresting cells within the G2/M phase of the cell cycle. Cells treated with nocodazole typically enter mitosis but are unable to progress through cytokinesis, leading to either apoptosis in cells remaining in arrested mitosis for an extended period of time, or subsequent "mitotic slippage" into G0/G1 phase followed

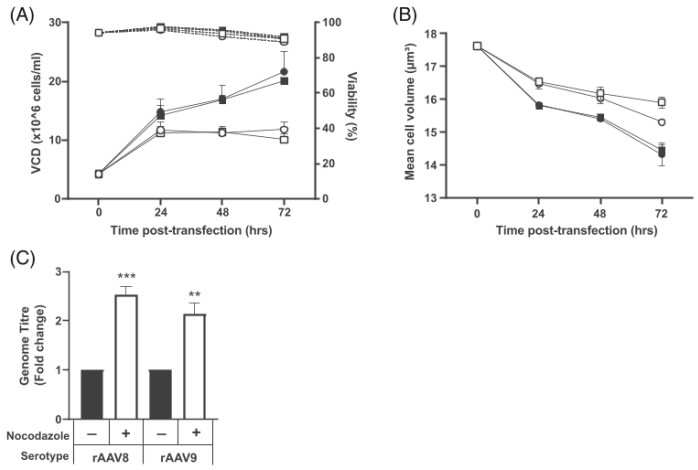


FIGURE 3 Increased genome titre in nocodazole treated cells is maintained in larger scale cultures and across two separate serotypes. 30 ml shake flask cultures were transfected with rAAV8 or rAAV9 producing plasmids and cultured for 72 HPT in the presence (rAAV8: ○; rAAV9: □) or absence (rAAV8: ●; rAAV9: ■) of 4 μM nocodazole added 4 HPT. Cultures were measured daily for VCD (A – solid lines), viability (A – dashed lines) and mean cell volume (B). Fold change analysis of genome titre for both serotypes in the presence (+) or absence (–) of 4 μM nocodazole at 72 HPT is shown in (C). Data shown are the mean ± SEM. $n = 3$ independent biological replicates. Fold change between nocodazole treated and untreated cultures was analysed for each serotype separately by Student's unpaired two-tailed *t*-test with respect to untreated cultures. $^{**}p < 0.01$, $^{***}p < 0.001$. HPT, hours post-transfection; rAAV, recombinant adeno-associated virus; SEM, standard error of mean; VCD, viable cell density

by apoptosis.[34–37] In addition to its use as a cell synchronisation agent, nocodazole has previously been shown to increase transient recombinant protein expression in a mammalian cell system.[38]

Flow cytometry was carried out to determine the cell cycle status of rAAV8-producing cultures treated with nocodazole. Untreated rAAV8 producing cells were found to be almost entirely in either G1 or S phase at 72 HPT (Figure 4A). Cells treated with 4 μM nocodazole at the point of transfection (0 HPT) resulted in a very significant proportion (38%) arresting in G2/M phase (Figure 3B). The ratio of cells in G2/M:G1 phase was found to decrease the later nocodazole was added (Figure 4C). Of note is the observation that a significant number (26%) of cells remain arrested in G2/M phase at 72 HPT when nocodazole is added as late as 24 HPT, yet the positive effect on rAAV8 genome titre is substantially lessened compared to earlier treatment. This would suggest a critical temporal component to the addition of nocodazole, such that later additions to cell culture are sub-optimal in terms of producing high rAAV titres, likely due to the rapid assembly process of cap proteins that largely occurred within 24 HPT.[30]

The nucleolus is a dynamic compartment within the nucleus which undergoes extensive remodelling during cell cycle progression, particularly during mitosis.[39] Previous immunocytochemical examination of nucleolar localisation throughout normal cell cycle progression shows a distinctive dispersal of nucleolar protein staining throughout the cytoplasm during prometaphase and metaphase – concomitant with the breakdown of the nuclear membrane during mitosis.[40–42] Nucleolar proteins in cells in interphase and prophase are typically found located within the nucleus itself. Of note, the nucleolus is considered to be a likely site for AAV capsid assembly and Rep-mediated loading of AAV genome into capsids,[43–46] as well as being linked more generally to viral replication in other human viruses.[47]

Cells from rAAV8 producing cultures – both with and without nocodazole addition 4 HPT – were harvested and fixed at 24 HPT. Cells were stained with a nucleolar marker (fibrillarin – a protein component of the nucleolus associated with ribosomal RNA processing[41]) and DAPI (a nuclear DNA stain). Widefield fluorescent microscopy was used to visualise the phenotypic changes induced by nocodazole addition. Untreated cells displayed highly localised or punctate staining of fibrillarin within the area of nuclear DNA staining (Figure 4D), indicative of cells in interphase or prophase. In contrast, cells treated with nocodazole exhibited a proportion displaying a disorganised nuclear phenotype – with fibrillarin distributed broadly throughout the cytoplasm, and with condensed nuclear DNA (Figure 4E) – indicative of cells either progressing through mitosis or arrested within G2/M phase.

3.5 | Small molecule additives can be combined to further enhance rAAV8 production

In order to evaluate whether specific combinations of effectors could act synergistically to further increase rAAV production, we utilised

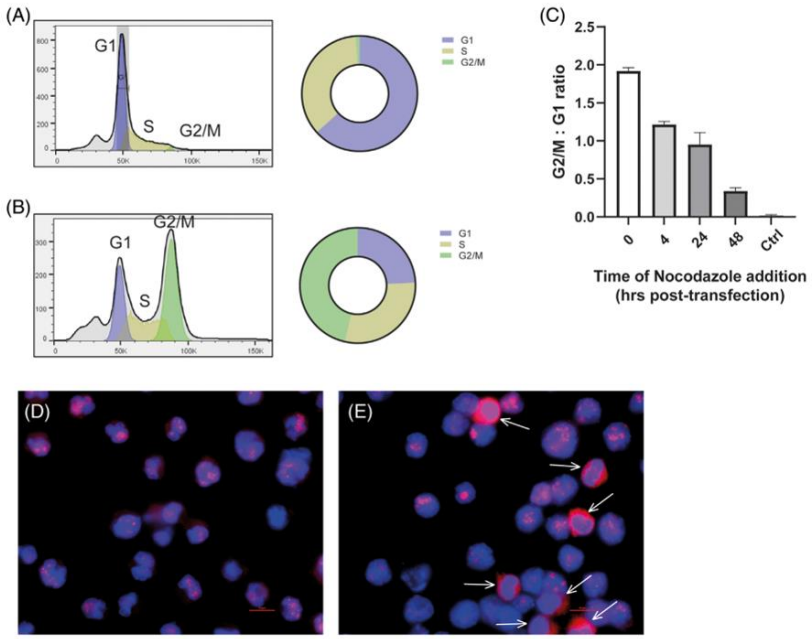


FIGURE 4 rAAV8 producing cells treated with nocodazole arrest in G2/M phase. Flow cytometry histograms triggered against propidium iodide signal for cells show the relative abundance of cells in either G1 (blue), S (orange) or G2/M (green) phases of the cell cycle in the absence (A) or presence (B) of 4 μM nocodazole addition at 0 HPT. Ratio of cells in G2/M phase to G1 phase at 72 HPT with 4 μM nocodazole added at 0, 4, 24 and 48 HPT or untreated (Ctrl) (C). Representative fluorescent composite images of rAAV-producing cells fixed 24 HPT show staining of the fibrillar component of the nucleolus with anti-fibrillarin (red) and nuclear DNA with DAPI (blue). Cells displaying disorganised nucleolar morphology are indicated by white arrows in non-treated cells (D) and cells treated 4 HPT with 4 μM nocodazole (E). Images taken using 100 × oil objective. Scale bar 10 μm. Data shown in (C) are mean ± SEM for three biological replicates. HPT, hours post-transfection; rAAV, recombinant adeno-associated virus; SEM, standard error of mean

nocodazole in combination with either z-VAD-fmk or M344 (see Figure 1). M344 is a synthetic analogue of the anti-fungal drug Trichostatin A, an inhibitor of Class I and IIB histone deacetylases and recently used as an enhancer of recombinant protein expression in mammalian cell culture.[10] Z-VAD-fmk is a pan-caspase inhibitor.[19] Each chemical was added to rAAV8 producing HEK293 cultures together with nocodazole at 4 HPT in 24-well microplates. The inclusion of z-VAD-fmk in the initial screening experiment was predicated on its anti-caspase activity, as we hypothesised caspase mediated apoptosis – linked to rAAV production and expression of AAV2 Rep proteins – would negatively affect final crude titre.[50,49] Additionally, caspase-mediated apoptosis induced by nocodazole-mediated cell cycle dysregulation would further increase cell death within the production cultures. Increased cell viability and VCD in cultures treated with nocodazole/z-VAD-fmk relative to untreated, nocodazole treated or nocodazole/M344 treated cells suggest that apoptosis was reduced (Figure 5A,B) but with an unexpected reduction in crude genome titre (Figure 5D). The observed reduction may be a consequence of off-target induction of autophagy via inactivation of *n*-glycanase 1 (NGLY1) that has been shown to occur in HEK293 cells treated with z-VAD-fmk.[50] While the role of autophagy in rAAV production is undetermined, inhibition of autophagy has been shown to increase recombinant protein expression in CHO cells[51,52] and unintentional upregulation of autophagy may result in a reduction in viral component proteins necessary for production of high viral titres. Interestingly, the highest mean cell volume, which appeared to correlate positively with genome titre, was measured in cells treated with both nocodazole and z-VAD-fmk (Figure 5C). This may be a result of reduced apoptotic cell death allowing cell size to increase but with the caveat that any benefit gained from this from a production perspective is attenuated by the potential negative off-target effects of z-VAD-fmk.

We observed that the combination of 2.5 μM M344 and 4 μM nocodazole produced an additive effect, increasing crude rAAV genome titre 2.6-fold compared to untreated cultures, an improvement on nocodazole alone (two-fold increase compared to untreated) (Figure 5D). This additive effect was also observed in 30 ml shake

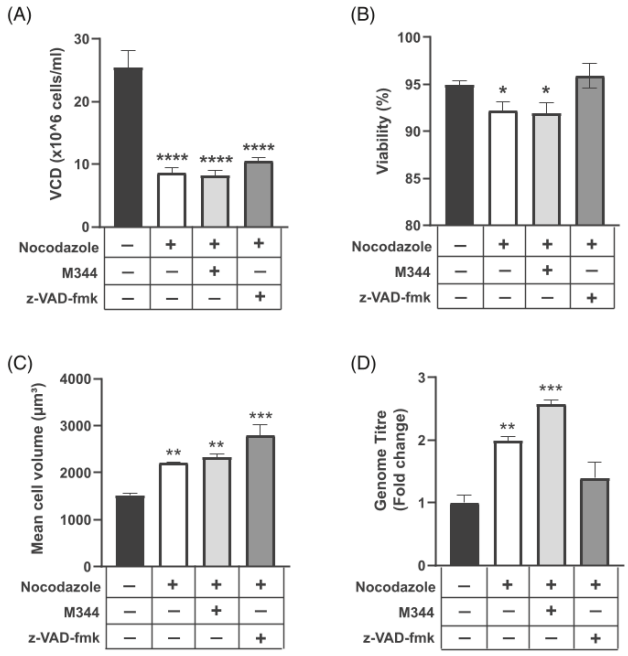


FIGURE 5 Combinatorial use of small molecule enhancers has an additive effect on genome titre. rAAV8 producing cells were treated 4 HPT with small molecule enhancers of genome titre or left untreated. Cells were harvested 72 HPT and measurements taken of VCD (A), viability (B), mean cell volume (C) and rAAV8 genome titre as a fold change with respect to untreated cells (–/–/–) (D). Data shown are the mean ± SEM. $n = 3$ independent biological replicates. Data were analysed by one-way ANOVA followed by post-hoc Holm-Šídák multiple comparisons test with respect to untreated cells (–/–/–). $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, $^{****}p < 0.0001$. ANOVA, analysis of variance; HPT, hours post-transfection; SEM, standard error of mean; VCD, viable cell density

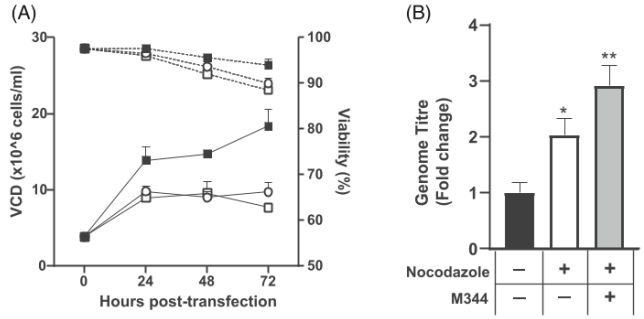


FIGURE 6 Genome titre improvements are retained in larger scale cultures. 30 ml shake flask cultures were transfected with rAAV8 producing plasmids with no chemical additives (■), with the addition of 4 μM nocodazole at 4 HPT (○) or with the addition of 4 μM nocodazole and 2.5 μM M344 at 4 HPT (□). Cultures were measured daily for VCD (A – solid lines) and viability (A – dashed lines). Genome titre was measured 72 HPT and is shown as fold change with respect to untreated cultures (B). Data shown are the mean ± SEM. $n = 3$ independent biological replicates. Data were analysed by one-way ANOVA followed by post-hoc Holm-Šídák multiple comparisons test with respect to untreated cells (–/–). $^{*}p < 0.05$, $^{**}p < 0.01$. ANOVA, analysis of variance; HPT, hours post-transfection; SEM, standard error of mean; VCD, viable cell density

flask cultures, whereby measurements of VCD and viability closely replicated those observed in the screening assay (Figure 6A), together with an improvement in genome titre with nocodazole/M344 from a 2.6-fold increase in the screening assay to a three-fold increase compared to untreated cells in larger volume cultures (Figure 6B). AAV8 intact capsid-specific ELISA analysis of total capsid titre showed a significant increase in intact capsids after addition of nocodazole (4.3-fold increase) and nocodazole/M344 (11.3-fold increase) compared to

untreated controls (data not shown). While this increase in the relative ratio of empty to full capsids is suboptimal in terms of commonly assessed product quality attributes, recent improvements in downstream processing of crude AAV lysate will likely reduce the impacts of excess empty capsids during final product formulation, whilst retaining the benefits of increased genome titre.[53] Additionally, we anticipate that rAAV vector development could complement the process engineering strategy to further maximise rAAV titre (e.g., high intact capsid levels are likely to benefit from hybrid Rep with improved genome packaging efficiency)[54] while combinatorial empirical modelling will enable systematic determination of optimal chemical dosage and timing.

Taken together, these data show that nocodazole, either alone or in combination with select small molecules, can reproducibly boost both genome and total viral particle titre in a transient suspension HEK293 rAAV production system and that it is both scalable and applicable to production of rAAVs derived from two phylogenetically distinct and clinically translatable pseudotyped capsids.

4 | DISCUSSION

Improving the yield of intact, genome-containing rAAV particles during viral vector production is a critical step to reducing overall production costs. Here, we describe a simple and robust method by which viral vector genome titre, a fundamental quality attribute of viral vector production, may be quickly improved in an established and previously optimised suspension culture system. Small molecule enhancers of recombinant protein expression have been extensively used across a wide range of mammalian cell systems to improve transient production performance.[8,55–56,21,38,57,58] The ease of use and low cost of small molecule enhancers (particularly for chemicals with pronounced biological activity at low dosages) makes them an attractive solution to improving rAAV yield.

Investigation of the cellular mechanisms underpinning increased viral genome titre point to the arrest of rAAV producing cells in mitosis soon after PEI-mediated transient transfection of rAAV producing plasmids. Prior studies have shown a strong preference for wildtype AAV replication within the G2/M phase of the cell cycle.[59,60] Nocodazole has been shown to significantly improve transient transfection efficiency in CHO cells,[38] which may stem from increased nuclear permissibility of transfection complexes due to the breakdown of the nuclear membrane during mitosis. A reduction in cell proliferation caused by nocodazole addition may also benefit viral production by reducing plasmid copy number dilution and maintaining mRNA transcript levels.[56] A recent study utilising a CRISPR-mediated genome wide screening strategy identified two target genes (ITPRIP and SKA2) that when modulated in cells increased rAAV genome titre and improved full/empty capsid ratios, with both target genes (strongly, in the case of SKA2) associated with cell cycle modulation.[61] Further to this, a proteomic study of HEK cells during AAV5 production highlighted a number of proteins involved in cell cycle and proliferation as being strongly downregulated during production.[62] The molecular effects of cell cycle arrest within the G2/M phase on rAAV production are unknown, but there is a potential correlation between the loss of essential nucleolar functions and the nucleolar localisation of viral proteins, while the volume increase could possibly minimise crowding effects due to accumulation of viral proteins.[63] This apparent link between the cell cycle and AAV production, and the abundance of cell cycle modulating molecules, necessitates further investigation into the use of cell cycle modulators for both improving rAAV production yields and ultimately improving our understanding of the underlying biological processes governing rAAV production.

The apparent important temporal aspect of cell cycle regulation within the production process may also provide avenues for non-chemical interventions to improve vector yield as this relationship becomes better understood. Whilst we have not investigated the mechanism behind M344-mediated titre enhancement, HDAC6 (of which M344 is a selective inhibitor) has been shown to bind to, and regulate clearance of, ubiquitinated proteins via induction of the heat-shock cellular response.[64] AAV capsid proteins are a known target of ubiquitination post-viral entry,[65] and the ubiquitin-proteasome pathway (UPP) has been suggested to play an active role in AAV capsid monomer degradation,[66] therefore we posit a link between M344-mediated UPP dysregulation and increased crude viral titre. Due to the robustness of the results between small-scale plate-based cultures and larger scale shake flasks, we believe that the screening process shown here could be further scaled down and automated to increase throughput, due to the availability of instrumentation that can rapidly and accurately dispense very small volumes of drugs into culture. As minimal volumes are required for ddPCR analysis of genome titre, identification of novel enhancers of rAAV could be rapidly incorporated into existing rAAV production platforms with minimal changes to existing protocols.

In summary, we show that the use of readily available small molecule enhancers can significantly improve rAAV production yield. We show that small molecule enhancers of rAAV production are amenable to optimisation in an existing suspension HEK293 cell system and that positive hits from initial small-scale screening of enhancer molecules are translatable to improvements in genome titre up to at least 30 ml shake flask scale, with the potential for translation to commercially viable volumetric production (>10 L), as has been demonstrated for a similar production system.[67] We also show that increased titre resulting from nocodazole treatment is consistent across two different serotypes, suggesting broad applicability in rAAV manufacturing.

AUTHOR CONTRIBUTIONS

Joseph M. Scarrott: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Writing – original draft. Yusuf B. Johari: Formal analysis; Investigation; Methodology; Writing – review & editing. Thilo H. Pohle: Formal analysis; Investigation; Methodology; Writing – review & editing. Ping Liu: Conceptualization; Project administration; Resources; Supervision; Writing – review & editing. Ayda Mayer: Conceptualization; Project administration; Resources; Supervision; Writing – review & editing. David C. James: Conceptualization;

Funding acquisition; Methodology; Project administration; Resources; Supervision; Writing – review & editing.

ACKNOWLEDGEMENTS

This study was supported by REGENXBIO Inc., U.S.A. The authors thank Susan Clark (University of Sheffield) for technical assistance in flow cytometry analysis and Lynne Baxter (University of Sheffield) for access to widefield fluorescence microscopy equipment.

CONFLICTS OF INTEREST

Part of this work was conducted under a research agreement between the University of Sheffield and REGENXBIO Inc. Ping Liu and Ayda Mayer are employed by REGENXBIO Inc. and hold shares in the company.

DATA AVAILABILITY STATEMENT

The data that support the findings of this study are available from the lead author upon request. Please contact j.scarrott@sheffield.ac.uk for data.

ORCID

Joseph M. Scarrott https://orcid.org/0000-0002-6046-7687
Yusuf B. Johari https://orcid.org/0000-0001-9933-5764

REFERENCES

1. Bulaklak, K., & Gersbach, C. A. (2020). The once and future gene therapy. *Nature Communications*, *11*, 11–14.
2. Kuzmin, D. A., Shutova, M. V., Johnston, N. R., Smith, O. P., Fedorin, V. V., Kukushkin, Y. S., van der Loo, J. C. M., & Johnstone, E. C. (2021). The clinical landscape for AAV gene therapies. *Nature Reviews Drug Discovery*, *20*, 173–174.
3. . Mendell, J. R., Al-Zaidy, S., Shell, R., Arnold, W. D., Rodino-Klapac, L. R., Prior, T. W., Lowes, L., Alfano, L., Berry, K., Church, K., Kissel, J. T., Nagendran, S., L'Italien, J., Sproule, D. M., Wells, C., Cardenas, J. A., Heitzer, M. D., Kaspar, A., Corcoran, S., … Kaspar, B. K. (2017). Single-dose gene-replacement therapy for spinal muscular atrophy. *New England Journal of Medicine*, *377*, 1713–1722.
4. Xiao, X., Li, J., & Samulski, R. J. (1998). Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *Journal of Virology*, *72*, 2224–2232.
5. Maurer, A. C., & Weitzman, M. D. (2020). Adeno-associated virus genome interactions important for vector production and transduction. *Human Gene Therapy*, *31*, 499–511.
6. Meier, A. F., Fraefel, C., & Seyffert, M. (2020). The interplay between adeno-associated virus and its helper viruses. *Viruses*, *12*, 8–12.
7. Sha, S., Maloney, A. J., Katsikis, G., Nguyen, T. N. T., Neufeld, C., Wolfrum, J., Barone, P. W., Springs, S. L., Manalis, S. R., Sinskey, A. J., & Braatz, R. D. (2021). Cellular pathways of recombinant adeno-associated virus production for gene therapy. *Biotechnology Advances*, *49*, 107764.
8. Chang, J., Chen, X., Wang, R., Shi, R., Wang, X., Lu, W., Ma, S., & Xia, Q. (2020). High-throughput screening identifies two novel small molecule enhancers of recombinant protein expression. *Molecules (Basel, Switzerland)*, *25*, 353.
9. Zhao, H., Lee, K. J., Daris, M., Lin, Y., Wolfe, T., Sheng, J., Plewa, C., Wang, S., & Meisen, W. H. (2020). Creation of a high-yield AAV vector production platform in suspension cells using a design-of-experiment approach. *Molecular Therapy – Methods & Clinical Development*, *18*, 312–320.
10. Yu, C., Trivedi, P. D., Chaudhuri, P., Bhake, R., Johnson, E. J., Caton, T., Potter, M., Byrne, B. J., & Clément, N. (2021). NaCl and KCl mediate log increase in AAV vector particles and infectious titers in a specific/timely manner with the HSV platform. *Molecular Therapy – Methods & Clinical Development*, *21*, 1–13.
11. Hildinger, M., Baldi, L., Stettler, M., & Wurm, F. M. (2007). High-titer, serum-free production of adeno-associated virus vectors by polyethyleneimine-mediated plasmid transfection in mammalian suspension cells. *Biotechnology Letters*, *29*, 1713–1721.
12. Duetz, W. A., Rüedi, L., Hermann, R., O'Connor, K., Büchs, J., & Witholt, B. (2000). Methods for intense aeration, growth, storage, and replication of bacterial strains in microtiter plates. *Applied and Environmental Microbiology*, *66*, 2641–2646.
13. Wang, Q., Nambiar, K., & Wilson, J. M. (2021). Isolating natural adeno-associated viruses from primate tissues with a high-fidelity polymerase. *Human Gene Therapy*, *32*, 1439–1449.
14. Kusaczuk, M. (2019). Tauroursodeoxycholate—bile acid with chaperoning activity: Molecular and cellular effects and therapeutic perspectives. *Cells*, *8*, 1471.
15. Zou, Q., Bennion, B. J., Daggett, V., & Murphy, K. P. (2002). The molecular mechanism of stabilization of proteins by TMAO and its ability to counteract the effects of urea. *Journal of the American Chemical Society*, *124*, 1192–1202.
16. Roth, S. D., Schüttrumpf, J., Milanov, P., Abriss, D., Ungerer, C., Quade-Lyssy, P., Simpson, J. C., Pepperkok, R., Seifried, E., & Tonn, T. (2012). Chemical chaperones improve protein secretion and rescue mutant factor VIII in mice with hemophilia A. *PLoS ONE*, *7*, e44505.
17. Blajeski, A. L., Phan, V. A., Kottke, T. J., & Kaufmann, S. H. (2002). G1 and G2 cell-cycle arrest following microtubule depolymerization in human breast cancer cells. *Journal of Clinical Investigation*, *110*, 91–99.
18. Steegmaier, M., Hoffmann, M., Baum, A., Lénárt, P., Petronczki, M., Krššák, M., Gürtler, U., Garin-Chesa, P., Lieb, S., Quant, J., Grauert, M., Adolf, G. R., Kraut, N., Peters, J. M., & Rettig, W. J. (2007). BI 2536, a potent and selective inhibitor of polo-like kinase 1, inhibits tumor growth in vivo. *Current Biology*, *17*, 316–322.
19. Li, X., Yao, X., Zhu, Y., Zhang, H., Wang, H., Ma, Q., Yan, F., Yang, Y., Zhang, J., Shi, H., Ning, Z., Dai, J., Li, Z., Li, C., Su, F., Xue, Y., Meng, X., Dong, G., & Xiong, H. (2019). The caspase inhibitor Z-VAD-FMK alleviates endotoxic shock via inducing macrophages necroptosis and promoting MDSCs-mediated inhibition of macrophages activation. *Frontiers in Immunology*, *10*, 1824.
20. Mimura, Y., Lund, J., Church, S., Dong, S., Li, J., Goodall, M., & Jefferis, R. (2001). Butyrate increases production of human chimeric IgG in CHO-K1 cells whilst maintaining function and glycoform profile. *Journal of Immunological Methods*, *247*, 205–216.
21. Yang, W. C., Lu, J., Nguyen, N. B., Zhang, A., Healy, N. V., Kshirsagar, R., Ryll, T., & Huang, Y. M. (2014). Addition of valproic acid to CHO cell fed-batch cultures improves monoclonal antibody titers. *Molecular Biotechnology*, *56*, 421–428.
22. Riessland, M., Brichta, L., Hahnen, E., & Wirth, B. (2006). The benzamide M344, a novel histone deacetylase inhibitor, significantly increases SMN2 RNA/protein levels in spinal muscular atrophy cells. *Human Genetics*, *120*, 101–110.
23. Wong, V. V. T., Ho, K. W., & Yap, M. G. S. (2004). Evaluation of insulin-mimetic trace metals as insulin replacements in mammalian cell cultures. *Cytotechnology*, *45*, 107–115.
24. Mitchell, A. M., & Samulski, R. J. (2013). Mechanistic insights into the enhancement of adeno-associated virus transduction by proteasome inhibitors. *Journal of Virology*, *87*, 13035–13041.
25. Jennings, K., Miyamae, T., Traister, R., Marinov, A., Katakura, S., Sowders, D., Trapnell, B., Wilson, J. M., Gao, G., & Hirsch, R. (2005). Proteasome inhibition enhances AAV-mediated transgene expression in human synoviocytes in vitro and in vivo. *Molecular Therapy*, *11*, 600–607.

26. Wu, M., Chen, P., Liu, F., Lv, B., Ge, M., Jiang, P., Xu, W., Liu, X., & Yang, D. (2021). ONX0912, a selective oral proteasome inhibitor, triggering mitochondrial apoptosis and mitophagy in liver cancer. *Biochemical and Biophysical Research Communications*, *547*, 102–110.
27. Fernandez-Martell, A., Johari, Y. B., & James, D. C. (2018). Metabolic phenotyping of CHO cells varying in cellular biomass accumulation and maintenance during fed-batch culture. *Biotechnology and Bioengineering*, *115*, 645–660.
28. Lloyd, D. R., Holmes, P., Jackson, L. P., Emery, A. N., & Al-Rubeai, M. (2000). Relationship between cell size, cell cycle and specific recombinant protein productivity. *Cytotechnology*, *34*, 59–70.
29. . Pan, X., Dalm, C., Wijffels, R. H., & Martens, D. E. (2017). Metabolic characterization of a CHO cell size increase phase in fed-batch cultures. *Applied Microbiology and Biotechnology*, *101*, 8101–8113.
30. Nguyen, T. N. T., Sha, S., Hong, M. S., Maloney, A. J., Barone, P. W., Neufeld, C., Wolfrum, J., Springs, S. L., Sinskey, A. J., & Braatz, R. D. (2021). Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. *Molecular Therapy—Methods & Clinical Development*, *21*, 642–655.
31. Trempe, J. P., & Carter, B. J. (1988). Regulation of adeno-associated virus gene expression in 293 cells: Control of mRNA abundance and translation. *Journal of Virology*, *62*, 68–74.
32. Beswick, R. W., Ambrose, H. E., & Wagner, S. D. (2006). Nocodazole, a microtubule depolymerising agent, induces apoptosis of chronic lymphocytic leukaemia cells associated with changes in Bcl-2 phosphorylation and expression. *Leukemia Research*, *30*, 427–436.
33. Bernard, D., Mondesert, O., Gomes, A., Duthen, Y., & Lobjois, V. (2019). A checkpoint-oriented cell cycle simulation model. *Cell Cycle*, *18*, 795–808.
34. Uetake, Y., & Sluder, G. (2007). Cell-cycle progression without an intact microtubule cytoskeleton. *Current Biology*, *17*, 2081–2086.
35. Vasquez, R. J., Howell, B., Yvon, A. M. C., Wadsworth, P., & Cassimeris, L. (1997). Nanomolar concentrations of nocodazole alter microtubule dynamic instability in vivo and in vitro. *Molecular Biology of the Cell*, *8*, 973–985.
36. Rieder, C. L., & Maiato, H. (2004). Stuck in division or passing through. *Developmental Cell*, *7*, 637–651.
37. Quignon, F., Rozier, L., Lachages, A. M., Bieth, A., Simili, M., & Debatisse, M. (2007). Sustained mitotic block elicits DNA breaks: One-step alteration of ploidy and chromosome integrity in mammalian cells. *Oncogene*, *26*, 165–172.
38. Tait, A. S., Brown, C. J., Galbraith, D. J., Hines, M. J., Hoare, M., Birch, J. R., & James, D. C. (2004). Transient production of recombinant proteins by Chinese hamster ovary cells using polyethyleneimine/DNA complexes in combination with microtubule disrupting anti-mitotic agents. *Biotechnology and Bioengineering*, *88*, 707–721.
39. Hernandez-Verdun, D. (2011). Assembly and disassembly of the nucleolus during the cell cycle. *Nucleus*, *2*, 189–194.
40. Ma, N., Matsunaga, S., Takata, H., Ono-Maniwa, R., Uchiyama, S., & Fukui, K. (2007). Nucleolin functions in nucleolus formation and chromosome congression. *Journal of Cell Science*, *120*, 2091–2105.
41. Amin, M. A., Matsunaga, S., Ma, N., Takata, H., Yokoyama, M., Uchiyama, S., & Fukui, K. (2007). Fibrillarin, a nucleolar protein, is required for normal nuclear morphology and cellular growth in HeLa cells. *Biochemical and Biophysical Research Communications*, *360*, 320–326.
42. Stenström, L., Mahdessian, D., Gnann, C., Cesnik, A. J., Ouyang, W., Leonetti, M. D., Uhlén, M., Cuylen-Haering, S., Thul, P. J., & Lundberg, E. (2020). Mapping the nucleolar proteome reveals a spatiotemporal organization related to intrinsic protein disorder. *Molecular Systems Biology*, *16*, 1–16.
43. Wistuba, A., Kern, A., Weger, S., Grimm, D., & Kleinschmidt, J. A. (1997). Subcellular compartmentalization of adeno-associated virus type 2 assembly. *Journal of Virology*, *71*, 1341–1352.
44. Sonntag, F., Schmidt, K., & Kleinschmidt, J. A. (2010). A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proceedings of the National Academy of Sciences of the United States of America*, *107*, 10220–10225.
45. Qiu, J., & Brown, K. E. (1999). A 110-kDa nuclear shuttle protein, nucleolin, specifically binds to adeno-associated virus type 2 (AAV-2) capsid. *Virology*, *257*, 373–382.
46. . Bevington, J. M., Needham, P. G., Verrill, K. C., Collaco, R. F., Basrur, V., & Trempe, J. P. (2007). Adeno-associated virus interactions with B23/Nucleophosmin: Identification of sub-nucleolar virion regions. *Virology*, *357*, 102–113.
47. Greco, A. (2009). Involvement of the nucleolus in replication of human viruses. *Reviews in Medical Virology*, *19*, 201–214.
48. Saudan, P., Vlach, J., & Beard, P. (2000). Inhibition of S-phase progression by adeno-associated virus Rep78 protein is mediated by hypophosphorylated pRb. *The EMBO Journal*, *19*, 4351–4361.
49. Berthet, C., Raj, K., Saudan, P., & Beard, P. (2005). How adeno-associated virus Rep78 protein arrests cells completely in S phase. *Proceedings of the National Academy of Sciences of the United States of America*, *102*, 13634–13639.
50. Needs, S. H., Bootman, M. D., Grotzke, J. E., Kramer, H. B., & Allman, S. A. (2022). Offâtarget inhibition of NGLY1 by the polycaspase inhibitor Z-VAD-fmk induces cellular autophagy. *The FEBS Journal*, 1–17. https://doi.org/10.1111/febs.16345
51. Jardon, M. A., Sattha, B., Braasch, K., Leung, A. O., Côté, H. C. F., Butler, M., Gorski, S. M., & Piret, J. M. (2012). Inhibition of glutamine-dependent autophagy increases t-PA production in CHO Cell fed-batch processes. *Biotechnology and Bioengineering*, *109*, 1228–1238.
52. Nasseri, S. S., Ghaffari, N., Braasch, K., Jardon, M. A., Butler, M., Kennard, M., Gopaluni, B., & Piret, J. M. (2014). Increased CHO cell fed-batch monoclonal antibody production using the autophagy inhibitor 3-MA or gradually increasing osmolality. *Biochemical Engineering Journal*, *91*, 37–45.
53. Dickerson, R., Argento, C., Pieracci, J., & Bakhshayeshi, M. (2020). Separating empty and full recombinant adeno-associated virus particles using isocratic anion exchange chromatography. *Biotechnology Journal*, 2000015, 1–9.
54. Mietzsch, M., Eddington, C., Jose, A., Hsi, J., Chipman, P., Henley, T., Choudhry, M., McKenna, R., & Agbandje-McKenna, M. (2021). Improved genome packaging efficiency of adeno-associated virus vectors using rep hybrids. *Journal of Virology*, *95*, 1–19.
55. Johari, Y. B., Estes, S. D., Alves, C. S., Sinacore, M. S., & James, D. C. (2015). Integrated cell and process engineering for improved transient production of a 'difficult-to-express' fusion protein by CHO cells. *Biotechnology and Bioengineering*, *112*, 2527–2542.
56. Meyer, H. J., Turincio, R., Ng, S., Li, J., Wilson, B., Chan, P., Zak, M., Reilly, D., Beresini, M. H., & Wong, A. W. (2017). High throughput screening identifies novel, cell cycle-arresting small molecule enhancers of transient protein expression. *Biotechnology Progress*, *33*, 1579–1588.
57. Allen, M. J., Boyce, J. P., Trentalange, M. T., Treiber, D. L., Rasmussen, B., Tillotson, B., Davis, R., & Reddy, P. (2008). Identification of novel small molecule enhancers of protein production by cultured mammalian cells. *Biotechnology and Bioengineering*, *100*, 1193–1204.
58. Johari, Y. B., Jaffé, S. R. P., Scarrott, J. M., Johnson, A. O., Mozzanino, T., Pohle, T. H., Maisuria, S., Bhayat-Cammack, A., Lambiase, G., Brown, A. J., Tee, K. L., Jackson, P. J., Wong, T. S., Dickman, M. J., Sargur, R. B., & James, D. C. (2021). Production of trimeric SARS-CoV-2 spike protein by CHO cells for serological COVID-19 testing. *Biotechnology and Bioengineering*, *118*, 1013–1021.
59. Franzoso, F. D., Seyffert, M., Vogel, R., Yakimovich, A., de Andrade Pereira, B., Meier, A. F., Sutter, S. O., Tobler, K., Vogt, B., Greber, U. F., Büning, H., Ackermann, M., & Fraefel, C. (2017). Cell cycle-dependent expression of adeno-associated virus 2 (AAV2) rep in coinfections with herpes simplex virus 1 (HSV-1) gives rise to a mosaic of cells replicating either AAV2 or HSV-1. *Journal of Virology*, *91*, 1–19.
60. Raj, K., Ogston, P., & Beard, P. (2001). Virus-mediated killing of cells that lack p53 activity. *Nature*, *412*, 914–917.

61. Barnes, C. R., Lee, H., Ojala, D. S., Lewis, K. K., Limsirichai, P., & Schaffer, D. V. (2021). Genome-wide activation screens to increase adeno-associated virus production. *Molecular Therapy Nucleic Acids*, *26*, 94–103.
62. Strasser, L., Boi, S., Guapo, F., Donohue, N., Barron, N., Rainbow-fletcher, A., & Bones, J. (2021). Proteomic landscape of adeno-associated virus (AAV)-producing HEK293 cells. *International Journal of Molecular Sciences*, *22*, 11499.
63. Matthews, D., Emmott, E., & Hiscox, J. (2011). Viruses and the nucleolus. *Protein Reviews*, *15*, 321–345.
64. Boyault, C., Zhang, Y., Fritah, S., Caron, C., Gilquin, B., So, H. K., Garrido, C., Yao, T. P., Vourc'h, C., Matthias, P., & Khochbin, S. (2007). HDAC6 controls major cell response pathways to cytotoxic accumulation of protein aggregates. *Genes & Development*, *21*, 2172–2181.
65. Yan, Z., Zak, R., Luxton, G. W. G., Ritchie, T. C., Bantel-Schaal, U., & Engelhardt, J. F. (2002). Ubiquitination of both adeno-associated virus type 2 and 5 capsid proteins affects the transduction efficiency of recombinant vectors. *Journal of Virology*, *76*, 2043–2053.
66. Maurer, A. C., Pacouret, S., Cepeda Diaz, A. K., Blake, J., Andres-Mateos, E., & Vandenberghe, L. H. (2018). The assembly-activating protein promotes stability and interactions between AAV's viral proteins to nucleate capsid assembly. *Cell Reports*, *23*, 1817–1830.
67. Grieger, J. C., Soltys, S. M., & Samulski, R. J. (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. *Molecular Therapy*, *24*, 287–297.

SUPPORTING INFORMATION

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

How to cite this article: Scarrott, J. M., Johari, Y. B., Pohle, T. H., Liu, P., Mayer, A., & James, D. C. (2023). Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives. *Biotechnology Journal*, *18*, e2200450. https://doi.org/10.1002/biot.202200450

8.2.4 Molecular design of controllable recombinant AAV expression systems for enhanced vector production -Johari and Pohle et al.

Received: 3 December 2023 | Revised: 28 May 2024 | Accepted: 31 May 2024

DOI: 10.1002/biot.202300685

Biotechnology Journal

RESEARCH ARTICLE

Molecular design of controllable recombinant adeno-associated virus (AAV) expression systems for enhanced vector production

Yusuf B. Johari[1] | Thilo H. Pohle[1,3] | Jared Whitehead[1] | Joseph M. Scarrott[1] | Ping Liu[2] | Ayda Mayer[2] | David C. James[1,3]

[1]Department of Chemical and Biological Engineering, University of Sheffield, Sheffield, UK

[2]Cell Line Development, REGENXBIO Inc., Rockville, Maryland, USA

[3]Syngensys Ltd., Sheffield, UK

Correspondence
David C. James, Department of Chemical and Biological Engineering, University of Sheffield, Mappin St., Sheffield, S1 3JD, UK.
Email: d.c.james@sheffield.ac.uk
Present address
Yusuf B. Johari, Lonza Biologics, Cambridge, UK.

Funding information
REGENXBIO

Abstract

Recombinant adeno-associated virus (rAAV) is the leading vector for the delivery of gene therapies. However, low viral genome (VG) titers are common and the proportion of "full" capsids containing the therapeutic gene payload can be highly variable. The coordinated molecular design of plasmids encoding viral components and Helper functions remains a major challenge for rAAV manufacturing. Here we present the design of improved Rep/Cap and Helper plasmids for rAAV2/8 production, (i) a Rep/Cap expression vector harboring independently controllable *rep* and *cap* genes and (ii) an improved Helper plasmid harboring E4 gene deletion variants. First, an optimized Rep/Cap vector utilized a truncated p5 promoter, a p5 *cis*-regulatory element at the 3′ end in combination with a heterologous promoter to drive Cap expression and an additional copy of the *rep52/40* gene to overexpress short Rep proteins. We demonstrate that Rep78 is essential for efficient rAAV2/8 production in HEK293 cells, and a higher ratio of short Rep to long Rep proteins enhances genome packaging. Second, we identified regulators and open reading frames within the Helper plasmid that contribute to increased rAAV2/8 production. While L4-33k/22k is integral to optimal production, the use of E4orf6–6/7 subset significantly enhanced VG titer. Together, an optimal combination of engineered Rep/Cap and Helper plasmid variants increased VG titer by 3.1-fold. This study demonstrates that configuring and controlling the expression of the different AAV genetic elements contributes toward high rAAV production and product quality (full/empty capsid ratio).

KEYWORDS
adeno-associated virus, HEK293 cells, promoter, vector design, viral vector production

Abbreviations: Cap, capsid; DBP, DNA binding protein; ddPCR, digital droplet PCR; DoE, design-of-experiment; HEK, human embryonic kidney; ORF/orf, open reading frame; rAAV, recombinant adeno-associated virus; Rep, replication; TF, transcription factor; VCD, viable cell density; VG, viral genome.

Yusuf B. Johari and Thilo H. Pohle contributed equally to this work.

1 | INTRODUCTION

Despite the evidence of notable clinical efficacy, the high cost of manufacture and vector dosages required has limited the economic viability of recombinant adeno-associated virus (rAAV) mediated gene therapies.[1,2] Three plasmid transient transfection of HEK293 cells is currently the most widely utilized method for producing rAAV, where one of the main challenges in creating high-yielding AAV expression systems and generating AAV packaging cell lines is the E1A-mediated transactivation of promoters p5 and p19.[3] The former promoter controls the gene expression of AAV replication proteins Rep78/68 that are known to be cytostatic/cytotoxic[4,5] but are also required for transactivation of promoters p19 (transcribing Rep52/40) and p40 (transcribing capsid (Cap)).[6] On the other hand, overexpression of Cap has been shown to be advantageous for the optimal production of rAAV.[7,8] To this end, Rep78/68 (large Rep), Rep52/40 (small Rep), and Cap expression have to be regulated independently. Furthermore, Rep78 represses adenovirus Helper promoters E1A, E2A, and E4 and therefore needs to be controlled tightly during rAAV production and cell growth.[9] Engineering AAV expression vectors for improved manufacturability using mammalian cell factories remains a highly desirable objective.

In its natural context, AAVs achieve a precise expression stoichiometry of multiple genes within a compact genome (4.7 kb) using a combination of internal (within open-reading frame (ORF/orf)) promoters, overlapping ORFs, differential mRNA splicing, alternative translation start sites (with varying initiation rates) and feedback loops (using transactivators or repressors).[3] Various vector engineering strategies have been utilized to improve rAAV expression in cell hosts. Examples include the use of inducible promoters,[4] intron insertion,[10] Kozak/start codon mutations[11,12] and a four-plasmid system[13] to modulate the *rep* and *cap* gene expression, as well as a hybrid Rep to improve genome packaging efficiency.[14] Further efforts to boost rAAV yields targeted the Helper plasmid via utilization of human bocavirus 1 Helper genes[15] or design-of-experiment (DoE) approach to optimize the Helper, packaging, and transgene plasmid ratios.[16] Despite these improvements, there remain limited reports on the impact of individual (sub)components that need to be considered when designing an AAV vector and how they can be controlled and enhanced.

In this study, we identify the components and regulators of rAAV2/8 transient expression in HEK293 cells by mechanistically dissecting the packaging and Helper plasmids. We systematically determined the impact of p5 *cis*-regulatory element, endogenous and heterologous promoters, introns, removal of ORFs, and up/downregulation of specific genes on rAAV product titer. Using optimized split Rep/Cap and Helper plasmids, we further demonstrate that it is possible to control genome titer and product quality (full/empty capsid ratio) in a transient rAAV expression system.

2 | MATERIALS AND METHODS

2.1 | Plasmid construction

Proprietary Rep/Cap (pAAV2/8) and Helper (Helper 1.0; Figure S1) plasmids were provided by REGENXBIO. The Rep/Cap and Helper plasmid variants were constructed by PCR amplification (Q5 High-Fidelity 2× Master Mix; NEB), site-directed mutagenesis (Q5 Site-Directed Mutagenesis kit; NEB), and/or gene synthesis (Eurofins Genomics). PCR products were purified using QIAquick PCR Purification kit (Qiagen), and gel extraction was performed using QIAquick Gel Extraction kit (Qiagen). Restriction enzymes were obtained from NEB. Ligation was performed using T4 DNA ligase (NEB), and assembly of multiple DNA fragments was performed using NEBuilder HiFi DNA Assembly Master Mix (NEB). The sequence regions of the relevant promoters are detailed in Table S1. The sequence mutations are detailed in Table S2. Rep/Cap plasmids were amplified in DH5α competent cells (Thermo Fisher), and Helper plasmids were amplified in NEB Stable competent cells (NEB). Clonally derived plasmids were purified using QIAprep Spin Miniprep kit (Qiagen) or QIAGEN Plasmid Plus kit (Qiagen). The sequence of all plasmid constructs was confirmed by restriction enzyme analysis and DNA sequencing (Eurofins Genomics).

2.2 | HEK293 cultures

Suspension-adapted HEK293 cells were provided by REGENXBIO and cultured in Dynamis medium (Thermo Fisher) supplemented with L-glutamine (Thermo Fisher). Cells were maintained in Erlenmeyer flasks (Corning) at 37°C, 140 rpm under 5% CO_2, 85% humidity, and were subcultured every 3–4 days by seeding at 3×10^5 viable cells/mL. Cell viability and viable cell density (VCD) were measured using a Vi-CELL XR (Beckman Coulter).

2.3 | PEI-mediated transient vector production

rAAV2/8 production was performed by triple transfection and in shallow-well 24-well plates (Corning) using the Deutz system as previously described.[17,18] Briefly, cells were subcultured in an Erlenmeyer flask and grown to 4×10^6 cells/mL. Prior to transfection, aliquots of 700 µL were added to each well of 24-well plate. Plasmid DNA (weight ratio of 1:2:0.1 for packaging, Helper, transgene plasmids)[18] and PEIpro (Polyplus-transfection) were each prediluted in Dynamis medium, combined and incubated at room temperature for 10 min before being added into the culture. Transfected cells were cultured for 72 h at 37°C, 230 rpm under 5% CO_2, 85% humidity.

2.4 | Quantification of viral genome titer by ddPCR

Intra and extracellular rAAV2/8 titer was quantified by digital droplet PCR (ddPCR) as previously described.[17] Briefly, 10× cell lysis buffer containing 1× cOmplete EDTA-free Protease Inhibitor Cocktail (Roche) was added to cell culture and incubated at 37°C for 1 h. Samples were centrifuged to remove cell debris and the supernatant was treated with DNase I (Roche), followed by dilution in GeneAmp PCR Buffer I (Thermo Fisher) containing 0.02% UltraPure Salmon Sperm DNA Solution (Thermo Fisher) and 0.1% Pluronic F-68 Nonionic Surfactant (Thermo Fisher). Viral genome (VG) titer was quantified using QX200 Droplet Digital PCR system (Bio-Rad) and primers and a probe (Table S3) targeting the poly A sequence of the transgene plasmid harboring a CAG promoter and a GFP gene. The absolute VG titer was determined using the Quantasoft analysis software (Bio-Rad).

2.5 | Measurement of recombinant mRNA copy numbers

3×10^6 viable cells were collected at 72 h post-transfection by centrifugation at 300 × *g* for 5 min. Cell pellets were resuspended in 150 µL of RNAlater (Sigma-Aldrich). Total RNA was extracted using RNeasy Plus Mini kit in combination with QIAshredder homogenizer (Qiagen) according to the manufacturer's instructions. gDNA-free RNA was converted to cDNA and quantified using One-Step RT-ddPCR Advanced Kit (Bio-Rad) and QX200 Droplet Digital PCR system (Bio-Rad) according to the manufacturer's instructions. Primers and probes used are detailed in Table S3. mRNA copy number was determined using the Quantasoft analysis software (Bio-Rad).

2.6 | Measurement of intracellular proteins by Western blotting

Cells were harvested at 72 h post-transfection by centrifugation at 300 × *g* for 5 min and lysed using RIPA buffer supplemented with Halt Protease Inhibitor Cocktail (Thermo Fisher) according to the manufacturer's instructions. Protein concentration of cell lysates was determined by BCA assay (Pierce) and SDS-PAGE was performed using 10% Novex Tris–Glycine gels (Thermo Fisher) loaded with ≈40 µg and ≈10 µg of lysate for Rep and Cap Western blot, respectively. For Rep, proteins were transferred to nitrocellulose membranes using the miniblot module system (Thermo Fisher), blocked with 5% (w/v) milk-PBST for 1 h at room temperature, and then probed using the anti-AAV2 replicase antibody (1:200; 303.9, Progen) in 2% (w/v) milk-PBST at 4°C overnight, followed by anti-mouse IgG HRP antibody (1:2,000; 7076, CST) in 2% (w/v) milk-PBST for 1 h at room temperature. For Cap, proteins were transferred to PVDF membranes, blocked with 5% (w/v) BSA-PBST (Sigma-Aldrich) for 1 h at room temperature, and then probed using the anti-AAV VP1/VP2/VP3 antibody (1:200; B1, Progen) in 2% (w/v) BSA-PBST at 4°C overnight, followed by anti-mouse IgG HRP antibody (1:2,000) in 2% (w/v) BSA-PBST. The anti-vinculin HRP antibody (1:2,000; E1E9V, CST) was used similarly to the other primary antibodies. Membranes were exposed to ECL substrate (Pierce) for imaging by iBright CL1500 (Thermo Fisher).

2.7 | Quantification of intact capsid titer

Total capsid titer quantification was performed using the AAV8 titration ELISA (Progen) according to the manufacturer's instructions. Total cell lysis supernatant diluted in 1× ASSB assay buffer (Progen). OD values at 450 and 650 nm (background absorbance) were measured using a SpectraMax iD5 microplate reader (Molecular Devices).

2.8 | Statistics

Microsoft Excel 2016 (Microsoft) was used to analyze the difference between the means (normalized titers or intact capsids) of a plasmid construct and the control. As there were multiple batches of transfection, titers were normalized to the mean of control from the same batch to correct for possible differences in cell number and growth. Analysis was performed using an unpaired Student's *t*-test with p-value < 0.05 was considered significant.

3 | RESULTS

3.1 | A heterologous promoter and inclusion of Cap intron enable a controllable Rep/Cap plasmid system

Previous studies showed that unregulated overexpression of Rep78/68 inhibited rAAV production, while reduced levels of Rep78/68 enhanced rAAV titers.[12,19–21] Further, promoter p5 acts a *cis*-regulatory element where its deletion was shown to cause downregulation of promoters p19 and p40.[6,22,23] Our standard Rep/Cap plasmid comprised two modifications; (i) truncation of the p5 promoter to attenuate expression of the large Rep proteins, and (ii) introduction of a p5 promoter downstream of the AAV *cap* region to retain the expression of p19 and p40 (Figure 1A).[24] As shown in Figure 1B, the improved Rep/Cap plasmid (Rep/Cap 1.2; Control) displayed a 32% increase in rAAV2/8 titer over the conventional packaging plasmid ($p < 0.01$), reaching 10^{14} VG/L in serum-free media. We note that triple transfection was performed at a weight ratio of 1:2:0.1 (packaging, Helper, transgene plasmids).[18] While it is relatively easy to regulate *rep78/68*, control of *rep52/40* and *cap* expression would involve the complex multimeric gene assembly in which the p19 and p40 promoters are located within the *rep* coding sequences.

In order to enable control of *cap* gene expression, we split the *rep* and *cap* genes by cloning the p40 promoter and *cap* open reading frame (ORF) downstream of the *rep* gene stop codons (Figure 1C). To prevent expression of truncated viral gene products, the TATA box and Initiator (Inr) of the internal p40 promoter as well as the start codon

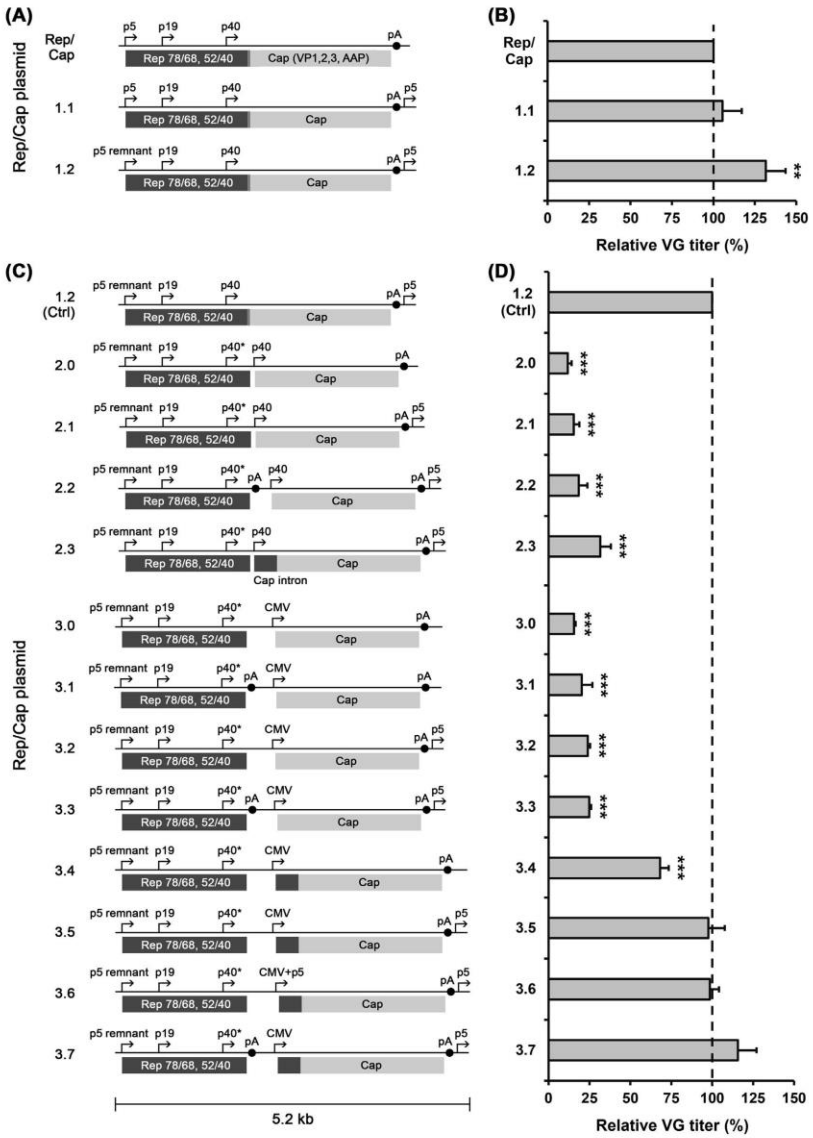


FIGURE 1 Functional evaluation of the p5 *cis*-regulatory element and split Rep/Cap plasmid constructs for rAAV2/8 production. (A, C) Schematic depiction of Rep/Cap plasmid constructs. Replication (Rep) and capsid (Cap) open reading frames are indicated. Arrow denotes a promoter, asterisk denotes sequence mutation, black circle denotes a poly A. All components are drawn to approximate scale. (B, D) HEK293 cells were triple transfected with each Rep/Cap, Helper and transgene plasmid at 1:2:0.1 weight ratio. rAAV2/8 crude viral genome (VG) titers were analyzed 72 h post-transfection, expressed as a percentage compared to the conventional Rep/Cap plasmid (B) or Rep/Cap 1.2 (D). Data shown are the mean ± SD of three independent transfections. Data were analyzed using unpaired Student's *t*-test with respect to the conventional Rep/Cap plasmid or Rep/Cap 1.2. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

of the internal *cap* remnant were inactivated by mutations while preserving the functionality of the Rep proteins (i.e., without altering the encoded amino acid; Table S2). Measurement of rAAV2/8 titer at 72 h post-transfection is shown in Figure 1D. Separation of the *rep* and *cap* genes, either without or with a downstream p5 promoter (Rep/Cap 2.0 and 2.1, respectively) dramatically reduced the rAAV titer to <15% compared to the control. The addition of a poly A for the *rep* ORFs (Rep/Cap 2.2) for independent termination of transcription did not result in a noticeable increase in rAAV titer. While the inclusion of *cap* intron (Rep/Cap 2.3) for efficient post-transcriptional processes increased the titer by two-fold compared to Rep/Cap 2.1 (indicative of a critical element), it was only one-third of the control titer. Accordingly, we surmised that the transcription factor (TF) binding sites within the upstream (inactivated) p40 promoter acted as competing binding elements[25] resulting in reduced *cap* transcriptional activity.

In order to evaluate whether the upstream p40 promoter corresponded to "TF decoy sites," we replaced the downstream p40 with a heterologous promoter, the human CMV.[26] As shown in Figure 1D, employing the CMV promoter to drive *cap* expression without its intron (Rep/Cap 3.0–3.3) resulted in slight increases in rAAV titer compared to the Rep/Cap 2.0–2.2 constructs, with Rep/Cap 3.3 exhibiting 25% of the control titer. The inclusion of the *cap* intron (Rep/Cap 3.4) significantly increased the rAAV production to 70% of the control titer, and the addition of a p5 promoter downstream (Rep/Cap 3.5) restored the titer to the control level—consistent with the enhancer function associated with p5.[22,23] While the addition of a second copy of p5 *cis*-regulatory element directly upstream of the CMV promoter (Rep/Cap 3.6) did not increase the titer relative to the Rep/Cap control, introducing a poly A downstream of the *rep* gene (Rep/Cap 3.7) further enhanced the rAAV yield to 116% (albeit statistically insignificant, $p = 0.075$). ddPCR and Western blot analyses on Rep/Cap 3.7 at 72 h post-transfection demonstrated that *cap* expression was upregulated, with similar VP1–3 stoichiometry compared to the Rep/Cap 1.2 control (Figure 2A,B). The analyses also showed that the codon mutations in the split *cap* system indirectly attenuated long Rep expression at the post-transcriptional level (Figure 2C,D). Combining all observations made above, we inferred that (i) a heterologous promoter is required to drive efficient transcription of the split *cap* gene, (ii) the *cap* intron is a key regulator of Cap expression, and (iii) a poly A can be used to enable independent control of the *rep* genes for improved rAAV production. The split Rep/Cap system will permit attunement of the expression level of both the Rep proteins and the capsids to increase production of rAAV for use in gene therapy, for example, through modification of promoters and Kozak sequences, rearrangement of the genes, as well as codon (de)optimization.

3.2 | Rep78 protein is essential while higher ratio of Rep52/40 to Rep78/68 proteins enhances rAAV2/8 production

Previous studies have suggested that the functions of Rep78 and Rep68 are the same, as reported for Rep52 and Rep40.[27,28] Further, Rep52/40 proteins (in contrast to Rep78/68) were found not to inhibit the growth of primary, transformed, and immortalized cells.[29] Therefore, we hypothesized that rAAV production can be increased via (i) complete ablation of highly cytotoxic replication protein Rep78, and (ii) overexpression of Rep52/40 to enhance the packaging and accumulation of single-stranded viral genome[30] without inducing cytotoxicity. With regard to the latter, the constraint in regulating p19 is due to the position of this promoter which is located within the protein-coding sequence of Rep78/68. To illustrate this, we mutated the weak Kozak sequence (TACATGG, start codon underlined) of Rep/Cap 3.5 to promote the short Rep expression (see Figure S2). Measurement of rAAV2/8 titer after transient transfection of HEK293 cells showed that mutating the TAC (tyrosine) to ATC (isoleucine) within the Kozak diminished the rAAV production by ≈1000-fold (Rep/Cap 4.0; Figure 3). As all Rep/Cap mRNAs and proteins were expressed as expected (Figure 2C,D), the data implies a loss in Rep78/68 functionality.

To test the hypothesis that rAAV2/8 can be produced in HEK293 cells using only one large Rep protein and one small Rep protein, we modified the Rep/Cap 3.5 plasmid (Figure 1C) with deleted alternate splice sites within the *rep* codons to produce only Rep68 and Rep40 proteins (Rep/Cap 4.1; Figure 3A) according to Emmerling et al.[13] Subsequent Western blotting confirmed that only Rep68 and Rep40 proteins were present (Figure 2D). As shown in Figure 3B, the removal of *rep78* and *rep52* markedly reduced rAAV titer by 73% compared to the Rep/Cap 1.2 control. We observed no significant differences in cell viability or VCD between Rep/Cap 3.5 (or 3.7) and Rep/Cap 4.1–4.4 constructs (data not shown) although this was not entirely unexpected due to the very low level of Rep78 protein using the split *cap* system (see Figure 1). To determine whether the absence of rep52 was responsible for the titer reduction, we introduced a rep52/40 gene downstream of cap (Rep/Cap 4.2). To further overexpress the short Rep proteins, we added a second copy of p5 (*cis*-regulatory element for the p19 promoter) upstream of the CMV (Rep/Cap 4.3), or substituted the p19 with p5 promoter and mutated the weak Kozak sequence (Rep/Cap 4.4). We note that the p5 promoter is approximately twice as active as the p19 promoter (Figure S2), while ddPCR and Western blotting confirmed that Rep52 was reintroduced (Figure 2C,D). The results showed that neither reintroduction of *rep52* nor attempts to augment Rep52/40 expression resulted in noticeable improvements in rAAV titer compared to Rep/Cap 4.1. To understand the impact of Rep78 and Rep52 removal on rAAV production, we measured the AAV-associated mRNA and protein levels in Rep/Cap 4.1 and 4.4 (Figure 2). These analyses revealed that while the Cap mRNA levels were comparable to Rep/Cap 3.7, VP expression appeared to be dependent on Rep78 and Rep52 where the former has been reported to augment the splicing of Cap premRNA.[31] Additional studies are needed to determine the mechanism responsible for the reduction of VP expression.

Based on the above observations, we created a library of Rep/Cap plasmids with all four *rep* coding regions including an additional copy of *rep52/40* downstream of *cap* to enhance the packaging and accumulation of single-stranded viral genome. The second copy of *rep52/40* was driven by either promoter p19 or p5 and with or without a mutated Kozak sequence (Rep/Cap 5.0–5.4; Figure 3A). ddPCR and

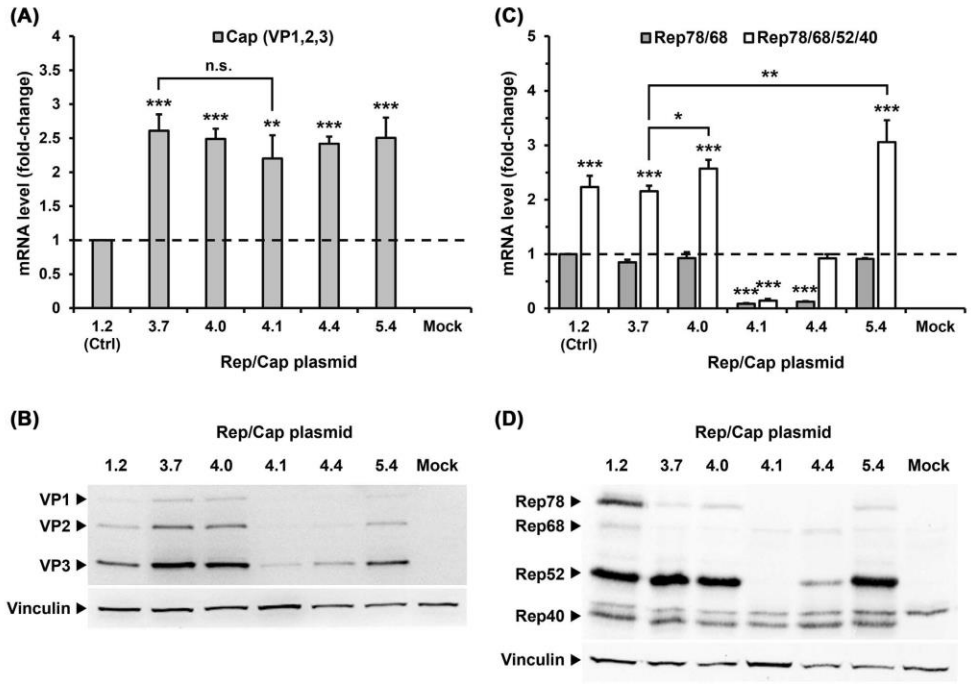


FIGURE 2 Comparative analysis of Rep/Cap mRNAs and proteins during rAAV2/8 production using the split Rep/Cap plasmid. Cells triple-transfected with a subset of Rep/Cap plasmids in Figures 1 and 3, or mock-transfected with no plasmid were harvested at 72 h post-transfection. (A) Cap (VP1, VP2, and VP3) and (C) Rep mRNA transcript levels were analyzed by ddPCR, expressed as a fold-change compared to the Cap or Rep78/68 mRNA of Rep/Cap 1.2 control. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$, n.s. not significant. (B) Representative immunoblots of VP and (D) Rep proteins in rAAV-producing cells transfected with the different Rep/Cap plasmids. Vinculin was used as an internal standard.

Western blot analyses indicate that Rep/Cap 5.4 had increased short Rep expression levels compared to Rep/Cap 3.7 (Figure 2C,D). As shown in Figure 3B, analysis of rAAV titer confirmed that p5 promoter at the 3′ end was critical for maximal titers (see Rep/Cap 5.1 and 5.2 vs. 5.0). Importantly, the data demonstrated that the reintroduction of *rep78* reinstated the rAAV production system with two constructs (Rep/Cap 5.3 and 5.4) exceeded the Rep/Cap 1.2 control titer (≈118%; $p < 0.05$). No further increase in rAAV titer was observed when a poly A was added between the long *rep* and *cap* genes (Rep/Cap 5.3 vs. 5.4). In summary, even though either Rep78 or Rep68 alone may be sufficient for AAV DNA replication,[27] our study shows that rAAV production in HEK293 cells is critically regulated by the full-length Rep78—corroborating previous studies suggesting that Rep78 and Rep52 proteins are necessary for efficient viral production.[32,33] We deduce that (low level) Rep78 is required for optimal rAAV DNA replication, and a higher ratio of short Rep (Rep52/40) to long Rep (Rep78/68) proteins may enhance genome packaging without inducing cytotoxicity in transient rAAV expression systems.

3.3 | L4-33k/22k proteins are required for optimal rAAV2/8 production

The Helper plasmid (Helper 1.0; Figure S1) utilized in this study was composed of the E2A (encoding DNA binding protein (DBP)), E4, and VA RNA regions derived from the adenovirus-5 genome. The E2A is transcribed by two promoters, namely E2-early and E2-late (Table S1), with DBP mRNA levels coming from the E2-early promoter being dominant. Additionally, the E2-early promoter/intron sequence encodes the L4-33k/22k proteins on the opposite strand, driven by L4 promoter (Figure 4A).[34,35] Even though the L4-22k/33k proteins have been indicated to play a role in adenovirus assembly, gene expression, and viral DNA packaging,[36,37] their significance in rAAV production is largely undetermined. This element represents a potential engineering target or possibly a redundant motif that could be eliminated from the vector. With regard to the latter, minimizing plasmid size is desirable for enhanced transient production by increased transfection efficiency and copy numbers of required genes per DNA weight.

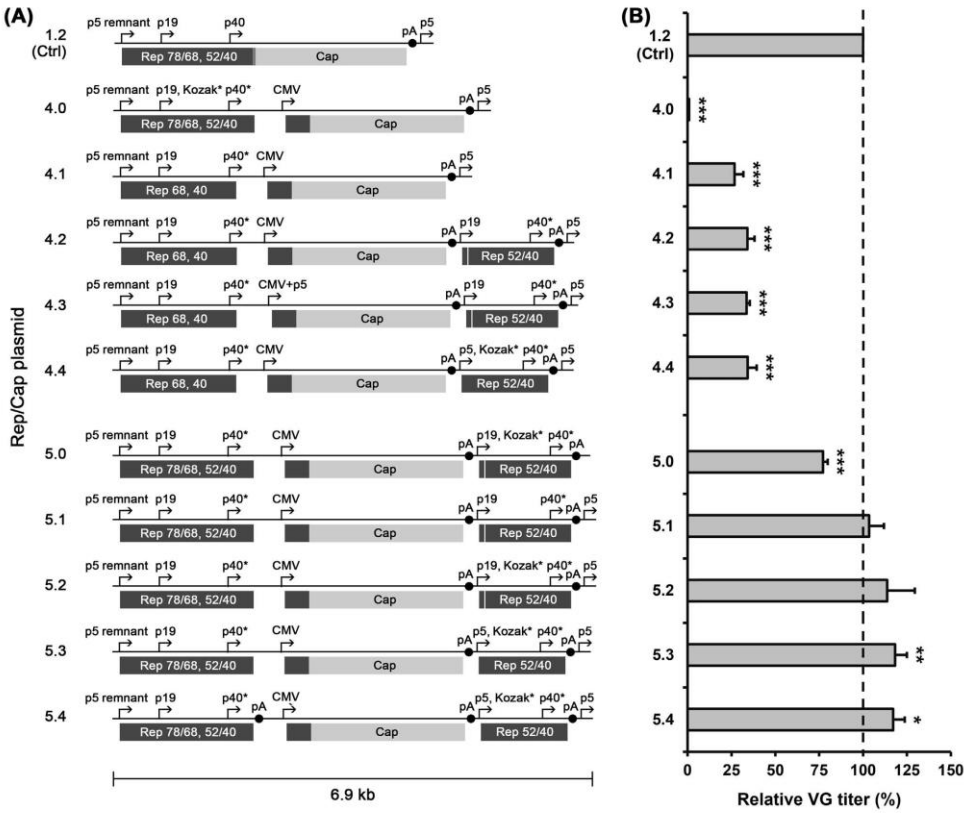


FIGURE 3 Functional evaluation of different Rep proteins within the split Rep/Cap plasmid construct for rAAV2/8 production. (A) Schematic depiction of Rep/Cap plasmid constructs. Replication (Rep) and capsid (Cap) open reading frames are indicated. Rep78 and Rep52 were removed by deleting the alternate splice site within the *rep* codons to produce only Rep68 and Rep40 proteins. The Rep52/40 Kozak sequence was optimized by mutating TACATGG → ATCATGG (start codon underlined). Arrow denotes a promoter, asterisk denotes sequence mutation, black circle denotes a poly A. All components are drawn to approximate scale. (B) HEK293 cells were triple transfected with each Rep/Cap, Helper and transgene plasmid at 1:2:0.1 weight ratio. rAAV2/8 crude viral genome (VG) titers were analyzed 72 h post-transfection, expressed as a percentage compared to the Rep/Cap 1.2. Data shown are the mean ± SD of three independent transfections. Data were analyzed using unpaired Student's *t*-test with respect to the Rep/Cap 1.2. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

In order to evaluate the function of the E2A region, we constructed Helper plasmids containing a partially or fully deleted E2A intron (including a 77 bp exon contained within),[34,35] with the former retaining the E2-late promoter and L4-33k/22k coding sequence ($E2A_{min1}$ and $E2A_{min2}$ Helper; Figure 4A). Triple transfection was performed as described above and rAAV2/8 titer (Figure 4B) and E2A mRNA level (Figure 5A) were measured 72 h post-transfection. These data show that the $E2A_{min1}$ and $E2A_{min2}$ reduced the titer to 82% and 37% ($p < 0.05$) of that deriving from E2A control, respectively, despite the latter exhibiting similar E2A mRNA level. Indeed, the L4-33k/22k ORFs located in the E2A promoter/intron fragment did not allow the conclusion that only the DBP contributed to the rAAV Helper function. To elucidate this, we constructed an L4 promoter-driven plasmid expressing only the L4-33k/22k proteins and co-transfected it with the $E2A_{min2}$ Helper plasmid at an equal molar ratio. This analysis demonstrated that the L4-33k/22k single gene co-expression ($E2A_{min2}$ + L4-33k/22k) resulted in a significant increase in rAAV titer compared to $E2A_{min2}$ (53% increase, $p < 0.05$; Figure 4B) with no significant difference in the E2A mRNA level (Figure 5A). As L4-33k mutant virus has been shown to produce only empty adenoviral capsids,[37] we hypothesized that the L4-33k-deficient rAAV production suffered from a defect in viral DNA packaging resulting in a lower titer. We inferred that the

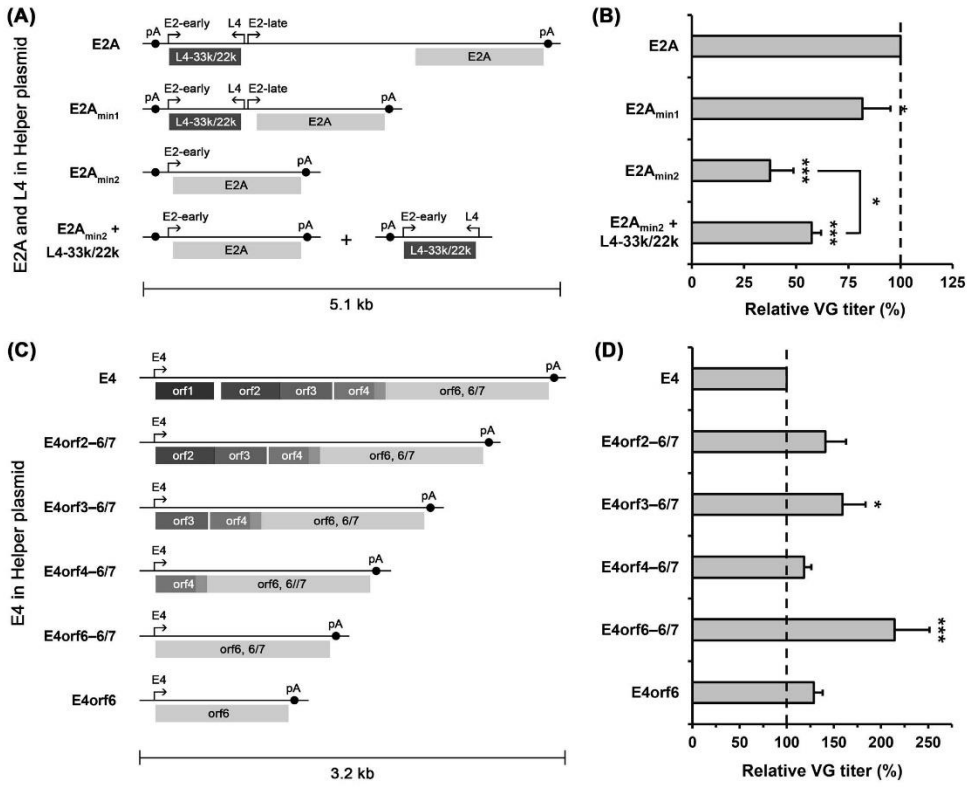


FIGURE 4 Functional evaluation of the E2A, L4-33k/22k and E4 Helper components for rAAV2/8 production. (A) Schematic depiction of the E2A and L4-33k/22k open reading frames within Helper plasmid constructs (E4 and VA RNA are not indicated). A L4 promoter-driven plasmid expressing only the L4-33k/22k protein was also constructed (L4-33k/22k plasmid). Arrow denotes a promoter, black circle denotes a poly A. All components are drawn to approximate scale. (B) HEK293 cells were triple transfected with each Helper, Rep/Cap 1.2 and transgene plasmid at 2:1:0.1 weight ratio. The L4-33k/22k plasmid was spiked at equal molar ratio to the E2A$_{min2}$ Helper plasmid. rAAV2/8 crude viral genome (VG) titers were analyzed 72 h post-transfection, expressed as a percentage compared to the Helper plasmid consisting the complete E2A and L4-33k/22k components. (C) Schematic depiction of the E4 orfs (open reading frames) within the Helper plasmid constructs (E2A, L4-33k/22k, and VA RNA are not indicated). (D) HEK293 cells were triple transfected with each Helper, Rep/Cap 1.2 and transgene plasmid at 2:1:0.1 weight ratio. rAAV2/8 crude VG titers were analyzed 72 h post-transfection, expressed as a percentage compared to the full-length E4 gene. Data shown are the mean ± SD of three independent transfections. Data were analyzed using unpaired Student's *t*-test with respect to the complete E2A or E4 Helper plasmid. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

L4-33k/22k proteins were integral components for optimal, high-yield rAAV2/8 production and that it represents a potential engineering target (e.g., via its overexpression) to enhance rAAV production.

3.4 | Helper plasmid comprising E4 orf6 and 6/7 subset enhances rAAV2/8 production

The E4 gene of adenovirus encodes seven proteins, namely E4 orf (open reading frame) 1, 2, 3, 3/4, 4, 6, and 6/7, each with different functions including promoting viral gene expression and replication as well as modulation of TF activities.[38] Among these, only the E4orf6 protein was thought to contribute to rAAV production and solely employed in a number of Helper plasmid variants.[39–41] Despite the minimal observed effect of other E4orfs on adenovirus growth in cultured cells,[41,42] we hypothesized that rAAV production could be optimized by specific combinations of the E4orf proteins.

In order to specifically determine the functional contribution of different E4orfs, we dissected the E4 gene by constructing Helper plasmids containing different subsets of the orfs (Figure 4C). This set of

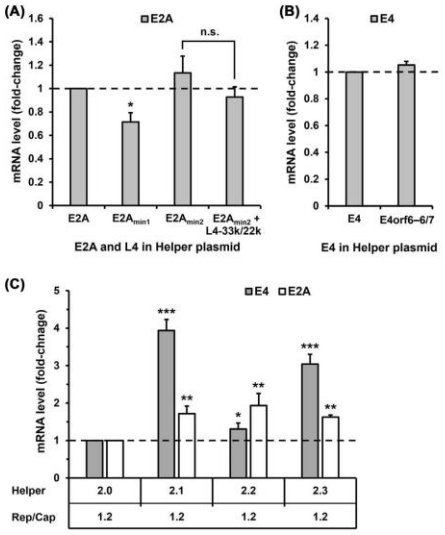


FIGURE 5 Comparative analysis of E2A and E4 mRNAs during rAAV2/8 production using the engineered Helper plasmids. (A, B) Cells triple-transfected with truncated E2A and E4orf6–6/7 Helper plasmids in Figure 4 (all utilizing Rep/Cap 1.2), and (C) CMV-driven Helper plasmids in Figure 6 were harvested at 72 h post-transfection. mRNA transcript levels were analyzed by ddPCR, expressed as a fold-change compared to the complete E2A or E4 Helper, or Helper 2.0 plasmid. Data were analyzed using unpaired Student's *t*-test with respect to the complete E2A or E4 Helper, or Helper 2.0 plasmid. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$, n.s., not significant.

six plasmids was then tested for their ability to mediate AAV2/8 vector production. Measurement of rAAV titer after triple transfection of HEK293 cells with each Helper variation is shown in Figure 4D. As anticipated, the E4orf6 was capable of producing rAAV help equivalent to a full-length E4 gene ($p > 0.05$). Moreover, the data show that the removal of orf1 and 2 (i.e., E4orf2–6/7 and E4orf3–6/7 subsets) increased the rAAV titer by 41% and 59% ($p < 0.05$), respectively, compared to the E4 control. In this regard, we conjecture that the deletion of these two redundant orfs (where their functions are largely undefined)[37] increased the abundance of other orf mRNAs spliced from the same precursor mRNA transcript. Deletion of orf1–3 (E4orf4–6/7 subset) decreased the rAAV titer to the control level—this was not unexpected considering that orf3 (similar to orf6) functions in promoting viral gene expression and replication.[41] Importantly, our data shows that further deletion of orf4 while retaining orf6/7 (E4orf6–6/7 subset) significantly enhanced the titer to 214% of the E4 control titer ($p < 0.001$). This result accords with previous studies that identified orf4 as a negative regulator of E1A and E4 transcription[43] while orf6/7 modulates the activity of the cellular transcription factor E2F.[44] Measurement of E4 mRNA at 72 h post-transfection showed that the E4orf6–6/7 construct did not result in a higher overall E4 transcript level (Figure 5B). We conclude that rAAV production in HEK293 cells can be enhanced via removal of redundant E4orfs.

3.5 | Engineered Rep/Cap and Helper plasmids can be used together to control rAAV2/8 gene expression

To evaluate whether the controllable Rep/Cap system could complement the engineered Helper to enable efficient rAAV production, we utilized E4orf6–6/7 Helper plasmid in combination with either Rep/Cap 3.7 or 5.4 plasmid. Additionally, previous studies showed that regulatory loops exist in which E1A, DBP and E4orf6/7 proteins positively or negatively regulate promoters p5, E1A, E2-early, and E4 as well as transcription factor E2F, among others.[44,45] Therefore, we evaluated whether substitution of constitutively active CMV promoter sequences for the E2A and/or E4 regulatory sequences in the E4orf6–6/7 Helper plasmid (Figure 6A) have positive/negative effects on rAAV vector production.

Measurement of rAAV2/8 titer after triple transfection of HEK293 cells with different Helper and Rep/Cap variant combinations is shown in Figure 6B. The result showed that both Rep/Cap 3.7 and 5.4 were incompatible with Helper E4orf6–6/7 (Helper 2.0) where they displayed 42% and 26% reduction in rAAV titer, respectively, relative to the Rep/Cap 1.2 control ($p < 0.01$). In this regard, we postulate that the use of the strong, highly complex CMV promoter in Rep/Cap 3.7 and 5.4 plasmids titrated away the limited pool of available TF molecules from the E4 promoter resulting in E4orf6–6/7 downregulation.[25,26] The use of CMV promoter to drive E4orf6–6/7 transcription (Helper 2.1) restored the rAAV titer comparable to Helper 2.0 with Rep/Cap 1.2 (83%–110%). In contrast, substituting the E2-early promoter with the CMV promoter (Helper 2.2) led to reduced titers especially when used in conjunction with Rep/Cap 1.2 and 5.4. Further decreases in rAAV level were observed with Helper 2.3 plasmid that harbored the CMV promoter to drive both E4orf6–6/7 and E2A expression. Measurement of mRNA levels (Figure 5C) showed that the use of Helper 2.1 resulted in a 3.94-fold increase in E4orf6–6/7 transcript level as well as a 1.72-fold increase in E2A transcript level. This is expected considering that the E4 promoter is inhibited by DBP whereas E4orf6/7 stimulates the activity of E2-early promoter.[44] E4 transcript level was also slightly increased when CMV was used to drive E2A expression, which can be attributed to the CMV acting as a "downstream enhancer" to the E4 promoter.[46] Slightly lower E4orf6–6/7 and E2A transcript levels were observed when CMV was utilized to simultaneously drive E4orf6–6/7 and E2A expression (Helper 2.3) compared to when it was used separately (Helper 2.1 and 2.2)—in general agreement with our view of competing TF binding sites. Nevertheless, our data did not show any correlation between rAAV titer (Figure 6B) and the E4 or E2A transcript level (Figure 5C), thus illustrating the highly complex interactions between various components during the viral production process.

The data in Figure 6B also shows that Rep/Cap 5.4 with Helper 2.0 or 2.1 yielded higher titers compared to Rep/Cap 3.7. To expound this observation, we selected a subpanel of different Rep/Cap and

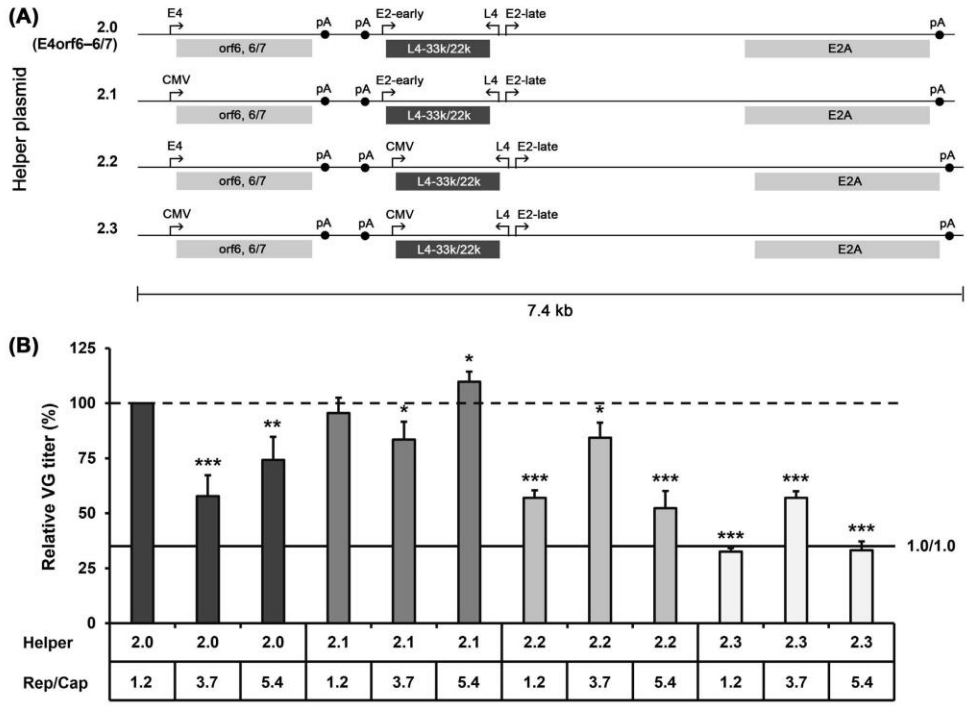


FIGURE 6 Evaluation of engineered Helper and Rep/Cap plasmid combinations for rAAV2/8 production. (A) Schematic depiction of the Helper plasmid constructs (VA RNA is not indicated). The E4orf6–6/7 Helper plasmid (Figure 4A) is denoted as Helper 2.0. Arrow denotes a promoter, black circle denotes a poly A. All components are drawn to approximate scale. (B) HEK293 cells were triple transfected with the Helper, Rep/Cap and transgene plasmids at 2:1:0.1 weight ratio. rAAV2/8 crude viral genome (VG) titers were analyzed 72 h post-transfection, expressed as a percentage compared to the Helper 2.0 and Rep/Cap 1.2 plasmid combination. Solid horizontal line represents the titer level of the conventional Helper 1.0 and Rep/Cap 1.0 (1.0/1.0) in Figure 1B. Data shown are the mean ± SD of three independent transfections. Data were analyzed using unpaired Student's *t*-test with respect to the Helper 2.0 and Rep/Cap 1.2 plasmid combination. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$.

Helper combinations from Figure 6B as well as the Rep/Cap 1.2 control from Figure 1D (utilizing the original Helper plasmid; Helper 1.0), and quantified fully assembled, intact capsids to determine the ratio of full to empty particles. This analysis indicated that the use of the CMV promoter to drive Cap expression (Rep/Cap 3.7 and 5.4) with either Helper 2.0 or 2.1 boosted total capsids by an average of ≈3-fold compared to Rep/Cap 1.2 (Figure 7A), resulting in full/empty capsid ratio of <9% (Figure 7B). Importantly, the analysis also revealed that all transfections utilizing Rep/Cap 5.4 (overexpressing Rep52/40 proteins) yielded relatively higher full-to-empty capsid ratios compared to Rep/Cap 3.7, indicating a higher rate of packaging and accumulation of single-stranded DNA progeny genomes (see also the accompanied increase in VG titer for Helper 2.0 and 2.1; Figure 7C). Very high full/empty ratios (up to 66%) were achieved using Helper 2.2 and 2.3 although this was largely due to considerable reductions in intact capsid abundance compared to other Helper variants. Taken together, these data demonstrate that it is possible to control both genome and total viral particle titer in a transient rAAV expression system. We anticipate that the novel library of CMV promoter sequences with variable strengths,[26] combined with reoptimization of the triple plasmid ratio for the new vector design,[47] would enable systematic determination of the optimal Cap expression for maximal rAAV product titer and quality.

4 | DISCUSSION

In this study we have characterized the diverse components (ORFs) and regulators (promoters, introns) of the AAV transient triple transfection plasmid system underpinning the biomanufacturing processes, for example, of how the abundance (or absence) of the four Rep proteins affects the efficiency of rAAV production yield. Specifically, our

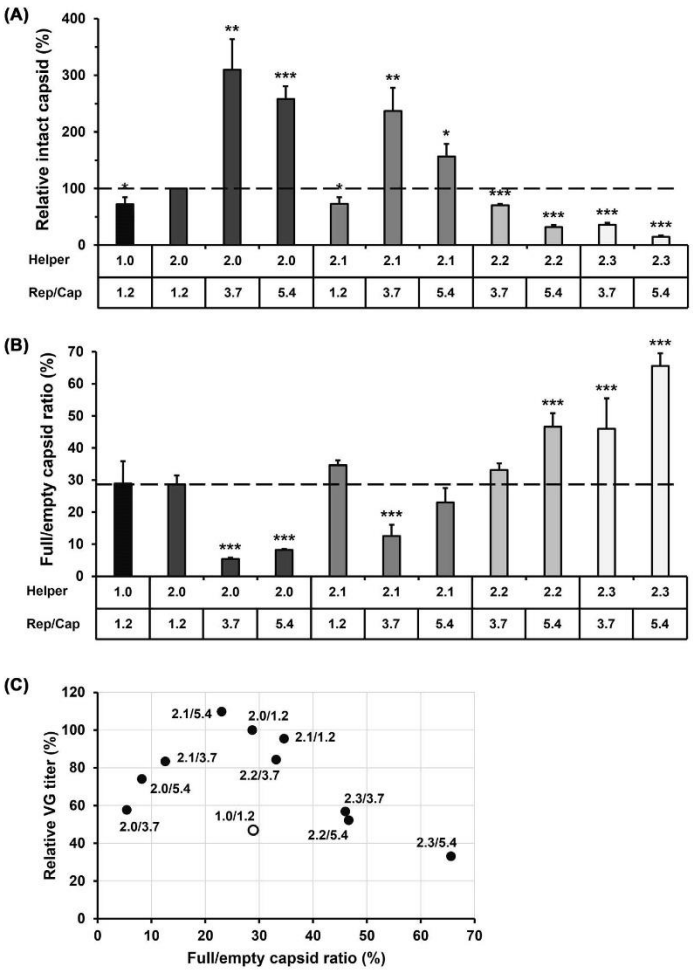


FIGURE 7 Determination of product quality of the engineered Helper and Rep/Cap plasmids for rAAV2/8 production. The original Helper plasmid (Figure S1) is denoted as Helper 1.0. (A) Intact capsids were quantified at 72 h post-transfection using rAAV8-specific capsid ELISA and expressed as a percentage compared to the Helper 2.0 and Rep/Cap 1.2 plasmid combination. (B) The full/empty capsid ratio was calculated from the measured intact capsids in (A) and its viral genome (VG) titer. Data shown are the mean ± SD of three independent transfections. Data were analyzed using unpaired Student's *t*-test with respect to the Helper 2.0 and Rep/Cap 1.2 plasmid combination. $^{*}p < 0.05$, $^{**}p < 0.01$, $^{***}p < 0.001$. (C) Scatter plot of VG titer (Figure 6B) and full/empty capsid ratio (B) of the engineered Rep/Cap and Helper plasmids. Number refers to Helper/Rep/Cap constructs. Open circle denotes the Helper 1.0 and Rep/Cap 1.2 control in Figures 1B and 3B.

data indicated a suboptimal rAAV2/8 production state when *rep78* was removed. This finding is in line with the previous reports in which Rep78 was shown to be more efficient than Rep68 in producing infectious Rep-negative AAV,[27] and hence indirectly favors the vector design strategy that omitted *rep68*[48] rather than *rep78*[13] to mitigate Rep toxicity effects. Furthermore, this study identified the L4-33k/22k gene as an integral rAAV component, corroborating a recent study that showed a >20-fold decrease in *rep* and *cap* DNA in HeLa cells transfected with the 33k/22k-targeting siRNA[49]—suggesting a cell engineering opportunity for increased rAAV production, for example,

via overexpression of these adenovirus Helper proteins. More generally, while previous studies demonstrated the need to lower long Rep expression,[12,19–21] further overexpression of short Rep (e.g., by adding a second gene copy) could promote enhancement of rAAV titer and product quality. Overall, the Helper 2.1 and Rep/Cap 5.4 plasmid system presented in this study enabled a ≈3.1-fold increase in titer compared to the conventional Rep/Cap and Helper plasmids system (Figure 6B; we note that similar results are achieved in TubeSpin or Erlenmeyer flask[17]).

Another key finding of our work is that E4 orf1, 2, and 4 are functionally redundant within the AAV expression system, consistent with the existing notion that a single E4orf6 protein is needed to produce rAAV vectors in HEK293 cells.[39–41] Critically, our results reveal that rAAV titer could be enhanced by specific combinations of the E4orf proteins particularly orf6 and 6/7. The removal of redundant orfs also likely improved the expression of other orfs due to reduced splice sites. With regard to the latter protein, E4orf6/7 modulates the activity of the E2-early promoter by forming a direct complex with transcription factor E2F and stabilizing the DNA-bound form.[44] As the E4 promoter is inhibited by the E2A product,[45] a regulatory loop exists in which E4orf6/7 protein increases E2A transcription while DBP negatively regulates E4 transcription. From a mechanistic perspective, we assume that the advantage of this temporal coordination is restricted DBP "toxic" effects[50,51] during the bioproduction process thus ensuring maximal productivity. Replacing the E2-early promoter with the constitutive, highly active promoter CMV could result in uncontrollable E2A gene expression and therefore rapid cellular accumulation of DBP. However, our cell concentration/viability data showed no differences between the endogenous E2-early and CMV promoter-driven E2A constructs (data not shown). We speculate that the detrimental effects of the CMV-driven E2A in HEK293 cells were via negative regulation of specific AAV components rather than direct exertion of cytotoxicity on the host cells.

Even though the heterologous CMV promoter is beneficial for the production of rAAV from the split packaging vector system, product quality analysis showed that most of the capsids generated from these vectors were empty and therefore were unable to provide therapeutic benefits. This remains the case even when a second short *rep* gene was introduced to enhance packaging and accumulation of single-stranded viral genome. Nevertheless, it may be possible to circumvent this drawback by using specific *cis*-regulatory modules within the CMV promoter architecture (i.e., specific strengths) we previously reported[26] for defined capsid expression levels. Moreover, promoter activity in a given cell host is governed by a system-specific combination of interactions between the promoter's constituent TF binding sites and the availability of endogenous TFs.[26] Accordingly, the use of the CMV promoter to drive *cap* gene expression likely resulted in the titration of TFs away from the endogenous E4 promoter affecting the expression level of E4orf proteins (and consequently rAAV titer; Figure 6). Expectedly, further bioinformatic analysis of regulatory elements within these promoters indicated significant (active) TF binding site overlaps between them (Figure S3). In this regard, vectors utilizing synthetic promoters designed de novo using specific TF binding site building blocks[52,53] are likely to be the solution for predictable stoichiometries of different AAV vector components in transient as well as stable systems.[54]

Lastly, our study highlights the complexity of rAAV vector expression systems and that coordinated optimization of a variety of linked dynamic processes (within and between packaging and Helper plasmids) is ultimately necessary to maximize volumetric rAAV product yield and quality. Systematic optimization study could be achieved via DoE-based co-transfection of multiple plasmids each carrying a specific AAV gene(s). This approach explores a large design space and theoretically enables the identification of the "ideal" gene expression stoichiometry for a given system. However, such experimental design discounts the spatial relationship underpinning promoter behavior that may prevent rational improvement or confident prediction of their functionality—thus necessitating testing of each component directly in the final packaging/Helper plasmid constructs. The constraint of the latter method is the difficulty in constructing and screening hundreds of possible vector variants to identify the optimal vector design(s). Nevertheless, given the availability of high-throughput screening techniques,[55] and as gene synthesis costs are becoming cheaper[56] and the emerging technology of DNA-synthesizing enzymes (for long genes and whole vectors) is becoming more efficient,[57] rapid parallel evaluation of rAAV vector designs may indeed be tractable.

AUTHOR CONTRIBUTIONS

Jared Whitehead: Data curation (supporting); formal analysis (supporting); investigation (supporting); methodology (supporting); validation (supporting); writing—review and editing (supporting). Ping Liu: Conceptualization (supporting); funding acquisition (supporting); methodology (supporting); project administration (supporting); resources (supporting); supervision (supporting); validation (supporting); writing—review and editing (supporting). Ayda Mayer: Conceptualization (supporting); funding acquisition (supporting); methodology (supporting); project administration (supporting); resources (supporting); supervision (supporting); validation (supporting); writing—review and editing (supporting).

ACKNOWLEDGMENTS

This study was supported by REGENXBIO, U.S.A.

CONFLICT OF INTEREST STATEMENT

The authors have patent applications filed based on the work in this paper.

DATA AVAILABILITY STATEMENT

Data is available in the article's supplementary material.

ORCID

Yusuf B. Johari https://orcid.org/0000-0001-9933-5764
Thilo H. Pohle https://orcid.org/0000-0003-4437-3231
Jared Whitehead https://orcid.org/0000-0002-5311-4197
Joseph M. Scarrott https://orcid.org/0000-0002-6046-7687
David C. James https://orcid.org/0000-0002-1697-151X

REFERENCES

1. Bulaklak, K., & Gersbach, C. A. (2020). The once and future gene therapy. *Nature Communications*, *11*, 11–14.
2. Dobrowsky, T., Gianni, D., Pieracci, J., & Suh, J. (2021). AAV manufacturing for clinical use: Insights on current challenges from the upstream process perspective. *Current Opinion in Biomedical Engineering*, *20*, 100353.
3. Sha, S., Maloney, A. J., Katsikis, G., Nguyen, T. N. T., Neufeld, C., Wolfrum, J., Barone, P. W., Springs, S. L., Manalis, S. R., Sinskey, A. J., & Braatz, R. D. (2021). Cellular pathways of recombinant adeno-associated virus production for gene therapy. *Biotechnology Advances*, *49*, 107764.
4. Yang, Q., Chen, F., & Trempe, J. P. (1994). Characterization of cell lines that inducibly express the adeno-associated virus Rep proteins. *The Journal of Virology*, *68*, 4847–4856.
5. Schmidt, M., Afione, S., & Kotin, R. M. (2000). Adeno-associated virus type 2 Rep78 induces apoptosis through caspase activation independently of p53. *The Journal of Virology*, *74*, 9441–9450.
6. McCarty, D. M., Christensen, M., & Muzyczka, N. (1991). Sequences required for coordinate induction of adeno-associated virus p19 and p40 promoters by Rep protein. *The Journal of Virology*, *65*, 2936–2945.
7. Mathews, L. C., Gray, J. T., Gallagher, M. R., & Snyder, R. O. (2002). Recombinant adeno-associated viral vector production using stable packaging and producer cell lines. *Methods in Enzymology*, *346*, 393–413.
8. Ogasawara, Y., Urabe, M., Kogure, K., Kume, A., Colosi, P., Kurtzman, G. J., & Ozawa, K. (1999). Efficient production of adeno-associated virus vectors using split-type helper plasmids. *Japanese Journal of Cancer Research*, *90*, 476–483.
9. Jing, X. J., Kalman-Maltese, V., Cao, X., Yang, Q., & Trempe, J. P. (2001). Inhibition of adenovirus cytotoxicity, replication, and E2a gene expression by adeno-associated virus. *Virology*, *291*, 140–151.
10. Cao, L., Liu, Y., During, M. J., & Xiao, W. (2000). High-titer, wild-type free recombinant adeno-associated virus vector production using intron-containing helper plasmids. *The Journal of Virology*, *74*, 11456–11463.
11. Wilmes, G. M., Carey, K. L., Hicks, S. W., Russell, H. H., Stevenson, J. A., Kocjan, P., Lutz, S. R., Quesenberry, R. S., Shulga-Morskoy, S. V., Lewis, M. E., Clark, E., Medik, V., Cooper, A. B., & Reczek, E. E. (2014). Non-viral adeno-associated virus-based platform for stable expression of antibody combination therapeutics. *Monoclonal Antibodies*, *6*, 957–967.
12. Li, J., Samulski, R. J., & Xiao, X. (1997). Role for highly regulated rep gene expression in adeno-associated virus vector production. *The Journal of Virology*, *71*, 5236–5243.
13. Emmerling, V. V., Pegel, A., Milian, E. G., Venereo-Sanchez, A., Kunz, M., Wegele, J., Kamen, A. A., Kochanek, S., & Hoerer, M. (2016). Rational plasmid design and bioprocess optimization to enhance recombinant adeno-associated virus (AAV) productivity in mammalian cells. *Biotechnology Journal*, *11*, 290–297.
14. Mietzsch, M., Eddington, C., Jose, A., His, J., Chipman, P., Henley, T., Choudhry, M., McKenna, R., & Agbandje-McKenna, M. (2021). Improved genome packaging efficiency of adeno-associated virus vectors using Rep hybrids. *The Journal of Virology*, *95*, e0077321.
15. Wang, Z., Cheng, F., Engelhardt, J. F., Yan, Z., & Qiu, J. (2018). Development of a novel recombinant adeno-associated virus production system using human bocavirus 1 Helper genes. *Molecular Therapy – Methods, & Clinical Development*, *11*, 40–51.
16. Zhao, H., Lee, K. J., Daris, M., Lin, Y., Wolfe, T., Sheng, J., Plewa, C., Wang, S., & Meisen, W. H. (2020). Creation of a high-yield AAV vector production platform in suspension cells using a design-of-experiment approach. *Molecular Therapy – Methods, & Clinical Development*, *18*, 312–320.
17. Scarrott, J. M., Johari, Y. B., Pohle, T. H., Liu, P., Mayer, A., & James, D. C. (2023). Increased recombinant adeno-associated virus production by HEK293 cells using small molecule chemical additives. *Biotechnology Journal*, *13*, e2200450.
18. Wang, Q., Nambiar, K., & Wilson, J. M. (2021). Isolating natural adeno-associated viruses from primate tissues with a high-fidelity polymerase. *Human Gene Therapy*, *32*, 1439–1449.
19. Ogasawara, Y., Urabe, M., & Ozawa, K. (1998). The use of heterologous promoters for adeno-associated virus (AAV) protein expression in AAV vector production. *Microbiology and Immunology*, *42*, 177–185.
20. Vincent, K. A., Piraino, S. T., & Wadsworth, S. C. (1997). Analysis of recombinant adeno-associated virus packaging and requirements for rep and cap gene products. *The Journal of Virology*, *71*, 1897–1905.
21. Grimm, D., Kern, A., Rittner, K., & Kleinschmidt, J. A. (1998). Novel tools for production and purification of recombinant adenoassociated virus vectors. *Human Gene Therapy*, *9*, 2745–2760.
22. Pereira, D. J., McCarty, D. M., & Muzyczka, N. (1997). The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. *The Journal of Virology*, *71*, 1079–1088.
23. Pereira, D. J., & Muzyczka, N. (1997). The adeno-associated virus type 2 p40 promoter requires a proximal Sp1 interaction and a p19 CArG-like element to facilitate Rep transactivation. *The Journal of Virology*, *71*, 4300–4309.
24. Xiao, X., Li, J., & Samulski, R. J. (1998). Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *The Journal of Virology*, *72*, 2224–2232.
25. Karreth, F. A., Tay, Y., & Pandolfi, P. P. (2014). Target competition: Transcription factors enter the limelight. *Genome Biology*, *15*, 114.
26. Johari, Y. B., Scarrott, J. M., Pohle, T. H., Liu, P., Mayer, A., Brown, A. J., & James, D. C. (2022). Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells. *Biotechnology Journal*, *17*, e2200062.
27. Hölscher, C., Kleinschmidt, J. A., & Bürkle, A. (1995). High-level expression of adeno-associated virus (AAV) Rep78 or Rep68 protein is sufficient for infectious-particle formation by a rep-negative AAV mutant. *The Journal of Virology*, *69*, 6880–6885.
28. Urabe, M., Ding, C., & Kotin, R. M. (2002). Insect cells as a factory to produce adeno-associated virus type 2 vectors. *Human Gene Therapy*, *13*, 1935–1943.
29. Saudan, P., Vlach, J., & Beard, P. (2000). Inhibition of S-phase progression by adeno-associated virus Rep78 protein is mediated by hypophosphorylated pRb. *The EMBO Journal*, *19*, 4351–4361.
30. King, J. A., Dubielzig, R., Grimm, D., & Kleinschmidt, J. A. (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *The EMBO Journal*, *20*, 3282–3291.
31. Qiu, J., & Pintel, D. J. (2002). The adeno-associated virus type 2 Rep protein regulates RNA processing via interaction with the transcription template. *Molecular and Cellular Biology*, *22*, 3639–3652.
32. Farris, K. D., & Pintel, D. J. (2008). Improved splicing of adeno-associated viral (AAV) capsid protein-supplying pre-mRNAs leads to increased recombinant AAV vector production. *Human Gene Therapy*, *19*, 1421–1427.
33. Di Pasquale, G., & Chiorini, J. A. (2003). PKA/PrKX activity is a modulator of AAV/adenovirus interaction. *The EMBO Journal*, *22*, 1716–1724.
34. Donovan-Banfield, I., Turnell, A. S., Hiscox, J. A., Leppard, K. N., & Matthews, D. A. (2020). Deep splicing plasticity of the human adenovirus type 5 transcriptome drives virus evolution. *Communications Biology*, *3*, 124.
35. Westergren Jakobsson, A., Segerman, B., Wallerman, O., Lind, S. B., Zhao, H., Rubin, C. J., Pettersson, U., & Akusjarvi, G. (2021). The human adenovirus 2 transcriptome: An amazing complexity of alternatively spliced mRNAs. *The Journal of Virology*, *95*, e01869–e01920.
36. Wu, K., Orozco, D., & Hearing, P. (2012). The adenovirus L4-22K protein is multifunctional and is an integral component of crucial aspects of infection. *The Journal of Virology*, *86*, 10474–10483.

37. Wu, K., Guimet, D., & Hearing, P. (2013). The adenovirus L4-33K protein regulates both late gene expression patterns and viral DNA packaging. *The Journal of Virology*, *87*, 6739–6747.
38. Leppard, K. N. (1997). E4 gene function in adenovirus, adenovirus vector and adeno-associated virus infections. *Journal of General Virology*, *78*, 2131–2138.
39. Matsushita, T., Elliger, S., Elliger, C., Podsakoff, G., Villarreal, L., Kurtzman, G. J., Iwaki, Y., & Colosi, P. (1998). Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Therapy*, *5*, 938–945.
40. Allen, J. M., Halbert, C. L., & Miller, A. D. (2000). Improved adeno-associated virus vector production with transfection of a single helper adenovirus gene, E4orf6. *Molecular Therapy*, *1*, 88–95.
41. Huang, M. M., & Hearing, P. (1989). Adenovirus early region 4 encodes two gene products with redundant effects in lytic infection. *The Journal of Virology*, *63*, 2605–2615.
42. Halbert, D. N., Cutt, J. R., & Shenk, T. (1985). Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. *The Journal of Virology*, *56*, 250–257.
43. Bondesson, M., Ohman, K., Manervik, M., Fan, S., & Akusjärvi, G. (1996). Adenovirus E4 open reading frame 4 protein autoregulates E4 transcription by inhibiting E1A transactivation of the E4 promoter. *The Journal of Virology*, *70*, 3844–3851.
44. Huang, M. M., & Hearing, P. (1989). The adenovirus early region 4 open reading frame 6/7 protein regulates the DNA binding activity of the cellular transcription factor, E2F, through a direct complex. *Genes and Development*, *3*, 1699–1710.
45. Chang, L. S., & Shenk, T. (1990). The adenovirus DNA-binding protein stimulates the rate of transcription directed by adenovirus and adeno-associated virus promoters. *The Journal of Virology*, *64*, 2103–2109.
46. Watanabe, M., Sakaguchi, M., Kinoshita, R., Kaku, H., Ariyoshi, Y., Ueki, H., Tanimoto, R., Ebara, S., Ochiai, K., Futami, J., Li, S., Huang, P., Nasu, Y., Huh, N., & Kumon, H. (2014). A novel gene expression system strongly enhances the anticancer effects of a REIC/Dkk-3-encoding adenoviral vector. *Oncology Reports*, *31*, 1089–1095.
47. Park, S., Shin, S., Lee, H., Jang, J. H., & Lee, G. M. (2024). Enhancing the production of adeno-associated virus (AAV)2 and AAV9 with high full capsid ratio in HEK293 cells through design-of-experiment optimization of triple plasmid ratio. *Biotechnology Journal*, *19*, e2300667.
48. Carbaugh, D., Fernandez, A., Suarez, L., & Samulski, R. J. (2021). Development of a split Rep/Cap system to improve AAV capsid production. *Molecular Therapy*, *29*, 96.
49. Su, W., Seymour, L. W., & Cawood, R. (2023). AAV production in stable packaging cells requires expression of adenovirus 22/33K protein to allow episomal amplification of integrated rep/cap genes. *Scientific Reports*, *13*, 21670.
50. Zhou, H., & Beaudet, A. L. (2000). A new vector system with inducible E2a cell line for production of higher titer and safer adenoviral vectors. *Virology*, *275*, 348–357.
51. Klessig, D. F., Brough, D. E., & Cleghon, V. (1984). Introduction, stable integration, and controlled expression of a chimeric adenovirus gene whose product is toxic to the recipient human cell. *Molecular and Cellular Biology*, *4*, 1354–1362.
52. Johari, Y. B., Brown, A. J., Alves, C. S., Zhou, Y., Wright, C. M., Estes, S. D., Kshirsagar, R., & James, D. C. (2019). CHO genome mining for synthetic promoter design. *Journal of Biotechnology*, *294*, 1–13.
53. Johari, Y. B., Mercer, A. C., Liu, Y., Brown, A. J., & James, D. C. (2021). Design of synthetic promoters for controlled expression of therapeutic genes in retinal pigment epithelial cells. *Biotechnology and Bioengineering*, *118*, 2001–2015.
54. Patel, Y. D., Brown, A. J., Zhu, J., Rosignoli, G., Gibson, S. J., Hatton, D., & James, D. C. (2021). Control of multigene expression stoichiometry in mammalian cells using synthetic promoters. *ACS Synthetic Biology*, *10*, 1155–1165.
55. Quan, D. N., & Shiloach, J. (2022). rAAV production and titration at the microscale for high-throughput screening. *Human Gene Therapy*, *33*, 94–102.
56. Riolo, J., & Steckl, A. J. (2022). Comparative analysis of genome code complexity and manufacturability with engineering benchmarks. *Scientific Reports*, *12*, 2808.
57. Eisenstein, M. (2020). Enzymatic DNA synthesis enters new phase. *Nature Biotechnology*, *38*, 1113–1115.

SUPPORTING INFORMATION

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

How to cite this article: Johari, Y. B., Pohle, T. H., Whitehead, J., Scarrott, J. M., Liu, P., Mayer, A., & James, D. C. (2024). Molecular design of controllable recombinant adeno-associated virus (AAV) expression systems for enhanced vector production. *Biotechnology Journal*, *19*, e2300685. https://doi.org/10.1002/biot.202300685

8.2.4.1 Supplementary Data Johari and Pohle et al.

SUPPLEMENTARY MATERIALS

Table S1. Sequences of AAV2 endogenous promoters and the CMV promoter. Nucleotide numbers refer to NCBI GenBank.

Promoter	Sequence	Reference
p5	nt 189–320, GenBank AF043303.1	Murphy et al.[S1]
p5 remnant	nt 275–320, GenBank AF043303.1	–
p19	nt 720–899, GenBank AF043303.1	McCarty et al.[S2]
p40	nt 1700–1879, GenBank AF043303.1	McCarty et al.[S2]
E2-early	nt 27325–27023, GenBank AY339865.1	Casper et al.[S3]
E2-late	nt 26078–25988, GenBank AY339865.1	Bhat et al.[S4]
L4	nt 26125–25887, GenBank AY339865.1	Morris et al.[S5]
E4	nt 35933–35565, GenBank AY339865.1	Gilardi & Perricaudet[S6]
CMV	nt 174143–174740, GenBank MN920393.1	Johari et al.[S7]

Table S2. Sequence mutations for inactivation or optimization of Rep/Cap elements.

Mutation	Sequence
Inactivation of p40 TATA box	TATAAGTGA → CATCAGTGA
Inactivation of p40 Inr	TCAGTT → AGCGTT
Optimization of *rep52/40* Kozak[S8]	TACATGG → ATCATGG (start codon underlined)
Inactivation of *cap* start codon	ATG → CTT

1

Table S3. AAV2/8 ddPCR primers and probes (5'→3').

Target	Primer/ Probe	Sequence
GFP poly A	Fwd	GGACATCATGAAGCCCCTT
	Rev	TCCAACACACTATTGCAATGAAAA
	Probe	FAM/AGCATCTGA/ZEN/CTTCTGGCTAATAAAGGAA/IBFQ
Rep78/68	Fwd	AACAAGGTGGTGGATGAGT
	Rev	CGTTTACGCTCCGTGAGATT
	Probe	FAM/ACTGTTCCA/ZEN/TATTAGTCCACGCCCAC/IBFQ
Rep78/68/52/40	Fwd	GGCCTCATACATCTCCTTCAAT
	Rev	AGTCAGGCTCATAATCTTTCCC
	Probe	HEX/TCCAACTCG/ZEN/CGGTCCCAAAT/IBFQ
Cap	Fwd	CCATTTGGCATGACACTACG
	Rev	AGGATCCCGTTACTGGGAAA
	Probe	FAM/ACACAAAGA/ZEN/CGACGAGGAGCGTTT/IBFQ
E4orf6, 6/7	Fwd	GTACCGGGAGGTGGTGAATTAC
	Rev	CATTGTCAAAGTGTTACATTCGG
	Probe	HEX/CTTTTGAGA/ZEN/CAGAAACCCGCGCTAC/IBFQ
E2A	Fwd	GCTCAGGTGGCTTTTAAGC
	Rev	GTAGCTGCCTTCCCAAAAAG
	Probe	FAM/CTTTTGATG/ZEN/CCACTACGGTGCGAG/IBFQ

Figure S1. Schematic depiction of the adenovirus-5 genome-derived Helper 1.0 plasmid (12.1 kb including backbone) utilized in this study. The E4 (encoding orf1–6/7 proteins), L4 (encoding 33k/22k proteins), E2A (encoding DNA binding protein) and VA RNA regions are indicated. The L5 region between E4 poly A and L4 poly A is deleted. Arrow denotes a promoter, black circle denotes a poly A. All components are drawn to approximate scale.

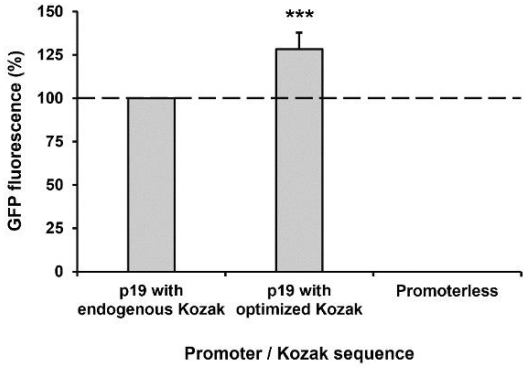


Figure S2. The Kozak sequence from *rep52/42* gene can be optimized to increase gene expression. 8 μg plasmid DNA encoding maxGFP, driven by the p19 promoter with either the endogenous or optimized Kozak sequence (Table S2), were transfected into 1×10^7 HEK293 cells using PEImax as previously described.[S7] A promoterless vector was used as a negative control. Cells were cultured in TubeSpin bioreactors at 37°C and GFP fluorescence was measured at 48 h post-transfection as previously described.[S7] Data are expressed as a percentage compared to the GFP fluorescence of the endogenous Kozak sequence. Data shown are the mean ± S.D. of three independent transfections. Data were analyzed using unpaired Student's *t*-test with respect to the endogenous Kozak. *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

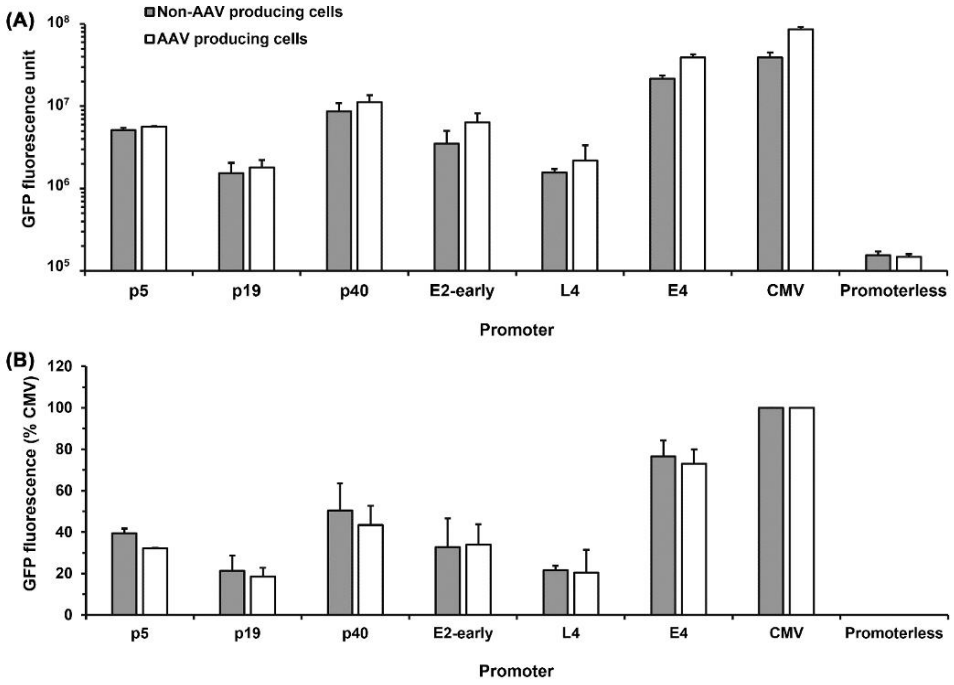


Figure S2. Endogenous AAV and the CMV promoters drive varying gene expression levels in non-AAV producing (■) and AAV producing (□; empty capsids) HEK293 cells. Vectors encoding maxGFP, driven by an endogenous AAV or the CMV promoter (Table S1), were co-transfected with Rep/Cap and Helper into 1×10^7 HEK293 cells at 4:1:2 weight ratio (GFP, Rep/Cap, Helper; 10 μg total DNA) using PEImax as previously described.[S7] Rep/Cap and Helper plasmids with deleted/inactivated p5, p19, p40, E2-early and E4 promoters/introns (Tables S1 and S2) were used for the non-AAV producing condition. A promoterless vector was used as a negative control. Cells were cultured in TubeSpin bioreactors at 37°C and GFP fluorescence was measured at 48 h post-transfection **(A)** and normalized against the CMV promoter **(B)** as previously described.[S7] Data shown are the mean ± S.D. of three independent transfections.

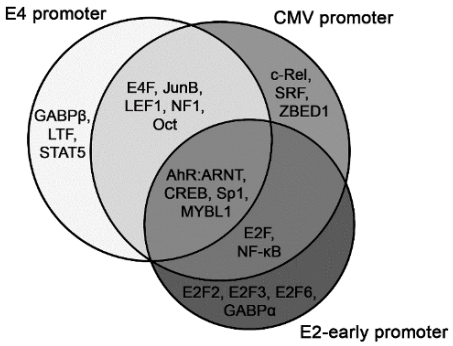


Figure S3. Distribution of discrete transcription factor (TF) binding sites within the E4, E2-early and CMV promoters. Promoters were surveyed for the presence of discrete TF binding sites using Genomatix Gene Regulation software as previously described.[S7] Active transcription factors in HK293 cells include E4F, JunB, Oct, AhR:ARNT, CREB, Sp1, MYBL1, c-Rel, ZBED1 and NF-κB.[S7]

References

S1. Murphy, M., Gomos-Klein, J., Stankic, M., & Falck-Pedersen, E. (2007). Adeno-associated virus type 2 p5 promoter: a rep-regulated DNA switch element functioning in transcription, replication, and site-specific integration. *Journal of Virology, 81,* 3721–3730.

S2. McCarty, D. M., Christensen, M., & Muzyczka, N. (1991). Sequences required for coordinate induction of adeno-associated virus p19 and p40 promoters by Rep protein. *Journal of Virology, 65,* 2936–2945.

S3. Casper, J. M., Timpe, J. M., Dignam, J. D., & Trempe, J. P. (2005). Identification of an adeno-associated virus Rep protein binding site in the adenovirus E2a promoter. *Journal of Virology, 79,* 28–38.

S4. Bhat, G., SivaRaman, L., Murthy, S., Domer, P., & Thimmappaya, B. (1987). In vivo identification of multiple promoter domains of adenovirus EIIA-late promoter. *The EMBO Journal, 6,* 2045–2052.

S5. Morris, S. J., Scott, G. E., & Leppard, K. N. (2010). Adenovirus late-phase infection is controlled by a novel L4 promoter. *Journal of Virology, 84,* 7096–7104.

S6. Gilardi, P., & Perricaudet, M. (1986). The E4 promoter of adenovirus type 2 contains an E1A dependent cis-acting element. *Nucleic Acids Research, 14,* 9035–9049.

S7. Johari, Y. B., Scarrott, J. M., Pohle, T. H., Liu, P., Mayer, A., Brown, A. J., & James, D. C. (2022). Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells. *Biotechnology Journal, 17,* e2200062.

S8. Kozak, M. (1997). Recognition of AUG and alternative initiator codons is augmented by G in position +4 but is not generally affected by the nucleotides in positions +5 and +6. *The EMBO Journal, 16,* 2482–2492.

8.3 Lists of Equipment and Materials

8.3.1 Laboratory Equipment

Table 6: Laboratory devices used during this work.

Device	Model / type designation	Manufacturer
Agarose gel electrophoresis tank	Fisherbrand Midi Plus Horizontal Gel System	Thermo Fisher Scientific
Autoclave	Top loading 100L	Prioclave
Automated cell counter	Vi-Cell XR	Beckman Coulter
Automated cell counter	Countess 3 FL	Thermo Fisher Scientific
Blue light LED table	DR-46B	Clare Chemical Research
Centrifuge	Pico 17 Microfuge	Thermo Fisher Scientific
Centrifuge, refrigerated	Mega Star 1.6R	VWR
Centrifuge, refrigerated	5424R	Eppendorf
CO2 Incubator with orbital shaking non-humidified	Multitron II	Infors HT
CO2 Incubator with orbital shaking humidified	Multitron Cell	Infors HT
Electrophoresis power supply	CS-300V Lab	Thermo Fisher Scientific
Flow cytometer	Attune Acoustic Focusing Cytometer	Applied Biosystems
Freezer / -20 °C	GG4060	Liebherr
Freezer / -80 °C	New Brunswick U410	Eppendorf
Mixing heating / cooling block	Thermomixer comfort 5355	Eppendorf
Laminar flow biosafety cabinet	Safe 2020 1.2 Class II Biological Safety Cabinet	Thermo Fisher Scientific
Microbiology incubation shaker	Multitron II	Infors HT
pH meter	Seven Compact	Mettler Toledo
Pipette, multi-channel	12 channels	StarLab
Pipettes	2.5 µL – 1 mL	Eppendorf
Pipetting aid	Pipetboy 2	Integra Bioscience
Refrigerator / +4 °C	LCSM1545	Beko
Static microbiology incubator	HCP 108	Memmert
Scale, analytical	SI-114	Denver instrument
Multi-Mode microplate reader	SpectraMax iD5	Molecular Devices
Spectrophotometer (UV-Vis)	Nanodrop Onec	Thermo-Scientific
PCR Thermocycler	SimpliAmp Thermo Cycler	Applied Biosystems
Vaccum Manifold	QIAvac 24 Plus	Qiagen
Vacuum pump	V-700	Büchi
Vortex	Vortex Genie 2	Scientific Industries
Water purifier	Arium mini	Sartorius

Device	Model / type designation	Manufacturer
Droplet generator	QX200	BioRad
PCR plate sealer	PX1	BioRad
Droplet Digital PCR reader	QX200	BioRad
ddPCR Thermocycler	C1000	BioRad

8.3.2 Laboratory Consumable Materials

Table 7: Consumables used in the laboratory during this work.

Consumable material	Model / type designation	Manufacturer
96-well ddPCR microtiter plate	ddPCR Plates 96-Well, Semi-Skirted, 25/BX	BioRad
ddPCR Supermix	ddPCR Supermix for Probes, 2 x 1 ml	BioRad
ddPCR cartridges	DG8 Cartridges, QX100/QX200, 24/PK	BioRad
ddPCR gaskets	DG8 Gaskets, QX100/QX200, 24/PK	BioRad
ddPCR droplet generation oil	Droplet Gen Oil for Probes, 10 x 7 ml	BioRad
Countess single use cell counting slides	Countess Cell Counting Chamber Slides	Invitrogen
Centrifuge tube	Falcon 15, 50 mL	Corning
Conical flask		Schott
Disposable gloves	StarGuard	Starlab
Glass plating beads	ColiRollers Plating Beads	Novagen
Microtiter plate	96-well, non-binding, flat bottom, poly sterol	Thermo Scientific
PCR reaction vessel	Thin-walled, plastic, 0.2µL	Starlab
Pipette tip	TipOne, plastic, graduated, with and w/o filter, 10, 20, 1000 µL	Starlab
Pipette tip low retention	TipOne, Low retention (RPT) plastic, graduated, with and w/o filter, 10, 20, 1000 µL	Starlab
Reaction tube	0.5, 1.5, 2.0, 5 mL	Starlab / Eppendorf
Sterile filter	Minisart syringe filter, pore sizes 0.2 µm	Sartorius
Test tube flow cytometry	5 mL, polystyrene	Hartenstein
Precision wipes	Kimtech Science Professional	Kimberly-Clark

8.3.3 Kits

Table 8: Kits used during this work.

Name	Manufacturer	Catalogue-ID
QIAprep Spin Miniprep Kit	Qiagen	7106
QIAGEN Plasmid Plus Midi Sample Kit	Qiagen	12945
QIAGEN Plasmid Plus Maxi Sample Kit	Qiagen	12963
QIAquick PCR Purification Kit	Qiagen	28106
QIAquick Gel Extraction Kit	Qiagen	28706
AAV8 Titration ELISA	Progen	PRAAV8
Q5 Site-Directed Mutagenesis Kit	NEB	E0554S
NEBuilder HiFi DNA Assembly Master Mix	NEB	E2621L

8.3.4 Chemicals

Table 9: DNA molecular weight markers used in agarose gels.

Name	Manufacturer	Cataloge-ID
HyperLadder 1 kb	Bioline	BIO-33053
HyperLadder 100 bp	Bioline	BIO-33056

Table 10: Chemicals used during this work.

Chemical	Manufacturer	ID No.
Acetic acid	Fisher Scientific	CAS 64-19-7
Agar-Agar	Sigma-Aldrich	CAS 9002-18-0
Agarose	AppliChem	CAS 9002-18-0
$CaCl_2 \cdot 2\ H2O$	Merck	CAS 10035-04-8
EDTA disodium salt dihydrate	Sigma	CAS 6381-92-6
Ethanol (p.a.)	VWR Chemicals	CAS 64-17-5
Ethanol (IMS)	Fisher Scientific	CAS 64-17-5
HCl	VWR Chemicals	CAS 7647-01-0
Isopropanol	VWR Chemicals	CAS 67-63-0
KCl	VWR Chemicals	CAS 7447-40-7
KH2PO4	Fluka	CAS 7778-77-0
$Mg_2Cl_2\ 6\ H2O$	Carl Roth	CAS 7791-18-6
$MgSO_4\ 7\ H2O$	Fluka	CAS 10034-99-8
NaCl	VWR Chemicals	CAS 7647-14-5
Pluronic F-68	Gibco	CAS 9003-11-6
TRIS	Carl Roth	CAS 77-86-1
TRIS HCl	Carl Roth	CAS 1185-53-1
Triton X-100	Sigma	CAS 9002-93-1
Tween 20	Carl Roth	CAS 9005-64-5

8.3.5 Enzymes

Table 11: Enzyme buffers and supplements.

Buffer name	Manufacturer	Cataloge-ID
10x 2.1 Buffer	NEB	B7202S
10× 3.1 Buffer	NEB	B7203S
10× CutSmart Buffer	NEB	B7204S
10× T4 DNA ligase reaction buffer	NEB	B0202S
DMSO	NEB	B0515A
dNTP mix	NEB	N0447S
Gel loading dye, purple (6×)	NEB	B7024S
PCR buffer 10x, GeneAmp	Applied Biosystems	4379876

Table 12: Enzymes and enzyme kits, excluding restriction enzymes.

Name	Manufacturer	Cat. No.
Benzonase nuclease (≥250 U/µL)	Sigma-Aldrich	E1014-25KU
DNase I (RNase-free) (10 U/µL)	Roche	04716728001
2x Q5 High-Fidelity 2x Master Mix	NEB	M0492S
Proteinase K	NEB	P8107S
T4 DNA Ligase (1 U/µL)	NEB	M0202S

Table 13: Restriction enzymes.

Enzyme name	Restriction site	Manufacturer	Cat. No.
AatII	5'-GACGT|C-3'	NEB	R0117S
AgeI-HF	5'-A|CCGGT-3'	NEB	R3552S
AclI	5'-AA|CGTT-3'	NEB	R0598S
BamHI-HF	5'-G|GATCC-3'	NEB	R3136S
BclI-HF	5'-T|GATCA-3'	NEB	R3160S
BlpI	5'-GC|TNAGC-3'	NEB	R0585S
BsiWI-HF	5'-C|GTACG-3'	NEB	R3553S
BsrGI-HF	5'-T|GTACA-3'	NEB	R3575S
BstEII-HF	5'-G|GTNACC-3'	NEB	R3162S
BstZ17I-HF	5'-GTA'TAC-3'	NEB	R3594S
EagI-HF	5'-C|GGCCG-3'	NEB	R3505S
EcoRI-HF	5'-G|AATC-3'	NEB	R3101S
EcoRV-HF	5'-GAT|ATC-3'	NEB	R3195S
HindIII-HF	5'-A|AGCTT-3'	NEB	R3104S
KpnI-HF	5'-GGTAC'C-3'	NEB	R3142S
MfeI-HF	5'-C|AATTG-3'	NEB	R3589S
MluI-HF	5'-A|CGCGT-3'	NEB	R3198S
NcoI-HF	5'-C|CATGG-3'	NEB	R3193S
NdeI	5'-CA|TATG-3'	NEB	R0111S
NgoMIV	5'-G|CCGGC-3'	NEB	R0564S
NheI-HF	5'-G|CTAGC-3'	NEB	R3131S
NotI-HF	5'-GC|GGCCGC-3'	NEB	R3189S
PsiI	5'-TTA|TAA-3'	NEB	R0657S
PspOMI	5'-G|GGCCC-3'	NEB	R0653S
PstI-HF	5'-CTGCA|G-3'	NEB	R3140S
PvuI-HF	5'-CGAT|CG-3'	NEB	R3151S
SacI-HF	5'-GAGCT|C-3'	NEB	R3156S
SacII	5'-CCGC|GG-3'	NEB	R0157S
SalI-HF	5'-G|TCGAC-3'	NEB	R3138S
ScaI-HF	5'-AGT|ACT-3'	NEB	R3122S
SgrAI	5'-CR|CCGGYG-3'	NEB	R0603S
SpeI-HF	5'-A|CTAGT-3'	NEB	R3133S
SphI-HF	5'-GCATG|C-3'	NEB	R3182S
SrfI	5'-GCCC|GGGC-3'	NEB	R3132S
SwaI	5'-ATTT|AAAT-3'	NEB	R0604
XbaI	5'-T|CTAGA-3'	NEB	R0145S
XmaI	5'-C|CCGGG-3'	NEB	R0180S

8.3.6 Antibiotics

Table 14: Antibiotics utilised for bacterial selection pressure.

Antibiotic	Stock solution concentration	Working concentration
Ampicillin / Amp	100 mg/mL in MilliQ	100 µg/mL
Kanamycin / Kan	40 mg/mL in MilliQ	40 µg/mL

8.3.7 Oligonucleotides

Table 15: Primers and probes used for ddPCR experiments.

Primer name	Sequence	Length / bp
dd_RGX_polyA.F	GGACATCATGAAGCCCCTT	25
dd_RGX_polyA.R	TCCAACACACTATTGCAATGAAAA	20
ddPCR_E2A_fwd	GCTCAGGTGGCTTTTAAGC	19
ddPCR_E2A_rev	GTAGCTGCCTTCCCAAAAAG	20
ddPCR_E2A_probe	CTTTTGATGCCACTACGGTGCGAG/56-FAM/CTTTTGATG/ZEN/CCACTACGGTGCGAG/3IABKFQ	24
ddPCR_E4_fwd	CCATTTGGCATGACACTACG	20
ddPCR_E4_rev	CATTGTCAAAGTGTTACATTCGG	23
ddPCR_E4_probe	CTTTTGAGACAGAAACCCGCGCTAC/5HEX/CTTTTGAGA/ZEN/CAGAAACCCGCGCTAC/3IABKFQ	25
ddPCR_orf6_fwd	GTACCGGGAGGTGGTGAATTAC	22
ddPCR_orf6_rev	CAAGGCGCTGTATCCAAAG	19
ddPCR_orf6_probe	CACTTAATCTACCTGCGCTTGTGGTATG/56-FAM/CACTTAATC/ZEN/TACCTGCGCTTGTGGTATG/3IABKFQ	28

Table 16: Sequencing primers used and created.

ID	Primer name	Sequence	Length / bp
Seq1	SeqE2a_1_Fwd	TGAGTGTGCCGATCGTGTCTG	21
Seq2	SeqE2a_2_Rev	GTTCGCCTTTAGATCGTTATCCACG	25
Seq3	SeqE2a_3_Fwd	TATGCAGGCGCTGTATCCTAACG	23
Seq4	SeqE2aPro_1_Fwd	TTATACCCTGCCCGGGCGAC	20
Seq5	SeqE2aPro_2_Rev	ACGACACGTCCTCCATGGTTG	21
Seq6	SeqE4_1_Fwd	AGCGCTGTATGTTGTTCTGGAGC	23
Seq7	SeqE4_2_Rev	GCAGGCGTAGAGACAACATTACAGC	25
Seq8	SeqE4_3_Fwd	GGTTAGCATAGCTCCGAGTATGCG	24
Seq9	SeqE4_4_Fwd	GTGGTGGATGTTATCAGGGCAGC	23
Seq10	SeqE4_5_Fwd	GGCTCTCCACTGTCATTGTTCCAG	24
Seq11	SeqE4_6_Fwd	ACCACTGCCATGTTGTATTCCTGC	24
Seq12	SeqVA_1_Fwd	GCATCACCTGGATGTCCAGGTAC	23
Seq13	SeqVA_2_Rev	AGTCCGCACCAGGTACTGGTATCC	24
Seq14	SeqCap_Fwd_6	CCCGAGATCCAGTACACCTCCAAC	24
Seq14b	SeqCap_Rev_6	GTTGGAGGTGTACTGGATCTCGGG	24
Seq15	SeqCap8_1_Rev	ACTCGCGAATGCCCTCAGAG	20
Seq16	SeqCap8_2_Rev	GCCTTCGTTATTGTCTGCCATTGG	24
Seq17	SeqCap8_3_Fwd	TTGGTCAGACTGGCGACTCAGAG	23

Seq18	SeqCap8_4_Fwd	CAAGACCATCGCCAATAACCTCAC	24
Seq19	SeqCap8_5_Fwd	TAGCAACTTTGCCTGGACTGCTG	23
Seq20	SeqColE1_1_Fwd	TTCCGCTTCCTCGCTCACTG	20
Seq20b	SeqColE1_1_Rev	CAGTGAGCGAGGAAGCGGAA	20
Seq21	SeqKan_1_Rev	TCCGCTCATGAGACAATAACCCTG	24
Seq22	SeqRep_1_Rev	GCGCCATTTCTGGTCTTTGTG	21
Seq22b	SeqRep_1_Fwd	CACAAAGACCAGAAATGGCGC	21
Seq23	SeqRep_2_Fwd	CGTTTCCTGAGTCAGATTCGCG	22
Seq24	SeqRep_3_Fwd	AATATGCGGCTTCCGTCTTTCTG	23
Seq25	SeqRep_4_Fwd	CGTCTGGATCATGACTTTGGGAAG	24
Seq26	SeqRep_2_Rev	ACTGTTCCATATTAGTCCACGCCC	24
Seq27	SeqCap8_1b_Rev	TGTAGCCAGGAAGCACCAGACC	22
Seq28	SeqE2a_intron_3	CGGTACTTAATGGGCAC AAAGTCG	25
Seq29	SeqE2a_intron_4	TCATTAGTTTGCGTCGCTCCTCC	23
Seq30	SeqE2aPro_1_Rev	GGAGCTGCCCAAGACTA CTCAACC	25
Seq31	SeqE4_7_Fwd	AGGTACGGTGATCTGTATAAGCTATGTGG	29
Seq32	Seq_BB_1_Fwd	GAGATTTTGAGACACCGTT CTTCGG	26
Seq33	Seq_VA_3_Rev	GTTGACGCTCTAGCGTGCA AAAG	24
Seq34	Seq_E2a_4_Fwd	GCGGATCTGATCTCCGACA AGAG	24
Seq35	Seq_E2a_5_Rev	CGCGTCTTGTGATGAGTCT TCC	23
Seq36	Seq_E4_7_Fwd2	GATGAGCGTTTGGCTCGAC AG	22
Seq37	SeqE2a_intron_2	CTGACTGTCCCAGTATTCCTCCTCG	25
Seq38	SeqE2a_intron_1	AATCCTGTTTCTAAGCTCGCGGG	23
Seq39	Seq_E4_6_Rev	CCAGCTGGCCAAAACCTG	18
Seq40	SeqKan_2_Fwd	CATGTTGGAATTTAATCGCGG	21
Seq41	SeqE2a_intron_5_Rev	GCATGTTGCAGAACTTTAGG	20
Seq42	SeqE4_3_Rev	CGCATACTCGGAGCTATGCTAACC	24
Seq43	SeqE4_4_Rev	GCTGCCCTGATAACATCCACCAC	23
Seq44	SeqAmpRprom_Fwd	CAAATAGGGGTTCCGCGCAC	20
Seq45	YJ_Seq_Prev_PrimerD	CTCTTGAAGTGCATGTTATGGGAC	24
Seq46	YJ_OnGFP_seq_4_Fwd	CGAGGACAGCGTGATCTTCACC	22
Seq47	YJ_CMVcore_3'seq	GCGATCTGACGGTTCAC	17
Seq48	YJ_Seq_PFwd_OnCMV	GCCCAGTACATGACCTTATGGGAC	24
Seq49	PFwd_Multi	GAAGCTCGAGTCGGTTG	17
Seq50	Prev_Multi	TCGACGTTGACCAAGC	16

Table 17: Primers used and created for genetic engineering.

No	Primer name	Sequence	bp
26	Clone_Cap_Rev	TCTAGATAAGATCAATTGCTTCGCAGAGA CCAAAG	35
27	Clone_OriKan_Fwd	CAATTGATCTTATCTAGAGTATTGGGCGCT CTTCC	35
28	Clone_OriKan_Fwd_pHelp	ATAGTCGACGTATTGGGCGCTCTTC	25
29	Clone_OriKan_Rev_pHelp	CGGTGATCACTTACAATTTAGGTGGCAC	28
30	Clone_p5_Fwd	ACGCAATTGGAGGTCCTGTATTAGAGG	27
31	Clone_p5_Rev	TGATCTAGACAGAATTCGGCTTGGC	25
32	Clone_p40_Rev	GATCATATGTGACGTCGATGGCTGC	25
33	Clone_Rep_Fwd	TGTACAAATCTAGCTAGCATGCCGGGGTTT	33

No	Primer name	Sequence	bp
		TAC	
34	Clone_Rep_Rev	CATATGAATGATACGCGTCAGAGAGGTTG TCCTC	34
35	E2a_fwd_NdeI_XbaI	CATATGTACGCTAGAAGGAGAAGGAAATG GCC	32
36	E2aProm_fwd_NdeI	CATCGAGCATATGGTTATACCCTGCCCGG	29
37	E2aProm_rev_XbaI	AATTATCTAGAGCGGAGGCTCTCTTCAGT	29
38	E2aPromIntron_rev_XbaI	AATTAGTCTAGATTCCTTCTCCTATAGGCA G	31
39	E4_fwd_BclI_BsrGI	TGATCACGCTTGTACAATGGCTGCCGCT	28
40	KanR_rev_BsrGI_BclI	TGTACAAGCGTGATCATGTGCGCGGAAC	28
41	Orf6_fwd_BclI_BsrGI	TGATCACGCTTGTACAATGACTACGTCCGG C	31
42	Orf6_rev_XbaI_NdeI	TCTAGACGTACATATGCTACATGGGGGTA GAGTC	34
43	VAI_rev_SalI	GCTAGTTGTCGACGGGCACTCTTCCGT	27
44	VARNA_fwd_SphI_PspOMI	GCATGCACTGATGGGCCCTAATGCTTTCGC TTTCC	35
45	Clone_Cap_Fwd	ACGCGTATCATTCATATGGGAAGCTTCGAT CAACTAC	37
46	Clone_OriKan_Rev	GCTAGCTAGATTTGTACACTTACAATTTAG GTGGCAC	37
47	Clone_p40_Fwd	ATTACGCGTTAATATACTAGTGGTCACCAA GCAGGAAG	38
48	E2a_rev_PspOMI_SphI	GGGCCCATCAGTGCATGCTTTAAAAATCAA AGGGGTTCTGC	41
49	E4_rev_XbaI_NdeI	TCTAGACGTACATATGTTTCAATTGCAGA AAATTTCAAGTC	41
50	VAI_fwd_SphI_PspOMI	TATGCATGCACTGATGGGCCCAAAAGGAG CACTCCCC	37
51	Clone_p40-Cap_mut_Fwd	GTCAATGGTCACCAAGCAGGAAG	23
52	Clone_p40-Cap_mut_Rev	CATATGAGCTCAAGTACTACTAAGGTTGTC CTCGA	35
53	Clone_p40_Fwd-v2	TCAAGTACTTACTATACTAGTGGTCACCAA GCAGGAAG	38
54	newE2a_rev_PspOMI_SphI	GGGCCCATCAGTGCATGCCGTTTAAAAATC AAAGGG	36
55	newE4_fwd_ScaI_BsrGI	AGTACTCGCTTGTACAATGGCTGCCGCT	28
56	newE4prom_rev_BsrGI	GACAGATGTACAAACAGTCAGCCTTACC	28
57	newOrf6_fwd_ScaI_BsrGI	AGTACTCGCTTGTACAATGACTACGTCCGG C	31
58	polyA_E2a_fwd_Synt	CATAGCGGTGCGAAGCTCGAGTCGGTTG	28
59	newKanR_rev_BsrGI_ScaI	TGTACAAGCGAGTACTTGTGCGCGGAAC	28
60	newE4prom_fwd_ScaI	CGTCGGAGTACTCAATCATCAATAATATAC CTTATTTTG	39
61	polyA_E2a_rev_XbaI_SrfI	GACTTCTAGACATTAAGCCCGGGCGACCAC ATCCTCTTACAC	42
62	newE2a_fwd_NdeI_XbaI	CATATGTACGTCTAGACTATAGGAGAAGG AAATGGCC	37
63	BBpHelper_rev_BsrGI_ScaI	TGTACAAGCGAGTACTGAAATGTGCGCGG AAC	32
64	BBpHelper_fwd_SphI_PspOMI	GCATGCACTGATGGGCCCACTTAATGCTTT CGCTTTC	37

No	Primer name	Sequence	bp
68	BBpHelper_fwd_V2	TAAACGGCATGCACTGATGGGCCCACTTAATGCTTTCGCTTTC	43
69	newBBpHelper_rev_V2	GCAGCCATTGTACAAGCGAGTACTGAAATGTGCGCGGAAC	40
70	newE4_fwd_v2	CACATTTCAGTACTCGCTTGTACAATGGCTGCCGCT	36
71	E4_rev_V2	TCCTATAGTCTAGACGTACAT…TTGCAGAAAATTTCAAGTC	49
72	newE2a_fwd_v2	AATTGAAACATATGTACFTCTAGACTATAGGAGAAGGAAATGGCC	45
73	newE2a_rev_v2	TTAAGTGGGCCCATCAGTGCATGCCGTTTAAAAATCAAAGGG	42
74	CMV_fwd_BB8_ScaI	GTGCCACCTAAATTGTAAGCGAAGTACTTGTACAGTTGACATTGATTATTGAC	53
75	CMV_rev_E2a_NdeI_XbaI	TCTAGACAGTACATATGCGTAGTCGCTGGTGTCTTCTATGGAGGTCAAAACAG	53
76	BB8_fwd_E2a_SalI	GAGCATGGAGTTGGGTACGTCGACACGTATTGGGCGCTCTTC	42
77	BB8_rev_E2aProm_ScaI	CCGGGCAGGGTATAACAGTACTTCGCTTACAATTTAGGTGGCAC	44
78	BB8_rev_CMV_ScaI	GTCAATAATCAATGTCAACTGTACAAGTACTTCGCTTACAATTTAGGTGGCAC	53
79	E2aProm_fwd_BB8_ScaI	GTGCCACCTAAATTGTAAGCGAAGTACTGTTATACCCTGCCCGG	44
80	E2aProm_rev_E2a_XbaI_NdeI	TCTAGACAGTACATATGCGTAGTCGCTGGTGGCAGAAAAAGATCATGGAGTCAG	54
81	E2a_fwd_E2aProm_NdeI_XbaI	ACCAGCGACTACGCATATGTACTGTCTAGACTATAGGAGAAGGAAATGGCC	51
82	E2a_rev_BB8_SalI	GAAGAGCGCCCAATACGTGTCGACGTACCCAACTCCATGCTC	42
83	E4prom_fwd_BB8_ScaI	GTGCCACCTAAATTGTAAGCGAAGTACTAATCATCAATAATATACCTTATTTTG	54
84	E4prom_rev_E4orf6	CAAATGGAACGCCGGACGTAGTCATGAACAGTCAGCCTTACC	42
85	CMV_rev_E4orf6	CAAATGGAACGCCGGACGTAGTCATGGTCTTCTATGGAGGTCAAAACAG	49
86	E4orf6_fwd_CMV	CTGTTTTGACCTCCATAGAAGACCATGACTACGTCCGGCGTT	42
87	E4Orf6_rev_XbaI_NdeI_BB8	GAAGAGCGCCCAATACGTTCTAGACGTACATATGCTACATGGGGGTAGAGTC	52
88	E4orf6_fwd_ E4prom	CTTTTTTACTGGTAAGGCTGACTGTTCATGACTACGTCCGGCGTT	45
89	E4Orf6_rev_XbaI_NdeI_BB8	GAAGAGCGCCCAATACGTTCTAGACGATACATATGCTACATGGGGGTAGAGTC	53
90	BB8_fwd_E4orf6_NdeI_XbaI	CGATTATGACTCTACCCCCATGTAGCATATGTATCGTCTAGAACGTATTGGGCGCTCTTC	60
91	BB8_rev_E4prom_ScaI	CAAAATAAGGTATATTATTGATGATTAGTACTTCGCTTACAATTTAGGTGGCAC	54
92	VARNA_fwd_BB8_ScaI	GCATTCAGTACTCCACTTAATGCTTTCGCTTTC	33
93	VARNA_rev_XbaI_NdeI_BB8	GCAAGTTCTAGACGATACATATGCTGCAGGAATTCGATATCAAGC	45
94	E4_rev_XbaI_NdeI_BB8	GGTACGTTCTAGACGATACATATGCCACAT	46

No	Primer name	Sequence	bp
		CCTCTTACACTTTTTC	
95	E4whole_fwd_ScaI_BsrGI	GGTTGAAGTACTCGCTTGTACAATGGCTGC CGCTGT	36
96	E4orf2plus_fwd_ScaI_BsrGI	GGTTGAAGTACTCGCTTGTACAATGTTTG AGAGAAAAATGGTGTC	45
97	E4orf3plus_fwd_ScaI_BsrGI	GGTTGAAGTACTCGCTTGTACATCATGATT CGCTGCTTGAGG	42
98	E4orf4plus_fwd_ScaI_BsrGI	GGTTGAAGTACTCGCTTGTACACATGGTTC TTCCAGCTCTTCC	43
99	E4orf6/7_fwd_ScaI_BsrGI	GGTTGAAGTACTCGCTTGTACAATGACTAC GTCCGGCGTT	40
100	CMV_rev_BsrGI	GACAGATGTACAGGTCTTCTATGGAGGTC AAAACAG	36
101	CMV_fwd_BB8_ScaI	GAGCGAAGTACTCAGTTGACATTGATTAT TGACTAG	36
102	BBexchange SDM 1 Fwd	GCCAGTGTTACAACCAATTAAC	22
103	BBexchange SDM 1 Rev	GTCAGACCCCGTAGAAAAG	19
104	BBexchange SDM 2 Fwd	ACTCTTCCTTTTTCAATATTATTG	24
105	BBexchange SDM 2 Rev	ATGAGCCATATTCAACGG	18
106	33k ATG mut SMD 1 Fwd	TCACCTGGGAAGCAAGGGCCC	21
107	33k ATG mut SMD 2 Rev	GCACCCAAAAAGAAGCTGCAGCTG	24
108	010 BB SpeI SDM Fwd	GGGGGATCCAGTAGTTCTAGC	21
109	010 BB SpeI SDM Rev	GGGCTGCAGGAATTCGAT	18
110	PFwd_Multi	GAAGCTCGAGTCGGTTG	17
111	PRev_Multi	TCGACGTTGACCAAGC	16
112	CMV_fwd_MSCe2a_SrfI	GAGCGAGCCCGGGCCAGTTGACATTGATTA TTGACTAG	38
113	CMV_rev_MSCe2a_AclI	GACAGAAACGTTGGTCTTCTATGGAGGTC AAAACAG	36
114	CMV_rev_XbaI	GACAGATCTAGAGGTCTTCTATGGAGGTC AAAACAG	36
115	E4polyAextensionSDM_Fwd	GTTTGATTAAGGTACGGTG	19
116	E4polyAextensionSDM_Rev	CTACATGGGGGTAGAGTC	18
117	CAGwIntron_fwd	GCTAGTCGACATTGATTATTGAC	23
118	CAGwIntron_rev	GCTGTAGGAAAAAGAAGAAGGC	22
119	CAG_fwd	CCAGGGTAATGGGGATCCTCTAGAACTAT AGC	32
120	CAG_rev	AAGGCAGCGCGCAGCGACT	19
121	rBetaGlobPA_fwd	GACGGGGTGAACTAC	15
122	rBetaGlobPA_rev	GCCGAGTGAGAGACAC	16
123	SV40prom_AB_fwd	GTCTCACACAGGTGTG	16
124	SV40prom_AB_rev	GTTTGCAAAAGCCTAGGC	18
125	EF1aProm_Fwd	CAAATCATGGCTCCGGTG	18
126	EF1aProm_Rev	TCACGACACCTGAAATGG	18
127	bGHpA_Fwd	CGACTGTGCCTTCTAGTTGCCAGC	24
128	bGHpA_Rev	TCCCCAGCATGCCTGCTAT	19
129	hGHpA_Fwd	GATCTACGGGTGGCATCCCTGTGACCC	27
130	hGHpA_Rev	CTACAAAATCAGAAGGACAGGGAAGG	26
131	SV40pAlateRGX_Fwd	TGAGTTTGGACAAACCAC	18
132	SV40pAlateRGX_Rev	TACCACATTTGTAGAGGTTTTAC	23
133	BB8_EcoRI_SDM_fwd	ATTCCCAAGGACATGGTGCTTCTG	24
134	BB8_EcoRI_SDM_rev	TCCGGCCTGCTGAATGCG	18
135	E2ApA_FseI_SDM_fwd	GGCCGCATGCCGGCGCAGACGG	22

No	Primer name	Sequence	bp
136	E2ApA_FseI_SDM_rev	GGCCCGTTTAAAAATCAAAGGGGTTCTGCCGCG	33
137	orf67_MCSext_fwd	CATTGTCCTAGGACCAGCACGGTCCGTGGACTCGTACGCTTGCGAAGTACTCAATCATCAATAATATACC	70
138	orf67_MCSext_rev	CAGCAATAGGCCTTGCGTACTTATAACCAATGCGCGGCCGCAACGCTAGGATCCTCAATTCGACCACATCCTCTTACAC	79
139	E2A_MCSext_fwd	GTTATCAGCCCGGGCATTCACAGTATACTTGTATCAACGTTAGGCGACCGCACCCTGTGACGAAA	65
140	E2A_MCSext_rev	GATATAGGCGCGCCTACAGGAGTTTAAACTCCTGTATTCGAAAGGACTAGGGCCCACTTTCAATAAAGGCAAATGC	76
141	VARNA_newMCSs_fwd	CGAGGTGTTAATTAACGACTTAATGCTTTCGCTTTCCAG	39
142	VARNA_newMCSs_rev	GTAACAGTCGACAGAAGCACCATG	24
143	bareBB8_GA_fwd	ACGTATTGGGCGCTCTTC	18
144	bareBB8_GA_rev	TCGCTTACAATTTAGGTGGCAC	22
145	GA_1_BB8_fwd	GTCGACTGTTACACGTATTGGGCGCTCTTC	30
146	GA_1_BB8_rev	CCTAGGACAATGTCGCTTACAATTTAGGTGGC	32
147	GA_1_orf67_fwd	AAATTGTAAGCGACATTGTCCTAGGACCAGC	31
148	GA_1_orf67_rev	CGGGCTGATAACCAGCAATAGGCCTTGCGTAC	32
149	GA_1_E2A_fwd	AGGCCTATTGCTGGTTATCAGCCCGGGCATTC	32
150	GA_1_E2A_rev	ATTAACACCTCGGATATAGGCGCGCCTACAG	31
151	GA_1_VARNA_fwd	GCGCGCCTATATCCGAGGTGTTAATTAACGAC	32
152	GA_1_VARNA_rev	GCGCCCAATACGTGTAACAGTCGACAGAAGC	31
153	GA_2_BB8_fwd	GTCGACTGTTACACGTATTGGGCGCTCTTC	30
154	GA_2_BB8_rev	CGGGCTGATAACTCGCTTACAATTTAGGTGGC	32
155	GA_2_E2A_fwd	AAATTGTAAGCGAGTTATCAGCCCGGGCATTC	32
156	GA_2_E2A_rev	TCCTAGGACAATGGATATAGGCGCGCCTACAG	32
157	GA_2_orf67_fwd	CGCGCCTATATCCATTGTCCTAGGACCAGC	30
158	GA_2_orf67_rev	ATTAACACCTCGCAGCAATAGGCCTTGCGTAC	32
159	GA_2_VARNA_fwd	AGGCCTATTGCTGCGAGGTGTTAATTAACGAC	32
160	GA_2_VARNA_rev	GCGCCCAATACGTGTAACAGTCGACAGAAGC	31
161	GA_3_BB8_fwd	GGCCTATTGCTGACGTATTGGGCGCTCTTC	30
162	GA_3_BB8_rev	CGGGCTGATAACTCGCTTACAATTTAGGTGGC	32
163	GA_3_E2A_fwd	AAATTGTAAGCGAGTTATCAGCCCGGGCATTC	32

No	Primer name	Sequence	bp
164	GA_3_E2A_rev	ATTAACACCTCGGATATAGGCGCGCCTACAG	31
165	GA_3_VARNA_fwd	GCGCGCCTATATCCGAGGTGTTAATTAACGAC	32
166	GA_3_VARNA_rev	TCCTAGGACAATGGTAACAGTCGACAGAAGC	31
167	GA_3_orf67_fwd	GTCGACTGTTACCATTGTCCTAGGACCAGC	30
168	GA_3_orf67_rev	GCGCCCAATACGTCAGCAATAGGCCTTGCGTAC	33
169	GA_4_BB8_fwd	CGCGCCTATATCACGTATTGGGCGCTCTTC	30
170	GA_4_BB8_rev	CCTAGGACAATGTCGCTTACAATTTAGGTGGC	32
171	GA_4_orf67_fwd	AAATTGTAAGCGACATTGTCCTAGGACCAGC	31
172	GA_4_orf67_rev	ATTAACACCTCGCAGCAATAGGCCTTGCGTAC	32
173	GA_4_VARNA_fwd	AGGCCTATTGCTGCGAGGTGTTAATTAACGAC	32
174	GA_4_VARNA_rev	CCGGGCTGATAACGTAACAGTCGACAGAAGC	31
175	GA_4_E2A_fwd	GTCGACTGTTACGTTATCAGCCCGGGCATTC	31
176	GA_4_E2A_rev	GCGCCCAATACGTGATATAGGCGCGCCTACAG	32
177	GA_5_BB8_fwd	CGCGCCTATATCACGTATTGGGCGCTCTTC	30
178	GA_5_BB8_rev	ATTAACACCTCGTCGCTTACAATTTAGGTGGC	32
179	GA_5_VARNA_fwd	AAATTGTAAGCGACGAGGTGTTAATTAACGAC	32
180	GA_5_VARNA_rev	TCCTAGGACAATGGTAACAGTCGACAGAAGC	31
181	GA_5_orf67_fwd	GTCGACTGTTACCATTGTCCTAGGACCAGC	30
182	GA_5_orf67_rev	CGGGCTGATAACCAGCAATAGGCCTTGCGTAC	32
183	GA_5_E2A_fwd	AGGCCTATTGCTGGTTATCAGCCCGGGCATTC	32
184	GA_5_E2A_rev	GCGCCCAATACGTGATATAGGCGCGCCTACAG	32
185	GA_6_BB8_fwd	CGCGCCTATATCACGTATTGGGCGCTCTTC	30
186	GA_6_BB8_rev	GGCCTATTGCTGTCGCTTACAATTTAGGTGGC	32
187	GA_6_rOrf67_fwd	AAATTGTAAGCGACAGCAATAGGCCTTGCGTAC	33
188	GA_6_rOrf67_rev	ATTAACACCTCGCATTGTCCTAGGACCAGC	30
189	GA_6_VARNA_fwd	TCCTAGGACAATGCGAGGTGTTAATTAACGAC	32
190	GA_6_VARNA_rev	CCGGGCTGATAACGTAACAGTCGACAGAAGC	31
191	GA_6_E2A_fwd	GTCGACTGTTACGTTATCAGCCCGGGCATTC	31
192	GA_6_E2A_rev	GCGCCCAATACGTGATATAGGCGCGCCTACAG	32
193	GA_7_BB8_fwd	CCTAGGACAATGACGTATTGGGCGCTCTTC	30

No	Primer name	Sequence	bp
194	GA_7_BB8_rev	CGGGCTGATAACTCGCTTACAATTTAGGTGGC	32
195	GA_7_E2A_fwd	AAATTGTAAGCGAGTTATCAGCCCGGGCATTC	32
196	GA_7_E2A_rev	ATTAACACCTCGGATATAGGCGCGCCTACAG	31
197	GA_7_VARNA_fwd	GCGCGCCTATATCCGAGGTGTTAATTAACGAC	32
198	GA_7_VARNA_rev	AGGCCTATTGCTGGTAACAGTCGACAGAAGC	31
199	GA_7_rOrf67_fwd	GTCGACTGTTACCAGCAATAGGCCTTGCGTAC	32
200	GA_7_rOrf67_rev	GCGCCCAATACGTCATTGTCCTAGGACCAGC	31
201	GA_8_BB8_fwd	CCTAGGACAATGACGTATTGGGCGCTCTTC	30
202	GA_8_BB8_rev	ATTAACACCTCGTCGCTTACAATTTAGGTGGC	32
203	GA_8_VARNA_fwd	AAATTGTAAGCGACGAGGTGTTAATTAACGAC	32
204	GA_8_VARNA_rev	CCGGGCTGATAACGTAACAGTCGACAGAAGC	31
205	GA_8_E2A_fwd	GTCGACTGTTACGTTATCAGCCCGGGCATTC	31
206	GA_8_E2A_rev	AGGCCTATTGCTGGATATAGGCGCGCCTACAG	32
207	GA_8_rOrf67_fwd	CGCGCCTATATCCAGCAATAGGCCTTGCGTAC	32
208	GA_8_rOrf67_rev	GCGCCCAATACGTCATTGTCCTAGGACCAGC	31
209	GA_9_BB8_fwd	GTCGACTGTTACACGTATTGGGCGCTCTTC	30
210	GA_9_BB8_rev	CGGGCTGATAACTCGCTTACAATTTAGGTGGC	32
211	GA_9_E2A_fwd	AAATTGTAAGCGAGTTATCAGCCCGGGCATTC	32
212	GA_9_E2A_rev	AGGCCTATTGCTGGATATAGGCGCGCCTACAG	32
213	GA_9_rOrf67_fwd	CGCGCCTATATCCAGCAATAGGCCTTGCGTAC	32
214	GA_9_rOrf67_rev	ATTAACACCTCGCATTGTCCTAGGACCAGC	30
215	GA_9_VARNA_fwd	TCCTAGGACAATGCGAGGTGTTAATTAACGAC	32
216	GA_9_VARNA_rev	GCGCCCAATACGTGTAACAGTCGACAGAAGC	31
217	GA_10_BB8_fwd	CGGGCTGATAACACGTATTGGGCGCTCTTC	30
218	GA_10_BB8_rev	CCTAGGACAATGTCGCTTACAATTTAGGTGGC	32
219	GA_10_orf67_fwd	AAATTGTAAGCGACATTGTCCTAGGACCAGC	31
220	GA_10_orf67_rev	ATTAACACCTCGCAGCAATAGGCCTTGCGTAC	32
221	GA_10_VARNA_fwd	AGGCCTATTGCTGCGAGGTGTTAATTAACGAC	32
222	GA_10_VARNA_rev	GCGCGCCTATATCGTAACAGTCGACAGAAG	31

No	Primer name	Sequence	bp
223	GA_10_rE2A_fwd	CGTCGACTGTTACGATATAGGCGCGCCTACAG	31
224	GA_10_rE2A_rev	GCGCCCAATACGTGTTATCAGCCCGGGCATTC	32
225	GA_11_BB8_fwd	GGCCTATTGCTGACGTATTGGGCGCTCTTC	30
226	GA_11_BB8_rev	CGCGCCTATATCTCGCTTACAATTTAGGTGGC	32
227	GA_11_rE2A_fwd	AAATTGTAAGCGAGATATAGGCGCGCCTACAG	32
228	GA_11_rE2A_rev	ATTAACACCTCGGTTATCAGCCCGGGCATTC	31
229	GA_11_VARNA_fwd	CCGGGCTGATAACCGAGGTGTTAATTAACGAC	32
230	GA_11_VARNA_rev	TCCTAGGACAATGGTAACAGTCGACAGAAGC	31
231	GA_11_orf67_fwd	GTCGACTGTTACCATTGTCCTAGGACCAGC	30
232	GA_11_orf67_rev	GCGCCCAATACGTCAGCAATAGGCCTTGCGTAC	33
233	SDM_shortVARNA_Fwd	GGAACTCGTAGGAATTCCCAAGGACATG	28
234	SDM_shortVARNA_Rev	GGATGTACCAACCTCGAGTCAAATACGTAG	30
235	CAREfull_fwd	CTTCAAGTCGACGTCCTGTATTAGAGGTCACGTGAGTG	38
236	CAREfull_rev	CTTTGTAACTAGTCAATAATCAATGGGCCTTACTCACACGGC	42
237	SDM_CAREssential_Fwd	AGGGTCTCCATTTTGAAGCGGGAGGTTGACATTGATTATTGACTAGTTATTAATAG	56
238	SDM_CAREssential_Rev	GCGTGCTCACTCGGGCTTAAATACCCAGGACTAGCTATAGTTCTAGAG	48
239	rbGlobRev_099to101_fwd	TTGTATCAACGTTAGGCCCGGATCTCCATAAG	32
240	rbGlobRev_099to101_rev	CTTAAGGACTAGTTGAATTCACTCCTCAGGTGC	33
241	33K_099to104_fwd	GAGGAGTGAATTCAACTAGTCCTTAAGAGTCAGCGC	36
242	33K_099and102_rev	ATAATCAATGTCAATGACGAAAAGTCCGCGGC	32
243	CMV_099and102_fwd	GGACTTTTCGTCATTGACATTGATTATTGACTAGTTATTAATAG	44
244	CMV_099and102_rev	CCTATAGTCTAGAGTCTTCTATGGAGGTCAAAAC	34
245	E2A_099and102_fwd	CCTCCATAGAAGACTCTAGACTATAGGAGAAGGAATGGCCAGTCGGG	47
246	E2A_099and100_rev	GCAATAAACAAGTTGCATGCGGCCGGCCCGTT	32
247	SV40pA_099and100_fwd	GCCGGCCGCATGCAACTTGTTTATTGCAGCTTATAATG	38
248	SV40pA_099and100_rev	AGGACTAGGGCCCTAAGATACATTGATGAGTTTGG	35
249	E4varnaAndBB_099and100_fwd	ATCAATGTATCTTAGGGCCCTAGTCCTTTCGAATAC	36
250	E4varnaAndBB_099to101_rev	GGAGATCCGGGCCTAACGTTGATACAAGT	41

No	Primer name	Sequence	bp
		ATACTGTGAATG	
251	33K_100_101_103_104_rev	CCTAACTGACACACTGACGAAAAGTCCGCG GC	32
252	SV40prom_100_101_103_104_f wd	GGACTTTTCGTCAGTGTGTCAGTTAGGGTG TG	32
253	SV40prom_100_101_103_104_r ev	CCTATAGTCTAGATTTGCAAAAGCCTAGGC C	31
254	E2A_100_103_fwd	TAGGCTTTTGCAAATCTAGACTATAGGAG AAGGAATGGCCAGTCGGG	47
255	E2AwPA_101_104_fwd	TAGGCTTTTGCAAATCTAGACTATAGGAG AAGG	33
256	E2AwPA_101_104_rev	AAGGACTAGGGCCCACTTTCAATAAAGGC AAATG	34
257	E4varnaAndBB_101_104_fwd	CTTTATTGAAAGTGGGCCCTAGTCCTTTCG AATAC	35
258	SV40pArev_102_103_fwd	AGGACTAGGGCCCAACTTGTTTATTGCAGC TTATAATG	38
259	SV40pArev_102_103_rev	GCCGGCCGCATGCTAAGATACATTGATGAG TTTGG	35
260	E2Arev_102_103_fwd	ATCAATGTATCTTAGCATGCGGCCGGCCCG TT	32
261	rbGlob_102to104_fwd	CTTAAGGACTAGTTGAATTCACTCCTCAGG TG	32
262	rbGlob_102to104_rev	TTGTATCAACGTTGTCGAGGGATCTCCATA AG	32
263	VA_E4_BB_102to104_fwd	GGAGATCCCTCGACAACGTTGATACAAGT ATACTGTGAATG	41
264	VA_E4_BB_102_103_rev	GCAATAAACAAGTTGGGCCCTAGTCCTTTC GAATAC	36
265	Fwd_SDMdelete_ITRstuffer	GGCCGAGTTGAGCGGTTC	18
266	Rev_SDMdelete_ITRstuffer	CCCAGGCCGTTCTATGATTC	20
267	E4_106_fwd	AAATTGTAAGCGACATTGTCCTAGGACCA GC	31
268	E4_106_rev	AGGGAGAGGGGCGCATATGTTTCAATTGC AGAAAATTTC	39
269	IRES_106_fwd	AATTGAAACATATGCGCCCCTCTCCCTCCC CC	32
270	IRES_106_rev	CCCGACTGGCCATGGTATTATCGTGTTTTT CAAAGGAAAACCACGTCCCC	50
271	E2A_106_fwd	AACACGATAATACCATGGCCAGTCGGGAA GAGGAGC	36
272	E2A_106_rev	TTATCGATAAGCTTGCATGCGGCCGGCCCG TT	32
273	WPREhGHpA_106_fwd	GCCGGCCGCATGCAAGCTTATCGATAATCA ACC	33
274	WPREhGHpA_rev	AAGGACTAGGGCCCCACGTGGTTACCTACA AAATC	35
275	VA_BB_106_fwd	AGGTAACCACGTGGGGCCCTAGTCCTTTCG AATAC	35
276	VA_BB_106_rev	GTCCTAGGACAATGTCGCTTACAATTTAGG TGGC	34
277	CMV_107_fwd	AAATTGTAAGCGATTGACATTGATTATTG ACTAGTTATTAATAG	44

No	Primer name	Sequence	bp
278	CMV_107_rev	TAGTCATTGTACAGTCTTCTATGGAGGTCAAAAC	34
279	E4noProm_107_fwd	CCTCCATAGAAGACTGTACAATGACTACGTCC	32
280	E4noProm_107_rev	AGGGAGAGGGGCGCATATGTTTCAATTGCAGAAAATTTC	39
281	IRES_107_fwd	AATTGAAACATATGCGCCCCTCTCCCTCCCCC	32
282	IRES_107_rev	CCCGACTGGCCATGGTATTATCGTGTTTTTCAAAGGAAAACCACGTCCCC	50
283	E2A_107_fwd	AACACGATAATACCATGGCCAGTCGGGAAGAGGAGC	36
284	E2A_107_rev	TTATCGATAAGCTTGCATGCGGCCGGCCCGTT	32
285	WPREhGHpA_107_fwd	GCCGGCCGCATGCAAGCTTATCGATAATCAACC	33
286	WPREhGHpA_107_rev	AAGGACTAGGGCCCCACGTGGTTACCTACAAAATC	35
287	VA_BB_107_fwd	AGGTAACCACGTGGGGCCCTAGTCCTTTCGAATAC	35
288	VA_BB_107_rev	ATAATCAATGTCAATCGCTTACAATTTAGGTGGC	34
289	CMV_108_fwd	AAATTGTAAGCGATTGACATTGATTATTGACTAGTTATTAATAG	44
290	CMV_108_rev	TAGTCATTGTACAGTCTTCTATGGAGGTCAAAAC	34
291	E4noProm_108_fwd	CCTCCATAGAAGACTGTACAATGACTACGTCC	32
292	E4noProm_108_rev	TGCTACCGAATTCCAGAACCCTAGTATTCAAC	32
293	P2Av_fwd	ATACTAGGGTTCTGGAATTCGGTAGCAAGCTTGGTAGTGGAG	42
294	P2A_108_rev	CCCGACTGGCCATCTTAAGGGGGGCGGGGCC	31
295	E2A_108_fwd	CCGCCCCCCTTAAGATGGCCAGTCGGGAAGAGGAGC	36
296	E2A_108_rev	TTATCGATAAGCTTGCATGCGGCCGGCCCGTT	32
297	WPREhGHpA_108_fwd	GCCGGCCGCATGCAAGCTTATCGATAATCAACC	33
298	WPREhGHpA_108_rev	AAGGACTAGGGCCCCACGTGGTTACCTACAAAATC	35
299	VA_BB_108_fwd	AGGTAACCACGTGGGGCCCTAGTCCTTTCGAATAC	35
300	VA_BB_108_rev	ATAATCAATGTCAATCGCTTACAATTTAGGTGGC	34
301	E2AwIntronProm_109_fwd	CTTGTATCAACGTTAGGCGACCGCACCCTGTG	32
302	E2AwIntronProm_109_rev	GAGGGAGAGGGGCGGCATGCGGCCGGCCCGTT	32
303	IRES_109_fwd	GCCGGCCGCATGCCGCCCCTCTCCCTCCCCC	31
304	IRES_109_rev	CCGGACGTAGTCATGGTATTATCGTGTTTTTCAAAGGAAAACCACGTCCCC	51

No	Primer name	Sequence	bp
305	E4noProm_109_fwd	ACACGATAATACCATGACTACGTCCGGCGTTC	32
306	E4noProm_109_rev	TATCGATAAGCTTCATATGTTTCAATTGCAGAAAATTTCAAGTC	44
307	WPREhGHpA_109_fwd	AATTGAAACATATGAAGCTTATCGATAATCAACC	34
308	WPREhGHpA_109_rev	CAACGCTAGGATCCCACGTGGTTACCTACAAAATC	35
309	VA_BB_109_fwd	AGGTAACCACGTGGGATCCTAGCGTTGCGGC	31
310	VA_BB_109_rev	GGTGCGGTCGCCTAACGTTGATACAAGTATACTGTGAATGC	41
311	CMV_110_fwd	TTGTATCAACGTTTTGACATTGATTATTGACTAGTTATTAATAG	44
312	CMV_110_rev	CCTATAGTCTAGAGTCTTCTATGGAGGTCAAAAC	34
313	E2A_110_fwd	CCTCCATAGAAGACTCTAGACTATAGGAGAAGGAATGGCCAGTCGGG	47
314	E2A_110_rev	GAGGGAGAGGGGCGGCATGCGGCCGGCCCGTT	32
315	IRES_110_fwd	GCCGGCCGCATGCCGCCCCTCTCCCTCCCCC	31
316	IRES_110_rev	CCGGACGTAGTCATGGTATTATCGTGTTTTTCAAAGGAAAACCACGTCCCC	51
317	E4noProm_110_fwd	ACACGATAATACCATGACTACGTCCGGCGTTC	32
318	E4noProm_110_rev	TATCGATAAGCTTCATATGTTTCAATTGCAGAAAATTTCAAGTC	44
319	WPREhGHpA_110_fwd	AATTGAAACATATGAAGCTTATCGATAATCAACC	34
320	WPREhGHpA_110_rev	CAACGCTAGGATCCCACGTGGTTACCTACAAAATC	35
321	VA_BB_110_fwd	AGGTAACCACGTGGGATCCTAGCGTTGCGGC	31
322	VA_BB_110_rev	ATAATCAATGTCAAAACGTTGATACAAGTATACTGTGAATGC	42
323	CMV_111_fwd	TTGTATCAACGTTTTGACATTGATTATTGACTAGTTATTAATAG	44
324	CMV_111_rev	CCTATAGTCTAGAGTCTTCTATGGAGGTCAAAAC	34
325	E2A_111_fwd	CCTCCATAGAAGACTCTAGACTATAGGAGAAGG	33
326	E2A_111_rev	TTGCTACCGAATTCAAAATCAAAGGGGTTCTG	32
327	P2A_111_fwd	CCCCTTTGATTTTGAATTCGGTAGCAAGCTTGGTAGTGGAG	41
328	P2A_111_rev	CCGGACGTAGTCATCTTAAGGGGGGCGGGGCC	32
329	E4noProm_111_fwd	CGCCCCCCTTAAGATGACTACGTCCGGCGTTC	32
330	E4noProm_111_rev	TATCGATAAGCTTCATATGTTTCAATTGCAGAAAATTTCAAGTC	44
331	WPREhGHpA_111_fwd	AATTGAAACATATGAAGCTTATCGATAATCAACC	34

No	Primer name	Sequence	bp
332	WPREhGHpA_111_rev	CAACGCTAGGATCCCACGTGGTTACCTACAAAATC	35
333	VA_BB_111_fwd	AGGTAACCACGTGGGATCCTAGCGTTGCGGC	31
334	VA_BB_111_rev	ATAATCAATGTCAAAACGTTGATACAAGTATACTGTGAATGC	42
335	SDM_delOrf4_Fwd	ATGACTACGTCCGGCGTT	18
336	SDM_delOrf4_Rev	GTTTTTTTTTTTATTCCAAAAGATTATCCAAAACC	35
337	114VA_BB_L4_CMV_E2A_fwd	AGTACTTCGCAAGCGTACGAGTCCACGGACC	31
338	114VA_BB_L4_CMV_E2A_rev	CTTATCATGTCTGAGGCCTATTGCTGGATATAGG	34
339	114SV40rev_fwd	AGCAATAGGCCTCAGACATGATAAGATACATTGATG	36
340	114SV40rev_rev	TTGAAACATATGCAACTTGTTTATTGCAGC	30
341	114CMV_orf6/7_rev_fwd	AATAAACAAGTTGCATATGTTTCAATTGCAGAAAATTTC	39
342	114CMV_orf6/7_rev_rev	TGGACTCGTACGCTTGCGAAGTACTTTGTAC	31
343	E1fpGFP_CMV_BB_fwd	AGGTACTGAAGATGAGTTTGGACAAACCACAACTAGAATGCAG	43
344	E1fpGFP_CMV_BB_rev	TATGTCTCATGGTGGCGACCGGTAGCGC	28
345	E1full_fwd	GGTCGCCACCATGAGACATATTATCTGCC	29
346	E1full_rev	CAAACTCATCTTCAGTACCTCAATCTGTATC	31
347	E1ApGFP_CMV_BB_fwd	GCTGAATGAGGATGAGTTTGGACAAACCACAACTAGAATGCAG	43
348	E1A_rev	CAAACTCATCCTCATTCAGCAAACAAAG	28
349	TPcore_fwd18.1	TGTACAAATTTCCGGGCGGGAACGGGACTTT	31
350	TPcore_rev18.1	CCGGCATGCTAGCGTCTTCTATGGAGGCTACGTCC	35
351	Rep_fwd18.1	TCCATAGAAGACGCTAGCATGCCGGGGTTTTAC	33
352	Rep_rev18.1	TCCACGCCCATTGCTTCGCAGAGACCAAAGTTCAAC	36
353	CRM1_fwd18.1	GTCTCTGCGAAGCAATGGGCGTGGATAGCG	30
354	CRM1_rev18.1	GCTTCCCATATGGTCTTCTATGGAGGTCAAAACAG	35
355	Cap_fwd18.1	CTCCATAGAAGACCATATGGGAAGCTTCGATC	32
356	Cap_rev18.1	CATATGACATCATCTTCGCAGAGACCAAAGTTC	33
357	TP14_fwd18.1	GTCTCTGCGAAGATGATGTCATATGACTCATTG	33
358	TP14_rev18.1	GCTCCATCCCGGGGTCTTCTATGGAGGTCAAAAC	34
359	smallRep_BB_fwd18.1	TCCATAGAAGACCCCGGGATGGAGCTGGTC	30
360	smallRep_BB_rev18.1	CCGTTCCCGCCCGGAAATTTGTACACTTACAATTTAGGTGGCAC	44
361	CRM1_fwd18.2	TGTACAAATTTCCCAATGGGCGTGGATAGC	31

No	Primer name	Sequence	bp
362	Cap_rev18.2	GCCGTTCCCGCCCCTTCGCAGAGACCAAAGTTC	32
363	TPcore_fwd18.2	GGTCTCTGCGAAGGGGCGGGAACGGGACTTT	31
364	Rep_rev18.2	CATATGACATCATCTTCGCAGAGACCAAAGTTCAAC	36
365	smallRep_BB_rev18.2	CCACGCCCATTGGGAAATTTGTACACTTACAATTTAGGTGGCAC	44
366	Cap_rev18.3	TGGCACTTTTCGGCTTCGCAGAGACCAAAGTTC	33
367	p5remRep_fwd18.3	GTCTCTGCGAAGCCGAAAAGTGCCACCTAAATTG	34
368	Cap_rev18.4	GGTCTCTGCGAAGCTTCGCAGAGACCAAAGTTC	33
369	Repp5remR_fwd18.4	GTCTCTGCGAAGCTTCGCAGAGACCAAAGTTC	32
370	Repp5remR_rev18.4	CATATGACATCATCCGAAAAGTGCCACCTAAATTG	35
371	TP14_fwd18.4	GGCACTTTTCGGATGATGTCATATGACTCATTG	33
372	TPcore_fwd19.1	GTACAACATTTCCGGGCGGGAACGGGACTTT	31
373	Cap_rev19.1	GGACCTCCAATTGCTTCGCAGAGACCAAAGTTC	33
374	p5_BB_fwd19.1	GTCTCTGCGAAGCAATTGGAGGTCCTGTATTAG	33
375	p5_BB_rev19.1	CCGTTCCCGCCCGGAAATGTTGTACACTTACAATTTAG	38
376	Rep_rev19.2	GTCTCTGCGAAGCTTCGCAGAGACCAAAGTTCAAC	35
377	revCRM1_rev19.2	GGACCTCCAATTGCAATGGGCGTGGATAGCG	31
378	p5_BB_fwd19.2	CCACGCCCATTGCAATTGGAGGTCCTGTATTAG	33
379	117E4_BB_rE2A_fwd	ACCTCCATAGAAGACTGTACATCATGATTCGCTG	34
380	117E4_BB_rE2A_rev	ACCTCCATAGAAGACTCTAGACTATAGGAGAAGG	34
381	117rTP17_fwd	CTCCTATAGTCTAGAGTCTTCTATGGAGGTCAAAACAGC	39
382	117rTP17_rev	GGACTTTTCGTCATTTTTGGCGCCAAATGCGTG	33
383	117L4-VA_fwd	GCATTTGGCGCCAAAAATGACGAAAAGTCCGCG	33
384	117L4-VA_rev	GGTAATAGCGATGACTGAGTACTTCGCAAGCGTAC	35
385	117CRM1+2_fwd	CTTGCGAAGTACTCAGTCATCGCTATTACCATG	33
386	117CRM1+2_rev	GAATCATGATGTACAGTCTTCTATGGAGGTCAAAAC	36
387	118rCRM1+4_fwd	CTCCTATAGTCTAGAGTCTTCTATGGAGGTCAAAAC	36

No	Primer name	Sequence	bp
388	118rCRM1+4_rev	GGACTTTTCGTCATTCATATGCCAAGTACGCCC	33
389	118L4-VA_fwd	CGTACTTGGCATATGAATGACGAAAAGTCCGCG	33
390	118L4-VA_rev	AACGCGGAACTCCATTGAGTACTTCGCAAGCGTAC	35
391	118CMV4_01_fwd	CTTGCGAAGTACTCAATGGAGTTCCGCGTTACATAAC	37
392	20.1smallRep_p5BB_p5remRep_fwd	ACCTCCATAGAAGACCCCGGGATGGAGCTGGTC	33
393	20.1smallRep_p5BB_p5remRep_rev	TATCCACGCCCATTGCTTCGCAGAGACCAAAGTTCAACTG	40
394	20.1CRM1_Cap_fwd	TTGGTCTCTGCGAAGCAATGGGCGTGGATAGCG	33
395	20.1CRM1_Cap_rev	GTAGGGGATTCCCCTCTTCGCAGAGACCAAAGTTC	35
396	20.1TP11_fwd	TTGGTCTCTGCGAAGAGGGGAATCCCCTACCGTTAAC	37
397	20.1TP11_rev	CAGCTCCATCCCGGGGTCTTCTATGGAGGTCAAAACAGC	39
398	20.2CRM1_Cap_rev	GAACTCCATATATGGCTTCGCAGAGACCAAAGTTC	35
399	20.2CMV2_01_fwd	TTGGTCTCTGCGAAGCCATATATGGAGTTCCGC	33
400	20.2CMV2_01_rev	CAGCTCCATCCCGGGGTCTTCTATGGAGGTCAAAAC	36
401	20.3smallRep_p5BB_p5remRep_rev	CATCACGCACACGCACTTCGCAGAGACCAAAGTTCAACTG	40
402	20.3TP19_fwd	TTGGTCTCTGCGAAGTGCGTGTGCGTGATGATG	33
403	20.3TP19_rev	GAAGCTTCCCATATGGTCTTCTATGGAGGTCAAAACAG	38
404	20.3Cap_fwd	ACCTCCATAGAAGACCATATGGGAAGCTTCGATC	34
405	20.3Cap_rev	GTAGGGGATTCCCCTCAATTGCTTCGCAGAGAC	33
406	20.3TP11_fwd	TCTGCGAAGCAATTGAGGGGAATCCCCTACCGTTAAC	37
407	20.4Cap_rev	GAACTCCATATATGGCAATTGCTTCGCAGAGAC	33
408	20.4CMV2_01_fwd	TCTGCGAAGCAATTGCCATATATGGAGTTCCGC	33
409	21.1smallRep_p5BB_rev	ACGCAATGAGTCATTCTTACAATTTAGGTGGCACTTTTCGGG	42
410	21.1TP16_fwd	CACCTAAATTGTAAGAATGACTCATTGCGTGTTTG	35
411	21.1TP16_rev	GTAAAACCCCGGCATGTCTTCTATGGAGGTCAAAAC	36
412	21.1Rep_CRM1_Cap_fwd	ACCTCCATAGAAGACATGCCGGGGTTTTACGAG	33
413	21.3Rep_rev	CATCACGCACACGCACTTCGCAGAGACCAAAGTTC	35
414	22.2smallRep_p5BB_rev	GTAGGGGATTCCCCTCTTACAATTTAGGTG	42

No	Primer name	Sequence	bp
415	22.2TP11tp_fwd	GCACTTTTCGGG CACCTAAATTGTAAGAGGGGAATCCCCTAC CGTTAAC	37
416	22.2TP11tp_rev	GTAAAACCCCGGCATGTCTTCTATGGAGGC TACGTC	36
417	22.2Rep_CRM1_Cap_fwd	GCCTCCATAGAAGACATGCCGGGGTTTTAC GAG	33
418	BB20_1fix_rev	GAAGAGTATGAGCCATATTCAACGGGAAA CG	31
419	BB20_1fix_fwd	CGTTTCCCGTTGAATATGGCTCATACTCTT C	31
420	RepCRM1CapFix_fwd	CAGAAAAAGACTCAATTTTGGTCAGACTG	29
421	RepCRM1CapFix_rev	CAGTCTGACCAAAATTGAGTCTTTTTCTG	29

8.3.8 Plasmids

Table 18: Helper plasmids used and created.

Plasmid	Sequence ID	Content Name	Length / bp
0.1	pADΔF6	Parental Plasmid	15770
0.2	pRGX005	Promoter less Helper	8012
1.0	pRGX010	12kb Helper	12060
1.1	pRGX009	9kb Helper	9060
pE2A	pRGX011	pE2A	7306
	pRGX012	pE2Acmv	4471
pE4orf6	pRGX015	pE4orf6	3320
	pRGX016	pE4orf6cmv	3565
pVA	pRGX022	pVARNAonly	3004
	pRGX072	pVAIonly	2551
	pRGX023	E4whole_E4prom	5135
	pRGX024	E4whole_CMV	5382
	pRGX025	E4orf2plus_E4prom	4701
	pRGX026	E4orf2plus_CMV	4947
	pRGX027	E4orf3plus_E4prom	4314
	pRGX028	E4orf3plus_CMV	4560
	pRGX029	E4orf4plus_E4prom	3952
	pRGX030	E4orf4plus_CMV	4198
	pRGX031	E4orf67plus_E4prom	3686
	pRGX032	E4orf67plus_CMV	3932
	pRGX035	010Orf2plusE4prom	11626
	pRGX037	010Orf3plusE4prom	11239
	pRGX039	010Orf4plusE4prom	10877
2.0	pRGX041	010Orf67E4prom	10611
	pRGX046	010orf6	10240
	pRGX049	010_33Kintron	10230

Plasmid	Sequence ID	Content Name	Length / bp
	pRGX050	P33Konly	3584
	pRGX051	33KmutATG049	20230
1.2	pRGX056	12kb_E2aCMV	12314
1.3	pRGX057	9kbHelper009_E2aCMV	9258
2.2	pRGX060	Orf67_CMVe2a	10865
2.3	pRGX061	Orf67_doubleCMV	11106
2.4	pRGX062	Orf6_7_E2aCMV	7783
	pRGX066	SDMdeletedBBSpeIof052	8141
	pRGX065	010Orf6_extendetpolyA	10332
2.5	pRGX067	Orf67_E2aCMV_orf67CMV	8061
	pRGX068	VARNAtoVAIexchange066	7274
	pRGX071	Orf67_E2aCMV_E2aMCS	7807
	pRGX073	Orf67_BB8	9907
3.0	pRGX076	E4-E2A-VA	10133
3.2	pRGX077	E2A-E4-VA	10133
3.6	pRGX078	E2A-VA-E4	10133
3.4	pRGX079	E4-VA-E2A	10133
3.3	pRGX080	VA-E4-E2A	10133
3.5	pRGX081	E4rev-VA-E2A	10133
3.1	pRGX082	E2A-VA-E4rev	10133
3.9	pRGX083	VA-E2A-E4rev	10133
3.10	pRGX084	E2A-E4rev-VA	10133
3.7	pRGX085	E4-VA-E2Arev	10133
3.8	pRGX086	E2Arev-VA-E4	10133
-	pRGX087	10withL5Fiber	13162
-	pRGX088	041Orf67withL5Fiber	14611
2.1	pRGX089	12kbCMVorf67	10857
3.0 E2Amin1	pRGX091	orf67_E2A33Kintron_VARNA	8303
3.2 E2Amin1	pRGX092	E2A33Kintron_orf67_VARNA	8303
3.10 E2Amin1	pRGX093	E2A33Kintron_revorf67_VARNA	8903
3.8 E2Amin1	pRGX094	revE2A33Kintron_orf67_VARNA	8903
3.10 E2Amin1	pRGX095	E2A33Kintron_revorf67_VAshort	7901
3.8 E2Amin1	pRGX096	revE2A33Kintron_VAshort_orf67	7901
3.10.1	pRGX099	33Krev_CMV_E2A_SV40pA_orf67rev_VA	8698
3.10.2	pRGX100	33Krev_SV40_E2A_SV40pA_orf67rev_VA	8372
3.10.3	pRGX101	33Krev_SV40_E2A_orf67rev_VA	8326
3.8.1	pRGX102	SV40pArev_E2Arev_CMVrev_33K_VA_orf67	8698
3.8.2	pRGX103	E2Arev_SV40pArev_SV40rev_33K_VA_orf67	8372
3.0 IRES	pRGX106	orf67_IRES_E2A_WPREhGH_VA	8075
3.0cmvIRES	pRGX107	CMV_orf67_IRES_E2A_WPREhGH_VA	8267

Plasmid	Sequence ID	Content Name	Length / bp
3.0cmvP2A	pRGX108	CMV_orf67noStop_P2A_E2A_WPREhGH_VA	7685
3.2cmvIRES	pRGX110	CMV_E2A_IRES_orf67_WPREhGH_VA	8321
3.2cmvP2A	pRGX111	CMV_E2AnoStop_P2A_orf67_WPREhGH_VA	7820
4.0	pRGX114	084doubleCMVSV40pA	8951
	pRGX115	Orf3Orf6Orf67	10974
	pRGX116	114E4prom	8705
4.1	pRGX117	rE2A_rTP17_L4VA_CRM1n2_E4	8785
4.2	pRGX118	rE2A_rCRM1n4_L4VA_CMV4_01_E4	8812
4.3	pRGX119	rE2A_rCRM1n4_L4VA_CRM1n2_E4	8638
4.4	pRGX120	rTP17_L4VA_CMV4_01_E4	8959
	pRGX121	115n_orf3_6_67_086origin	10496

Table 19: Rep/Cap and other plasmids used and created.

Plasmid	Sequence ID	Content changes	Length / bp
v2.1	pAAV8	-	7291
v3.7	pAAV8v13.6	polyA, CMVshort-intron, p5	7176
v5.5	pAAV8v16.9	mutp40pACMVshortp5smallRep5	8723
	RepCap_v18.1		8932
	RepCap_v18.2		8932
	RepCap_v18.3		8837
	RepCap_v18.4		8837
	RepCap_v19.1		7111
	RepCap_v19.2		7111
	v20.1	p5rem_Rep_CRM1_Cap_TP11_smallRep_p5	8840
	v20.2	p5rem_Rep_CRM1_Cap_CMV2_01_smallRep_p5	8951
	v20.3	p5rem_Rep_TP19_Cap_TP11_smallRep_p5	9149
	v20.4	p5rem_Rep_TP19_Cap_CMV2_01_smallRep_p5	9260
	v21.1	TP16_Rep_CRM1_Cap_TP11_smallRep_p5	9177
	v21.2	TP16_Rep_CRM1_Cap_CMV2_01_smallRep_p5	9288
	v21.3	TP16_Rep_TP19_Cap_TP11_smallRep_p5	9486
	v21.4	TP16_Rep_TP19_Cap_CMV2_01_smallRep_p5	9597
	v22.2	TP11tp_Rep_CRM1_Cap_CMV2_01_smallRep_p5	9327

Plasmid	Sequence ID	Content changes	Length / bp
pE1	pE1full_CMVsv40pA		5579
pE1A	pE1A_CMVsv40pA		3658
ITR/GOI			5855
pGFP	pGFP-hCMV-IE		3299

8.3.9 Synthesised Genes and Gene Fragments

Table 20: Ordered gene syntheses.

ID	Synthesis name	Sequence	Length / bp
GS1	polyA_E4	GAAGCTCGAGTCGGTTGCATATGAATAAACACGTTGAAACATAACACAAACGATTCTTTATTCTTGGGCAATGTATGAAAAAGTGTAAGAGGATGTGGTCTAGAGCTTGGTCAACGTCGA	120
GS2	polyA_E2a	GAAGCTCGAGTCGGTTGCATGCCGGCGCAGACGGCAAGGGTGGGGGTAAATAATCACCCGAGAGTGTACAAATAAAAGCATTTGCCTTTATTGAAAGTGGGCCCGCTTGGTCAACGTCGA	120
GS3	p5remnant	GAAGCTCGAGTCGGTTGTACAACATTTCCCCGAAAAGTGCCACCTAAATTGTAACGCAGGGTCTCCATTTTGAAGCGGGAGGTTTGAACGCGCAGCCGCCGCTAGCTTGGTCAACGTCGA	120
GS4	E4_promoter	TGATCAATCATCAATAATATACCTTATTTTGGATTGAAGCCAATATGATAATGAGGGGGTGGAGTTTGTGACGTGGCGCGGGGCGTGGGAACGGGGCGGGTGACGTAGGTTTTAGGGCGGAGTAACTTGTATGTGTTGGGAATTGTAGTTTTCTTAAAATGGGAAGTGACGTAACGTGGGAAAACGGAAGTGACGATTTGAGGAAGTTGTGGGTTTTTTGGCTTTCGTTTCTGGGCGTAGGTTCGCGTGCGGTTTTCTGGGTGTTTTTTGTGGACTTTAACCGTTACGTCATTTTTTAGTCCTATATATACTCGCTCTGCACTTGGCCCTTTTTTACACTGTGACTGATTGAGCTGGTGCCGTGTCGAGTGGTGTTTTTTTAATAGGTTTTCTTTTTTACTGGTAAGGCTGA	423

GS5	p40-Cap_mutated	CTGTTTGTACA GGTCACCAAGCAGGAAGTCAAAGACTTTTTCCGG TGGGCAAAGGATCACGTGGTTGAGGTGGAGCATG AATTCTACGTCAAAAAGGGTGGAGCCAAGAAAAG ACCCGCCCCCAGTGACGCAGACATCAGTGAGCCCA AACGGGTGCGCGAGAGCGTTGCGCAGCCATCGACG TCAGACGCGGAAGCTTCGATCAACTACGCAGACA GGTACCAAAACAAATGTTCTCGTCACGTGGGCAT GAATCTGATGCTGTTTCCCTGCAGACAATGCGAG AGAATGAATCAGAATTCAAATATCTGCTTCACTC ACGGACAGAAAGACTGTTTAGAGTGCTTTCCCGT GTCAGAATCTCAACCCGTTTCTGTCGTCAAAAAGG CGTATCAGAAACTGTGCTACATTCATCATATCAT GGGAAAGGTGCCAGACGCTTGCACTGCCTGCGATC TGGTCAATGTGGATTTGGATGACTGCATCTTTGA ACAATAAATGACTTAAACCAGGTCTTGCTGCCGA TGGTTATCTTCCAGATTGGCTCGAGGACAACCTTA GTACGCGT	557
GS6	polyA_E2a_mutBsrGI	GAAGCTCGAGTCGGTTGCATGCCGGCGCAGACGGC AAGGGTGGGGGTAAATAATCACCCGAGAGTGTTC AAATAAAAGCATTTGCCTTTATTGAAAGTGGGCC CGCTTGGTCAACGTCGA	120
GS7	MSC_E2aProm	GAAGCTCGAGTCGGTTGGAACGCCCGGGCTCTAGC GTCAGTATACCCTGAAGTCATAACGTTCTTCAGTA AGACTAGTCTGGCTTGGTCAACGTCGA	97
GS8	p40-Cap_mutate-polyA	GGTCACCAAGCAGGAAGTCAAAGACTTTTTCCGG TGGGCAAAGGATCACGTGGTTGAGGTGGAGCATG AATTCTACGTCAAAAAGGGTGGAGCCAAGAAAAG ACCCGCCCCCAGTGACGCAGACATCAGTGAGCCCA AACGGGTGCGCGAGAGCGTTGCGCAGCCATCGACG TCAGACGCGGAAGCTTCGATCAACTACGCAGACA GGTACCAAAACAAATGTTCTCGTCACGTGGGCAT GAATCTGATGCTGTTTCCCTGCAGACAATGCGAG AGAATGAATCAGAATTCAAATATCTGCTTCACTC ACGGACAGAAAGACTGTTTAGAGTGCTTTCCCGT GTCAGAATCTCAACCCGTTTCTGTCGTCAAAAAGG CGTATCAGAAACTGTGCTACATTCATCATATCAT GGGAAAGGTGCCAGACGCTTGCACTGCCTGCGATC TGGTCAATGTGGATTTGGATGACTGCATCTTTGA ACAATAAATGATTTAAATCAGGTCTTGCTGCCGA TGGTTATCTTCCAGATTGGCTCGAGGACAACCTCT CTGAAGTACTAGCTCAACGCGTTTGCCTGTTAATC	649

	AATAAACCGGTTGATTCGTTTCAGTTGAACTTTG
	GTCTCTGCGAAGACTAGTTGAGGCTCATATG

8.3.10 Cultivated Cells

Table 21: Bacterial strains.

Species	Name	Manufacturer
Escherichia coli	DH5α	Invitrogen
Escherichia coli	NEBstable	NEB

Table 22: Mammalian cell lines.

Cell line	Cultivation medium
HEK293 SKMB	Dynamis
HEK 4B9-11A4	Dynamis
293F	Freestyle
Expi293F	Expi
CHO K1	CD CHO

8.3.11 Culture Media

8.3.11.1 Bacterial Culture Media

Table 23: Bacterial cultivation media.

Name	Component	Concentration	Manufacturer
LB-broth, Miller	Tryptone	10 g/L	BD Life Science
Medium	Yeast extract	5 g/L	
	NaCl	5 g/L	
	pH	7.0 ±0.2 g/L	
LB-Agar, Miller	LB-Medium	20 g/L	Fisher Scientific
	Agar-Agar	1.5 % (w/v)	
SOC-Medium	Peptone	2 % (w/v)	NEB/MP Biomedicals
	Yeast extract	0.5 % (w/v)	
	Glucose	20 mM	
	NaCl	10 mM	
	KCl	2.5 mM	
	MgCl2	10 mM	
	MgSO4	10 mM	

8.3.11.2 Mammalian Cell Culture Media

Table 24: Mammalian cell culture media.

Name	Additional components	Concentration	Manufacturer
Dynamis	L-Glutamine	6 mM	Gibco
CD CHO	L-Glutamine	8 mM	Gibco
Freestyle 293	-	-	Gibco
Expi293	-	-	Gibco

Table 25: Cell culture media additives and transfection chemicals.

Name	Manufacturer	Catalogue-ID
Anti-Clumping Agent	Gibco	0010057AE
L-Glutamine (200 mM)	Gibco	25030081
NaCl (150 mM	Polyplus-Transfection	201000002.
PEImax (Polyethyleneimine "MAX", MW 40000)	Polyscience Inc.	9002-98-6
PEIpro	Polyplus-Transfection	101000026

8.3.12 Software

Table 26: Software tools used.

Name	Version	Developer
SnapGene	6.1	Insightful Science
GraphPad Prism	10.0	GraphPad Software
Microsoft Office 365	2211	Microsoft
Mendeley Desktop	1.19.8	Mendeley Ltd
Genomatix	MatInspector 8.4 / MatBase 11.2	Intrexon Bioinformatics

8.4 List of Abbreviations

Table 27: Abbreviations used in this work.

Abbreviation	Meaning
AAP	Assembly activating protein
AAV	Adeno-associated virus
AdV	Adenovirus
ASGCT	American society of gene and cell therapy
ATCC	American Type Culture Collection
ATP	Adenosine triphosphate
Bp	Base pair
CAR T	Chimeric antigen receptor T cells
CARE	*Cis*-acting replication element
Cas	CRISPR associated protein
Cat. No.	Catalogue number
CDS	Coding sequence
cGMP	Current Good Manufacturing Practice
CGT	Cell and gene therapy
CMV	Human Cytomegalovirus intermediate early promoter
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
CRM	*Cis*-regulatory modules
DMSO	Dimethyl sulfoxide
DNA	Desoxyribonuclein acid

Abbreviation	Meaning
DOE	Design of experiment
dsDNA	Double stranded DNA
E. coli	Escherichia coli
E2E	E2 early promoter
E2L	E2 late promoter
E4P	E4 promoter
EDTA	Ethylenediaminetetraacetic acid
GOI	Gene of interest
GT	Gene therapy
HCP	Host cell protein
HEK293	Human embryonic kidney 293 cells
hGH	Human growth hormone
HPW	Highly purified water
HSV	Herpes simplex virus
ITR	Inverted terminal repeat
Kb	Kilobase (1000 base pairs)
kDa	Kilodalton
L4P	L4 promoter
LB	Luria-Bertani
MAAP	Membrane associated accessory protein
mAb	Monoclonal Antibody
MLP	Major late promoter
MW	Molecular weight
n/a	Not available
NLS	Nuclear localisation sequence
NPC	Nuclear pore complex
nt	Nucleotide
OD	Optical density
ORF	Open reading frame
PBS	Phosphate buffered saline
PCR	Polymerase chain reaction
PEI	Polyethylenimin
polyA / pA	Polyadenylation signal
PTM	Post translational modification
qPCR	Quantitative real time PCR
rAAV	Recombinant AAV
RBD	Receptor binding domain

Abbreviation	Meaning
RBE/RBS	Rep binding element / site
rbGlob	Rabbit beta-globin polyadenylation signal
Rep	AAV replicase
RNA	Ribonucleic acid
SDS	Sodiumdodecylsulfate
SMA	Spinal muscular atrophy
SOC	Super optimal broth with catabolite repression
ssDNA	Single stranded DNA
SV40	Simian virus 40
TAE	Tris acetate EDTA buffer
TALEN	Transcription activator-like effector nuclease
TF	Transcription factor
TFRE	Transcription factor regulatory element
Tm	Melting temperature
Tris	Tris-(hydroxymethyl)-aminomethane
TSS	Transcription start site
U	Units
V	Volume
v/v	Volume per volume
vc	Viable cells
VG	Virus genome
VP	Viral capsid protein
VRC	Viral replication center
w/v	Mass per volume
wt	Wild type
X-SCID	X-linked severe combined immunodeficiency

8.5 Supplementary Data

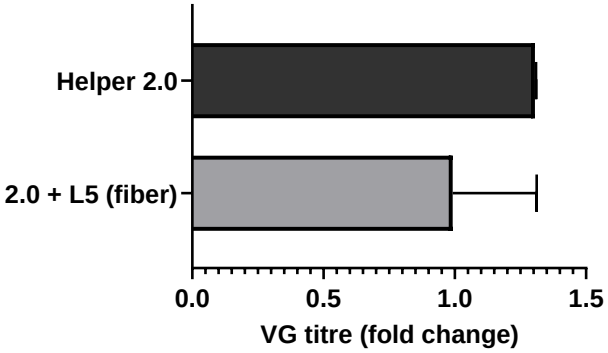


Figure 48: Analysis of the L5 (fiber) gene including 2.6 kb fragment, reinserted in Helper 2.0. Virus genome titre fold change related to the transfection of control Helper plasmid 0.1. Display of mean and standard deviation as error bars of two independent biological replicates.

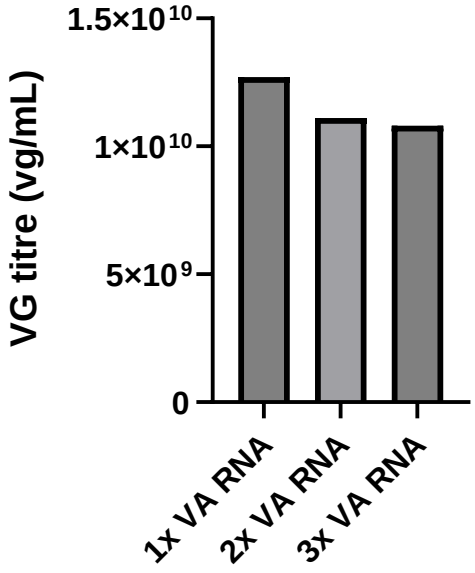


Figure 49: Test of increased pVA quantities in the 6-plasmid system (n=1).

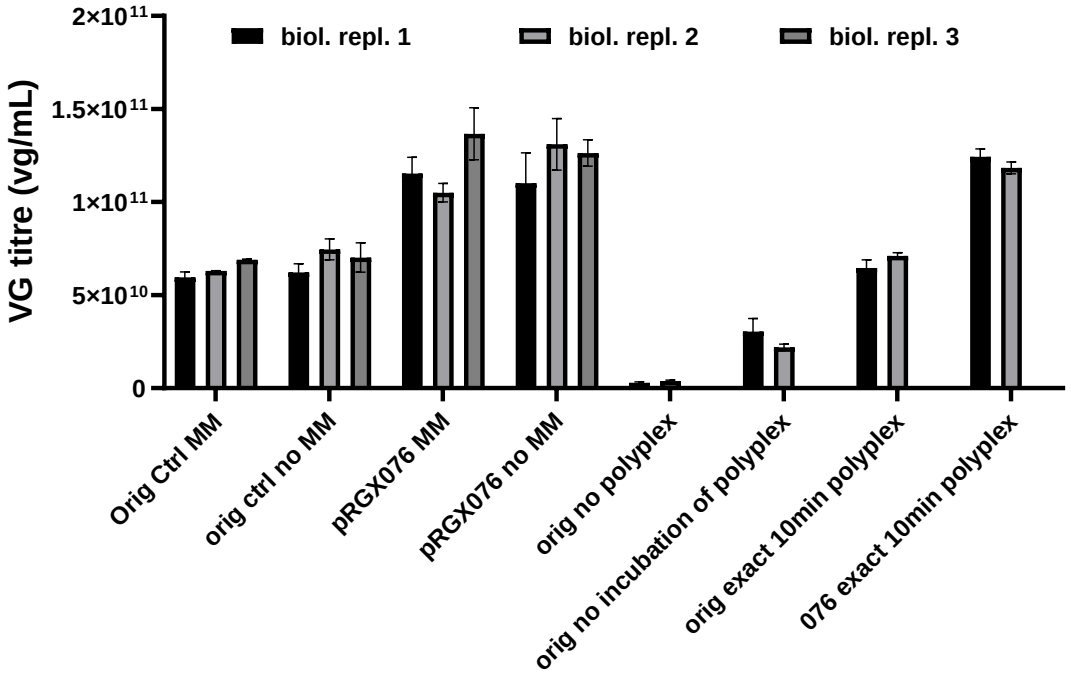


Figure 50: Experiments with different polyplexing times and methods confirming the used polyplex formation method for rAAV transient triple transfections. Displayed are means and standard deviations of three technical replicates (individual transfections of the same culture in different vessels) and three biological replicates.

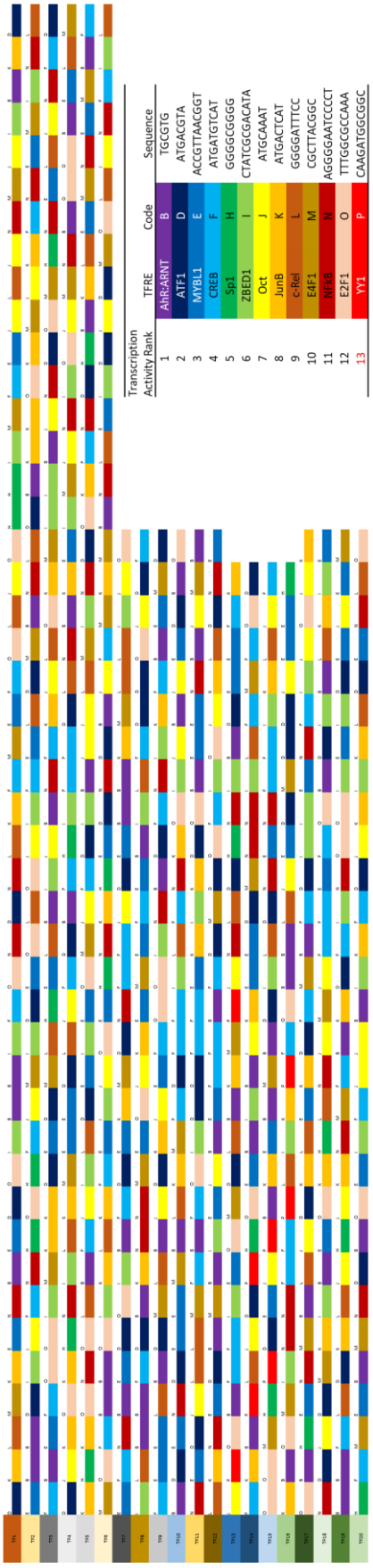


Figure 51: TFRE composition of synthetic promoters TP1 to TP20. Table depicts colour coding of TFREs, which were ranked based on their transcription activity in (Johari et al., 2022).

9 References

Agbandje-McKenna, M., & Kleinschmidt, J. (2012). AAV Capsid Structure and Cell Interactions. In *Methods in Molecular Biology* (Vol. 807, pp. 47–92). https://doi.org/10.1007/978-1-61779-370-7_3

Ahi, Y. S., & Mittal, S. K. (2016). Components of adenovirus genome packaging. *Frontiers in Microbiology*, *7*(SEP), 214495. https://doi.org/10.3389/FMICB.2016.01503/BIBTEX

Akusjärvi, G. (1985). Anatomy of region L1 from adenovirus type 2. *Journal of Virology*, *56*(3), 879–886. https://doi.org/10.1128/jvi.56.3.879-886.1985

Al-Husini, N., Medler, S., & Ansari, A. (2020). Crosstalk of promoter and terminator during RNA polymerase II transcription cycle. *Biochimica et Biophysica Acta (BBA) - Gene Regulatory Mechanisms*, *1863*(12), 194657. https://doi.org/10.1016/j.bbagrm.2020.194657

Allen, J. M., Debelak, D. J., Reynolds, T. C., & Miller, A. D. (1997). Identification and elimination of replication-competent adeno-associated virus (AAV) that can arise by nonhomologous recombination during AAV vector production. *Journal of Virology*, *71*(9), 6816–6822. https://doi.org/10.1128/jvi.71.9.6816-6822.1997

Allen, J. M., Halbert, C. L., & Miller, A. D. (2000). Improved Adeno-Associated Virus Vector Production with Transfection of a Single Helper Adenovirus Gene, E4orf6. *Molecular Therapy*, *1*(1), 88–95. https://doi.org/10.1006/mthe.1999.0010

Allen, M. J., Boyce, J. P., Trentalange, M. T., Treiber, D. L., Rasmussen, B., Tillotson, B., Davis, R., & Reddy, P. (2008). Identification of novel small molecule enhancers of protein production by cultured mammalian cells. *Biotechnology and Bioengineering*, *100*(6), 1193–1204. https://doi.org/10.1002/bit.21839

Anderson, W. F., Blaese, R. M., & Culver, K. (2008). Points to Consider Response with Clinical Protocol, July 6, 1990. *Https://Home.Liebertpub.Com/Hum*, *1*(3), 331–362. https://doi.org/10.1089/HUM.1990.1.3-331

Andersson, M. G., Haasnoot, P. C. J., Xu, N., Berenjian, S., Berkhout, B., & Akusjärvi, G. (2005). Suppression of RNA Interference by Adenovirus Virus-Associated RNA. *Journal of Virology*, *79*(15), 9556–9565. https://doi.org/10.1128/jvi.79.15.9556-9565.2005

Antoni, B. A., Rabson, A. B., Miller, I. L., Trempe, J. P., Chejanovsky, N., & Carter, B. J. (1991). Adeno-associated virus Rep protein inhibits human immunodeficiency

virus type 1 production in human cells. *Journal of Virology*, *65*(1), 396–404. https://doi.org/10.1128/jvi.65.1.396-404.1991

Aparicio, O., Carnero, E., Abad, X., Razquin, N., Guruceaga, E., Segura, V., & Fortes, P. (2010). Adenovirus VA RNA-derived miRNAs target cellular genes involved in cell growth, gene expression and DNA repair. *Nucleic Acids Research*, *38*(3), 750–763. https://doi.org/10.1093/nar/gkp1028

Arabi, F., Mansouri, V., & Ahmadbeigi, N. (2022). Gene therapy clinical trials, where do we go? An overview. *Biomedicine & Pharmacotherapy*, *153*, 113324. https://doi.org/10.1016/J.BIOPHA.2022.113324

Araujo, F. D., Stracker, T. H., Carson, C. T., Lee, D. V., & Weitzman, M. D. (2005). Adenovirus Type 5 E4orf3 Protein Targets the Mre11 Complex to Cytoplasmic Aggresomes. *Journal of Virology*, *79*(17), 11382–11391. https://doi.org/10.1128/JVI.79.17.11382-11391.2005

ARM. (2021). *Regenerative Medicine: Disrupting The Status Quo*. https://alliancerm.org/sector-report/2021-annual-report/

ARM. (2022, September). *Regenerative Medicine: The Pipeline Momentum Builds H1*. https://alliancerm.org/sector-report/h1-2022-report/

ASGCT, & Citeline. (2023). *Gene, Cell, & RNA Therapy Landscape Q1 2023 Quarterly Data Report*.

ASGCT, & PharmaIntelligence. (2021). *Gene, Cell, & RNA Therapy Landscape*.

Asokan, A., Hamra, J. B., Govindasamy, L., Agbandje-McKenna, M., & Samulski, R. J. (2006). Adeno-Associated Virus Type 2 Contains an Integrin 5 1 Binding Domain Essential for Viral Cell Entry. *Journal of Virology*, *80*(18), 8961–8969. https://doi.org/10.1128/JVI.00843-06

Atchison, R. W., Casto, B. C., & Hammon, W. McD. (1965). Adenovirus-Associated Defective Virus Particles. *Science*, *149*(3685), 754–756. https://doi.org/10.1126/science.149.3685.754

Aucoin, M. G., Perrier, M., & Kamen, A. A. (2008). Critical assessment of current adeno-associated viral vector production and quantification methods. *Biotechnology Advances*, *26*(1), 73–88. https://doi.org/10.1016/j.biotechadv.2007.09.001

Avantaggiati, M. L., Carbone, M., Graessmann, A., Nakatani, Y., Howard, B., & Levine, A. S. (1996). The SV40 large T antigen and adenovirus E1a oncoproteins interact with distinct isoforms of the transcriptional co-activator, p300. *The*

EMBO Journal, *15*(9), 2236–2248. https://doi.org/10.1002/j.1460-2075.1996.tb00577.x

Babiss, L. E. (1989). The cellular transcription factor E2f requires viral E1A and E4 gene products for increased DNA-binding activity and functions to stimulate adenovirus E2A gene expression. *Journal of Virology*, *63*(6), 2709–2717. https://doi.org/10.1128/jvi.63.6.2709-2717.1989

Backström, E., Kaufmann, K. B., Lan, X., & Akusjärvi, G. (2010). Adenovirus L4-22K stimulates major late transcription by a mechanism requiring the intragenic late-specific transcription factor-binding site. *Virus Research*, *151*(2), 220–228. https://doi.org/10.1016/j.virusres.2010.05.013

Bagchi, S., Raychaudhuri, P., & Nevins, J. R. (1990). Adenovirus E1A proteins can dissociate heteromeric complexes involving the E2F transcription factor: a novel mechanism for E1A trans-activation. *Cell*, *62*(4), 659–669. https://doi.org/10.1016/0092-8674(90)90112-R

Balakrishnan, B., & Jayandharan, G. (2014). Basic Biology of Adeno-Associated Virus (AAV) Vectors Used in Gene Therapy. *Current Gene Therapy*, *14*(2), 86–100. https://doi.org/10.2174/1566523214666140302193709

Barnes, C. R., Lee, H., Ojala, D. S., Lewis, K. K., Limsirichai, P., & Schaffer, D. V. (2021). Genome-wide activation screens to increase adeno-associated virus production. *Molecular Therapy. Nucleic Acids*, *26*, 94–103. https://doi.org/10.1016/J.OMTN.2021.06.026

Bartlett, J. S., Wilcher, R., & Samulski, R. J. (2000). Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *Journal of Virology*, *74*(6), 2777–2785. http://www.ncbi.nlm.nih.gov/pubmed/10684294

Bayley, S., & Mymryk, J. (1994). Adenovirus E1A Proteins and Transformation. *International Journal of Oncology*, *5*(3). https://doi.org/10.3892/ijo.5.3.425

bcc Research. (2022). *Global Cell and Gene Therapy Market*.

Beaton, A., Palumbo, P., & Berns, K. I. (1989). Expression from the adeno-associated virus p5 and p19 promoters is negatively regulated in trans by the rep protein. *Journal of Virology*, *63*(10), 4450–4454. https://doi.org/10.1128/jvi.63.10.4450-4454.1989

Bellutti, F., Kauer, M., Kneidinger, D., Lion, T., & Klein, R. (2015). Identification of RISC-associated adenoviral microRNAs, a subset of their direct targets, and global changes in the targetome upon lytic adenovirus 5 infection. *Journal of Virology*, *89*(3), 1608–1627. https://doi.org/10.1128/JVI.02336-14

Berk, A. J. (2005). Recent lessons in gene expression, cell cycle control, and cell biology from adenovirus. In *Oncogene* (Vol. 24, Issue 52, pp. 7673–7685). Oncogene. https://doi.org/10.1038/sj.onc.1209040

Berk, A. J., Lee, F., Harrison, T., Williams, J., & Sharp, P. A. (1979). Pre-early adenovirus 5 gene product regulates synthesis of early viral messenger RNAs. *Cell*, *17*(4), 935–944. https://doi.org/10.1016/0092-8674(79)90333-7

Berns, K. I., & Adler, S. (1972). Separation of Two Types of Adeno-Associated Virus Particles Containing Complementary Polynucleotide Chains. *Journal of Virology*, *9*(2), 394–396. https://doi.org/10.1128/jvi.9.2.394-396.1972

Berns, K. I., & Labow, M. A. (1987). Parvovirus gene regulation. In *Journal of General Virology* (Vol. 68, Issue 3, pp. 601–614). Microbiology Society. https://doi.org/10.1099/0022-1317-68-3-601

Berns, K. I., & Linden, R. M. (1995). The cryptic life style of adenoassociated virus. *BioEssays*, *17*(3), 237–245. https://doi.org/10.1002/bies.950170310

Berns, K. I., & Muzyczka, N. (2017). AAV: An Overview of Unanswered Questions. In *Human Gene Therapy* (Vol. 28, Issue 4, pp. 308–313). Mary Ann Liebert Inc. https://doi.org/10.1089/hum.2017.048

Berthet, C., Raj, K., Saudan, P., & Beard, P. (2005). How adeno-associated virus Rep78 protein arrests cells completely in S phase. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(38), 13634–13639. https://doi.org/10.1073/pnas.0504583102

Bevington, J. M., Needham, P. G., Verrill, K. C., Collaco, R. F., Basrur, V., & Trempe, J. P. (2007). Adeno-associated virus interactions with B23/Nucleophosmin: identification of sub-nucleolar virion regions. *Virology*, *357*(1), 102–113. https://doi.org/10.1016/j.virol.2006.07.050

Bhat, R. A., & Thimmappaya, B. (1984). Adenovirus mutants with DNA sequence perturbations in the intragenic promoter of VAI RNA gene allow the enhanced transcription of VAII RNA gene in HeLa cells. *Nucleic Acids Research*, *12*(19), 7377–7388. https://doi.org/10.1093/nar/12.19.7377

Biancalana, M., Natan, E., Lenardo, M. J., & Fersht, A. R. (2021). NF-κB Rel subunit exchange on a physiological timescale. *Protein Science : A Publication of the Protein Society*, *30*(9), 1818–1832. https://doi.org/10.1002/pro.4134

Biasiotto, R., & Akusjärvi, G. (2015). Regulation of Human Adenovirus Alternative RNA Splicing by the Adenoviral L4-33K and L4-22K Proteins. *International*

Journal of Molecular Sciences 2015, Vol. 16, Pages 2893-2912, *16*(2), 2893–2912. https://doi.org/10.3390/IJMS16022893

Blackford, A. N., & Grand, R. J. A. (2009). Adenovirus E1B 55-Kilodalton Protein: Multiple Roles in Viral Infection and Cell Transformation. *Journal of Virology*, *83*(9), 4000–4012. https://doi.org/10.1128/jvi.02417-08

Blackford, A. N., Patel, R. N., Forrester, N. A., Theil, K., Groitl, P., Stewart, G. S., Taylor, A. M. R., Morgan, I. M., Dobner, T., Grand, R. J. A., & Turnell, A. S. (2010). Adenovirus 12 E4orf6 inhibits ATR activation by promoting TOPBP1 degradation. *Proceedings of the National Academy of Sciences*, *107*(27), 12251–12256. https://doi.org/10.1073/pnas.0914605107

Blaese, R. M., Culver, K. W., Miller, A. D., Carter, C. S., Fleisher, T., Clerici, M., Shearer, G., Chang, L., Chiang, Y., Tolstoshev, P., Greenblatt, J. J., Rosenberg, S. A., Klein, H., Berger, M., Mullen, C. A., Ramsey, W. J., Muul, L., Morgan, R. A., & Anderson, W. F. (1995). T Lymphocyte-Directed Gene Therapy for ADA− SCID: Initial Trial Results After 4 Years. *Science*, *270*(5235), 475–480. https://doi.org/10.1126/SCIENCE.270.5235.475

Bleker, S., Sonntag, F., & Kleinschmidt, J. A. (2005). Mutational Analysis of Narrow Pores at the Fivefold Symmetry Axes of Adeno-Associated Virus Type 2 Capsids Reveals a Dual Role in Genome Packaging and Activation of Phospholipase A2 Activity. *Journal of Virology*, *79*(4), 2528–2540. https://doi.org/10.1128/jvi.79.4.2528-2540.2005

Bochkov, Y. A., & Palmenberg, A. C. (2006). Translational efficiency of EMCV IRES in bicistronic vectors is dependent upon IRES sequence and gene location. *BioTechniques*, *41*(3), 283–292. https://doi.org/10.2144/000112243

Borrelli, E., Hen, R., & Chambon, P. (1984). Adenovirus-2 E1A products repress enhancer-induced stimulation of transcription. *Nature*, *312*(5995), 608–612. https://doi.org/10.1038/312608a0

Bouard, D., Alazard-Dany, N., & Cosset, F. L. (2009). Viral vectors: from virology to transgene expression. *British Journal of Pharmacology*, *157*(2), 153. https://doi.org/10.1038/BJP.2008.349

Boye, C., Arpag, S., Francis, M., DeClemente, S., West, A., Heller, R., & Bulysheva, A. (2022). Reduction of plasmid vector backbone length enhances reporter gene expression. *Bioelectrochemistry (Amsterdam, Netherlands)*, *144*, 107981. https://doi.org/10.1016/j.bioelechem.2021.107981

Brightwell, G., Poirier, V., Cole, E., Ivins, S., & Brown, K. W. (1997). Serum-dependent and cell cycle-dependent expression from a cytomegalovirus-based

mammalian expression vector. *Gene*, *194*(1), 115–123. https://doi.org/10.1016/S0378-1119(97)00178-9

Brimble, M. A., Cheng, P.-H., Winston, S. M., Reeves, I. L., Souquette, A., Spence, Y., Zhou, J., Wang, Y.-D., Morton, C. L., Valentine, M., Thomas, P. G., Nathwani, A. C., Gray, J. T., & Davidoff, A. M. (2022). Preventing packaging of translatable P5-associated DNA contaminants in recombinant AAV vector preps. *Molecular Therapy - Methods & Clinical Development*, *24*, 280–291. https://doi.org/10.1016/j.omtm.2022.01.008

Brister, J. R., & Muzyczka, N. (2000). Mechanism of Rep-Mediated Adeno-Associated Virus Origin Nicking. *Journal of Virology*, *74*(17), 7762–7771. https://doi.org/10.1128/jvi.74.17.7762-7771.2000

Brooks, A. R., Harkins, R. N., Wang, P., Qian, H. S., Liu, P., & Rubanyi, G. M. (2004). Transcriptional silencing is associated with extensive methylation of the CMV promoter following adenoviral gene delivery to muscle. *The Journal of Gene Medicine*, *6*(4), 395–404. https://doi.org/10.1002/jgm.516

Brown, A. J., Gibson, S. J., Hatton, D., & James, D. C. (2017). In silico design of context-responsive mammalian promoters with user-defined functionality. *Nucleic Acids Research*, *45*(18), 10906–10919. https://doi.org/10.1093/nar/gkx768

Brown, A. J., Sweeney, B., Mainwaring, D. O., & James, D. C. (2014). Synthetic promoters for CHO cell engineering. *Biotechnology and Bioengineering*, *111*(8), 1638–1647. https://doi.org/10.1002/bit.25227

Bulaklak, K., & Gersbach, C. A. (2020). The once and future gene therapy. *Nature Communications 2020 11:1*, *11*(1), 1–4. https://doi.org/10.1038/s41467-020-19505-2

Büning, H., Perabo, L., Coutelle, O., Quadt-Humme, S., & Hallek, M. (2008). Recent developments in adeno-associated virus vector technology. *The Journal of Gene Medicine*, *10*(7), 717–733. https://doi.org/10.1002/jgm.1205

Büning, H., & Srivastava, A. (2019). Capsid Modifications for Targeting and Improving the Efficacy of AAV Vectors. *Molecular Therapy. Methods & Clinical Development*, *12*, 248. https://doi.org/10.1016/J.OMTM.2019.01.008

Bürck, C., Mund, A., Berscheminski, J., Kieweg, L., Müncheberg, S., Dobner, T., & Schreiner, S. (2016). KAP1 Is a Host Restriction Factor That Promotes Human Adenovirus E1B-55K SUMO Modification. *Journal of Virology*, *90*(2), 930. https://doi.org/10.1128/JVI.01836-15

Calcedo, R., Vandenberghe, L. H., Gao, G., Lin, J., & Wilson, J. M. (2009). Worldwide Epidemiology of Neutralizing Antibodies to Adeno-Associated Viruses. *The Journal of Infectious Diseases*, *199*(3), 381–390. https://doi.org/10.1086/595830

Cameau, E., Pedregal, A., & Glover, C. (2019). Cost modelling comparison of adherent multi-trays with suspension and fixed-bed bioreactors for the manufacturing of gene therapy products. *Cell and Gene Therapy Insights*, *5*(11), 1663–1674. https://doi.org/10.18609/cgti.2019.175

Cao, M., Chiriva-Internati, M., & Hermonat, P. L. (2015). AAV2 X increases AAV6 rep/cap-driven rAAV production. *Virology*, *482*, 84–88. https://doi.org/10.1016/j.virol.2015.03.007

Cao, M., You, H., & Hermonat, P. L. (2014). The X Gene of Adeno-Associated Virus 2 (AAV2) Is Involved in Viral DNA Replication. *PLOS ONE*, *9*(8), e104596. https://doi.org/10.1371/journal.pone.0104596

Carter, B. J., Antoni, B. A., & Klessig, D. F. (1992). Adenovirus containing a deletion of the early region 2A gene allows growth of adeno-associated virus with decreased efficiency. *Virology*, *191*(1), 473–476. https://doi.org/10.1016/0042-6822(92)90213-9

Carter, B. J., Marcus-Sekura, C. J., Laughlin, C. A., & Ketner, G. (1983). Properties of an adenovirus type 2 mutant, Ad2d/807, having a deletion near the right-hand genome terminus: Failure to help AAV replication. *Virology*, *126*(2), 505–516. https://doi.org/10.1016/S0042-6822(83)80008-7

Carvalho, T., Seeler, J. S., Ohman, K., Jordan, P., Pettersson, U., Akusjärvi, G., Carmo-Fonseca, M., & Dejean, A. (1995). Targeting of adenovirus E1A and E4-ORF3 proteins to nuclear matrix-associated PML bodies. *The Journal of Cell Biology*, *131*(1), 45–56. https://doi.org/10.1083/jcb.131.1.45

Casper, J. M., Timpe, J. M., Dignam, J. D., & Trempe, J. P. (2005). Identification of an Adeno-Associated Virus Rep Protein Binding Site in the Adenovirus E2a Promoter. *Journal of Virology*, *79*(1), 28–38. https://doi.org/10.1128/jvi.79.1.28-38.2005

Castro-Mondragon, J. A., Riudavets-Puig, R., Rauluseviciute, I., Berhanu Lemma, R., Turchi, L., Blanc-Mathieu, R., Lucas, J., Boddie, P., Khan, A., Manosalva Pérez, N., Fornes, O., Leung, T. Y., Aguirre, A., Hammal, F., Schmelter, D., Baranasic, D., Ballester, B., Sandelin, A., Lenhard, B., … Mathelier, A. (2022). JASPAR 2022: the 9th release of the open-access database of transcription factor binding profiles.

Nucleic Acids Research, *50*(D1), D165–D173. https://doi.org/10.1093/nar/gkab1113

Cecchini, S., Virag, T., & Kotin, R. M. (2011). Reproducible High Yields of Recombinant Adeno-Associated Virus Produced Using Invertebrate Cells in 0.02- to 200-Liter Cultures. *Human Gene Therapy*, *22*(8), 1021–1030. https://doi.org/10.1089/hum.2010.250

Cevec. (2020). *ELEVECTA - Establishment of a Scalable Production Process Using Stable Helper-Virus Free AAV Producer Cell Lines Based on Human Suspension Cells - ASGCT 23rd Annual Meeting*. https://cslide-us.ctimeetingtech.com/asgct23/attendee/eposter/file/1094#1

Chadeuf, G., Ciron, C., Moullier, P., & Salvetti, A. (2005). Evidence for Encapsidation of Prokaryotic Sequences during Recombinant Adeno-Associated Virus Production and Their in Vivo Persistence after Vector Delivery. *Molecular Therapy*, *12*(4), 744–753. https://doi.org/10.1016/J.YMTHE.2005.06.003

Chai, S., Wakefield, L., Norgard, M., Li, B., Enicks, D., Marks, D. L., & Grompe, M. (2023). Strong ubiquitous micro-promoters for recombinant adeno-associated viral vectors. *Molecular Therapy - Methods & Clinical Development*, *29*, 504–512. https://doi.org/10.1016/j.omtm.2023.05.013

Chamberlain, K., Riyad, J. M., & Weber, T. (2016). Expressing transgenes that exceed the packaging capacity of adeno-associated virus capsids. *Human Gene Therapy Methods*, *27*(1), 1–12. https://doi.org/10.1089/hgtb.2015.140

Chang, J., Chen, X., Wang, R., Shi, R., Wang, X., Lu, W., Ma, S., & Xia, Q. (2020). High-Throughput Screening Identifies Two Novel Small Molecule Enhancers of Recombinant Protein Expression. *Molecules*, *25*(2), 353. https://doi.org/10.3390/molecules25020353

Chang, L. S., & Shenk, T. (1990). The adenovirus DNA-binding protein stimulates the rate of transcription directed by adenovirus and adeno-associated virus promoters. *Journal of Virology*, *64*(5), 2103–2109. https://doi.org/10.1128/jvi.64.5.2103-2109.1990

Chang, L. S., Shi, Y., & Shenk, T. (1989). Adeno-associated virus P5 promoter contains an adenovirus E1A-inducible element and a binding site for the major late transcription factor. *Journal of Virology*, *63*(8), 3479–3488. https://doi.org/10.1128/jvi.63.8.3479-3488.1989

Chao, L. C., Jamil, A., Kim, S. J., Huang, L., & Martinson, H. G. (1999). Assembly of the Cleavage and Polyadenylation Apparatus Requires About 10 Seconds In Vivo

and Is Faster for Strong than for Weak Poly(A) Sites. *Molecular and Cellular Biology*, *19*(8), 5588. https://doi.org/10.1128/MCB.19.8.5588

Cheng, C. Y., Gilson, T., Wimmer, P., Schreiner, S., Ketner, G., Dobner, T., Branton, P. E., & Blanchette, P. (2013). Role of E1B55K in E4orf6/E1B55K E3 Ligase Complexes Formed by Different Human Adenovirus Serotypes. *Journal of Virology*, *87*(11), 6232. https://doi.org/10.1128/JVI.00384-13

Chiorini, J. A., Zimmermann, B., Yang, L., Smith, R. H., Ahearn, A., Herberg, F., & Kotin, R. M. (1998). Inhibition of PrKX, a Novel Protein Kinase, and the Cyclic AMP-Dependent Protein Kinase PKA by the Regulatory Proteins of Adeno-Associated Virus Type 2. *Molecular and Cellular Biology*, *18*(10), 5921–5929. https://doi.org/10.1128/mcb.18.10.5921

Chrivia, J. C., Kwok, R. P. S., Lamb, N., Hagiwara, M., Montminy, M. R., & Goodman, R. H. (1993a). Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature*, *365*(6449), 855–859. https://doi.org/10.1038/365855a0

Chrivia, J. C., Kwok, R. P. S., Lamb, N., Hagiwara, M., Montminy, M. R., & Goodman, R. H. (1993b). Phosphorylated CREB binds specifically to the nuclear protein CBP. *Nature*, *365*(6449), 855–859. https://doi.org/10.1038/365855a0

Clément, N. (2019). Large-Scale Clinical Manufacturing of AAV Vectors for Systemic Muscle Gene Therapy. *Muscle Gene Therapy, Second Edition*, 253–273. https://doi.org/10.1007/978-3-030-03095-7_15/COVER

Clément, N., & Grieger, J. C. (2016). Manufacturing of recombinant adeno-associated viral vectors for clinical trials. *Molecular Therapy - Methods & Clinical Development*, *3*, 16002. https://doi.org/10.1038/mtm.2016.2

Coban, C., Ishii, K. J., Gursel, M., Klinman, D. M., & Kumar, N. (2005). Effect of plasmid backbone modification by different human CpG motifs on the immunogenicity of DNA vaccine vectors. *Journal of Leukocyte Biology*, *78*(3), 647–655. https://doi.org/10.1189/jlb.1104627

Cockett, M. I., Bebbington, C. R., & Yarranton, G. T. (1991). The use of engineered E1A genes to transactivate the hCMV-MIE promoter in permanent CHO cell lines. *Nucleic Acids Research*, *19*(2), 319–325. https://doi.org/10.1093/nar/19.2.319

Colella, P., Ronzitti, G., & Mingozzi, F. (2018). Emerging Issues in AAV-Mediated In Vivo Gene Therapy. *Molecular Therapy - Methods & Clinical Development*, *8*, 87–104. https://doi.org/10.1016/j.omtm.2017.11.007

Collaco, R. F., Kalman-Maltese, V., Smith, A. D., Dignam, J. D., & Trempe, J. P. (2003). A Biochemical Characterization of the Adeno-associated Virus Rep40 Helicase.

Journal of Biological Chemistry, *278*(36), 34011–34017. https://doi.org/10.1074/jbc.M301537200

Colosi. (1996). *Adenovirus helper-free recombinant AAV Virion production*.

Colosi. (2008). *High-efficiency wild-type-free AAV helper functions*.

Coulon, A., Chow, C. C., Singer, R. H., & Larson, D. R. (2013). Eukaryotic transcriptional dynamics: from single molecules to cell populations. *Nature Reviews Genetics*, *14*(8), 572–584. https://doi.org/10.1038/nrg3484

Cuconati, A., Mukherjee, C., Perez, D., & White, E. (2003). DNA damage response and MCL-1 destruction initiate apoptosis in adenovirus-infected cells. *Genes & Development*, *17*(23), 2922–2932. https://doi.org/10.1101/GAD.1156903

Dallaire, F., Schreiner, S., Blair, G. E., Dobner, T., Branton, P. E., & Blanchette, P. (2016). The Human Adenovirus Type 5 E4orf6/E1B55K E3 Ubiquitin Ligase Complex Enhances E1A Functional Activity. *MSphere*, *1*(1). https://doi.org/10.1128/MSPHERE.00015-15/ASSET/DED3A35F-A3FE-4ACC-A01F-1C6322B6CF09/ASSETS/GRAPHIC/SPH0011600130005.JPEG

Darweesh, M., Kamel, W., Gavrilin, M. A., Akusjärvi, G., & Svensson, C. (2019). Adenovirus VA RNAI Blocks ASC Oligomerization and Inhibits NLRP3 Inflammasome Activation. *Frontiers in Immunology*, *10*, 485240. https://doi.org/10.3389/FIMMU.2019.02791/BIBTEX

Das, A. T., Tenenbaum, L., & Berkhout, B. (2016). Tet-On Systems For Doxycycline-inducible Gene Expression. *Current Gene Therapy*, *16*(3), 156. https://doi.org/10.2174/1566523216666160524144041

Dash, S., Sharon, D. M., Mullick, A., & Kamen, A. A. (2022). Only a small fraction of cells produce assembled capsids during transfection-based manufacturing of adeno-associated virus vectors. *Biotechnology and Bioengineering*, *119*(6), 1685–1690. https://doi.org/10.1002/bit.28068

Daya, S., & Berns, K. I. (2008). Gene therapy using adeno-associated virus vectors. *Clinical Microbiology Reviews*, *21*(4), 583–593. https://doi.org/10.1128/CMR.00008-08

Deng, J., Zhang, J., Shi, K., & Liu, Z. (2022). Drug development progress in duchenne muscular dystrophy. *Frontiers in Pharmacology*, *13*. https://doi.org/10.3389/FPHAR.2022.950651/FULL

Dery, C. V, Herrmann, C. H., & Mathews, M. B. (1987). Response of individual adenovirus promoters to the products of the E1A gene. *Oncogene*, *2*(1), 15–23. http://www.ncbi.nlm.nih.gov/pubmed/2963989

Di Pasquale, G., & Stacey, S. N. (1998). Adeno-Associated Virus Rep78 Protein Interacts with Protein Kinase A and Its Homolog PRKX and Inhibits CREB-Dependent Transcriptional Activation. *Journal of Virology*, *72*(10), 7916–7925. https://doi.org/10.1128/JVI.72.10.7916-7925.1998/ASSET/C5481273-6523-48B9-8ED0-BAEEF72AF2A4/ASSETS/GRAPHIC/JV1080737007.JPEG

Dix, I., & Leppard, K. N. (1993). Regulated splicing of adenovirus type 5 E4 transcripts and regulated cytoplasmic accumulation of E4 mRNA. *Journal of Virology*, *67*(6), 3226–3231. https://doi.org/10.1128/jvi.67.6.3226-3231.1993

Dix, I., & Leppard, K. N. (1995). Expression of adenovirus type 5 E4 Orf2 protein during lytic infection. *The Journal of General Virology*, *76 (Pt 4)*(4), 1051–1055. https://doi.org/10.1099/0022-1317-76-4-1051

Dong, J.-Y., Fan, P.-D., & Frizzell, R. A. (1996). Quantitative Analysis of the Packaging Capacity of Recombinant Adeno-Associated Virus. *Human Gene Therapy*, *7*(17), 2101–2112. https://doi.org/10.1089/hum.1996.7.17-2101

Donovan-Banfield, I., Turnell, A. S., Hiscox, J. A., Leppard, K. N., & Matthews, D. A. (2020). Deep splicing plasticity of the human adenovirus type 5 transcriptome drives virus evolution. *Communications Biology*, *3*(1), 124. https://doi.org/10.1038/s42003-020-0849-9

Douin, V., Bornes, S., Creancier, L., Rochaix, P., Favre, G., Prats, A.-C., & Couderc, B. (2004). Use and comparison of different internal ribosomal entry sites (IRES) in tricistronic retroviral vectors. *BMC Biotechnology 2004 4:1*, *4*(1), 1–12. https://doi.org/10.1186/1472-6750-4-16

Duan, D., Sharma, P., Dudus, L., Zhang, Y., Sanlioglu, S., Yan, Z., Yue, Y., Ye, Y., Lester, R., Yang, J., Fisher, K. J., & Engelhardt, J. F. (1999). Formation of Adeno-Associated Virus Circular Genomes Is Differentially Regulated by Adenovirus E4 ORF6 and E2a Gene Expression. *Journal of Virology*, *73*(1), 161–169. https://doi.org/10.1128/JVI.73.1.161-169.1999

Dunbar, C. E., High, K. A., Joung, J. K., Kohn, D. B., Ozawa, K., & Sadelain, M. (2018). Gene therapy comes of age. *Science*, *359*(6372). https://doi.org/10.1126/science.aan4672

Durocher, Y., Pham, P. L., St-Laurent, G., Jacob, D., Cass, B., Chahal, P., Lau, C. J., Nalbantoglu, J., & Kamen, A. (2007). Scalable serum-free production of recombinant adeno-associated virus type 2 by transfection of 293 suspension cells. *Journal of Virological Methods*, *144*(1–2), 32–40. https://doi.org/10.1016/j.jviromet.2007.03.014

Dzananovic, E., Astha, Chojnowski, G., Deo, S., Booy, E. P., Padilla-Meier, P., McEleney, K., Bujnicki, J. M., Patel, T. R., & McKenna, S. A. (2017). Impact of the structural integrity of the three-way junction of adenovirus VAI RNA on PKR inhibition. *PLOS ONE*, *12*(10), e0186849. https://doi.org/10.1371/journal.pone.0186849

Edwards, A. S., & Scott, J. D. (2000). A-kinase anchoring proteins: Protein kinase A and beyond. In *Current Opinion in Cell Biology* (Vol. 12, Issue 2, pp. 217–221). Elsevier Current Trends. https://doi.org/10.1016/S0955-0674(99)00085-X

Egan, C., Bayley, S. T., & Branton, P. E. (1989). Binding of the Rb1 protein to E1A products is required for adenovirus transformation. *Oncogene*, *4*(3), 383–388. http://www.ncbi.nlm.nih.gov/pubmed/2523032

Ellsworth, D., Finnen, R. L., & Flint, S. J. (2001). Superimposed Promoter Sequences of the Adenoviral E2 Early RNA Polymerase III and RNA Polymerase II Transcription Units. *Journal of Biological Chemistry*, *276*(1), 827–834. https://doi.org/10.1074/jbc.M007036200

Elmore, Z. C., Patrick Havlik, L., Oh, D. K., Anderson, L., Daaboul, G., Devlin, G. W., Vincent, H. A., & Asokan, A. (2021). The membrane associated accessory protein is an adeno-associated viral egress factor. *Nature Communications*, *12*(1), 6239. https://doi.org/10.1038/s41467-021-26485-4

Emmerling, V. V., Pegel, A., Milian, E. G., Venereo-Sanchez, A., Kunz, M., Wegele, J., Kamen, A. A., Kochanek, S., & Hoerer, M. (2016). Rational plasmid design and bioprocess optimization to enhance recombinant adeno-associated virus (AAV) productivity in mammalian cells. *Biotechnology Journal*, *11*(2), 290–297. https://doi.org/10.1002/biot.201500176

Endter, C., Härtl, B., Spruss, T., Hauber, J., & Dobner, T. (2005). Blockage of CRM1-dependent nuclear export of the adenovirus type 5 early region 1B 55-kDa protein augments oncogenic transformation of primary rat cells. *Oncogene*, *24*(1), 55–64. https://doi.org/10.1038/sj.onc.1208170

Esposito, D., Mehalko, J., Drew, M., Snead, K., Wall, V., Taylor, T., Frank, P., Denson, J. P., Hong, M., Gulten, G., Sadtler, K., Messing, S., & Gillette, W. (2020). Optimizing high-yield production of SARS-CoV-2 soluble spike trimers for serology assays. *Protein Expression and Purification*, *174*, 105686. https://doi.org/10.1016/j.pep.2020.105686

Eszterhas, S. K., Bouhassira, E. E., Martin, D. I. K., & Fiering, S. (2002). Transcriptional Interference by Independently Regulated Genes Occurs in Any Relative Arrangement of the Genes and Is Influenced by Chromosomal Integration

Position. *Molecular and Cellular Biology*, *22*(2), 469–479. https://doi.org/10.1128/MCB.22.2.469-479.2002

Evans, J. D., & Hearing, P. (2005). Relocalization of the Mre11-Rad50-Nbs1 Complex by the Adenovirus E4 ORF3 Protein Is Required for Viral Replication. *Journal of Virology*, *79*(10), 6207–6215. https://doi.org/10.1128/JVI.79.10.6207-6215.2005

EveluatePharma, & BCG. (2020). *Global Gene Therapy Market: Present and Forecast – Creative Biolabs Gene Therapy Blog*. https://www.creative-biolabs.com/blog/gene-therapy/global-gene-therapy-market-present-and-forecast/

Ewing, S. G., Byrd, S. A., Christensen, J. B., Tyler, R. E., & Imperiale, M. J. (2007). Ternary Complex Formation on the Adenovirus Packaging Sequence by the IVa2 and L4 22-Kilodalton Proteins. *Journal of Virology*, *81*(22), 12450–12457. https://doi.org/10.1128/JVI.01470-07

Farley, D. C., Brown, J. L., & Leppard, K. N. (2004). Activation of the Early-Late Switch in Adenovirus Type 5 Major Late Transcription Unit Expression by L4 Gene Products. *Journal of Virology*, *78*(4), 1782–1791. https://doi.org/10.1128/jvi.78.4.1782-1791.2004

Farris, K. D., & Pintel, D. J. (2008). Improved Splicing of Adeno-Associated Viral (AAV) Capsid Protein-Supplying Pre-mRNAs Leads to Increased Recombinant AAV Vector Production. *Human Gene Therapy*, *19*(12), 1421–1427. https://doi.org/10.1089/hum.2008.118

Fattaey, A. R., & Consigli, R. A. (1989). Synthesis, posttranslational modifications, and nuclear transport of polyomavirus major capsid protein VP1. *Journal of Virology*, *63*(7), 3168–3175. https://doi.org/10.1128/jvi.63.7.3168-3175.1989

Fax, P., Lehmkühler, O., Kühn, C., Esche, H., & Brockmann, D. (2000). E1A12S-mediated activation of the adenovirus type 12 E2 promoter depends on the histone acetyltransferase activity of p300/CBP. *Journal of Biological Chemistry*, *275*(51), 40554–40560. https://doi.org/10.1074/jbc.M004626200

Fax, P., Lipinski, K. S., Esche, H., & Brockmann, D. (2000). cAMP-independent activation of the adenovirus type 12 E2 promoter correlates with the recruitment of CREB-1/ATF-1, E1A(12S), and CBP to the E2- CRE. *Journal of Biological Chemistry*, *275*(12), 8911–8920. https://doi.org/10.1074/jbc.275.12.8911

FDA. (2023). *Approved Cellular and Gene Therapy Products | FDA*. https://www.fda.gov/vaccines-blood-biologics/cellular-gene-therapy-products/approved-cellular-and-gene-therapy-products

Ferrari, F. K., Samulski, T., Shenk, T., & Samulski, R. J. (1996). Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *Journal of Virology*, *70*(5), 3227–3234. https://doi.org/10.1128/jvi.70.5.3227-3234.1996

Ferreon, J. C., Martinez-Yamout, M. A., Dyson, H. J., & Wright, P. E. (2009). Structural basis for subversion of cellular control mechanisms by the adenoviral E1A oncoprotein. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(32), 13260–13265. https://doi.org/10.1073/pnas.0906770106

Fessler, S. P., & Young, C. S. H. (1998). Control of Adenovirus Early Gene Expression during the Late Phase of Infection. *Journal of Virology*, *72*(5), 4049. https://doi.org/10.1128/JVI.72.5.4049-4056.1998

Finnen, R. L., Biddle, J. F., & Flint, J. (2001). Truncation of the Human Adenovirus Type 5 L4 33-kDa Protein: Evidence for an Essential Role of the Carboxy-Terminus in the Viral Infectious Cycle. *Virology*, *289*(2), 388–399. https://doi.org/10.1006/viro.2001.1130

Fisher, K. J., Gao, G. P., Weitzman, M. D., DeMatteo, R., Burda, J. F., & Wilson, J. M. (1996). Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *Journal of Virology*, *70*(1), 520–532. https://doi.org/10.1128/jvi.70.1.520-532.1996

Florea, M., Nicolaou, F., Pacouret, S., Zinn, E. M., Sanmiguel, J., Andres-Mateos, E., Unzu, C., Wagers, A. J., & Vandenberghe, L. H. (2023). High-efficiency purification of divergent AAV serotypes using AAVX affinity chromatography. *Molecular Therapy. Methods & Clinical Development*, *28*, 146. https://doi.org/10.1016/J.OMTM.2022.12.009

Forrester, N. A., Sedgwick, G. G., Thomas, A., Blackford, A. N., Speiseder, T., Dobner, T., Byrd, P. J., Stewart, G. S., Turnell, A. S., & Grand, R. J. A. (2011). Serotype-Specific Inactivation of the Cellular DNA Damage Response during Adenovirus Infection. *Journal of Virology*, *85*(5), 2201–2211. https://doi.org/10.1128/JVI.01748-10

Franzoso, F. D., Seyffert, M., Vogel, R., Yakimovich, A., Pereira, B. de A., Meier, A. F., Sutter, S. O., Tobler, K., Vogt, B., Greber, U. F., Büning, H., Ackermann, M., & Fraefel, C. (2017). Cell Cycle-Dependent Expression of Adeno-Associated Virus

2 (AAV2) Rep in Coinfections with Herpes Simplex Virus 1 (HSV-1) Gives Rise to a Mosaic of Cells Replicating either AAV2 or HSV-1. *Journal of Virology*, *91*(15). https://doi.org/10.1128/JVI.00357-17

Friedmann, T. (1992). A brief history of gene therapy. *Nature Genetics 1992 2:2*, *2*(2), 93–98. https://doi.org/10.1038/ng1092-93

Frisch, S. M., & Mymryk, J. S. (2002). Adenovirus-5 E1A: Paradox and paradigm. In *Nature Reviews Molecular Cell Biology* (Vol. 3, Issue 6, pp. 441–452). Nature Publishing Group. https://doi.org/10.1038/nrm827

Galibert, L., Hyvönen, A., Eriksson, R. A. E., Mattola, S., Aho, V., Salminen, S., Albers, J. D., Peltola, S. K., Weman, S., Nieminen, T., Ylä-Herttuala, S., Lesch, H. P., Vihinen-Ranta, M., & Airenne, K. J. (2021). Functional roles of the membrane-associated AAV protein MAAP. *Scientific Reports*, *11*(1), 21698. https://doi.org/10.1038/s41598-021-01220-7

Gallimore, P. H., & Turnell, A. S. (2001). Adenovirus E1A: remodelling the host cell, a life or death experience. *Oncogene*, *20*(54), 7824–7835. https://doi.org/10.1038/sj.onc.1204913

Gao, G., Vandenberghe, L. H., Alvira, M. R., Lu, Y., Calcedo, R., Zhou, X., & Wilson, J. M. (2004). Clades of Adeno-Associated Viruses Are Widely Disseminated in Human Tissues. *Journal of Virology*, *78*(12), 6381–6388. https://doi.org/10.1128/jvi.78.12.6381-6388.2004

Gao, G., Vandenberghe, L., & Wilson, J. (2005). New Recombinant Serotypes of AAV Vectors. *Current Gene Therapy*, *5*(3), 285–297. https://doi.org/10.2174/1566523054065057

Geisberg, J. V, Lee, W. S., Berk, A. J., & Ricciardi, R. P. (1994). The zinc finger region of the adenovirus E1A transactivating domain complexes with the TATA box binding protein. *Proceedings of the National Academy of Sciences*, *91*(7), 2488–2492. https://doi.org/10.1073/pnas.91.7.2488

Geiss-Friedlander, R., & Melchior, F. (2007). Concepts in sumoylation: A decade on. In *Nature Reviews Molecular Cell Biology* (Vol. 8, Issue 12, pp. 947–956). Nature Publishing Group. https://doi.org/10.1038/nrm2293

Georg-Fries, B., Biederlack, S., Wolf, J., & Zur Hausen, H. (1984). Analysis of proteins, helper dependence, and seroepidemiology of a new human parvovirus. *Virology*, *134*(1), 64–71. https://doi.org/10.1016/0042-6822(84)90272-1

Giacinti, C., & Giordano, A. (2006). RB and cell cycle progression. *Oncogene*, *25*(38), 5220–5227. https://doi.org/10.1038/sj.onc.1209615

Gibson, D. G., Young, L., Chuang, R. Y., Venter, J. C., Hutchison, C. A., & Smith, H. O. (2009). Enzymatic assembly of DNA molecules up to several hundred kilobases. *Nature Methods 2009 6:5*, *6*(5), 343–345. https://doi.org/10.1038/nmeth.1318

Gilardi, P., & Perricaudet, M. (1986). The E4 promoter of adenovirus type 2 contains an E1A dependent cis-actlng element. *Nucleic Acids Research*, *14*(22), 9035–9049. https://doi.org/10.1093/nar/14.22.9035

Giles, A. R., Sims, J. J., Turner, K. B., Govindasamy, L., Alvira, M. R., Lock, M., & Wilson, J. M. (2018). Deamidation of Amino Acids on the Surface of Adeno-Associated Virus Capsids Leads to Charge Heterogeneity and Altered Vector Function. *Molecular Therapy*, *26*(12), 2848–2862. https://doi.org/10.1016/j.ymthe.2018.09.013

Girod, A., Wobus, C. E., Zádori, Z., Ried, M., Leike, K., Tijssen, P., Kleinschmidt, J. A., & Hallek, M. (2002). The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. *Journal of General Virology*, *83*(5), 973–978. https://doi.org/10.1099/0022-1317-83-5-973

Gonçalves, M. A. (2005). Adeno-associated virus: from defective virus to effective vector. *Virology Journal*, *2*(1), 43. https://doi.org/10.1186/1743-422X-2-43

Görlich, D., Kostka, S., Kraft, R., Dingwall, C., Laskey, R. A., Hartmann, E., & Prehn, S. (1995). Two different subunits of importin cooperate to recognize nuclear localization signals and bind them to the nuclear envelope. *Current Biology*, *5*(4), 383–392. https://doi.org/10.1016/S0960-9822(95)00079-0

Graham, F. L., Smiley, J., Russell, W. C., & Nairn, R. (1977). Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *Journal of General Virology*, *36*(1), 59–72. https://doi.org/10.1099/0022-1317-36-1-59

Grand, R. J. A., Ibrahim, A. P., Taylor, A. M. R., Milner, A. E., Gregory, C. D., Gallimore, P. H., & Turnell, A. S. (1998). Human Cells Arrest in S Phase in Response to Adenovirus 12 E1A. *Virology*, *244*(2), 330–342. https://doi.org/10.1006/VIRO.1998.9102

Green, E. A., Hamaker, N. K., & Lee, K. H. (2023). Comparison of vector elements and process conditions in transient and stable suspension HEK293 platforms using SARS-CoV-2 receptor binding domain as a model protein. *BMC Biotechnology*, *23*(1), 1–13. https://doi.org/10.1186/S12896-023-00777-7/FIGURES/4

Greig, J. A., Jennis, M., Dandekar, A., Chorazeczewski, J. K., Smith, M. K., Ashley, S. N., Yan, H., & Wilson, J. M. (2021). Muscle-directed AAV gene therapy rescues the maple syrup urine disease phenotype in a mouse model. *Molecular Genetics and*

Metabolism, *134*(1–2), 139–146. https://doi.org/10.1016/j.ymgme.2021.08.003

Grieger, J. C., Johnson, J. S., Gurda-Whitaker, B., Agbandje-McKenna, M., & Samulski, R. J. (2007). Surface-Exposed Adeno-Associated Virus Vp1-NLS Capsid Fusion Protein Rescues Infectivity of Noninfectious Wild-Type Vp2/Vp3 and Vp3-Only Capsids but Not That of Fivefold Pore Mutant Virions. *Journal of Virology*, *81*(15), 7833–7843. https://doi.org/10.1128/JVI.00580-07

Grieger, J. C., & Samulski, R. J. (2005). Packaging Capacity of Adeno-Associated Virus Serotypes: Impact of Larger Genomes on Infectivity and Postentry Steps. *Journal of Virology*, *79*(15), 9933–9944. https://doi.org/10.1128/JVI.79.15.9933-9944.2005

Grieger, J. C., & Samulski, R. J. (2012). Adeno-Associated Virus Vectorology, Manufacturing, and Clinical Applications. In *Methods in Enzymology* (Vol. 507, pp. 229–254). https://doi.org/10.1016/B978-0-12-386509-0.00012-0

Grieger, J. C., Snowdy, S., & Samulski, R. J. (2006). Separate Basic Region Motifs within the Adeno-Associated Virus Capsid Proteins Are Essential for Infectivity and Assembly. *Journal of Virology*, *80*(11), 5199–5210. https://doi.org/10.1128/JVI.02723-05

Grieger, J. C., Soltys, S. M., & Samulski, R. J. (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. *Molecular Therapy*, *24*(2), 287–297. https://doi.org/10.1038/mt.2015.187

Grimm, D., Kern, A., Rittner, K., & Kleinschmidt, J. A. (1998). Novel Tools for Production and Purification of Recombinant Adenoassociated Virus Vectors. *Human Gene Therapy*, *9*(18), 2745–2760. https://doi.org/10.1089/hum.1998.9.18-2745

Grimm, D., Zhou, S., Nakai, H., Thomas, C. E., Storm, T. A., Fuess, S., Matsushita, T., Allen, J., Surosky, R., Lochrie, M., Meuse, L., McClelland, A., Colosi, P., & Kay, M. A. (2003). Preclinical in vivo evaluation of pseudotyped adeno-associated virus vectors for liver gene therapy. *Blood*, *102*(7), 2412–2419. https://doi.org/10.1182/blood-2003-02-0495

Grinstein, J. D. (2023). ASGCT News: Siren Biotechnology Hails Universal Gene Therapy for Oncology. *Https://Home.Liebertpub.Com/Genedge*, *5*(1), 382–386. https://doi.org/10.1089/GENEDGE.5.1.74

Grosse, S., Penaud-Budloo, M., Herrmann, A.-K., Börner, K., Fakhiri, J., Laketa, V., Krämer, C., Wiedtke, E., Gunkel, M., Ménard, L., Ayuso, E., & Grimm, D. (2017). Relevance of Assembly-Activating Protein for Adeno-associated Virus Vector Production and Capsid Protein Stability in Mammalian and Insect Cells. *Journal of Virology*, *91*(20), 1198–1215. https://doi.org/10.1128/JVI.01198-17

Gruda, M. C., & Alwine, J. C. (1991). Simian virus 40 (SV40) T-antigen transcriptional activation mediated through the Oct/SPH region of the SV40 late promoter. *Journal of Virology*, *65*(7), 3553–3558. https://doi.org/10.1128/jvi.65.7.3553-3558.1991

Gruda, M. C., Zabolotny, J. M., Xiao, J. H., Davidson, I., & Alwine, J. C. (1993). Transcriptional Activation by Simian Virus 40 Large T Antigen: Interactions with Multiple Components of the Transcription Complex. *Molecular and Cellular Biology*, *13*(2), 961–969. https://doi.org/10.1128/mcb.13.2.961-969.1993

Gu, B., Guenther, C. M., & Seth, A. (2020). *Adeno-associated virus (AAV) producer cell line and related methods*.

Gu, B., & Wang, J. (2022). *Adeno-Associated Virus (AAV) Production*.

Gwizdek, C., Ossareh-Nazari, B., Brownawell, A. M., Doglio, A., Bertrand, E., Macara, I. G., & Dargemont, C. (2003). Exportin-5 mediates nuclear export of minihelix-containing RNAs. *Journal of Biological Chemistry*, *278*(8), 5505–5508. https://doi.org/10.1074/jbc.C200668200

Halbert, D. N., Cutt, J. R., & Shenk, T. (1985). Adenovirus early region 4 encodes functions required for efficient DNA replication, late gene expression, and host cell shutoff. *Journal of Virology*, *56*(1), 250–257. https://doi.org/10.1128/jvi.56.1.250-257.1985

Hamilton, H., Gomos, J., Berns, K. I., & Falck-Pedersen, E. (2004). Adeno-Associated Virus Site-Specific Integration and AAVS1 Disruption. *Journal of Virology*, *78*(15), 7874–7882. https://doi.org/10.1128/jvi.78.15.7874-7882.2004

Hansen, J., Qing, K., Kwon, H. J., Mah, C., & Srivastava, A. (2000). Impaired intracellular trafficking of adeno-associated virus type 2 vectors limits efficient transduction of murine fibroblasts. *Journal of Virology*, *74*(2), 992–996. https://doi.org/10.1128/JVI.74.2.992-996.2000

Hasegawa, K., & Nakatsuji, N. (2002). Insulators prevent transcriptional interference between two promoters in a double gene construct for transgenesis. *FEBS Letters*, *520*(1–3), 47–52. https://doi.org/10.1016/S0014-5793(02)02761-8

Hastie, E., & Samulski, R. J. (2015). Adeno-Associated Virus at 50: A Golden Anniversary of Discovery, Research, and Gene Therapy Success—A Personal Perspective. *Human Gene Therapy*, *26*(5), 257. https://doi.org/10.1089/HUM.2015.025

Heilbronn, R., Bürkle, A., Stephan, S., & zur Hausen, H. (1990). The adeno-associated virus rep gene suppresses herpes simplex virus-induced DNA amplification. *Journal of Virology*, *64*(6), 3012–3018. https://doi.org/10.1128/jvi.64.6.3012-3018.1990

Heller, M. J. (2002). DNA Microarray Technology: Devices, Systems, and Applications. *Annual Review of Biomedical Engineering*, *4*(1), 129–153. https://doi.org/10.1146/annurev.bioeng.4.020702.153438

Hemström, C., Virtanen, A., Bridge, E., Ketner, G., & Pettersson, U. (1991). Adenovirus E4-dependent activation of the early E2 promoter is insufficient to promote the early-to-late-phase transition. *Journal of Virology*, *65*(3), 1440–1449. https://doi.org/10.1128/jvi.65.3.1440-1449.1991

Hermonat, P. L. (1994). Down-regulation of the human c-fos and c-myc proto-oncogene promoters by adeno-associated virus Rep78. *Cancer Letters*, *81*(2), 129–136. https://doi.org/10.1016/0304-3835(94)90193-7

Hermonat, P. L., & Muzyczka, N. (1984). Use of adeno-associated virus as a mammalian DNA cloning vector: Transduction of neomycin resistance into mammalian tissue culture cells. *Proceedings of the National Academy of Sciences of the United States of America*, *81*(20 I), 6466–6470. https://doi.org/10.1073/pnas.81.20.6466

Hidalgo, P., Ip, W. H., Dobner, T., & Gonzalez, R. A. (2019). The biology of the adenovirus E1B 55K protein. In *FEBS Letters* (Vol. 593, Issue 24, pp. 3504–3517). Wiley Blackwell. https://doi.org/10.1002/1873-3468.13694

Hiebert, S. W., Lipp, M., & Nevins, J. R. (1989). E1A-dependent trans-activation of the human MYC promoter is mediated by the E2F factor. *Proceedings of the National Academy of Sciences of the United States of America*, *86*(10), 3594–3598. https://doi.org/10.1073/PNAS.86.10.3594

Higashimoto, T., Urbinati, F., Perumbeti, A., Jiang, G., Zarzuela, A., Chang, L.-J., Kohn, D. B., & Malik, P. (2007). The woodchuck hepatitis virus post-transcriptional regulatory element reduces readthrough transcription from retroviral vectors. *Gene Therapy*, *14*(17), 1298–1304. https://doi.org/10.1038/sj.gt.3302979

Hildinger, M., Baldi, L., Stettler, M., & Wurm, F. M. (2007). High-titer, serum-free production of adeno-associated virus vectors by polyethyleneimine-mediated

plasmid transfection in mammalian suspension cells. *Biotechnology Letters*, *29*(11), 1713–1721. https://doi.org/10.1007/S10529-007-9441-3/TABLES/2

Hildinger, Wilson, & Auricchio. (2001). *Recombinant aav vectors with AAV5 capsids and AAV5 vectors pseudotyped in heterologous capsids*.

Hobson, D. J., Wei, W., Steinmetz, L. M., & Svejstrup, J. Q. (2012). RNA Polymerase II Collision Interrupts Convergent Transcription. *Molecular Cell*, *48*(3), 365–374. https://doi.org/10.1016/j.molcel.2012.08.027

Hood, I. V., Gordon, J. M., Bou-Nader, C., Henderson, F. E., Bahmanjah, S., & Zhang, J. (2019). Crystal structure of an adenovirus virus-associated RNA. *Nature Communications*, *10*(1), 2871. https://doi.org/10.1038/s41467-019-10752-6

Hoppe, A., Beech, S. J., Dimmock, J., & Leppard, K. N. (2006). Interaction of the Adenovirus Type 5 E4 Orf3 Protein with Promyelocytic Leukemia Protein Isoform II Is Required for ND10 Disruption. *Journal of Virology*, *80*(6), 3042–3049. https://doi.org/10.1128/JVI.80.6.3042-3049.2006

Hörer, M., Weger, S., Butz, K., Hoppe-Seyler, F., Geisen, C., & Kleinschmidt, J. A. (1995). Mutational analysis of adeno-associated virus Rep protein-mediated inhibition of heterologous and homologous promoters. *Journal of Virology*, *69*(9), 5485–5496. https://doi.org/10.1128/jvi.69.9.5485-5496.1995

Hornstein, B. D., Roman, D., Arévalo-Soliz, L. M., Engevik, M. A., & Zechiedrich, L. (2016). Effects of Circular DNA Length on Transfection Efficiency by Electroporation into HeLa Cells. *PLOS ONE*, *11*(12), e0167537. https://doi.org/10.1371/journal.pone.0167537

Horwitz, G. A., Zhang, K., McBrian, M. A., Grunstein, M., Kurdistani, S. K., & Berk, A. J. (2008). Adenovirus small e1a alters global patterns of histone modification. *Science (New York, N.Y.)*, *321*(5892), 1084–1085. https://doi.org/10.1126/science.1155544

Howe, S. J., Mansour, M. R., Schwarzwaelder, K., Bartholomae, C., Hubank, M., Kempski, H., Brugman, M. H., Pike-Overzet, K., Chatters, S. J., De Ridder, D., Gilmour, K. C., Adams, S., Thornhill, S. I., Parsley, K. L., Staal, F. J. T., Gale, R. E., Linch, D. C., Bayford, J., Brown, L., … Thrasher, A. J. (2008). Insertional mutagenesis combined with acquired somatic mutations causes leukemogenesis following gene therapy of SCID-X1 patients. *The Journal of Clinical Investigation*, *118*(9), 3143–3150. https://doi.org/10.1172/JCI35798

Huang, D. H., Horikoshi, M., & Roeder, R. G. (1988). Activation of the adenovirus EIIa late promoter by a single-point mutation which enhances binding of

transcription factor IID. *Journal of Biological Chemistry*, *263*(25), 12596–12601. https://doi.org/10.1016/S0021-9258(18)37796-2

Huang, M. M., & Hearing, P. (1989a). Adenovirus early region 4 encodes two gene products with redundant effects in lytic infection. *Journal of Virology*, *63*(6), 2605–2615. https://doi.org/10.1128/jvi.63.6.2605-2615.1989

Huang, M. M., & Hearing, P. (1989b). The adenovirus early region 4 open reading frame 6/7 protein regulates the DNA binding activity of the cellular transcription factor, E2F, through a direct complex. *Genes & Development*, *3*(11), 1699–1710. https://doi.org/10.1101/gad.3.11.1699

Hüser, D., Gogol-Döring, A., Lutter, T., Weger, S., Winter, K., Hammer, E. M., Cathomen, T., Reinert, K., & Heilbronn, R. (2010). Integration Preferences of Wildtype AAV-2 for Consensus Rep-Binding Sites at Numerous Loci in the Human Genome. *PLOS Pathogens*, *6*(7), e1000985. https://doi.org/10.1371/JOURNAL.PPAT.1000985

Hüser, D., Weger, S., & Heilbronn, R. (2002). Kinetics and Frequency of Adeno-Associated Virus Site-Specific Integration into Human Chromosome 19 Monitored by Quantitative Real-Time PCR. *Journal of Virology*, *76*(15), 7554–7559. https://doi.org/10.1128/jvi.76.15.7554-7559.2002

Imperiale, M. J., & Nevins, J. R. (1984). Adenovirus 5 E2 transcription unit: an E1A-inducible promoter with an essential element that functions independently of position or orientation. *Molecular and Cellular Biology*, *4*(5), 875–882. https://doi.org/10.1128/MCB.4.5.875

James, J. A., Escalante, C. R., Yoon-Robarts, M., Edwards, T. A., Linden, R. M., & Aggarwal, A. K. (2003). Crystal structure of the SF3 helicase from adeno-associated virus type 2. *Structure*, *11*(8), 1025–1035. https://doi.org/10.1016/S0969-2126(03)00152-7

Janik, J. E., Huston, M. M., Cho, K., & Rose, J. A. (1989). Efficient synthesis of adeno-associated virus structural proteins requires both adenovirus DNA binding protein and VA I RNA. *Virology*, *168*(2), 320–329. https://doi.org/10.1016/0042-6822(89)90272-9

Janik, J. E., Huston, M. M., & Rose, J. A. (1981). Locations of adenovirus genes required for the replication of adenovirus-associated virus. *Proceedings of the National Academy of Sciences of the United States of America*, *78*(3 I), 1925–1929. https://doi.org/10.1073/pnas.78.3.1925

Jiang, Z., & Sharfstein, S. T. (2008). Sodium butyrate stimulates monoclonal antibody over-expression in CHO cells by improving gene accessibility. *Biotechnology and Bioengineering*, *100*(1), 189–194. https://doi.org/10.1002/bit.21726

Jin, X., Liu, L., Nass, S., O'Riordan, C., Pastor, E., & Zhang, X. K. (2017). Direct Liquid Chromatography/Mass Spectrometry Analysis for Complete Characterization of Recombinant Adeno-Associated Virus Capsid Proteins. *Human Gene Therapy Methods*, *28*(5), 255–267. https://doi.org/10.1089/hgtb.2016.178

Jing, X. J., Kalman-Maltese, V., Cao, X., Yang, Q., & Trempe, J. P. (2001). Inhibition of adenovirus cytotoxicity, replication, and E2a gene expression by adeno-associated virus. *Virology*, *291*(1), 140–151. https://doi.org/10.1006/viro.2001.1192

Johari, Y. B., Brown, A. J., Alves, C. S., Zhou, Y., Wright, C. M., Estes, S. D., Kshirsagar, R., & James, D. C. (2019). CHO genome mining for synthetic promoter design. *Journal of Biotechnology*, *294*, 1–13. https://doi.org/10.1016/j.jbiotec.2019.01.015

Johari, Y. B., Estes, S. D., Alves, C. S., Sinacore, M. S., & James, D. C. (2015). Integrated cell and process engineering for improved transient production of a "difficult-to-express" fusion protein by CHO cells. *Biotechnology and Bioengineering*, *112*(12), 2527–2542. https://doi.org/10.1002/BIT.25687

Johari, Y. B., Jaffé, S. R. P., Scarrott, J. M., Johnson, A. O., Mozzanino, T., Pohle, T. H., Maisuria, S., Bhayat-Cammack, A., Lambiase, G., Brown, A. J., Tee, K. L., Jackson, P. J., Wong, T. S., Dickman, M. J., Sargur, R. B., & James, D. C. (2021). Production of trimeric SARS-CoV-2 spike protein by CHO cells for serological COVID-19 testing. *Biotechnology and Bioengineering*, *118*(2), 1013–1021. https://doi.org/10.1002/bit.27615

Johari, Y. B., Mercer, A. C., Liu, Y., Brown, A. J., & James, D. C. (2021). Design of synthetic promoters for controlled expression of therapeutic genes in retinal pigment epithelial cells. *Biotechnology and Bioengineering*, *118*(5), 2001–2015. https://doi.org/10.1002/BIT.27713

Johari, Y. B., Scarrott, J. M., Pohle, T. H., Liu, P., Mayer, A., Brown, A. J., & James, D. C. (2022). Engineering of the CMV promoter for controlled expression of recombinant genes in HEK293 cells. *Biotechnology Journal*, *17*(8), 2200062. https://doi.org/10.1002/biot.202200062

Johnson, A. O., Fowler, S. B., Webster, C. I., Brown, A. J., & James, D. C. (2022). Bioinformatic Design of Dendritic Cell-Specific Synthetic Promoters. *ACS*

Synthetic Biology, *11*(4), 1613–1626. https://doi.org/10.1021/acssynbio.2c00027

Johnson, D. G., Ohtani, K., & Nevins, J. R. (1994). Autoregulatory control of E2F1 expression in response to positive and negative regulators of cell cycle progression. *Genes & Development*, *8*(13), 1514–1525. https://doi.org/10.1101/gad.8.13.1514

Johnson, D. G., Schwarz, J. K., Cress, W. D., & Nevins, J. R. (1993). Expression of transcription factor E2F1 induces quiescent cells to enter S phase. *Nature*, *365*(6444), 349–352. https://doi.org/10.1038/365349a0

Johnson, J. S., & Samulski, R. J. (2009). Enhancement of Adeno-Associated Virus Infection by Mobilizing Capsids into and Out of the Nucleolus. *Journal of Virology*, *83*(6), 2632–2644. https://doi.org/10.1128/JVI.02309-08

Jones, N., & Shenk, T. (1979). An adenovirus type 5 early gene function regulates expression of other early viral genes. *Proceedings of the National Academy of Sciences*, *76*(8), 3665–3669. https://doi.org/10.1073/pnas.76.8.3665

Journal of Gene Medicine. (2023, June 1). *Gene Therapy Clinical Trials Worldwide - Wiley*. Wiley. https://a873679.fmphost.com/fmi/webd/GTCT

Junod, S. L., Saredy, J., & Yang, W. (2021). Nuclear Import of Adeno-Associated Viruses Imaged by High-Speed Single-Molecule Microscopy. *Viruses*, *13*(2), 167. https://doi.org/10.3390/v13020167

Juven-Gershon, T., Cheng, S., & Kadonaga, J. T. (2006). Rational design of a super core promoter that enhances gene expression. *Nature Methods*, *3*(11), 917–922. https://doi.org/10.1038/nmeth937

Kaji, E. H., & Leiden, J. M. (2001). Gene and stem cell therapies. *Journal of the American Medical Association*, *285*(5), 545–550. https://doi.org/10.1001/jama.285.5.545

Karreth, F. A., Tay, Y., & Pandolfi, P. P. (2014). Target competition: Transcription factors enter the limelight. *Genome Biology*, *15*(4), 1–3. https://doi.org/10.1186/GB4174/FIGURES/1

Kelich, J. M., Ma, J., Dong, B., Wang, Q., Chin, M., Magura, C. M., Xiao, W., & Yang, W. (2015). Super-resolution imaging of nuclear import of adeno-associated virus in live cells. *Molecular Therapy - Methods & Clinical Development*, *2*, 15047. https://doi.org/10.1038/mtm.2015.47

Khatwani, S. L., Pavlova, A., & Pirot, Z. (2021). Anion-exchange HPLC assay for separation and quantification of empty and full capsids in multiple adeno-

associated virus serotypes. *Molecular Therapy - Methods & Clinical Development*, *21*, 548–558. https://doi.org/10.1016/J.OMTM.2021.04.003

Kim, M., O'Callaghan, P. M., Droms, K. A., & James, D. C. (2011). A mechanistic understanding of production instability in CHO cell lines expressing recombinant monoclonal antibodies. *Biotechnology and Bioengineering*, *108*(10), 2434–2446. https://doi.org/10.1002/bit.23189

Kim, S.-Y., Lee, J.-H., Shin, H.-S., Kang, H.-J., & Kim, Y.-S. (2002). The human elongation factor 1 alpha (EF-1α) first intron highly enhances expression of foreign genes from the murine cytomegalovirus promoter. *Journal of Biotechnology*, *93*(2), 183–187. https://doi.org/10.1016/S0168-1656(01)00388-1

King, C. R., Zhang, A., Tessier, T. M., Gameiro, S. F., & Mymryk, J. S. (2018). Hacking the Cell: Network Intrusion and Exploitation by Adenovirus E1A. *MBio*, *9*(3). https://doi.org/10.1128/mBio.00390-18

King, J. A., Dubielzig, R., Grimm, D., & Kleinschmidt, J. A. (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *The EMBO Journal*, *20*(12), 3282–3291. https://doi.org/10.1093/emboj/20.12.3282

Kitada, T., DiAndreth, B., Teague, B., & Weiss, R. (2018). Programming gene and engineered-cell therapies with synthetic biology. *Science*, *359*(6376). https://doi.org/10.1126/SCIENCE.AAD1067

Kitajewski, J., Schneider, R. J., Safer, B., Munemitsu, S. M., Samuel, C. E., Thimmappaya, B., & Shenk, T. (1986). Adenovirus VAI RNA antagonizes the antiviral action of interferon by preventing activation of the interferon-induced eIF-2α kinase. *Cell*, *45*(2), 195–200. https://doi.org/10.1016/0092-8674(86)90383-1

Klamroth, R., Hayes, G., Andreeva, T., Gregg, K., Suzuki, T., Mitha, I. H., Hardesty, B., Shima, M., Pollock, T., Slev, P., Oldenburg, J., Ozelo, M. C., Stieltjes, N., Castet, S. M., Mahlangu, J., Peyvandi, F., Kazmi, R., Schved, J. F., Leavitt, A. D., … Wong, W. Y. (2022). Global Seroprevalence of Pre-existing Immunity Against AAV5 and Other AAV Serotypes in People with Hemophilia A. *Human Gene Therapy*, *33*(7–8), 432–441. https://doi.org/10.1089/HUM.2021.287/SUPPL_FILE/SUPPL_TABLES3.DOCX

Kleinberger, T. (2020). Biology of the adenovirus E4orf4 protein: from virus infection to cancer cell death. *FEBS Letters*, *594*(12), 1891–1917. https://doi.org/10.1002/1873-3468.13704

Kleinberger, T., & Shenk, T. (1993). Adenovirus E4orf4 protein binds to protein phosphatase 2A, and the complex down regulates E1A-enhanced junB

transcription. *Journal of Virology*, *67*(12), 7556–7560. https://doi.org/10.1128/jvi.67.12.7556-7560.1993

Klessig, D. F., Brough, D. E., & Cleghon, V. (1984). Introduction, Stable Integration, and Controlled Expression of a Chimeric Adenovirus Gene Whose Product Is Toxic to the Recipient Human Cell. *Molecular and Cellular Biology*, *4*(7), 1354–1362. https://doi.org/10.1128/mcb.4.7.1354-1362.1984

Koczot, F. J., Carter, B. J., Garon, C. F., & Rose, J. A. (1973). Self-complementarity of terminal sequences within plus or minus strands of adenovirus-associated virus DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *70*(1), 215–219. https://doi.org/10.1073/pnas.70.1.215

Köhler, M., Görlich, D., Hartmann, E., & Franke, J. (2001). Adenoviral E1A Protein Nuclear Import Is Preferentially Mediated by Importin α3 in Vitro. *Virology*, *289*(2), 186–191. https://doi.org/10.1006/viro.2001.1151

Kondo, S., Yoshida, K., Suzuki, M., Saito, I., & Kanegae, Y. (2014). Adenovirus-Encoding Virus-Associated RNAs Suppress HDGF Gene Expression to Support Efficient Viral Replication. *PLOS ONE*, *9*(10), e108627. https://doi.org/10.1371/JOURNAL.PONE.0108627

Kotin, R. M. (1994). Prospects for the use of adeno-associated virus as a vector for human gene therapy. In *Human Gene Therapy* (Vol. 5, Issue 7, pp. 793–801). Mary Ann Liebert Inc. https://doi.org/10.1089/hum.1994.5.7-793

Kotin, R. M., Menninger, J. C., Ward, D. C., & Berns, K. I. (1991). Mapping and direct visualization of a region-specific viral DNA integration site on chromosome 19q13-qter. *Genomics*, *10*(3), 831–834. https://doi.org/10.1016/0888-7543(91)90470-Y

Kotin, R. M., Siniscalco, M., Samulski, R. J., Zhu, X., Hunter, L., Laughlin, C. A., McLaughlin, S., Muzyczka, N., Rocchi, M., & Berns, K. I. (1990). Site-specific integration by adeno-associated virus. *Proceedings of the National Academy of Sciences of the United States of America*, *87*(6), 2211–2215. https://doi.org/10.1073/pnas.87.6.2211

Kovesdi, I., Reichel, R., & Nevins, J. R. (1987). Role of an adenovirus E2 promoter binding factor in E1A-mediated coordinate gene control. *Proceedings of the National Academy of Sciences*, *84*(8), 2180–2184. https://doi.org/10.1073/pnas.84.8.2180

Kreiss, P. (1999). Plasmid DNA size does not affect the physicochemical properties of lipoplexes but modulates gene transfer efficiency. *Nucleic Acids Research*, *27*(19), 3792–3798. https://doi.org/10.1093/nar/27.19.3792

Kusaczuk, M. (2019). Tauroursodeoxycholate-Bile Acid with Chaperoning Activity: Molecular and Cellular Effects and Therapeutic Perspectives. *Cells*, *8*(12). https://doi.org/10.3390/CELLS8121471

Kyöstiö, S. R., Owens, R. A., Weitzman, M. D., Antoni, B. A., Chejanovsky, N., & Carter, B. J. (1994). Analysis of adeno-associated virus (AAV) wild-type and mutant Rep proteins for their abilities to negatively regulate AAV p5 and p19 mRNA levels. *Journal of Virology*, *68*(5), 2947–2957. https://doi.org/10.1128/jvi.68.5.2947-2957.1994

Kyöstiö, S. R., Wonderling, R. S., & Owens, R. A. (1995). Negative regulation of the adeno-associated virus (AAV) P5 promoter involves both the P5 rep binding site and the consensus ATP-binding motif of the AAV Rep68 protein. *Journal of Virology*, *69*(11), 6787–6796. https://doi.org/10.1128/jvi.69.11.6787-6796.1995

Lackner, D. F., & Muzyczka, N. (2002). Studies of the Mechanism of Transactivation of the Adeno-Associated Virus p19 Promoter by Rep Protein. *Journal of Virology*, *76*(16), 8225–8235. https://doi.org/10.1128/jvi.76.16.8225-8235.2002

Lamarche, B. J., Orazio, N. I., & Weitzman, M. D. (2010). The MRN complex in double-strand break repair and telomere maintenance. In *FEBS Letters* (Vol. 584, Issue 17, pp. 3682–3695). NIH Public Access. https://doi.org/10.1016/j.febslet.2010.07.029

Lan Tee, K., Jackson, P. J., Scarrott, J. M., P Jaffe, S. R., Johnson, A. O., Johari, Y., Pohle, T. H., Mozzanino, T., Price, J., Grinham, J., Brown, A., Nicklin, M. J., James, D. C., Dickman, M. J., & Seng Wong, T. (2020). Purification of recombinant SARS-CoV-2 spike, its receptor binding domain, and CR3022 mAb for serological assay. *BioRxiv*, 2020.07.31.231282. https://doi.org/10.1101/2020.07.31.231282

Lanoix, J., & Acheson, N. H. (1988). A rabbit beta-globin polyadenylation signal directs efficient termination of transcription of polyomavirus DNA. *The EMBO Journal*, *7*(8), 2515. https://doi.org/10.1002/J.1460-2075.1988.TB03099.X

Lapteva, L., Purohit-Sheth, T., Serabian, M., & Puri, R. K. (2020). Clinical Development of Gene Therapies: The First Three Decades and Counting. *Molecular Therapy. Methods & Clinical Development*, *19*, 387. https://doi.org/10.1016/J.OMTM.2020.10.004

Laughlin, C. A., Jones, N., & Carter, B. J. (1982). Effect of deletions in adenovirus early region 1 genes upon replication of adeno-associated virus. *Journal of Virology*, *41*(3), 868–876. https://doi.org/10.1128/jvi.41.3.868-876.1982

Lee, J. S., Galvin, K. M., See, R. H., Eckner, R., Livingston, D., Moran, E., & Shi, Y. (1995). Relief of YY1 transcriptional repression by adenovirus E1A is mediated by E1A-associated protein p300. *Genes and Development*, *9*(10), 1188–1198. https://doi.org/10.1101/gad.9.10.1188

Leppard, K. N. (1997). E4 gene function in adenovirus, adenovirus vector and adeno-associated virus infections. *Journal of General Virology*, *78*(9), 2131–2138. https://doi.org/10.1099/0022-1317-78-9-2131

Leppard, K. N. (2014). Adenoviruses: Molecular Biology. In *Reference Module in Biomedical Sciences*. Elsevier. https://doi.org/10.1016/b978-0-12-801238-3.02525-3

Li, C., & Samulski, R. J. (2020). Engineering adeno-associated virus vectors for gene therapy. In *Nature Reviews Genetics* (Vol. 21, Issue 4, pp. 255–272). Nature Research. https://doi.org/10.1038/s41576-019-0205-4

Li, J., Huang, Z., Sun, X., Yang, P., & Zhang, Y. (2006). Understanding the enhanced effect of dimethyl sulfoxide on hepatitis B surface antigen expression in the culture of Chinese hamster ovary cells on the basis of proteome analysis. *Enzyme and Microbial Technology*, *38*(3–4), 372–380. https://doi.org/10.1016/j.enzmictec.2005.05.016

Li, J., Xiao, X., & Samulski, R. J. (1997). Role for Highly Regulated rep Gene Expression in Adeno-Associated Virus Vector Production Recent success achieving long-term in vivo gene transfer without a significant immune response by using adeno-associated virus (AAV) vectors (X. *Journal of Virology*, *71*(7), 5236–5243. http://jvi.asm.org/

Lienert, F., Lohmueller, J. J., Garg, A., & Silver, P. A. (2014). Synthetic biology in mammalian cells: next generation research tools and therapeutics. *NATURE REVIEWS | MOLECULAR CELL BIOLOGY*, *15*, 95. https://doi.org/10.1038/nrm3738

Lin, Y.-C., Boone, M., Meuris, L., Lemmens, I., Van Roy, N., Soete, A., Reumers, J., Moisse, M., Plaisance, S., Drmanac, R., Chen, J., Speleman, F., Lambrechts, D., Van de Peer, Y., Tavernier, J., & Callewaert, N. (2014). Genome dynamics of the human embryonic kidney 293 lineage in response to cell biology manipulations. *Nature Communications*, *5*(1), 4767. https://doi.org/10.1038/ncomms5767

Liu, C.-H., & Chen, L.-H. (2007). Promotion of recombinant macrophage colony stimulating factor production by dimethyl sulfoxide addition in Chinese

hamster ovary cells. *Journal of Bioscience and Bioengineering*, *103*(1), 45–49. https://doi.org/10.1263/jbb.103.45

Liu, P., Mayer, A., James, D. C., Pohle, T. H., & Johari, Y. B. (2023). *Compositions and Methods for Recombinant AAV Production*.

Liu, Z., Chen, O., Wall, J. B. J., Zheng, M., Zhou, Y., Wang, L., Ruth Vaseghi, H., Qian, L., & Liu, J. (2017). Systematic comparison of 2A peptides for cloning multi-genes in a polycistronic vector. *Scientific Reports*, *7*(1), 2193. https://doi.org/10.1038/s41598-017-02460-2

Lowe, S. W., Ruley, H. E., Jacks, T., & Housman, D. E. (1993). p53-dependent apoptosis modulates the cytotoxicity of anticancer agents. *Cell*, *74*(6), 957–967. https://doi.org/10.1016/0092-8674(93)90719-7

Lu, S., & Cullen, B. R. (2004). Adenovirus VA1 noncoding RNA can inhibit small interfering RNA and MicroRNA biogenesis. *Journal of Virology*, *78*(23), 12868–12876. https://doi.org/10.1128/JVI.78.23.12868-12876.2004

Lundblad, J. R., Kwok, R. P. S., Laurance, M. E., Harter, M. L., & Goodman, R. H. (1995). Adenoviral ElA-associated protein p300 as a functional homologue of the transcriptional co-activator CBP. *Nature*, *374*(6517), 85–88. https://doi.org/10.1038/374085a0

Lutz, P., & Kedinger, C. (1996). Properties of the adenovirus IVa2 gene product, an effector of late-phase-dependent activation of the major late promoter. *Journal of Virology*, *70*(3), 1396–1405. https://doi.org/10.1128/jvi.70.3.1396-1405.1996

Lutz, P., Rosa-Calatrava, M., & Kedinger, C. (1997). The product of the adenovirus intermediate gene IX is a transcriptional activator. *Journal of Virology*, *71*(7), 5102–5109. https://doi.org/10.1128/jvi.71.7.5102-5109.1997

Lux, K., Goerlitz, N., Schlemminger, S., Perabo, L., Goldnau, D., Endell, J., Leike, K., Kofler, D. M., Finke, S., Hallek, M., & Buning, H. (2005). Green Fluorescent Protein-Tagged Adeno-Associated Virus Particles Allow the Study of Cytosolic and Nuclear Trafficking. *Journal of Virology*, *79*(18), 11776–11787. https://doi.org/10.1128/JVI.79.18.11776-11787.2005

Lyle, A., Stamatis, C., Linke, T., Hulley, M., Schmelzer, A., Turner, R., & Farid, S. S. (2023). Process economics evaluation and optimization of adeno-associated virus downstream processing. *Biotechnology and Bioengineering*. https://doi.org/10.1002/BIT.28402

Lyons, R. H., Ferguson, B. Q., & Rosenberg, M. (1987). Pentapeptide nuclear localization signal in adenovirus E1a. *Molecular and Cellular Biology*, *7*(7), 2451–2456. https://doi.org/10.1128/MCB.7.7.2451

Ma, Y., & Mathews, M. B. (1996). Structure, function, and evolution of adenovirus-associated RNA: a phylogenetic approach. *Journal of Virology*, *70*(8), 5083–5099. https://doi.org/10.1128/jvi.70.8.5083-5099.1996

Mader, A., Prewein, B., Zboray, K., Casanova, E., & Kunert, R. (2013). Exploration of BAC versus plasmid expression vectors in recombinant CHO cells. *Applied Microbiology and Biotechnology*, *97*(9), 4049–4054. https://doi.org/10.1007/s00253-012-4498-x

Mannervik, M., Fan, S., Ström, A.-C., Helin, K., & Akusjärvi, G. (1999). Adenovirus E4 Open Reading Frame 4-Induced Dephosphorylation Inhibits E1A Activation of the E2 Promoter and E2F-1-Mediated Transactivation Independently of the Retinoblastoma Tumor Suppressor Protein. *Virology*, *256*(2), 313–321. https://doi.org/10.1006/viro.1999.9663

Marton, M. J., Baim, S. B., Ornelles, D. A., & Shenk, T. (1990). The adenovirus E4 17-kilodalton protein complexes with the cellular transcription factor E2F, altering its DNA-binding properties and stimulating E1A-independent accumulation of E2 mRNA. *Journal of Virology*, *64*(5), 2345–2359. https://doi.org/10.1128/jvi.64.5.2345-2359.1990

Mary, B., Maurya, S., Arumugam, S., Kumar, V., & Jayandharan, G. R. (2019). Post-translational modifications in capsid proteins of recombinant adeno-associated virus (<scp>AAV</scp>) 1-rh10 serotypes. *The FEBS Journal*, *286*(24), 4964–4981. https://doi.org/10.1111/febs.15013

Matsushita, T., Elliger, S., Elliger, C., Podsakoff, G., Villarreal, L., Kurtzman, G., Iwaki, Y., & Colosi, P. (1998). Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Therapy*, *5*(7), 938–945. https://doi.org/10.1038/sj.gt.3300680

Matsushita, T., Okada, T., Inaba, T., Mizukami, H., Ozawa, K., & Colosi, P. (2004). The adenovirus E1A and E1B19K genes provide a helper function for transfection-based adeno-associated virus vector production. *Journal of General Virology*, *85*(8), 2209–2214. https://doi.org/10.1099/vir.0.79940-0

Maucksch, C., Connor, B., & Rudolph, C. (2013). Plasmid DNA Concatemers: Influence of Plasmid Structure on Transfection Efficiency. In *Minicircle and Miniplasmid DNA Vectors* (pp. 59–69). Wiley. https://doi.org/10.1002/9783527670420.ch5

Maurer, A. C., Pacouret, S., Cepeda Diaz, A. K., Blake, J., Andres-Mateos, E., & Vandenberghe, L. H. (2018). The Assembly-Activating Protein Promotes Stability and Interactions between AAV's Viral Proteins to Nucleate Capsid Assembly. *Cell Reports*, *23*(6), 1817–1830. https://doi.org/10.1016/j.celrep.2018.04.026

Mayor, H. D., Torikai, K., Melnick, J. L., & Mandel, M. (1969). Plus and minus single-stranded DNA separately encapsidated in adeno-associated satellite virions. *Science*, *166*(3910), 1280–1282. https://doi.org/10.1126/science.166.3910.1280

McCarty, D. M., Young, S. M., & Samulski, R. J. (2004). Integration of adeno-associated virus (AAV) and recombinant AAV vectors. In *Annual Review of Genetics* (Vol. 38, Issue 1, pp. 819–845). Annual Reviews. https://doi.org/10.1146/annurev.genet.37.110801.143717

McCarty, D., Monahan, P., & Samulski, R. (2001). Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Therapy*, *8*(16), 1248–1254. https://doi.org/10.1038/sj.gt.3301514

McClements, M. E., & MacLaren, R. E. (2017). Adeno-associated Virus (AAV) Dual Vector Strategies for Gene Therapy Encoding Large Transgenes. *The Yale Journal of Biology and Medicine*, *90*(4), 611–623. http://www.ncbi.nlm.nih.gov/pubmed/29259525

McLaughlin, S. K., Collis, P., Hermonat, P. L., & Muzyczka, N. (1988). Adeno-associated virus general transduction vectors: analysis of proviral structures. *Journal of Virology*, *62*(6), 1963–1973. http://www.ncbi.nlm.nih.gov/pubmed/2835501

Meier, A. F., Fraefel, C., & Seyffert, M. (2020). The Interplay between Adeno-Associated Virus and Its Helper Viruses. *Viruses*, *12*(6), 662. https://doi.org/10.3390/v12060662

Melling, M. (2018). *The influence of SUMOylation on the adenoviral early region 4 protein Orf6/7*. https://ediss.sub.uni-hamburg.de/handle/ediss/7767

Merten, O.-W. (2016). AAV vector production: state of the art developments and remaining challenges. *Cell and Gene Therapy Insights*, *2*(5), 521–551. https://doi.org/10.18609/cgti.2016.067

Metcalf, J. P., Monick, M. M., Stinski, M. F., & Hunninghake, G. W. (1994). Adenovirus E1A 13S gene product up-regulates the cytomegalovirus major immediate early

promoter. *American Journal of Respiratory Cell and Molecular Biology*, *10*(4), 448–452. https://doi.org/10.1165/ajrcmb.10.4.8136160

Meyer, H. J., Turincio, R., Ng, S., Li, J., Wilson, B., Chan, P., Zak, M., Reilly, D., Beresini, M. H., & Wong, A. W. (2017). High throughput screening identifies novel, cell cycle-arresting small molecule enhancers of transient protein expression. *Biotechnology Progress*, *33*(6), 1579–1588. https://doi.org/10.1002/BTPR.2517

Michels, A., Ho, N., & Buchholz, C. J. (2022). Precision medicine: In vivo CAR therapy as a showcase for receptor-targeted vector platforms. *Molecular Therapy*, *30*(7), 2401–2415. https://doi.org/10.1016/J.YMTHE.2022.05.018

Mietzsch, M., Casteleyn, V., Weger, S., Zolotukhin, S., & Heilbronn, R. (2015). OneBac 2.0: Sf 9 Cell Lines for Production of AAV5 Vectors with Enhanced Infectivity and Minimal Encapsidation of Foreign DNA. *Human Gene Therapy*, *26*(10), 688–697. https://doi.org/10.1089/hum.2015.050

Mietzsch, M., Eddington, C., Jose, A., Hsi, J., Chipman, P., Henley, T., Choudhry, M., McKenna, R., & Agbandje-McKenna, M. (2021). Improved Genome Packaging Efficiency of Adeno-associated Virus Vectors Using Rep Hybrids. *Journal of Virology*, *95*(19), e0077321. https://doi.org/10.1128/JVI.00773-21

Miller, A. D. (1992). Human gene therapy comes of age. *Nature*, *357*(6378), 455–460. https://doi.org/10.1038/357455a0

Mitchell, A. M., & Samulski, R. J. (2013). Mechanistic Insights into the Enhancement of Adeno-Associated Virus Transduction by Proteasome Inhibitors. *Journal of Virology*, *87*(23), 13035–13041. https://doi.org/10.1128/JVI.01826-13

Mizuguchi, H., Xu, Z., Ishii-Watabe, A., Uchida, E., & Hayakawa, T. (2000). IRES-Dependent Second Gene Expression Is Significantly Lower Than Cap-Dependent First Gene Expression in a Bicistronic Vector. *Molecular Therapy*, *1*(4), 376–382. https://doi.org/10.1006/mthe.2000.0050

Morris, S. J., & Leppard, K. N. (2009). Adenovirus Serotype 5 L4-22K and L4-33K Proteins Have Distinct Functions in Regulating Late Gene Expression. *Journal of Virology*, *83*(7), 3049–3058. https://doi.org/10.1128/JVI.02455-08

Mouw, M. B., & Pintel, D. J. (2000). Adeno-associated virus RNAs appear in a temporal order and their splicing is stimulated during coinfection with adenovirus. *Journal of Virology*, *74*(21), 9878–9888. https://doi.org/10.1128/jvi.74.21.9878-9888.2000

Müller, U., Kleinberger, T., & Shenk, T. (1992). Adenovirus E4orf4 protein reduces phosphorylation of c-Fos and E1A proteins while simultaneously reducing the level of AP-1. *Journal of Virology*, *66*(10), 5867–5878. https://doi.org/10.1128/jvi.66.10.5867-5878.1992

Mulligan, R. C. (1993). The basic science of gene therapy. *Science*, *260*(5110), 926–932. https://doi.org/10.1126/science.8493530

Munteanu, A., Constante, M., Isalan, M., & Solé, R. V. (2010). Avoiding transcription factor competition at promoter level increases the chances of obtaining oscillation. *BMC Systems Biology*, *4*(1), 1–17. https://doi.org/10.1186/1752-0509-4-66/FIGURES/10

Murphy, M., Gomos-Klein, J., Stankic, M., & Falck-Pedersen, E. (2007). Adeno-Associated Virus Type 2 p5 Promoter: a Rep-Regulated DNA Switch Element Functioning in Transcription, Replication, and Site-Specific Integration. *Journal of Virology*, *81*(8), 3721–3730. https://doi.org/10.1128/jvi.02693-06

Muzyczka, N. (1992). Use of Adeno-Associated Virus as a General Transduction Vector for Mammalian Cells. In *Current topics in microbiology and immunology* (Vol. 158, pp. 97–129). Curr Top Microbiol Immunol. https://doi.org/10.1007/978-3-642-75608-5_5

Myers, M. W., & Carter, B. J. (1980). Assembly of adeno-associated virus. *Journal of Virology*, *102*(1), 71–82. http://www.ncbi.nlm.nih.gov/pubmed/6245509

Nada, S., & Trempe, J. P. (2002). Characterization of adeno-associated virus Rep protein inhibition of adenovirus E2a gene expression. *Virology*, *293*(2), 345–355. https://doi.org/10.1006/viro.2001.1286

Nakai, H., Storm, T. A., & Kay, M. A. (2000). Recruitment of Single-Stranded Recombinant Adeno-Associated Virus Vector Genomes and Intermolecular Recombination Are Responsible for Stable Transduction of Liver In Vivo. *Journal of Virology*, *74*(20), 9451–9463. https://doi.org/10.1128/jvi.74.20.9451-9463.2000

Nash, K., Chen, W., Salganik, M., & Muzyczka, N. (2009). Identification of Cellular Proteins That Interact with the Adeno-Associated Virus Rep Protein. *Journal of Virology*, *83*(1), 454–469. https://doi.org/10.1128/jvi.01939-08

Naso, M. F., Tomkowicz, B., Perry, W. L., & Strohl, W. R. (2017). Adeno-Associated Virus (AAV) as a Vector for Gene Therapy. *BioDrugs*, *31*(4), 317–334. https://doi.org/10.1007/s40259-017-0234-5

Natsoulis, G. (1997). *High efficiency helper system for AAV vector production*. https://patents.google.com/patent/US5622856A/en

Natsoulis, Kurtzman, & Colosi. (2005). *High-efficiency AAV helper functions*.

Nawaz, W., Huang, B., Xu, S., Li, Y., Zhu, L., Yiqiao, H., Wu, Z., & Wu, X. (2021). AAV-mediated in vivo CAR gene therapy for targeting human T-cell leukemia. *Blood Cancer Journal 2021 11:6*, *11*(6), 1–12. https://doi.org/10.1038/s41408-021-00508-1

Nayak, R., Farris, K. D., & Pintel, D. J. (2008). E4Orf6-E1B-55k-Dependent Degradation of De Novo-Generated Adeno-Associated Virus Type 5 Rep52 and Capsid Proteins Employs a Cullin 5-Containing E3 Ligase Complex. *Journal of Virology*, *82*(7), 3803–3808. https://doi.org/10.1128/JVI.02532-07

Nevins, J. R., Ginsberg, H. S., Blanchard, J.-M., Wilson, M. C., & Darnell, J. E. (1979). Regulation of the Primary Expression of the Early Adenovirus Transcription Units. *Journal of Virology*, *32*(3), 727–733. https://doi.org/10.1128/jvi.32.3.727-733.1979

Nguyen, T. N. T., Sha, S., Hong, M. S., Maloney, A. J., Barone, P. W., Neufeld, C., Wolfrum, J., Springs, S. L., Sinskey, A. J., & Braatz, R. D. (2021). Mechanistic model for production of recombinant adeno-associated virus via triple transfection of HEK293 cells. *Molecular Therapy - Methods & Clinical Development*, *21*, 642–655. https://doi.org/10.1016/j.omtm.2021.04.006

Nicholl, D. S. T. (2023). *An introduction to genetic engineering* (4th ed.). Cambridge University Press. https://books.google.com/books/about/An_Introduction_to_Genetic_Engineering.html?id=5SusEAAAQBAJ

Nicolson, S. C., Li, C., Hirsch, M. L., Setola, V., & Samulski, R. J. (2016). Identification and Validation of Small Molecules That Enhance Recombinant Adeno-associated Virus Transduction following High-Throughput Screens. *Journal of Virology*, *90*(16), 7019–7031. https://doi.org/10.1128/JVI.02953-15

Nicolson, S. C., & Samulski, R. J. (2014). Recombinant Adeno-Associated Virus Utilizes Host Cell Nuclear Import Machinery To Enter the Nucleus. *Journal of Virology*, *88*(8), 4132–4144. https://doi.org/10.1128/JVI.02660-13/ASSET/9E3E6DC3-567E-41B9-A380-CDF9D55E21DB/ASSETS/GRAPHIC/ZJV9990988450009.JPEG

Nieuwenhuis, B., Laperrousaz, E., Tribble, J. R., Verhaagen, J., Fawcett, J. W., Martin, K. R., Williams, P. A., & Osborne, A. (2023). Improving adeno-associated viral (AAV) vector-mediated transgene expression in retinal ganglion cells:

comparison of five promoters. *Gene Therapy*, *30*(6), 503–519. https://doi.org/10.1038/s41434-022-00380-z

Nirenberg, M. W. (1967). Will society be prepared? *Science (New York, N.Y.)*, *157*(3789), 633. https://doi.org/10.1126/SCIENCE.157.3789.633

Nonnenmacher, M., & Weber, T. (2012). Intracellular transport of recombinant adeno-associated virus vectors. *Gene Therapy*, *19*(6), 649–658. https://doi.org/10.1038/gt.2012.6

Nony, P., Tessier, J., Chadeuf, G., Ward, P., Giraud, A., Dugast, M., Linden, R. M., Moullier, P., & Salvetti, A. (2001). Novel *cis* -Acting Replication Element in the Adeno-Associated Virus Type 2 Genome Is Involved in Amplification of Integrated *rep-cap* Sequences. *Journal of Virology*, *75*(20), 9991–9994. https://doi.org/10.1128/JVI.75.20.9991-9994.2001

Ogasawara, Y., Urabe, M., & Ozawa, K. (1998). The Use of Heterologous Promoters for Adeno-Associated Virus (AAV) Protein Expression in AAV Vector Production. *Microbiology and Immunology*, *42*(3), 177–185. https://doi.org/10.1111/j.1348-0421.1998.tb02269.x

Ogawa, R., Kagiya, G., Kodaki, T., Fukuda, S., & Yamamoto, K. (2007). Construction of strong mammalian promoters by random *cis* -acting element elongation. *BioTechniques*, *42*(5), 628–633. https://doi.org/10.2144/000112436

Ogden, P. J., Kelsic, E. D., Sinai, S., & Church, G. M. (2019). Comprehensive AAV capsid fitness landscape reveals a viral gene and enables machine-guided design. *Science*, *366*(6469), 1139–1143. https://doi.org/10.1126/SCIENCE.AAW2900/SUPPL_FILE/AAW2900-OGDEN-SM.PDF

Ogston, P., Raj, K., & Beard, P. (2000). Productive Replication of Adeno-Associated Virus Can Occur in Human Papillomavirus Type 16 (HPV-16) Episome-Containing Keratinocytes and Is Augmented by the HPV-16 E2 Protein. *Journal of Virology*, *74*(8), 3494–3504. https://doi.org/10.1128/jvi.74.8.3494-3504.2000

Ohba, K., Sehara, Y., Enoki, T., Mineno, J., Ozawa, K., & Mizukami, H. (2023). Adeno-associated virus vector system controlling capsid expression improves viral quantity and quality. *IScience*, *26*(4), 106487. https://doi.org/10.1016/j.isci.2023.106487

Olanubi, O., Frost, J. R., Radko, S., & Pelka, P. (2017). Suppression of Type I Interferon Signaling by E1A via RuvBL1/Pontin. *Journal of Virology*, *91*(8). https://doi.org/10.1128/JVI.02484-16

O'Malley, R. (1986). A mechanism for the control of protein synthesis by adenovirus VA RNAI. *Cell*, *44*(3), 391–400. https://doi.org/10.1016/0092-8674(86)90460-5

Park, J. Y., Lim, B.-P., Lee, K., Kim, Y.-G., & Jo, E.-C. (2006). Scalable production of adeno-associated virus type 2 vectors via suspension transfection. *Biotechnology and Bioengineering*, *94*(3), 416–430. https://doi.org/10.1002/bit.20776

Patel, Y. D., Brown, A. J., Zhu, J., Rosignoli, G., Gibson, S. J., Hatton, D., & James, D. C. (2021). Control of Multigene Expression Stoichiometry in Mammalian Cells Using Synthetic Promoters. *ACS Synthetic Biology*, *10*(5), 1155–1165. https://doi.org/10.1021/ACSSYNBIO.0C00643/ASSET/IMAGES/LARGE/SB0C00643_0004.JPEG

Pechkovsky, A., Lahav, M., Bitman, E., Salzberg, A., & Kleinberger, T. (2013). E4orf4 induces PP2A- and Src-dependent cell death in Drosophila melanogaster and at the same time inhibits classic apoptosis pathways. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(19), E1724. https://doi.org/10.1073/PNAS.1220282110/-/DCSUPPLEMENTAL

Pechkovsky, A., Salzberg, A., & Kleinberger, T. (2013). The adenovirus E4orf4 protein induces a unique mode of cell death while inhibiting classical apoptosis. *Cell Cycle*, *12*(15), 2343. https://doi.org/10.4161/CC.25707

Pelka, P., Ablack, J. N. G., Fonseca, G. J., Yousef, A. F., & Mymryk, J. S. (2008). Intrinsic structural disorder in adenovirus E1A: a viral molecular hub linking multiple diverse processes. *Journal of Virology*, *82*(15), 7252–7263. https://doi.org/10.1128/JVI.00104-08

Penaud-Budloo, M., Le Guiner, C., Nowrouzi, A., Toromanoff, A., Chérel, Y., Chenuaud, P., Schmidt, M., von Kalle, C., Rolling, F., Moullier, P., & Snyder, R. O. (2008). Adeno-Associated Virus Vector Genomes Persist as Episomal Chromatin in Primate Muscle. *Journal of Virology*, *82*(16), 7875–7885. https://doi.org/10.1128/jvi.00649-08

Pereira, D. J., McCarty, D. M., & Muzyczka, N. (1997). The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. *Journal of Virology*, *71*(2), 1079–1088. http://www.ncbi.nlm.nih.gov/pubmed/8995628

Pereira, D. J., & Muzyczka, N. (1997a). The adeno-associated virus type 2 p40 promoter requires a proximal Sp1 interaction and a p19 CArG-like element to

facilitate Rep transactivation. *Journal of Virology*, *71*(6), 4300–4309. https://doi.org/10.1128/jvi.71.6.4300-4309.1997

Pereira, D. J., & Muzyczka, N. (1997b). The cellular transcription factor SP1 and an unknown cellular protein are required to mediate Rep protein activation of the adeno-associated virus p19 promoter. *Journal of Virology*, *71*(3), 1747–1756. https://doi.org/10.1128/jvi.71.3.1747-1756.1997

Petitclerc, D., Altai, J., Théron, M. C., Bearzotti, M., Bolifraud, P., Kann, G., Stinnakre, M. G., Pointu, H., Puissant, C., & Houdebine, L. M. (1995). The effect of various introns and transcription terminators on the efficiency of expression vectors in various cultured cell lines and in the mammary gland of transgenic mice. *Journal of Biotechnology*, *40*(3), 169–178. https://doi.org/10.1016/0168-1656(95)00047-T

Pham, P. L., Perret, S., Cass, B., Carpentier, E., St-Laurent, G., Bisson, L., Kamen, A., & Durocher, Y. (2005). Transient gene expression in HEK293 cells: Peptone addition posttransfection improves recombinant protein synthesis. *Biotechnology and Bioengineering*, *90*(3), 332–344. https://doi.org/10.1002/bit.20428

Pilder, S., Moore, M., Logan, J., & Shenk, T. (1986). The adenovirus E1B-55K transforming polypeptide modulates transport or cytoplasmic stabilization of viral and host cell mRNAs. *Molecular and Cellular Biology*, *6*(2), 470–476. https://doi.org/10.1128/mcb.6.2.470

Podsakoff, G., Wong, K. K., & Chatterjee, S. (1994). Efficient gene transfer into nondividing cells by adeno-associated virus-based vectors. *Journal of Virology*, *68*(9), 5656–5666. https://doi.org/10.1128/jvi.68.9.5656-5666.1994

Pollard, H., Remy, J.-S., Loussouarn, G., Demolombe, S., Behr, J.-P., & Escande, D. (1998). Polyethylenimine but Not Cationic Lipids Promotes Transgene Delivery to the Nucleus in Mammalian Cells. *Journal of Biological Chemistry*, *273*(13), 7507–7511. https://doi.org/10.1074/jbc.273.13.7507

Prösch, S., Stein, J., Staak, K., Liebenthal, C., Volk, H.-D., & Krüger, D. H. (1996). Inactivation of the Very Strong HCMV Immediate Early Promoter by DNA CpG Methylation *In Vitro*. *Biological Chemistry Hoppe-Seyler*, *377*(3), 195–202. https://doi.org/10.1515/bchm3.1996.377.3.195

Proudfoot, N. J. (2016). Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science (New York, N.Y.)*, *352*(6291), aad9926. https://doi.org/10.1126/SCIENCE.AAD9926

Punga, T., Darweesh, M., & Akusjärvi, G. (2020). Synthesis, Structure, and Function of Human Adenovirus Small Non-Coding RNAs. *Viruses*, *12*(10), 1182. https://doi.org/10.3390/v12101182

Qiao, C., Li, J., Skold, A., Zhang, X., & Xiao, X. (2002). Feasibility of Generating Adeno-Associated Virus Packaging Cell Lines Containing Inducible Adenovirus Helper Genes. *Journal of Virology*, *76*(4), 1904–1913. https://doi.org/10.1128/jvi.76.4.1904-1913.2002

Qin, J. Y., Zhang, L., Clift, K. L., Hulur, I., Xiang, A. P., Ren, B.-Z., & Lahn, B. T. (2010). Systematic Comparison of Constitutive Promoters and the Doxycycline-Inducible Promoter. *PLoS ONE*, *5*(5), e10611. https://doi.org/10.1371/journal.pone.0010611

Qiu, J., & Brown, K. E. (1999). A 110-kDa Nuclear Shuttle Protein, Nucleolin, Specifically Binds to Adeno-Associated Virus Type 2 (AAV-2) Capsid. *Virology*, *257*(2), 373–382. https://doi.org/10.1006/viro.1999.9664

Qiu, J., & Pintel, D. J. (2002). The Adeno-Associated Virus Type 2 Rep Protein Regulates RNA Processing via Interaction with the Transcription Template. *Molecular and Cellular Biology*, *22*(11), 3639–3652. https://doi.org/10.1128/MCB.22.11.3639-3652.2002

Qu, Y., Liu, Y., Noor, A., Tran, J., & Li, R. (2019). Characteristics and advantages of adeno-associated virus vector-mediated gene therapy for neurodegenerative diseases. *Neural Regeneration Research*, *14*(6), 931. https://doi.org/10.4103/1673-5374.250570

Querido, E., Marcellus, R. C., Lai, A., Charbonneau, R., Teodoro, J. G., Ketner, G., & Branton, P. E. (1997). Regulation of p53 levels by the E1B 55-kilodalton protein and E4orf6 in adenovirus-infected cells. *Journal of Virology*, *71*(5), 3788–3798. https://doi.org/10.1128/jvi.71.5.3788-3798.1997

Raguram, A., Banskota, S., & Liu, D. R. (2022). Therapeutic in vivo delivery of gene editing agents. *Cell*, *185*(15), 2806–2827. https://doi.org/10.1016/J.CELL.2022.03.045

Raj, K., Ogston, P., & Beard, P. (2001). Virus-mediated killing of cells that lack p53 activity. *Nature*, *412*(6850), 914–917. https://doi.org/10.1038/35091082

Rao, L., Debbas, M., Sabbatini, P., Hockenbery, D., Korsmeyer, S., & White, E. (1992). The adenovirus E1A proteins induce apoptosis, which is inhibited by the E1B 19-kDa and Bcl-2 proteins. *Proceedings of the National Academy of Sciences*, *89*(16), 7742–7746. https://doi.org/10.1073/pnas.89.16.7742

Raposo, V. L. (2019). The First Chinese Edited Babies: A Leap of Faith in Science. *JBRA Assisted Reproduction*, *23*(3), 197. https://doi.org/10.5935/1518-0557.20190042

Raup, A., Jérôme, V., Freitag, R., Synatschke, C. V., & Müller, A. H. E. (2016). Promoter, transgene, and cell line effects in the transfection of mammalian cells using PDMAEMA-based nano-stars. *Biotechnology Reports*, *11*, 53–61. https://doi.org/10.1016/j.btre.2016.05.003

Raychaudhuri, P., Bagchi, S., Neill, S. D., & Nevins, J. R. (1990). Activation of the E2F transcription factor in adenovirus-infected cells involves E1A-dependent stimulation of DNA-binding activity and induction of cooperative binding mediated by an E4 gene product. *Journal of Virology*, *64*(6), 2702–2710. https://doi.org/10.1128/jvi.64.6.2702-2710.1990

Reed Clark, K., Voulgaropoulou, F., & Johnson, P. R. (1996). A stable cell line carrying adenovirus-inducible rep and cap genes allows for infectivity titration of adeno-associated virus vectors. *Gene Therapy*, *3*(12), 1124–1132. https://europepmc.org/article/med/8986439

Ribet, D., & Cossart, P. (2010). Pathogen-Mediated Posttranslational Modifications: A Re-emerging Field. *Cell*, *143*(5), 694–702. https://doi.org/10.1016/j.cell.2010.11.019

Richardson, W. D., & Westphal, H. (1981). A cascade of adenovirus early functions is required for expression of adeno-associated virus. *Cell*, *27*(1 PART 2), 133–141. https://doi.org/10.1016/0092-8674(81)90367-6

Richardson, W. D., & Westphal, H. (1984). Requirement for either early region 1a or early region 1b adenovirus gene products in the helper effect for adeno-associated virus. *Journal of Virology*, *51*(2), 404–410. https://doi.org/10.1128/JVI.51.2.404-410.1984

Romanova, N., & Noll, T. (2018). Engineered and Natural Promoters and Chromatin-Modifying Elements for Recombinant Protein Expression in CHO Cells. *Biotechnology Journal*, *13*(3), 1700232. https://doi.org/10.1002/biot.201700232

Rose, J. A., Berns, K. I., Hoggan, M. D., & Koczot, F. J. (1969). Evidence for a single-stranded adenovirus-associated virus genome: formation of a DNA density hybrid on release of viral DNA. *Proceedings of the National Academy of Sciences of the United States of America*, *64*(3), 863–869. https://doi.org/10.1073/pnas.64.3.863

Rossini, M. (1983). The role of adenovirus early region 1A in the regulation of early regions 2A and 1B expression. *Virology*, *131*(1), 49–58. https://doi.org/10.1016/0042-6822(83)90532-9

Ruffing, M., Zentgraf, H., & Kleinschmidt, J. A. (1992). Assembly of viruslike particles by recombinant structural proteins of adeno-associated virus type 2 in insect cells. *Journal of Virology*, *66*(12), 6922–6930. http://www.ncbi.nlm.nih.gov/pubmed/1331503

Rumachik, N. G., Malaker, S. A., Poweleit, N., Maynard, L. H., Adams, C. M., Leib, R. D., Cirolia, G., Thomas, D., Stamnes, S., Holt, K., Sinn, P., May, A. P., & Paulk, N. K. (2020). Methods Matter: Standard Production Platforms for Recombinant AAV Produce Chemically and Functionally Distinct Vectors. *Molecular Therapy - Methods & Clinical Development*, *18*, 98–118. https://doi.org/10.1016/j.omtm.2020.05.018

Ryan, J. H., Zolotukhin, S., & Muzyczka, N. (1996). Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats. *Journal of Virology*, *70*(3), 1542–1553. https://doi.org/10.1128/jvi.70.3.1542-1553.1996

Sakaguchi, M., Watanabe, M., Kinoshita, R., Kaku, H., Ueki, H., Futami, J., Murata, H., Inoue, Y., Li, S. A., Huang, P., Putranto, E. W., Ruma, I. M. W., Nasu, Y., Kumon, H., & Huh, N. H. (2014). Dramatic increase in expression of a transgene by insertion of promoters downstream of the cargo gene. *Molecular Biotechnology*, *56*(7), 621–630. https://doi.org/10.1007/S12033-014-9738-0/FIGURES/5

Samulski, R. J., Berns, K. I., Tan, M., & Muzyczka, N. (1982). Cloning of adeno-associated virus into pBR322: rescue of intact virus from the recombinant plasmid in human cells. *Proceedings of the National Academy of Sciences*, *79*(6), 2077–2081. https://doi.org/10.1073/pnas.79.6.2077

Samulski, R. J., Chang, L. S., & Shenk, T. (1987). A recombinant plasmid from which an infectious adeno-associated virus genome can be excised in vitro and its use to study viral replication. *Journal of Virology*, *61*(10), 3096–3101. https://doi.org/10.1128/jvi.61.10.3096-3101.1987

Samulski, R. J., Chang, L. S., & Shenk, T. (1989). Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *Journal of Virology*, *63*(9), 3822–3828. https://doi.org/10.1128/jvi.63.9.3822-3828.1989

Samulski, R. J., & Muzyczka, N. (2014). AAV-Mediated Gene Therapy for Research and Therapeutic Purposes. *Annual Review of Virology*, *1*(1), 427–451. https://doi.org/10.1146/annurev-virology-031413-085355

Samulski, R. J., & Shenk, T. (1988). Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. *Journal of Virology*, *62*(1), 206–210. https://doi.org/10.1128/jvi.62.1.206-210.1988

Samulski, R. J., Srivastava, A., Berns, K. I., & Muzyczka, N. (1983). Rescue of adeno-associated virus from recombinant plasmids: Gene correction within the terminal repeats of AAV. *Cell*, *33*(1), 135–143. https://doi.org/10.1016/0092-8674(83)90342-2

Sangare, K., Helmold Hait, S., Moore, M., Hogge, C., Hoang, T., Rahman, M. A., Venzon, D. J., LaBranche, C., Montefiori, D., Robert-Guroff, M., & Thomas, M. A. (2022). E4orf1 Suppresses E1B-Deleted Adenovirus Vaccine-Induced Immune Responses. *Vaccines*, *10*(2), 295. https://doi.org/10.3390/vaccines10020295

Sangare, K., Tuero, I., Rahman, M. A., Hoang, T., Miller-Novak, L. K., Vargas-Inchaustegui, D. A., Venzon, D. J., LaBranche, C., Montefiori, D. C., Robert-Guroff, M., & Thomas, M. A. (2021). The Immunological Impact of Adenovirus Early Genes on Vaccine-Induced Responses in Mice and Nonhuman Primates. *Journal of Virology*, *95*(7). https://doi.org/10.1128/JVI.02253-20/ASSET/7B85EF10-AFA1-47B1-B73F-C77CD42A75F6/ASSETS/IMAGES/LARGE/JVI.02253-20-F0006.JPG

Sanlioglu, S., Benson, P. K., Yang, J., Atkinson, E. M., Reynolds, T., & Engelhardt, J. F. (2000). Endocytosis and Nuclear Trafficking of Adeno-Associated Virus Type 2 Are Controlled by Rac1 and Phosphatidylinositol-3 Kinase Activation. *Journal of Virology*, *74*(19), 9184–9196. https://doi.org/10.1128/JVI.74.19.9184-9196.2000

Sassone-Corsi, P. (1988). Cyclic AMP induction of early adenovirus promoters involves sequences required for E1A trans-activation. *Proceedings of the National Academy of Sciences*, *85*(19), 7192–7196. https://doi.org/10.1073/pnas.85.19.7192

Sassone-Corsi, P. (2012). The Cyclic AMP Pathway. *Cold Spring Harbor Perspectives in Biology*, *4*(12), a011148–a011148. https://doi.org/10.1101/cshperspect.a011148

Saudan, P., Vlach, J., & Beard, P. (2000). Inhibition of S-phase progression by adenoassociated virus Rep78 protein is mediated by hypophosphorylated pRb. *EMBO Journal*, *19*(16), 4351–4361. https://doi.org/10.1093/emboj/19.16.4351

Scarrott, J. M., Johari, Y. B., Pohle, T. H., Liu, P., Mayer, A., & James, D. C. (2023). Increased recombinant adeno-associated virus production by HEK293 cells

using small molecule chemical additives. *Biotechnology Journal*, *18*(3), 2200450. https://doi.org/10.1002/biot.202200450

Schaley, J., O'Connor, R. J., Taylor, L. J., Bar-Sagi, D., & Hearing, P. (2000). Induction of the Cellular E2F-1 Promoter by the Adenovirus E4-6/7 Protein. *Journal of Virology*, *74*(5), 2084–2093. https://doi.org/10.1128/JVI.74.5.2084-2093.2000

Scheller, E. L., & Krebsbach, P. H. (2009). Gene Therapy: Design and Prospects for Craniofacial Regeneration. *Journal of Dental Research*, *88*(7), 585. https://doi.org/10.1177/0022034509337480

Schlabach, M. R., Hu, J. K., Li, M., & Elledge, S. J. (2010). Synthetic design of strong promoters. *Proceedings of the National Academy of Sciences*, *107*(6), 2538–2543. https://doi.org/10.1073/pnas.0914803107

Schlehofer, J. R., Ehrbar, M., & Hausen, H. Zur. (1986). Vaccinia virus, herpes simplex virus, and carcinogens induce DNA amplification in a human cell line and support replication of a helpervirus dependent parvovirus. *Virology*, *152*(1), 110–117. https://doi.org/10.1016/0042-6822(86)90376-4

Schmidt, M., Afione, S., & Kotin, R. M. (2000). Adeno-Associated Virus Type 2 Rep78 Induces Apoptosis through Caspase Activation Independently of p53. *Journal of Virology*, *74*(20), 9441–9450. https://doi.org/10.1128/jvi.74.20.9441-9450.2000

Schnepp, B. C., Clark, K. R., Klemanski, D. L., Pacak, C. A., & Johnson, P. R. (2003). Genetic Fate of Recombinant Adeno-Associated Virus Vector Genomes in Muscle. *Journal of Virology*, *77*(6), 3495–3504. https://doi.org/10.1128/jvi.77.6.3495-3504.2003

Schreiner, S., Kinkley, S., Bürck, C., Mund, A., Wimmer, P., Schubert, T., Groitl, P., Will, H., & Dobner, T. (2013). SPOC1-Mediated Antiviral Host Cell Response Is Antagonized Early in Human Adenovirus Type 5 Infection. *PLoS Pathogens*, *9*(11), e1003775. https://doi.org/10.1371/journal.ppat.1003775

Schwartz, R. A., Palacios, J. A., Cassell, G. D., Adam, S., Giacca, M., & Weitzman, M. D. (2007). The Mre11/Rad50/Nbs1 Complex Limits Adeno-Associated Virus Transduction and Replication. *Journal of Virology*, *81*(23), 12936–12945. https://doi.org/10.1128/jvi.01523-07

Schwartz, S. L., & Conn, G. L. (2019). RNA regulation of the antiviral protein 2'-5'-oligoadenylate synthetase. *Wiley Interdisciplinary Reviews. RNA*, *10*(4), e1534. https://doi.org/10.1002/wrna.1534

Sha, S., Maloney, A. J., Katsikis, G., Nguyen, T. N. T., Neufeld, C., Wolfrum, J., Barone, P. W., Springs, S. L., Manalis, S. R., Sinskey, A. J., & Braatz, R. D. (2021). Cellular pathways of recombinant adeno-associated virus production for gene therapy. *Biotechnology Advances*, *49*, 107764. https://doi.org/10.1016/j.biotechadv.2021.107764

Shahryari, A., Jazi, M. S., Mohammadi, S., Nikoo, H. R., Nazari, Z., Hosseini, E. S., Burtscher, I., Mowla, S. J., & Lickert, H. (2019). Development and clinical translation of approved gene therapy products for genetic disorders. *Frontiers in Genetics*, *10*(SEP), 868. https://doi.org/10.3389/FGENE.2019.00868/BIBTEX

Shaw, A. R., & Ziff, E. B. (1980). Transcripts from the adenovirus-2 major late promoter yield a single early family of 3′ coterminal mRNAs and five late families. *Cell*, *22*(3), 905–916. https://doi.org/10.1016/0092-8674(80)90568-1

Shearwin, K. E., Callen, B. P., & Egan, J. B. (2005). Transcriptional interference--a crash course. *Trends in Genetics : TIG*, *21*(6), 339–345. https://doi.org/10.1016/j.tig.2005.04.009

Sheng, M., & Sala, C. (2001). PDZ Domains and the Organization of Supramolecular Complexes. *Annual Review of Neuroscience*, *24*(1), 1–29. https://doi.org/10.1146/annurev.neuro.24.1.1

Shepard, R. N., & Ornelles, D. A. (2004). Diverse Roles for E4orf3 at Late Times of Infection Revealed in an E1B 55-Kilodalton Protein Mutant Background. *Journal of Virology*, *78*(18), 9924–9935. https://doi.org/10.1128/JVI.78.18.9924-9935.2004

Shi, Y., Seto, E., Chang, L. S., & Shenk, T. (1991). Transcriptional repression by YY1, a human GLI-Krüppel-related protein, and relief of repression by adenovirus E1A protein. *Cell*, *67*(2), 377–388. https://doi.org/10.1016/0092-8674(91)90189-6

Sinzger, C., Digel, M., & Jahn, G. (2008). Cytomegalovirus Cell Tropism. In *Current topics in microbiology and immunology* (Vol. 325, pp. 63–83). Curr Top Microbiol Immunol. https://doi.org/10.1007/978-3-540-77349-8_4

Skopenkova, V. V., Egorova, T. V., & Bardina, M. V. (2021). Muscle-Specific Promoters for Gene Therapy. *Acta Naturae*, *13*(1), 47–58. https://doi.org/10.32607/actanaturae.11063

Smith, J., Grieger, J., & Samulski, R. J. (2018). Overcoming Bottlenecks in AAV Manufacturing for Gene Therapy. *Cell and Gene Therapy Insights*, *4*(8), 815–825. https://doi.org/10.18609/cgti.2018.083

Snyder, R. O., Im, D. S., Ni, T., Xiao, X., Samulski, R. J., & Muzyczka, N. (1993). Features of the adeno-associated virus origin involved in substrate recognition by the viral Rep protein. *Journal of Virology*, *67*(10), 6096–6104. https://doi.org/10.1128/jvi.67.10.6096-6104.1993

Söderlund, H., Pettersson, U., Vennström, B., Philipson, L., & Mathews, M. B. (1976). A new species of virus-coded low molecular weight RNA from cells infected with adenovirus type 2. *Cell*, *7*(4), 585–593. https://doi.org/10.1016/0092-8674(76)90209-9

Sogawa, K., Handa, H., Fujisawa-Sehara, A., Hiromasa, T., Yamane, M., & Fujii-Kuriyama, Y. (1989). Repression of cytochrome P-450c gene expression by cotransfection with adenovirus E1a DNA. *European Journal of Biochemistry*, *181*(3), 539–544. https://doi.org/10.1111/j.1432-1033.1989.tb14757.x

Somberg, M., Rush, M., Fay, J., Ryan, F., Lambkin, H., Akusjärvi, G., & Schwartz, S. (2009). Adenovirus E4orf4 induces HPV-16 late L1 mRNA production. *Virology*, *383*(2), 279–290. https://doi.org/10.1016/j.virol.2008.09.041

Song, C. Z., Loewenstein, P. M., Toth, K., & Green, M. (1995). Transcription factor TFIID is a direct functional target of the adenovirus E1A transcription-repression domain. *Proceedings of the National Academy of Sciences*, *92*(22), 10330–10333. https://doi.org/10.1073/pnas.92.22.10330

Sonntag, F., Bleker, S., Leuchs, B., Fischer, R., & Kleinschmidt, J. A. (2006). Adeno-Associated Virus Type 2 Capsids with Externalized VP1/VP2 Trafficking Domains Are Generated prior to Passage through the Cytoplasm and Are Maintained until Uncoating Occurs in the Nucleus. *Journal of Virology*, *80*(22), 11040–11054. https://doi.org/10.1128/JVI.01056-06

Sonntag, F., Köther, K., Schmidt, K., Weghofer, M., Raupp, C., Nieto, K., Kuck, A., Gerlach, B., Böttcher, B., Müller, O. J., Lux, K., Hörer, M., & Kleinschmidt, J. A. (2011). The Assembly-Activating Protein Promotes Capsid Assembly of Different Adeno-Associated Virus Serotypes. *Journal of Virology*, *85*(23), 12686–12697. https://doi.org/10.1128/JVI.05359-11

Sonntag, F., Schmidt, K., & Kleinschmidt, J. A. (2010). A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proceedings of the National Academy of Sciences*, *107*(22), 10220–10225. https://doi.org/10.1073/pnas.1001673107

Soria, C., Estermann, F. E., Espantman, K. C., & O'Shea, C. C. (2010). Heterochromatin silencing of p53 target genes by a small viral protein. *Nature*, *466*(7310), 1076–1081. https://doi.org/10.1038/nature09307

Soriano, A. M., Crisostomo, L., Mendez, M., Graves, D., Frost, J. R., Olanubi, O., Whyte, P. F., Hearing, P., & Pelka, P. (2019). Adenovirus 5 E1A Interacts with E4orf3 To Regulate Viral Chromatin Organization. *Journal of Virology*, *93*(10). https://doi.org/10.1128/JVI.00157-19

Spindler, K. R., Eng, C. Y., & Berk, A. J. (1985). An adenovirus early region 1A protein is required for maximal viral DNA replication in growth-arrested human cells. *Journal of Virology*, *53*(3), 742–750. https://doi.org/10.1128/jvi.53.3.742-750.1985

Srivastava, A. (2016). Advances and challenges in the use of recombinant AAV vectors for human gene therapy. *Cell and Gene Therapy Insights*, *2*(5), 553–575. https://doi.org/10.18609/cgti.2016.061

Steegenga, W. T., Laar, T. Van, Riteco, N., Mandarino, A., Shvarts, A., Van Der Eb, A. J., & Jochemsen, A. G. (1996). Adenovirus E1A Proteins Inhibit Activation of Transcription by p53. *Molecular and Cellular Biology*, *16*(5), 2101–2109. https://doi.org/10.1128/MCB.16.5.2101

Stephen Kemler, & Adam Lohr. (2022, February 21). *Cell Gene Therapies Investment Outlook In 2022 Beyond*. https://www.cellandgene.com/doc/cell-gene-therapies-investment-outlook-in-beyond-0001

Stinski, M. F., & Isomura, H. (2008). Role of the cytomegalovirus major immediate early enhancer in acute infection and reactivation from latency. *Medical Microbiology and Immunology*, *197*(2), 223–231. https://doi.org/10.1007/s00430-007-0069-7

Stracker, T. H., Carson, C. T., & Weitzman, M. D. (2002). Adenovirus oncoproteins inactivate the Mre11–Rad50–NBS1 DNA repair complex. *Nature*, *418*(6895), 348–352. https://doi.org/10.1038/nature00863

Stracker, T. H., Cassell, G. D., Ward, P., Loo, Y.-M., van Breukelen, B., Carrington-Lawrence, S. D., Hamatake, R. K., van der Vliet, P. C., Weller, S. K., Melendy, T., & Weitzman, M. D. (2004). The Rep Protein of Adeno-Associated Virus Type 2 Interacts with Single-Stranded DNA-Binding Proteins That Enhance Viral Replication. *Journal of Virology*, *78*(1), 441–453. https://doi.org/10.1128/JVI.78.1.441-453.2004

Stuiver, M. H., & van der Vliet, P. C. (1990). Adenovirus DNA-binding protein forms a multimeric protein complex with double-stranded DNA and enhances binding

of nuclear factor I. *Journal of Virology*, *64*(1), 379–386. https://doi.org/10.1128/jvi.64.1.379-386.1990

Stutika, C., Gogol-Döring, A., Botschen, L., Mietzsch, M., Weger, S., Feldkamp, M., Chen, W., & Heilbronn, R. (2016). A Comprehensive RNA Sequencing Analysis of the Adeno-Associated Virus (AAV) Type 2 Transcriptome Reveals Novel AAV Transcripts, Splice Variants, and Derived Proteins. *Journal of Virology*, *90*(3), 1278–1289. https://doi.org/10.1128/JVI.02750-15

Su, W. (2021). *Development of a self-silencing adenovirus for the efficient manufacture of adeno-associated virus vectors* [PhD Thesis]. University of Oxford.

Su, W., Patrício, M. I., Duffy, M. R., Krakowiak, J. M., Seymour, L. W., & Cawood, R. (2022). Self-attenuating adenovirus enables production of recombinant adeno-associated virus for high manufacturing yield without contamination. *Nature Communications 2022 13:1*, *13*(1), 1–14. https://doi.org/10.1038/s41467-022-28738-2

Subramanian, T., Vijayalingam, S., Kuppuswamy, M., & Chinnadurai, G. (2015). Interaction of cellular proteins with BCL-xL targeted to cytoplasmic inclusion bodies in adenovirus infected cells. *Virology*, *483*, 21–31. https://doi.org/10.1016/j.virol.2015.04.015

Summerford, C., & Samulski, R. J. (1998). Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *Journal of Virology*, *72*(2), 1438–1445. http://www.ncbi.nlm.nih.gov/pubmed/9445046

Sun, J., Li, D., Hao, Y., Zhang, Y., Fan, W., Fu, J., Hu, Y., Liu, Y., & Shao, Y. (2009). Posttranscriptional Regulatory Elements Enhance Antigen Expression and DNA Vaccine Efficacy. *DNA and Cell Biology*, *28*(5), 233. https://doi.org/10.1089/DNA.2009.0862

Surosky, R. T., Urabe, M., Godwin, S. G., Mcquiston, S. A., Kurtzman, G. J., Ozawa, K., & Natsoulis, G. (1997). Adeno-associated virus Rep proteins target DNA sequences to a unique locus in the human genome. *Journal of Virology*, *71*(10), 7951. https://doi.org/10.1128/JVI.71.10.7951-7959.1997

Tait, A. S., Brown, C. J., Galbraith, D. J., Hines, M. J., Hoare, M., Birch, J. R., & James, D. C. (2004). Transient production of recombinant proteins by Chinese hamster ovary cells using polyethyleneimine/DNA complexes in combination with microtubule disrupting anti-mitotic agents. *Biotechnology and Bioengineering*, *88*(6), 707–721. https://doi.org/10.1002/BIT.20265

Takako, K.-Y., Naomi, Y., & Shuichi, H. (1988). Activation of the adenovirus-2 E2a late promoter during inhibition of protein synthesis by cycloheximide. *Gene*, *69*(1), 165–169. https://doi.org/10.1016/0378-1119(88)90391-5

Tang, Q., Keeler, A. M., Zhang, S., Su, Q., Lyu, Z., Cheng, Y., Gao, G., & Flotte, T. R. (2020). Two-plasmid packaging system for recombinant adeno-associated virus. *BioResearch Open Access*, *9*(1), 219–228. https://doi.org/10.1089/BIORES.2020.0031/SUPPL_FILE/TANG_REV_SUPPLEMENTALTABLE1.DOCX

Thomas, M. A., Broughton, R. S., Goodrum, F. D., & Ornelles, D. A. (2009). E4orf1 Limits the Oncolytic Potential of the E1B - 55K Deletion Mutant Adenovirus . *Journal of Virology*, *83*(6), 2406–2416. https://doi.org/10.1128/JVI.01972-08/ASSET/4BA72820-8986-4988-8C36-2F516EB6DE3C/ASSETS/GRAPHIC/ZJV0060916210006.JPEG

Timpe, J. M., Verrill, K. C., & Trempe, J. P. (2006). Effects of Adeno-Associated Virus on Adenovirus Replication and Gene Expression during Coinfection. *Journal of Virology*, *80*(16), 7807–7815. https://doi.org/10.1128/jvi.00198-06

Toktay, Y., Dayanc, B., & Senturk, S. (2022). Engineering and validation of a dual luciferase reporter system for quantitative and systematic assessment of regulatory sequences in Chinese hamster ovary cells. *Scientific Reports*, *12*(1), 6050. https://doi.org/10.1038/s41598-022-09887-2

Tollefson, A. E., Ying, B., Doronin, K., Sidor, P. D., & Wold, W. S. M. (2007). Identification of a new human adenovirus protein encoded by a novel late l-strand transcription unit. *Journal of Virology*, *81*(23), 12918–12926. https://doi.org/10.1128/JVI.01531-07

Törmänen, H., Backström, E., Carlsson, A., & Akusjärvi, G. (2006). L4-33K, an Adenovirus-encoded Alternative RNA Splicing Factor. *Journal of Biological Chemistry*, *281*(48), 36510–36517. https://doi.org/10.1074/jbc.M607601200

Tratschin, J. D., Miller, I. L., Smith, M. G., & Carter, B. J. (1985). Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. *Molecular and Cellular Biology*, *5*(11), 3251–3260. https://doi.org/10.1128/mcb.5.11.3251

Tratschin, J. D., Tal, J., & Carter, B. J. (1986). Negative and positive regulation in trans of gene expression from adeno-associated virus vectors in mammalian cells by a viral rep gene product. *Molecular and Cellular Biology*, *6*(8), 2884–2894. https://doi.org/10.1128/mcb.6.8.2884

Tratschin, J. D., West, M. H., Sandbank, T., & Carter, B. J. (1984). A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. *Molecular and Cellular Biology*, *4*(10), 2072–2081. https://doi.org/10.1128/mcb.4.10.2072

Tsang, H.-F., Xue, V. W., Koh, S.-P., Chiu, Y.-M., Ng, L. P.-W., & Wong, S.-C. C. (2017). NanoString, a novel digital color-coded barcode technology: current and future applications in molecular diagnostics. *Expert Review of Molecular Diagnostics*, *17*(1), 95–103. https://doi.org/10.1080/14737159.2017.1268533

Tseng, Y.-S., & Agbandje-McKenna, M. (2014). Mapping the AAV Capsid Host Antibody Response toward the Development of Second Generation Gene Delivery Vectors. *Frontiers in Immunology*, *5*, 9. https://doi.org/10.3389/fimmu.2014.00009

Tsunoda, T., & Takagi, T. (1999). Estimating transcription factor bindability on DNA. *Bioinformatics (Oxford, England)*, *15*(7–8), 622–630. https://doi.org/10.1093/BIOINFORMATICS/15.7.622

Turnell, A. S. (2000). Regulation of the 26S proteasome by adenovirus E1A. *The EMBO Journal*, *19*(17), 4759–4773. https://doi.org/10.1093/emboj/19.17.4759

Vachon, V. K., Calderon, B. M., & Conn, G. L. (2015). A novel RNA molecular signature for activation of 2′-5′ oligoadenylate synthetase-1. *Nucleic Acids Research*, *43*(1), 544–552. https://doi.org/10.1093/NAR/GKU1289

Vachon, V. K., & Conn, G. L. (2016). Adenovirus VA RNA: An essential pro-viral non-coding RNA. *Virus Research*, *212*, 39–52. https://doi.org/10.1016/j.virusres.2015.06.018

van Breukelen, B., Brenkman, A. B., Holthuizen, P. E., & van der Vliet, P. C. (2003). Adenovirus Type 5 DNA Binding Protein Stimulates Binding of DNA Polymerase to the Replication Origin. *Journal of Virology*, *77*(2), 915–922. https://doi.org/10.1128/jvi.77.2.915-922.2003

van Lieshout, L. P., Rubin, M., Costa-Grant, K., Ota, S., Golebiowski, D., Panico, T., Wiberg, E., Szymczak, K., Gilmore, R., Stanvick, M., Burnham, B., Gagnon, J., Iwuchukwu, I., Yang, G., Ghazi, I., Meola, A., Dickerson, R., Thiers, T., Mustich, L., … Kelly, T. (2023). A novel dual-plasmid platform provides scalable transfection yielding improved productivity and packaging across multiple AAV serotypes and genomes. *Molecular Therapy - Methods & Clinical Development*, *29*, 426–436. https://doi.org/10.1016/j.omtm.2023.05.004

Vandenberghe, L. H., Wang, L., Somanathan, S., Zhi, Y., Figueredo, J., Calcedo, R., Sanmiguel, J., Desai, R. A., Chen, C. S., Johnston, J., Grant, R. L., Gao, G., & Wilson, J. M. (2006). Heparin binding directs activation of T cells against adeno-associated virus serotype 2 capsid. *Nature Medicine*, *12*(8), 967–971. https://doi.org/10.1038/nm1445

Vandenberghe, L. H., Xiao, R., Lock, M., Lin, J., Korn, M., & Wilson, J. M. (2010). Efficient Serotype-Dependent Release of Functional Vector into the Culture Medium During Adeno-Associated Virus Manufacturing. *Human Gene Therapy*, *21*(10), 1251–1257. https://doi.org/10.1089/hum.2010.107

Velcich, A., & Ziff, E. (1985). Adenovirus E1a proteins repress transcription from the SV40 early promoter. *Cell*, *40*(3), 705–716. https://doi.org/10.1016/0092-8674(85)90219-3

Vincent, K. A., Piraino, S. T., & Wadsworth, S. C. (1997). Analysis of recombinant adeno-associated virus packaging and requirements for rep and cap gene products. *Journal of Virology*, *71*(3), 1897–1905. https://doi.org/10.1128/jvi.71.3.1897-1905.1997

Virtanen, A., Gilardi, P., Näslund, A., LeMoullec, J. M., Pettersson, U., & Perricaudet, M. (1984). mRNAs from human adenovirus 2 early region 4. *Journal of Virology*, *51*(3), 822–831. https://doi.org/10.1128/jvi.51.3.822-831.1984

Vogel, R. (2013). Viral and Cellular Components of AAV2 Replication Compartments. *The Open Virology Journal*, *7*(1), 98–120. https://doi.org/10.2174/1874357901307010098

Walz, C., Deprez, A., Dupressoir, T., Dürst, M., Rabreau, M., & Schlehofer, J. R. (1997). Interaction of human papillomavirus type 16 and adeno-associated virus type 2 co-infecting human cervical epithelium. *Journal of General Virology*, *78*(6), 1441–1452. https://doi.org/10.1099/0022-1317-78-6-1441

Wang, D., Tai, P. W. L., & Gao, G. (2019). Adeno-associated virus vector as a platform for gene therapy delivery. *Nature Reviews Drug Discovery*, *18*(5), 358–378. https://doi.org/10.1038/s41573-019-0012-9

Wang, X.-S., Khuntirat, B., Qing, K., Ponnazhagan, S., Kube, D. M., Zhou, S., Dwarki, V. J., & Srivastava, A. (1998). Characterization of Wild-Type Adeno-Associated Virus Type 2-Like Particles Generated during Recombinant Viral Vector Production and Strategies for Their Elimination. *Journal of Virology*, *72*(7), 5472–5480. https://doi.org/10.1128/jvi.72.7.5472-5480.1998

Wang, Z., Cheng, F., Engelhardt, J. F., Yan, Z., & Qiu, J. (2018). Development of a Novel Recombinant Adeno-Associated Virus Production System Using Human

Bocavirus 1 Helper Genes. *Molecular Therapy - Methods & Clinical Development*, *11*, 40–51. https://doi.org/10.1016/j.omtm.2018.09.005

Wang, Z., Deng, X., Zou, W., Engelhardt, J. F., Yan, Z., & Qiu, J. (2017). Human Bocavirus 1 Is a Novel Helper for Adeno-associated Virus Replication. *Journal of Virology*, *91*(18). https://doi.org/10.1128/JVI.00710-17

Ward, P., Dean, F. B., O'Donnell, M. E., & Berns, K. I. (1998). Role of the Adenovirus DNA-Binding Protein in In Vitro Adeno-Associated Virus DNA Replication. *Journal of Virology*, *72*(1), 420–427. https://doi.org/10.1128/jvi.72.1.420-427.1998

Warrington, K. H., Gorbatyuk, O. S., Harrison, J. K., Opie, S. R., Zolotukhin, S., & Muzyczka, N. (2004). Adeno-Associated Virus Type 2 VP2 Capsid Protein Is Nonessential and Can Tolerate Large Peptide Insertions at Its N Terminus. *Journal of Virology*, *78*(12), 6595–6609. https://doi.org/10.1128/JVI.78.12.6595-6609.2004

Webster, K. A., Muscat, G. E. O., & Kedes, L. (1988). Adenovirus E1A products suppress myogenic differentiation and inhibit transcription from muscle-specific promoters. *Nature*, *332*(6164), 553–557. https://doi.org/10.1038/332553A0

Weger, S., Hammer, E., & Heilbronn, R. (2004). SUMO-1 modification regulates the protein stability of the large regulatory protein Rep78 of adeno associated virus type 2 (AAV-2). *Virology*, *330*(1), 284–294. https://doi.org/10.1016/j.virol.2004.09.028

Weger, S., Wendland, M., Kleinschmidt, J. A., & Heilbronn, R. (1999). The Adeno-Associated Virus Type 2 Regulatory Proteins Rep78 and Rep68 Interact with the Transcriptional Coactivator PC4. *Journal of Virology*, *73*(1), 260–269. https://doi.org/10.1128/jvi.73.1.260-269.1999

Weiss, R. S., Lee, S. S., Prasad, B. V, & Javier, R. T. (1997). Human adenovirus early region 4 open reading frame 1 genes encode growth-transforming proteins that may be distantly related to dUTP pyrophosphatase enzymes. *Journal of Virology*, *71*(3), 1857–1870. https://doi.org/10.1128/jvi.71.3.1857-1870.1997

Weitzman, M. D. (2005). Functions of the adenovirus E4 proteins and their impact on viral vectors. In *Frontiers in bioscience : a journal and virtual library* (Vol. 10, pp. 1106–1117). https://doi.org/10.2741/1604

Weitzman, M. D., Kyöstiö, S. R. M., Kotin, R. M., & Owens, R. A. (1994). Adeno-associated virus (AAV) Rep proteins mediate complex formation between AAV DNA and its integration site in human DNA. *Proceedings of the National*

Academy of Sciences of the United States of America, *91*(13), 5808. https://doi.org/10.1073/PNAS.91.13.5808

Weitzman, M. D., & Linden, R. M. (2012). Adeno-Associated Virus Biology. In *Methods in Molecular Biology* (Vol. 807, pp. 1–23). https://doi.org/10.1007/978-1-61779-370-7_1

West, M. H. P., Trempe, J. P., Tratschin, J. D., & Carter, B. J. (1987). Gene expression in adeno-associated virus vectors: The effects of chimeric mRNA structure, helper virus, and adenovirus VA, RNA. *Virology*, *160*(1), 38–47. https://doi.org/10.1016/0042-6822(87)90041-9

West, S., & Proudfoot, N. J. (2009). Transcriptional Termination Enhances Protein Expression in Human Cells. *Molecular Cell*, *33*(3–9), 354. https://doi.org/10.1016/J.MOLCEL.2009.01.008

Westergren Jakobsson, A., Segerman, B., Wallerman, O., Bergström Lind, S., Zhao, H., Rubin, C.-J., Pettersson, U., & Akusjärvi, G. (2021). The Human Adenovirus 2 Transcriptome: an Amazing Complexity of Alternatively Spliced mRNAs. *Journal of Virology*, *95*(4). https://doi.org/10.1128/JVI.01869-20

White, E. (2001a). Regulation of the cell cycle and apoptosis by the oncogenes of adenovirus. *Oncogene*, *20*(54), 7836–7846. https://doi.org/10.1038/sj.onc.1204861

White, E. (2001b). Regulation of the cell cycle and apoptosis by the oncogenes of adenovirus. *Oncogene*, *20*(54), 7836–7846. https://doi.org/10.1038/sj.onc.1204861

Whiteway, A., Deru, W., Prentice, H. G., & Anderson, R. (2003). Construction of adeno-associated virus packaging plasmids and cells that directly select for AAV helper functions. *Journal of Virological Methods*, *114*(1), 1–10. https://doi.org/10.1016/j.jviromet.2003.08.001

Williams, J. A., & Paez, P. A. (2023). Improving cell and gene therapy safety and performance using next-generation Nanoplasmid vectors. *Molecular Therapy - Nucleic Acids*, *32*, 494–503. https://doi.org/10.1016/j.omtn.2023.04.003

Wilson, & Xiao. (1998). *Methods and vector constructs useful for production of recombinant aav*.

Winter, K., von Kietzell, K., Heilbronn, R., Pozzuto, T., Fechner, H., & Weger, S. (2012). Roles of E4orf6 and VA I RNA in Adenovirus-Mediated Stimulation of Human Parvovirus B19 DNA Replication and Structural Gene Expression. *Journal of Virology*, *86*(9), 5099–5109. https://doi.org/10.1128/JVI.06991-11

Wistuba, A., Kern, A., Weger, S., Grimm, D., Rgen, J., & Kleinschmidt, A. (1997). Subcellular Compartmentalization of Adeno-Associated Virus Type 2 Assembly. *JOURNAL OF VIROLOGY*, *71*(2), 1341–1352. http://jvi.asm.org/content/71/2/1341.full.pdf

Wistuba, A., Weger, S., Kern, A., & Kleinschmidt, J. A. (1995). Intermediates of adeno-associated virus type 2 assembly: identification of soluble complexes containing Rep and Cap proteins. *Journal of Virology*, *69*(9), 5311–5319. https://doi.org/10.1128/JVI.69.9.5311-5319.1995

Wörner, T. P., Bennett, A., Habka, S., Snijder, J., Friese, O., Powers, T., Agbandje-McKenna, M., & Heck, A. J. R. (2021). Adeno-associated virus capsid assembly is divergent and stochastic. *Nature Communications 2021 12:1*, *12*(1), 1–9. https://doi.org/10.1038/s41467-021-21935-5

Wright, J. F. (2008). Manufacturing and characterizing AAV-based vectors for use in clinical studies. *Gene Therapy*, *15*(11), 840–848. https://doi.org/10.1038/gt.2008.65

Wright, J., & Leppard, K. N. (2013). The Human Adenovirus 5 L4 Promoter Is Activated by Cellular Stress Response Protein p53. *Journal of Virology*, *87*(21), 11617–11625. https://doi.org/10.1128/JVI.01924-13

Wu, K., Guimet, D., & Hearing, P. (2013). The Adenovirus L4-33K Protein Regulates both Late Gene Expression Patterns and Viral DNA Packaging. *Journal of Virology*, *87*(12), 6739–6747. https://doi.org/10.1128/jvi.00652-13

Wu, K., Orozco, D., & Hearing, P. (2012). The Adenovirus L4-22K Protein Is Multifunctional and Is an Integral Component of Crucial Aspects of Infection. *Journal of Virology*, *86*(19), 10474–10483. https://doi.org/10.1128/JVI.01463-12

Wustner, J. T., Arnold, S., Lock, M., Richardson, J. C., Himes, V. B., Kurtzman, G., & Peluso, R. W. (2002). Production of Recombinant Adeno-Associated Type 5 (rAAV5) Vectors Using Recombinant Herpes Simplex Viruses Containing rep and cap. *Molecular Therapy*, *6*(4), 510–518. https://doi.org/10.1006/MTHE.2002.0695

Xiao, X., Li, J., & Samulski, R. J. (1998). Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *Journal of Virology*, *72*(3), 2224–2232. http://www.ncbi.nlm.nih.gov/pubmed/9499080

Xin, T., Cheng, L., Zhou, C., Zhao, Y., Hu, Z., & Wu, X. (2022). In-Vivo Induced CAR-T Cell for the Potential Breakthrough to Overcome the Barriers of Current CAR-T

Cell Therapy. *Frontiers in Oncology*, *12*. https://doi.org/10.3389/FONC.2022.809754

Xu, Z.-J., Jia, Y.-L., Wang, M., Yi, D.-D., Zhang, W.-L., Wang, X.-Y., & Zhang, J.-H. (2019). Effect of promoter, promoter mutation and enhancer on transgene expression mediated by episomal vectors in transfected HEK293, Chang liver and primary cells. *Bioengineered*, *10*(1), 548–560. https://doi.org/10.1080/21655979.2019.1684863

Yakobson, B., Koch, T., & Winocour, E. (1987). Replication of adeno-associated virus in synchronized cells without the addition of a helper virus. *Journal of Virology*, *61*(4), 972–981. https://doi.org/10.1128/jvi.61.4.972-981.1987

Yalkinoglu, A. O., Heilbronn, R., Bürkle, A., Schlehofer, J. R., & zur Hausen, H. (1988). DNA amplification of adeno-associated virus as a response to cellular genotoxic stress. *Cancer Research*, *48*(11), 3123–3129. http://www.ncbi.nlm.nih.gov/pubmed/2835153

Yang, J., Zhou, W., Zhang, Y., Zidon, T., Ritchie, T., & Engelhardt, J. F. (1999). Concatamerization of Adeno-Associated Virus Circular Genomes Occurs through Intermolecular Recombination. *Journal of Virology*, *73*(11), 9468–9477. https://doi.org/10.1128/jvi.73.11.9468-9477.1999

Yang, W. C., Lu, J., Nguyen, N. B., Zhang, A., Healy, N. V., Kshirsagar, R., Ryll, T., & Huang, Y. M. (2014). Addition of valproic acid to CHO cell fed-batch cultures improves monoclonal antibody titers. *Molecular Biotechnology*, *56*(5), 421–428. https://doi.org/10.1007/S12033-013-9725-X

Yates, V. J., Dawson, G. J., & Pronovost, A. D. (1981). Serologic evidence of avian adeno-associated virus infection in an unselected human population and among poultry workers. *American Journal of Epidemiology*, *113*(5), 542–545. https://doi.org/10.1093/OXFORDJOURNALS.AJE.A113130

Yi, R., Qin, Y., Macara, I. G., & Cullen, B. R. (2003). Exportin-5 mediates the nuclear export of pre-microRNAs and short hairpin RNAs. *Genes & Development*, *17*(24), 3011–3016. https://doi.org/10.1101/GAD.1158803

Ying, B., Tollefson, A. E., & Wold, W. S. M. (2010). Identification of a Previously Unrecognized Promoter That Drives Expression of the UXP Transcription Unit in the Human Adenovirus Type 5 Genome. *Journal of Virology*, *84*(21), 11470–11478. https://doi.org/10.1128/JVI.01338-10

Yoneyama, M., & Fujita, T. (2010). Recognition of viral nucleic acids in innate immunity. *Reviews in Medical Virology*, *20*(1), 4–22. https://doi.org/10.1002/rmv.633

Yoon-Robarts, M., Blouin, A. G., Bleker, S., Kleinschmidt, J. A., Aggarwal, A. K., Escalante, C. R., & Linden, R. M. (2004). Residues within the B′ motif are critical for DNA binding by the superfamily 3 helicase Rep40 of adeno-associated virus type 2. *Journal of Biological Chemistry*, *279*(48), 50472–50481. https://doi.org/10.1074/jbc.M403900200

Yu, C., Trivedi, P. D., Chaudhuri, P., Bhake, R., Johnson, E. J., Caton, T., Potter, M., Byrne, B. J., & Clément, N. (2021). NaCl and KCl mediate log increase in AAV vector particles and infectious titers in a specific/timely manner with the HSV platform. *Molecular Therapy - Methods & Clinical Development*, *21*, 1–13. https://doi.org/10.1016/j.omtm.2021.02.015

Yue, Y., & Duan, D. (2002). Development of multiple cloning site cis-vectors for recombinant adeno-associated virus production. *Biotechniques*, *33*(3), 672, 674, 676–678. https://doi.org/10.2144/02333DD03

Zabet, N. R., & Adryan, B. (2013). The effects of transcription factor competition on gene regulation. *Frontiers in Genetics*, *4*(OCT), 65406. https://doi.org/10.3389/FGENE.2013.00197/ABSTRACT

Zarate-Perez, F., Mansilla-Soto, J., Bardelli, M., Burgner, J. W., Villamil-Jarauta, M., Kekilli, D., Samso, M., Linden, R. M., & Escalante, C. R. (2013). Oligomeric Properties of Adeno-Associated Virus Rep68 Reflect Its Multifunctionality. *Journal of Virology*, *87*(2), 1232–1241. https://doi.org/10.1128/jvi.02441-12

Zemke, N. R., & Berk, A. J. (2017). The Adenovirus E1A C Terminus Suppresses a Delayed Antiviral Response and Modulates RAS Signaling. *Cell Host & Microbe*, *22*(6), 789-800.e5. https://doi.org/10.1016/j.chom.2017.11.008

Zhao, H., Chen, M., & Pettersson, U. (2014). A new look at adenovirus splicing. *Virology*, *456–457*(1), 329–341. https://doi.org/10.1016/j.virol.2014.04.006

Zhao, H., Lee, K.-J., Daris, M., Lin, Y., Wolfe, T., Sheng, J., Plewa, C., Wang, S., & Meisen, W. H. (2020). Creation of a High-Yield AAV Vector Production Platform in Suspension Cells Using a Design-of-Experiment Approach. *Molecular Therapy - Methods & Clinical Development*, *18*, 312–320. https://doi.org/10.1016/j.omtm.2020.06.004

Zincarelli, C., Soltys, S., Rengo, G., & Rabinowitz, J. E. (2008). Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Molecular Therapy*, *16*(6), 1073–1080. https://doi.org/10.1038/mt.2008.76

10 Acknowledgements

At last, I wish to extend my heartfelt gratitude to all those who have played a pivotal role on my journey through this research endeavour and without whom this thesis would not have been the same or not even been possible at all.

First of all, I would like to thank REGENXBIO for their sponsorship of this research project. In particular, Ping Liu, Ayda Mayer, and their research group, not only for funding and materials, but also invaluable advice, fruitful research collaboration and the acknowledgement of our work in the patent(s).

My gratitude also goes to David James for granting me this great opportunity to conduct this research in his lab, providing an outstanding work atmosphere within his group, his mentorship, and the multifarious support over the years.

Thanks to all the members of the DCJ/AJB group, many of whom have become not just colleagues but also friends. The open-hearted and nurturing culture of this group exceeded my expectations and will forever hold a special place in my memory. I would like to extend a special thanks to Yusuf Johari and Joe Scarrott for shaping the RGX DCJ AAV dream team. All the help, the shared long hours in the lab and research output will always be remembered with great appreciation and fondness.

I cannot thank enough my family. Without my parents I would not be who I am. Their unconditional and endless love and support in so many ways have enabled me to reach this level of education, something for which I will be forever grateful.

My deepest and most heartfelt gratitude goes to the love of my life and future wife, Melinda! Words cannot describe the multitude of ways you have supported me during this endeavour, and I am unimaginably thankful for all of it! Thank you for coming here with me and for undertaking this journey together. Thank you for being the incredible person you are and for being my unwavering support. I am glad that I am and will never be without you at my side, (still, after all this time) it is my biggest joy and I am looking forward to every moment of it (…always).

11 Declaration of Originality

In accordance with the University regulations, I hereby declare that I, Thilo Hermann Pohle, have composed this thesis solely by myself. The research presented within is the result of my own efforts and achievements, unless acknowledged otherwise. It has not been submitted in part or whole for any other degree or personal qualification.

_____

(Thilo Pohle)