Development of Adeno-associated Virus Capsid Libraries Tailored Towards B Cell Engineering Applications



Henry Brigden

St Edmund Hall

The University of Oxford

A thesis submitted for the degree of Doctor of Philosophy to the Department of Oncology, University of Oxford

Michaelmas Term 2025

1 Abstract

Monoclonal antibody therapies have transformed the treatment of human disease, particularly cancer, but their reliance on repeated infusions limits accessibility. B lymphocytes, key effectors of humoral immunity uniquely adapted for regulated long-term antibody production, offer an attractive platform for therapeutic antibody delivery. Recently developed approaches for generating primary human B cells expressing monoclonal antibodies require extensive ex vivo manipulation and are unsuitable for large-scale clinical application. Furthermore, existing B cell engineering strategies were developed for infectious disease applications and require substantial adaptation for use in cancer therapy.

Adeno-associated virus (AAV) is the leading platform for safe and efficient gene delivery, yet its clinical utility is constrained by restricted tropism, pre-existing immunity, and poor transduction of many primary human cells, including B lymphocytes. The overarching aim of this thesis was to develop a platform for the engineering of anti-cancer B cell therapies, encompassing both vector optimisation and antibody knock-in template design. To this end, we developed AAV capsid libraries to identify variants capable of efficient gene delivery to primary human B cells. As part of this work, we evaluated established library production strategies and explored an inducible two-step library system enhance library quality. Finally, an approach was developed for engineering B cells with defined full-length antibody sequences, enabling the incorporation of optimised Fc domains.

We tested an inducible two-step library production platform intended to minimise cross-packaging of AAV genomes into incorrect capsids and prevent mosaic capsid formation during production. The system was designed such that AAV genomes encoding an inducible capsid library would first be packaged into wild-type AAV2 capsids, then re-packaged into genome-encoded capsids following transduction of producer cells at low multiplicity of infection and induction of capsid expression. Two alternative inducible systems, a protein transactivator system and a Cre recombinase system, achieved stringent repression under plasmid conditions but failed to maintain adequate control during AAV replication, highlighting the inherent challenges of inducible regulation within replicating AAV genomes. To develop AAV capsid libraries tailored towards B cell transduction, we examined viral binding and entry dynamics and found that AAV transduction in human B cells appears primarily constrained by post-entry steps. Accordingly, two complementary AAV capsid libraries were produced and screened with the aim of identifying variants capable of overcoming these post-entry barriers in primary human B cells. Finally, we demonstrated the simultaneous expression of membrane-bound and secreted anti-cancer antibodies from homology-directed repair templates designed for the knock-in at the human IgH locus.

2 Declaration of Authentication

The work in this thesis was performed under the supervision of Professor Leonard W. Seymour at the Department of Oncology, Medical Sciences Division, University of Oxford. The material presented in this thesis is entirely my own and any contribution made by others has been acknowledged accordingly. The work submitted here does not form part of another thesis in this or any other university.

Henry Brigden

October 2025

3 Acknowledgements

I would like to express my sincere appreciation to my supervisor, Professor Len Seymour, for his continuous guidance, support, and encouragement throughout my studies. Len's incredible generosity and compassion made this work possible, and I will forever be grateful.

I am also thankful to my industry supervisor, Dr Ryan Cawood, for his practical advice, and to all the members of the Seymour Group for their support. A special thanks goes to Hind Ghezraoui for her support and for her practical help with the work presented in Chapter 7.

To my family — thank you for your love and belief in me, particularly during the most challenging moments.

Finally, I dedicate this thesis to Dad, whose memory has been a constant source of strength.

4 Table of Contents

1 Abstract 2

2 Declaration of Authentication 4

3 Acknowledgements 5

4 Table of Contents 6

5 List of Abbreviations 9

6 Introduction 12

6.1 Gene therapy 12

6.2 Adeno-associated Viruses 13

6.2.1 A brief history of AAV 13

6.2.2 AAV Overview 15

6.2.3 Genome Organization 16

6.2.4 AAV Life Cycle 19

6.2.5 Entry and Transduction Pathway 21

6.3 AAV as a gene therapy vector 25

6.3.1 Vectorisation and Clinical Translation 25

6.3.2 Challenges in AAV Gene Therapy 28

6.3.3 AAV Manufacturing Systems 29

6.3.4 Inducible Systems in AAV Vector Engineering 31

6.4 AAV capsid Engineering 33

6.4.1 Motivation for Capsid Engineering 33

6.4.2 Rational Design 34

6.4.3 Directed Evolution 35

6.4.4 Computationally Guided Design 38

6.5 B Cell Biology 40

6.5.1 Central Role of B Cells in Immunity 40

6.5.2 B Cell Development 40

6.5.3 The Germinal Centre Reaction and B Cell Fates 41

6.5.4 Antibody effector functions 43

6.5.5 B Cell Receptor Signalling 45

6.5.6 Organization and Rearrangement of the Immunoglobulin Loci 46

6.6 Engineering B Cells to Express Therapeutic Antibodies 48

6.7 Cancer and the application of monoclonal antibodies 50

6.7.1 Cancer 50

6.7.2 Overview of Monoclonal Antibodies 51

6.7.3 Mechanisms of Action in Cancer Treatment 52

6.7.4 Clinical Impact 53

6.8 Thesis Overview 54

6.9 Overall Aims 56

7 Materials and Methods 57

7.1 Molecular Biology 57

7.1.1 Bacterial Transformation 57
7.1.2 DNA Purification 58
7.1.3 Polymerase Chain Reaction (PCR) 58
7.1.4 Restriction Enzyme Digestion 60
7.1.5 Agarose Gel Electrophoresis 61
7.1.6 PCR amplicon purification 61
7.1.7 Nucleic Acid Quantification 62
7.1.8 Illumina amplicon Sequencing 62
7.1.9 Sanger Sequencing 62
7.1.10 Whole plasmid nanopore sequencing 62
7.1.11 HiFi Assembly 62
7.1.12 Restriction-ligation cloning 63
7.1.13 Site directed mutagenesis of plasmid DNA 64
7.1.14 Assembly of Plasmid DNA Library 64

7.2 Cell Culture 66
7.2.1 Cell Line Maintenance 66
7.2.2 Peripheral Blood Mononuclear Cell (PBMC) Isolation 67
7.2.3 Primary B Cell Isolation and Culture 67
7.2.4 Primary Monocyte Isolation and Culture 68
7.2.5 Cryopreservation and recovery of cells 69
7.2.6 DNA Transfection 69
7.2.7 DNA Purification from mammalian cells 70
7.2.8 Production and purification of soluble Trastuzumab 71

7.3 Gene Editing 72
7.3.1 CRISPR RNA (crRNA) Design 72
7.3.2 Editing of Primary Human B Cells 73
7.3.3 Quantification of Double-Strand Break Efficiency 74

7.4 AAV Production and Characterisation 74
7.4.1 Small-scale adherent AAV productions 74
7.4.2 Production of nanobody-display AAVs 75
7.4.3 Large scale AAV production 75
7.4.4 Iodixanol Purification of AAV 76
7.4.5 Concentration of AAV 77
7.4.6 Measurement of AAV Physical Titre by qPCR 77
7.4.7 Assessment of AAV cross-packaging 79

7.5 Cell-based Assays 80
7.5.1 Flow cytometry 80
7.5.2 Microscopy 81
7.5.3 Fluorescent Reporter AAV Transduction Assay 81
7.5.4 AAV Binding/Internalisation Assay 82
7.5.5 Antibody-dependent Cell-mediated Phagocytosis (ADCP) Assay 84

7.6 AAV Library Screening 85
7.6.1 Variant Selection in Primary Human B Cells 85
7.6.2 Recovery of Selected Capsid Sequences 86

7.7 Common Methods 87
7.7.1 Polyacrylamide Gel Electrophoresis (PAGE) 87
7.7.2 Immunoblotting 88
7.7.3 SYPRO Ruby staining 89
7.7.4 Statistics 89

8 Development of a Two-Step AAV Capsid Library Platform 91

8.1 Introduction 91

8.2 Aims 95

8.3 Results 96

8.3.1 Characterisation of a CymR-VP16-inducible System. 96

8.3.2 Characterisation of a Cre-inducible System. 109

8.4 Discussion 119

9 Production and Screening of AAV Capsid Libraries Tailored Towards Primary human B Cells 127

9.1 Introduction 127

9.2 Aims 131

9.3 Results 132

9.3.1 Investigation into the bottlenecks impairing B Cell transduction by AAV. 132

9.3.2 Cloning of a shuffled AAV capsid library. 138

9.3.3 Cloning of a targeted random mutagenesis AAV capsid library. 144

9.3.4 Characterisation of viral AAV capsid library production 147

9.3.5 Measurement of AAV genome cross-packaging during viral library production. 156

9.3.6 Screening in primary human B Cells. 159

9.4 Discussion 165

10 Development of a B cell Engineering Strategy for Anti-cancer Antibody Expression 177

10.1 Introduction 177

10.2 Aims 181

10.3 Results 182

10.3.1 Validation of HDRT designs for the engineering of anti-cancer monoclonal-expressing primary human B cells. 182

10.3.2 Knock-in of DNA sequences into the IgHJ6 intron of primary human B cells 191

10.4 Discussion 199

11 Final Discussion 208

11.1 Rationale and Aims 208

11.2 Vectors to actualise human B cell therapies 209

11.3 B Cell Engineering for cancer 212

11.4 Future Directions 214

12 References 216

13 Appendix 241

13.1 AAV titration qPCR DNA standard sequence. 241

13.2 Binding/internalisation assay qPCR DNA standard sequence 241

13.3 Additional primer sequences 242

Trastuzumab-treated BT474 viability in ADCP assay 242

13.4 242

5 List of Abbreviations

AAP	Assembly-Activating Protein
AAV	Adeno-Associated Virus
AAVR	AAV Receptor
ADCP	Antibody-Dependent Cell-mediated Phagocytosis
ADCC	Antibody-Dependent Cell-mediated Cytotoxicity
ASCT1/2	Alanine-Serine-Cysteine Transporter 1/2
BaEV	Baboon Endogenous Retrovirus
bGH	Bovine Growth Hormone
BCR	B Cell Receptor
Cdc42	Cell Division Control Protein 42
cDNA	Complementary DNA
CFSE	Carboxyfluorescein Succinimidyl Ester
CH	Constant Heavy (region)
CL	Constant Light (region)
CLIC/GEEC	Clathrin-independent Carriers/GPI-enriched Early Endocytic Compartment
CLV	Capsid Library Vector
CMV	Cytomegalovirus (promoter)
CNS	Central Nervous System
crRNA	CRISPR RNA
Cs	Secretory Coding Sequence
CuO	Cumate Operator
CstF64K	Cleavage Stimulation Factor 64 kDa
CymR	Cumate Repressor
DC	Dendritic Cell
EF1α	Eukaryotic Translation Elongation Factor 1 alpha
EBV	Epstein-Barr Virus
eGFP	Enhanced Green Fluorescent Protein
FACS	Fluorescence-Activated Cell Sorting
Fc	Fragment Crystallisable region
FcγR	Fc γ Receptor
FLuc	Firefly Luciferase
FS	Frameshift
GC	Germinal Center
GC(/mL)	Genome Copies
CFSE	Carboxyfluorescein Succinimidyl Ester
GPR108	G Protein-coupled Receptor 108
GRAF1	GTPase Regulator Associated with Focal Adhesion Kinase-1
HDR	Homology-Directed Repair
HDRT	Homology-Directed Repair Template
HEK293	Human Embryonic Kidney 293 Cells
HER2	Human Epidermal Growth Factor Receptor 2

hr	Hour
HSPG	Heparan Sulfate Proteoglycan
HSV	Herpes Simplex Virus
ICE	Inference of CRISPR Edits
IDLV	Integrase-Deficient Lentiviral Vector
Ig	Immunoglobulin
IgH	Immunoglobulin Heavy Chain
ITR	Inverted Terminal Repeat
LLPC	Long-Lived Plasma Cell
LoxP	Locus of Crossover (in P1)
MAAP	Membrane-Associated Accessory Protein
MBC	Memory B Cell
MHC	Major Histocompatibility Complex
MiniCMV	Minimal CMV promoter
MOI	Multiplicity of Infection
mRNA	Messenger RNA
Nb	Nanobody
NGS	Next-Generation Sequencing
PAM	Protospacer Adjacent Motif
pAS	Polyadenylation Site
PBMC	Peripheral Blood Mononuclear Cells
PBS	Phosphate Buffered Saline
PCR	Polymerase Chain Reaction
PD-1	Programmed Cell Death Protein 1
PKD	Polycystic Kidney Disease (Domain)
PLA2	Phospholipase A2
Plad-B	Pladienolide B
PPP1R12C	Protein Phosphatase 1 Regulatory Subunit 12C
qPCR	Quantitative Polymerase Chain Reaction
rAAV	Recombinant AAV
RNP	Ribonucleoprotein
RT	Reverse Transcription
sHER2	Soluble HER2
shRNA	Short Hairpin RNA
SFFV	Spleen Focus-Forming Virus (promoter)
ssAAV	Single-Stranded Adeno-Associated Virus
scAAV	Self-Complementary AAV
SV40	Simian Virus 40 (polyA)
TESSA	Tetracycline-Enabled Self-Silencing Adenovirus
TFH	T Follicular Helper (cell)
TGN	Trans-Golgi Network
TM	Targeted Mutagenesis
TU	Transducing Units
UT	Untransfected
UTR	Untranslated Region
VG	Viral Genomes
VH	Variable Heavy

VL	Variable Light
VP1/2/3	Viral Protein 1, 2, 3
WT	Wild Type

6 Introduction

6.1 Gene therapy

Gene therapy, the modification of genetic material within human cells to treat disease, has evolved from a theoretical concept to clinical reality over five decades. Early proposals for somatic gene correction were articulated in the 1970s (Friedmann and Roblin, 1972). The first human gene therapy trial in 1990 used retroviral delivery of the adenosine deaminase (ADA) gene to treat children with severe combined immunodeficiency (Blaese et al., 1995).

The initial optimism of the 1990s was tempered by significant setbacks that reshaped regulatory and technical standards. A clinical trial for ornithine transcarbamylase deficiency using an adenoviral vector resulted in fatal inflammatory toxicity (Raper et al., 2003). Another trial for X-linked severe combined immunodeficiency (SCID-X1), involving transplantation of hematopoietic stem cells following *ex vivo* retroviral gene transfer, led to deregulated premalignant proliferation due to integration near proto-oncogenes (Hacein-Bey-Abina et al., 2003). These events prompted a reassessment of vector design and leading to the adoption of new vector systems.

Modern gene therapy encompasses a range of delivery platforms that introduce genetic material either transiently or stably. Non-viral approaches, such as electroporation, nanoparticle formulations, and lipid-based carriers enable transient expression but have limited efficiency in vivo. In contrast, viral vectors exploit evolved mechanisms for

cellular entry and genome delivery, in some cases achieving higher efficiency gene delivery and stable expression. Among these, adeno-associated virus (AAV), lentivirus and adenovirus have become central gene therapy vectors, each characterised by distinct integration profiles, tissue tropisms, and immunological properties.

AAV vectors have a favourable safety profile and provide durable gene expression, with persistent episomal genomes in slowly dividing cells. These attributes have enabled landmark clinical successes including the first FDA-approved AAV-based therapies: voretigene neparvovec (Luxturna, 2017) for retinal dystrophy, onasemnogene abeparvovec (Zolgensma, 2019) for spinal muscular atrophy, and etranacogene dezaparvovec (Hemgenix, 2022) for haemophilia B. A further four AAV gene therapies have since been approved for haemophilia A, Duchenne muscular dystrophy, haemophilia B and aromatic L-amino acid decarboxylase deficiency. Together, these advances have transformed AAV gene therapy into a clinically validated therapeutic platform capable of providing durable therapeutic benefit for monogenic diseases.

6.2 Adeno-associated Viruses

6.2.1 A brief history of AAV

In the mid-1960s, electron microscopy of human and simian adenovirus preparations revealed the presence of small, 20–25 nm particles that could be physically separated from adenovirus virions. These particles propagated only in the presence of adenovirus and were initially interpreted as either adenoviral fragments (Mayor et al., 1965) or defective viruses (Atchison et al., 1965). Subsequent studies established that they were

antigenically distinct from adenoviruses and capable of replication in the presence of heterologous adenoviruses (Hoggan et al., 1966), adding weight to the evidence that they represented distinct viruses. This study also established that different AAV isolates were serologically distinct. Clinical samples from children with adenoviral infections confirmed their presence in humans and suggested that they were non-pathogenic (Blacklow et al., 1968, Blacklow et al., 1967).

By 1969, AAV was shown to package plus- or minus-strand single-stranded DNA genomes (Rose et al., 1969) and in the 1970s, it was established that AAV could deliver DNA to the nucleus without the need for helper virus (Rose and Koczot, 1972). Furthermore, in the absence of adenovirus, AAV could persist in a latent state and be rescued upon superinfection with a helper virus (Berns et al., 1975). Later, integration studies demonstrated site-specific integration at the human chromosome 19 (Kotin et al., 1990).

The development of recombinant DNA technology in the early 1980s enabled sequencing and annotation of the AAV2 genome, including the inverted terminal repeats (ITRs) (Srivastava et al., 1983, Lusby et al., 1980). Cloning of AAV components into plasmids soon enabled the first demonstrations of AAV-mediated gene transfer into mammalian cells (Tratschin et al., 1984). Work in the late 1980s and 1990s clarified the functions of the AAV Rep proteins in replication, integration, and transcriptional regulation (Beaton et al., 1989a, Weitzman et al., 1996, Owens et al., 1993, Tratschin et al., 1986, Im and Muzyczka, 1990, Snyder et al., 1993, Snyder et al., 1990).

In the 2000s, the isolation of multiple naturally occurring AAV serotypes from primate tissues (Gao et al., 2002) expanded the known diversity of capsid sequences and tropisms. A landmark sequence analysis of AAV isolates from human tissues revealed that AAV sequences diverge into distinct clades, which are broadly distributed across different human tissues (Gao et al., 2004).

6.2.2 AAV Overview

Wild-type AAV is a small non-enveloped virus in the *Dependoparvovirus* genus of the *Parvoviridae* family. The virion measures approximately 20–25 nm in diameter and exhibits T = 1 icosahedral symmetry. The capsid is built from 60 subunits of the viral proteins VP1, VP2, and VP3 arranged in an approximate 1:1:10 stoichiometry (Rose et al., 1971). Structurally, each VP subunit core adopts the conserved parvoviral jelly-roll beta-barrel fold (Xie et al., 2002). Interconnecting loops between the beta-sheets define the surface topology of the capsid, which contains distinctive peaks and depressions around 2-, 3- and 5-fold axes of symmetry. Mutagenesis has mapped AAV2 heparan-sulfate binding and antigenic sites to these features, in particular the peaks around the 3-fold axis formed by interdigitating loops (Xie et al., 2002). The 5-fold axis consists of a pore through which genomes are loaded into the assembled capsids during packaging (Bleker et al., 2006). The interior of the capsid is lined by the β-barrel cores of the VPs and contains the structurally unresolved globular VP N-terminus (Kronenberg et al., 2005).

The AAV genome is a single-stranded DNA (ssDNA) of about 4.7 kilobases, and both positive- and negative-sense strands can be packaged with similar efficiency (Berns and

Rose, 1970). The genome is flanked by 145 nucleotide inverted terminal repeat (ITR) sequences (Figure 1b) which form T-shaped hairpin structures (Berns and Kelly, 1974, Lusby et al., 1980) that self-prime second-strand synthesis (Samulski et al., 1983) and contain elements required for genome replication and packaging (Im and Muzyczka, 1990). In fact, the ITRs are the only cis-acting AAV sequences required for packaging the genome into capsids (Hermonat and Muzyczka, 1984, Tratschin et al., 1984), a trait that is central to their success as recombinant vectors for gene therapy.

Wild-type AAV has long been considered non-pathogenic. Recent reports have found an association between AAV2 infection and clusters of cases of unexplained acute hepatitis in children (Servellita et al., 2023, Morfopoulou et al., 2023, Ho et al., 2023). However, no causative link between AAV and human disease has been identified.

6.2.3 Genome Organization

The wild-type AAV genome (Figure 1c) comprises two major genes, replication (Rep) and capsid (Cap) genes**,** each encoding multiple overlapping proteins. The Rep gene gives rise to four non-structural proteins, Rep78, Rep68, Rep52, and Rep40, which all share a 305–amino acid region containing a helicase/ATPase domain and nuclear localisation signals. Expression of Rep78 and Rep68 is driven by the p5 promoter, whereas the Rep gene-internal p19 promoter produces Rep52 and Rep40. Rep78 and Rep68 additionally contain a 224–amino acid N-terminal DNA-binding and endonuclease domain, enabling their roles in genome replication through interactions with the ITRs, as well as context-dependent positive or negative regulation of AAV transcription (Pereira et al., 1997). The

C-terminal ends of Rep78 and Rep52 differ from Rep68 and Rep40 through alternative splicing. Rep52 and Rep40 are involved in the accumulation of single-stranded genomes for packaging (Chejanovsky and Carter, 1989).

The p40 promoter produces a pre-mRNA transcript that utilises a single splice donor and alternative splice acceptors to produce two mRNAs encoding the three structural VPs, as well as out-of-frame accessory proteins. All three structural proteins share a common 533-amino acid region that makes up the entirety of VP3. VP1 and VP2 share 65 amino acids and VP1 contains a unique 137-amino acid N-terminal extension. The alternative splice acceptors flank the VP1 start codon so that the largest of the two transcripts encodes the full VP1 coding sequence and the smaller transcript encodes both the VP2 and VP3 coding sequences. VP2 utilises an inefficient ACG start codon, allowing frequent utilisation of the downstream VP3 AUG start codon (Trempe and Carter, 1988). This combination of alternative splicing and alternative start codon usage governs the abundance of each capsid protein.

The assembly-activating protein (AAP) is translated from a +1 reading frame within the VP2/3 coding sequence using a nonconventional CUG start codon (Sonntag et al., 2010). AAP has an N-terminal hydrophobic region and appears to bind newly synthesised VP proteins in the nucleolus, acting as a scaffold to facilitate proper capsid assembly (Naumer et al., 2012). Dependence on AAP varies by serotype but appears to have evolved to allow more freedom in the evolution of capsid sequences while maintaining capsid stability (Maurer et al., 2018).

More recently, two additional small proteins encoded within the wild-type AAV genome have been identified. The membrane-associated accessory protein (MAAP) is a ~12 kDa protein encoded within a +1 reading frame located mostly within the VP1-unique N-terminal region. MAAP contains a cationic amphipathic membrane anchoring N-terminus and appears to be important for viral egress (Elmore et al., 2021).

For AAV2, a small 155 amino-acid protein X is expressed from the p81 promoter, and its coding sequence overlaps the C-terminus of the VPs in a +1 reading frame. AAV2 protein X appears to enhance DNA replication and viral yield, especially in adenovirus-infected cells. However, the coding sequence is smaller in other serotypes and can be disrupted. The AAV2 protein X is not necessary for replication (Cao et al., 2014).

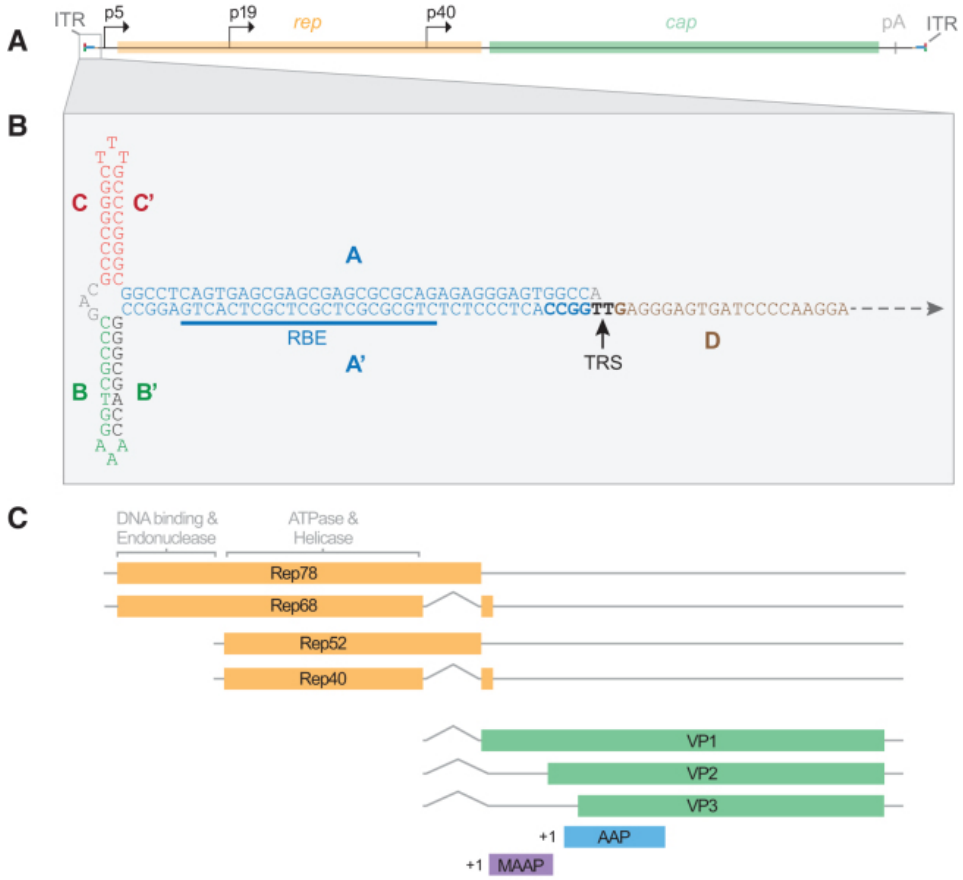


Figure 1 - Anatomy of the AAV ITR and AAV genome. A) Basic AAV genome organisation. B) ITR sequence and hairpin structure. C) AAV proteins. From: Maurer and Weitzman (2020).

6.2.4 AAV Life Cycle

In the absence of helper virus, wild-type AAV establishes a latent, non-productive infection in which its gene expression is autorepressed. During a latent infection, the AAV genome can persist episomally as a double stranded circular concatemer, or it can integrate into the host genome (Kotin et al., 1990, Yang et al., 1999). The integration of the AAV genome is directed by Rep78/68 using consensus genomic Rep-binding elements. An insertion site within the first exon of the human PPP1R12C gene (AAVS1) on

chromosome 19 was the first to be identified (Kotin et al., 1990); however, additional sites have since been identified and appear to occur with increased frequency in regions with high GC content and areas of open chromatin (Huser et al., 2010). Further studies have shown that integration patterns can vary by serotype (Janovitz et al., 2014) and cell type (Huser et al., 2014).

In the presence of a helper virus such as adenovirus, herpesvirus-family members (Atchison, 1970) and papillomaviruses (Ogston et al., 2000), AAV gene expression is activated and Rep liberates the AAV genome from the integration site or episomal concatemer, initiating genome replication. The large Rep proteins (Rep68 and Rep78) bind to the Rep-binding element (RBE) within each ITR and introduce a site-specific nick at the terminal resolution site (trs), generating a 3' hydroxyl that primes DNA synthesis (Figure 1b). Host DNA polymerase δ extends from this primer to produce a double-stranded replicative-form intermediate, which serves as a template for further replication rounds. Resolution of the terminal hairpin yields plus- and minus-strand single-stranded genomes that are packaged into pre-assembled capsids. The helper functions of human adenovirus type 5 (hAdV5) are the best studied and form the basis for most recombinant AAV producer systems. The adenoviral E1A/B, E2A, E4orf6 and VA RNA I genes are widely reported as the minimal set of helper genes for AAV production (Xiao et al., 1998, Matsushita et al., 1998); however, this has recently been refined to E2A, VA RNA I and L4-22K in HEK293 cells which provide E1A/B (Doshi et al., 2024).

6.2.5 Entry and Transduction Pathway

The earliest interaction between AAV and host cells is mediated by glycans on the cell surface. These interactions concentrate virions at the plasma membrane and increase the likelihood of receptor engagement. Glycan specificity varies between serotypes: AAV2, 3B, and 6 recognise heparan sulfate proteoglycans (HSPG) (Halbert et al., 2001, Handa et al., 2000, Summerford and Samulski, 1998b); AAV1, -4, -5, and -6 bind different sialic acid linkages (Kaludov et al., 2001, Wu et al., 2006); while AAV9 attaches to galactose moieties (Shen et al., 2011). Structural and functional analyses have mapped discrete binding regions, such as the Arg585–Arg588 heparin-binding domain of AAV2 (O'Donnell et al., 2009, Kern et al., 2003) and dual sialic acid-binding sites AAV5 (Afione et al., 2015).

These interactions are typically of low affinity and disruption does not necessarily abrogate transduction, indicating that glycans function more as attachment factors than true entry receptors. Even so, they strongly influence tissue tropism. For example, mutations that disrupt HSPG binding in AAV2 alter in vivo tropism (Kern et al., 2003), while disrupting sialic acid recognition in AAV5 abolishes lung transduction but enhances delivery to salivary glands and muscle (Afione et al., 2015). Glycans therefore play a critical role in concentrating virions and shaping tissue specificity, even if they are insufficient for productive infection on their own.

The identification of KIAA0319L, now termed the AAV receptor (AAVR), transformed understanding of AAV entry. Genetic screening demonstrated that AAVR is required for efficient transduction by most serotypes, with the exception of AAV4 and some divergent

simian strains (Dudek et al., 2018, Pillay et al., 2016). AAVR contains five extracellular polycystic kidney disease (PKD) domains (Pillay et al., 2016), and different serotypes show distinct binding preferences: AAV2 (Zhang et al., 2019a), AAV1, and AAV8 engage PKD2, whereas AAV5 binds PKD1 (Pillay et al., 2017).

Structural studies revealed that AAV2 interacts with PKD2 near the threefold spikes of the capsid (Zhang et al., 2019a). The binding footprints for AAV2 (Meyer et al., 2019) and AAV5 (Zhang et al., 2019b) overlap with epitopes targeted by neutralising antibodies, reinforcing the critical role of AAVR in AAV transduction. AAVR is not essential for the initial binding step at the cell surface (Dudek et al., 2018) but becomes critical after endocytosis. AAVR cycles between the plasma membrane and the Golgi (Pillay et al., 2016), and retrograde transport to the Golgi is critical for AAV2 transduction (Nonnenmacher et al., 2015b), suggesting that AAVR may play a role in guiding AAV towards the trans-Golgi network.

After attachment and receptor binding, AAV enters the cell by endocytosis. Early work suggested that AAV2 requires clathrin-mediated uptake (Nonnenmacher and Weber, 2011). However, knockdown of clathrin does not block uptake, suggesting that clathrin-mediated endocytosis is not a requirement for AAV2 transduction. Increasing evidence points to uptake through clathrin-independent carriers and the GPI-enriched early endocytic compartment (CLIC/GEEC). AAV internalisation is sensitive to cholesterol depletion and depends on Cdc42 and GRAF1, both central regulators of this pathway (Nonnenmacher and Weber, 2011).

Once inside the cell, AAV must traffic to the trans-Golgi network (TGN). Disruption of Golgi structure with agents such as Brefeldin A or Golgicide A abolishes transduction. Unlike classical endosome-to-TGN pathways that depend on Rab7/11 and syntaxins 6/16, AAV uses a non-canonical route that requires syntaxin 5. Knockdown of syntaxin 5 prevents productive infection, whereas loss of syntaxin 6 or 16 has little effect (Nonnenmacher et al., 2015b). GPR108, a Golgi-resident protein, is needed for transduction by nearly all serotypes except AAV5, but the precise role of GPR108 remains unknown (Dudek et al., 2020).

The N-terminal region of VP1 plays a pivotal role in AAV transduction by enabling critical post-entry steps. Although VP1 is dispensable for capsid assembly and genome packaging, mutations in its unique N-terminal region significantly impair transduction efficiency without affecting uptake or intracellular trafficking to perinuclear compartments (Girod et al., 2002). The VP1 phospholipase 2 (PLA2) domain, conserved across parvoviruses (Zadori et al., 2001), hydrolyses phospholipids to disrupt endosomal membranes, thereby promoting capsid escape into the cytosol. Mutations within the catalytic HDXXY motif or the Ca^{2+}-binding GXG loop abolish PLA2 activity and reduce infectivity by several orders of magnitude (Stahnke et al., 2011).

This region was predicted to be externalised through the 5-fold symmetry pores under endosomal acidification, exposing the PLA2 domain and clusters of basic residues functioning as nuclear localisation signals (NLS) (Sonntag et al., 2006). These signals facilitate subsequent nuclear import of the capsid or its genome. Cryo-electron microscopy of AAV9 confirmed that decreasing pH triggers conformational

rearrangements leading to N-terminal externalisation, with PLA2 exposure at lysosomal pH supporting membrane disruption (Penzes et al., 2021). Therefore, the VP1 unique region appears to catalyse endosomal escape and target intact AAV particles to the nucleus. Import into the nucleus occurs through nuclear pore complexes and depends on importin β (Nicolson and Samulski, 2014). Entry is relatively inefficient and imaging studies suggest that fewer than 20% of virions that dock at the pore successfully enter the nucleus (Kelich et al., 2015). Once inside, it appears that transient accumulation in the nucleolus is a requirement for efficient transduction (Johnson and Samulski, 2009). The uncoating step in the nucleus is poorly understood, but uncoating efficiency appears to correlate somewhat with capsid stability (Rossi et al., 2019, Thomas et al., 2004).

6.3 AAV as a gene therapy vector

6.3.1 Vectorisation and Clinical Translation

The transformation of adeno-associated virus (AAV) from a virological curiosity into a leading gene therapy vector was the result of successive advances in molecular virology and translational studies over four decades. In the 1980s, recombinant AAV2 was generated by cloning the viral genome into plasmids and replacing its coding sequences with foreign DNA flanked by inverted terminal repeats (ITRs). These early studies established the ITRs as the only cis-acting elements required for replication and packaging, demonstrated successful gene transfer into mammalian cells, and established methods for the production of recombinant AAV (rAAV) without wild-type AAV contamination (Hermonat and Muzyczka, 1984, McLaughlin et al., 1988, Samulski et al., 1989).

The 1990s marked the beginning of clinical translation. The cystic fibrosis transmembrane conductance regulator (CFTR) gene was one of the first therapeutic cassettes packaged in rAAV2. AAV2-CFTR was administered to the sinus epithelium, as a model for the lung, of patients with cystic fibrosis. This phase I study demonstrated the safety of airway delivery but revealed only transient expression, presumably due to turnover of epithelial cells. The lack of acute inflammation was seen as a significant advantage over the previous adenovirus clinical trials (Wagner et al., 1999).

Haemophilia B was among the next diseases targeted. In 2000, Kay and colleagues reported the first intramuscular trial of a rAAV2 encoding factor IX (FIX). The vector was

injected directly into skeletal muscle of patients with severe haemophilia B, leading to local expression detectable by muscle biopsy but no sustained increases in circulating FIX levels (Kay et al., 2000).

In 2006, Manno and colleagues reported results from a phase 1/2 trial in which they infused rAAV2-FIX via the hepatic artery in patients with severe haemophilia B. This trial provided the first evidence that hepatocytes could be transduced *in vivo* in humans, with transient rises in plasma FIX levels detected. However, expression was not durable and was lost after several weeks due to capsid-specific CD8+ T cell responses that eliminated transduced hepatocytes (Manno et al., 2006), an observation that had not been anticipated using mouse models (Mingozzi et al., 2007).

Improvements came with the development of vectors using newly discovered natural serotypes. AAV8 in particular showed markedly enhanced liver tropism in preclinical models (Gao et al., 2002), and when tested in humans, it achieved sustained factor IX activity at therapeutic levels following intravenous infusion (Nathwani et al., 2011). Expression was further boosted by using the hyperactive Padua FIX variant (Simioni et al., 2009), which enabled sustained benefits using the bioengineered capsid Spark100 and natural AAV5 (George et al., 2017, von Drygalski et al., 2020). These studies laid the groundwork for the approvals of fidanacogene elaparvovec (Beqvez) and etranacogene dezaparvovec (Hemgenix).

Haemophilia A posed a significant challenge due to the large size of the factor VIII coding sequence. This was addressed by constructing B-domain–deleted (BDD) variants of FVIII

that fit within the AAV packaging limit (Lind et al., 1995). The first liver-directed trial of an AAV5-BDD-FVIII vector showed sustained FVIII activity (Rangarajan et al., 2017). Subsequent larger studies confirmed that therapeutic levels could be achieved, but durability remained variable, with activity declining over years (Mahlangu et al., 2023, Ozelo et al., 2022). These pivotal trials underpinned the approval of valoctocogene roxaparvovec (Roctavian).

Success was also achieved in the retina. In 2008, three independent groups demonstrated that subretinal injection of rAAV2-RPE65 could partially restore retinal function in patients with Leber congenital amaurosis (Bainbridge et al., 2008, Cideciyan et al., 2008, Maguire et al., 2008). Follow-up studies confirmed long-term efficacy and safety, culminating in the 2017 FDA approval of voretigene neparvovec (Luxturna) (Russell et al., 2017).

AAV9 opened the path to neuromuscular and neurological applications. In 2017, Mendell and colleagues reported systemic delivery of rAAV9-SMN1 in infants with spinal muscular atrophy (SMA). A single intravenous infusion of onasemnogene abeparvovec (Zolgensma) resulted in dramatic improvements in survival and motor function compared to the untreated natural history (Mendell et al., 2017). These results revolutionised SMA therapy and secured regulatory approval in 2019.

For the treatment of Duchenne muscular dystrophy (DMD), micro-dystrophin constructs were developed to enable the packaging of dystrophin coding sequences into AAV. Intravenous delivery of an AAVrh74 vector encoding micro-dystrophin driven by the

MHCK7 promoter led to expression in muscle fibres and improvements in functional outcomes (Mendell et al., 2020), leading to the approval of delandistrogene moxeparvovec (Elevidys) in 2023.

6.3.2 Challenges in AAV Gene Therapy

Despite the clinical success of AAV gene therapies, several biological and translational challenges limit the broader application of AAV gene therapy. The most fundamental is the limited packaging capacity of ~4.7 kb, which precludes the delivery of very large genes such as full-length dystrophin or FVIII. Potential workarounds include deletion of non-critical domains and the use of dual vectors.

Host immunity presents another major barrier. Between 30–70% of humans carry neutralising antibodies to common AAV serotypes (Chhabra et al., 2024). Even in seronegative patients, administration can elicit the development of anti-capsid antibodies, preventing re-dosing (Mietzsch et al., 2025). In the first liver-directed haemophilia B trial, clearance of transduced hepatocytes by capsid-specific CD8+ T cells was demonstrated (Manno et al., 2006, Mingozzi et al., 2007).

Toxicities associated with high vector doses have emerged in systemic applications where some treatment groups can receive more than 1×10^{14} viral genomes/kg, particularly for neuromuscular trials. At such doses, patients have experienced acute hepatotoxicity, complement activation and thrombotic microangiopathy, with

occasional fatalities (Duan, 2023). Reducing required doses through more potent capsids and cassettes is therefore a priority.

Some concern of insertional mutagenesis has also been raised. While wild-type AAV integrates preferentially at AAVS1, rAAV vectors largely persist episomally due to the absence of ongoing Rep expression. Non-specific integration events do occur at very low frequency for rAAV (Russell et al., 2025), and a mouse study reported hepatocellular carcinoma (HCC) associated with rAAV integration at the Rian locus (Donsante et al., 2007). However, no HCC has been reported in non-human primates or humans. Recently, a deep analysis of rAAV integrations in patients treated with Roctavian found extremely low rates of integration and found that integration sites were widely distributed with a small bias towards open chromatin regions (Russell et al., 2025). Together, this suggests that the risk of insertional oncogenesis is extremely low for replacement gene therapies using rAAV. However, frequent insertion of AAV genomes has been observed in the presence of persistent CRISPR-Cas nuclease expression (Suchy et al., 2025, Hanlon et al., 2019).

6.3.3 AAV Manufacturing Systems

rAAV manufacturing comprises helper-free transfection, infection-based systems, stable producer lines, and downstream purification. In helper-free production, HEK293 cells are co-transfected with plasmids encoding an ITR-flanked vector genome, AAV Rep/Cap genes, and adenoviral helper genes. This approach was established to generate high-titre vector without infectious adenovirus (Xiao et al., 1998). Production modelling suggest

that inefficient nuclear delivery and poor coordination of genome replication with capsid supply represent major bottlenecks in rAAV production suggesting that adjusting the timing of Rep and Cap expression could reduce empty capsids (Nguyen et al., 2021). Stable producer lines aim to embed the rAAV genome, Rep/Cap, and helper cassettes under inducible control to improve scalability and cost, and have been constructed and subsequently optimized to increase per-cell yields and full-particle proportions (Lee et al., 2022, Lu and Hu, 2024).

Infection-based methods deliver helper functions via replicating vectors. Baculovirus expression systems infect Sf9 insect cells with baculoviruses encoding Rep/Cap and the rAAV genome to assemble vector at scale (Urabe et al., 2002). Recombinant HSV approaches use separate HSV vectors carrying Rep/Cap and the rAAV genome to infect mammalian cells. This platform supported clinical-grade rAAV1 used for α1-antitrypsin gene transfer (Conway et al., 1999, Flotte et al., 2011). However, these systems require specialized purification to eliminate helper-virus contaminants.

The tetracycline-enabled self-silencing adenovirus (TESSA) system was developed to enable the high yields and high quality of infection-based systems without the need for extensive downstream purification. By inserting tetracycline operator sequences downstream of the adenoviral major late promoter (MLP) and expressing the tetR repressor under MLP control, the expression of adenoviral late genes is self-repressed while AAV "helper" functions are retained. Combining a TESSA carrying the rAAV genome with a TESSA-RepCap vector increased rAAV2 yield up to 30-fold relative to helper-free transfection. Furthermore, the AAV particles appeared to exhibit superior transduction

efficiency. Only trace levels of replication attenuated adenoviral contamination was measured in crude lysates. The approach extended to rAAV6/8/9 with higher yields and, for some serotypes, lower proportions of empty particles (Su et al., 2022a).

6.3.4 Inducible Systems in AAV Vector Engineering

Switchable gene control systems in adeno-associated virus (AAV) vectors enable temporal regulation of therapeutic transgene expression.

The tetracycline (Tet) system is an established switch that has been used with AAV vectors. Reverse Tet transactivators (rtTA) bind tet operators to activate a minimal promoter in the presence of doxycycline, and compact, autoregulatory architectures can co-package regulator and payload. A liver-specific bidirectional Tet-On AAV8 achieved low off-state expression and approximately 250-fold induction with rapid reversibility (approximately 48 h) after doxycycline withdrawal (Vanrell et al., 2011). Advantages of Tet-like frameworks include high dynamic range and drug titratability, whereas payload burden, off-state leak, and potential immunogenicity of bacterial-derived components remain concerns (Domenger and Grimm, 2019).

RNA-based switches (aptazymes/riboswitches) provide compact, protein-free regulation suited to AAV. A tetracycline-responsive K19 aptazyme controlled AAV-mediated expression in mice (Strobel et al., 2020). Dual riboswitch modules embedded in an AAV-VEGF-B cassette produced rapid induction with reduced leakiness (Eriksson et al., 2022). These designs minimize immunogenic burden and conserve genome space, though their

dynamic range is typically lower than protein-based systems and ligand delivery can be limiting in vivo (Eriksson et al., 2022).

Cre/loxP recombinase logic enables irreversible or semi-reversible control. Cre-dependent “FLEX/DIO” cassettes support stringent on-switching by recombinase-mediated inversion (Atasoy et al., 2008) and are widely used for cell-type-defined activation. Tamoxifen-inducible CreERT2 adds an additional layer of temporal control where recombination follows ligand-driven nuclear translocation (Hirrlinger et al., 2009). These strategies eliminate basal transcription from a silent orientation and can create permanent “on” states, but they preclude dose titration and require careful design to prevent unwanted recombination during production.

Additional modalities are emerging. Small-molecule-induced proximity can recruit endogenous activators to rAAV genomes for tuneable, reversible activation without expressing bulky transactivators (Wasserman et al., 2025). Finally, miRNA detargeting and synthetic promoters layer tissue or state specificity onto inducible logic (Domenger and Grimm, 2019). Genetic control systems must balance fold-induction, leak, reversibility, ligand pharmacology, sequence size limits, and immunogenicity.

6.4 AAV capsid Engineering

6.4.1 Motivation for Capsid Engineering

Serotypes display several limitations that restrict their therapeutic utility, including suboptimal tissue tropism, limited transduction efficiency in human target cells, and susceptibility to neutralizing antibodies prevalent in human populations. These constraints necessitate high vector doses in clinical applications, which raises production costs and can trigger immune toxicity. For example, in the ASPIRO trial for X-linked myotubular myopathy, administration of a high-dose AAV8 vector led to two fatal cases of hepatotoxicity (Wilson and Flotte, 2020). Such findings underscore the need for vectors that achieve potent and selective transduction at reduced doses. In addition, pre-existing antibodies against common serotypes such as AAV2, AAV8, and AAV9 reduce efficacy and prevent re-administration (Calcedo et al., 2009, Chhabra et al., 2024).

Because the structural determinants of receptor binding, immune recognition, and intracellular trafficking are encoded by the AAV capsid, engineering the capsid itself provides a direct route to improve performance. Multiple complementary strategies, including rational modification, directed evolution, and computational design, have been utilised to expand the toolkit of available AAVs. Rational design uses structural and biochemical knowledge to introduce specific sequence changes, while directed evolution screens large libraries of variants under defined selective pressures. More recently, computational or machine-learning (ML)–guided approaches have been developed with the aim to optimise the available sequence diversity in capsid libraries

for maximum functional diversity. Collectively, these methods aim to generate variants that achieve targeted delivery, immune evasion, and improved manufacturability.

6.4.2 Rational Design

Early capsid engineering exploited known structural features to modify receptor interactions and intracellular stability. One of the earliest examples was substitution of surface-exposed tyrosine residues in AAV2, which prevented phosphorylation and ubiquitin-mediated degradation of virions, thereby enhancing nuclear trafficking and transduction efficiency up to 30-fold in mice (Zhong et al., 2008).

The determination of the crystal structures across natural serotypes have contributed significantly to capsid engineering. Wang *et al* constructed the chimeric AAV9.HR capsid by introducing two substitutions that map to the 3-fold protrusions (H527Y and R533S) from a chimpanzee-derived AAV9 variant sequence into the AAV9 capsid. This variant retained the ability of AAV9 to traverse the blood-brain barrier and to transduce the central nervous system (CNS), while detargeting the liver in mice (Wang et al., 2018), demonstrating that subtle modifications can significantly retarget AAV vectors.

Insertion of heterologous peptides has provided another rational route to alter tropism. Incorporation of a cell-penetrating peptide (TVSALK) between residues 588 and 589 of AAV9 generated AAV.CPP.16, which showed up to 249-fold and 5-fold enhanced transduction in the mouse and non-human primate CNS, respectively (Yao et al., 2022). Furthermore, insertion of a fibroblast-activation-protein-specific nanobody into

structurally tolerant capsid positions such as N262 yielded variants capable of selectively transducing FAP-positive cells (Hoffmann et al., 2024).

Rational approaches have also employed covalent chemical modification to improve immune evasion and targeting. Polyethylene glycol (PEG) conjugation to AAV2 capsids modestly increased resistance to antibody neutralization (2.3-fold) but excessive conjugation impaired infectivity (Lee et al., 2005). Attachment of N-acetylgalactosamine ligands, which binds the asialoglycoprotein receptor on hepatocytes, enhanced AAV2 uptake 2.4–5.6-fold while simultaneously reducing neutralization four-fold *in vitro*. However, no advantage in transduction efficiency was measured *in vivo* in mice (Mével et al., 2020).

Cryo-electron microscopy and comparative modelling have enabled the mapping of antigenic determinants and receptor-binding sites across serotypes. Inserting or replacing residues within these motifs has produced vectors such as CAM130 with reduced antibody recognition (Tse et al., 2017) and AAV-TT, an AAV2 variant combining conserved residues from seropositive human AAV2 isolates that displays improved thermal stability and strong CNS tropism in mice (Tordo et al., 2018).

6.4.3 Directed Evolution

Rational design is limited by incomplete understanding of AAV biology. Directed evolution provides a complementary, empirical approach in which large libraries of variant capsids are generated by mutagenesis, peptide insertion, or recombination and subjected to

functional screening to enrich variants with desired traits. Although this approach is often referred to as directed evolution, successive rounds of screening are not necessarily accompanied by re-diversification of capsid sequences and therefore may not mimic an evolutionary process.

Seminal work by Maheshri et al. (2006) used random mutagenesis of AAV2 to create a diverse library that was selected for altered receptor affinity and resistance to neutralising serum. The resulting variant required 96-fold higher antibody concentrations to neutralise compared to AAV2 (Maheshri et al., 2006) . Subsequent strategies have expanded library diversity through random peptide insertion at the three-fold protrusion (Perabo et al., 2003) and DNA shuffling among parental serotypes (Koerber et al., 2008b), producing novel capsids with unique tropisms. Dalkara et al. performed an *in vivo* screen using a mixed library of shuffled variants and peptide insertion variants, with error-prone PCR re-diversification between rounds, to identify variants capable of penetrating the inner limiting membrane of the retina after intravitreal injection. The resulting 7m8 vector was capable of photoreceptor transduction in mice (Dalkara et al., 2013).

Directed evolution has also generated muscle- and heart-tropic vectors. Tabebordbar *et al* inserted random 7-mer peptides into AAV9's hypervariable loop VIII and applied iterative in vivo selection in mice and primates to yield MyoAAV variants with 10–29-fold enhanced muscle transduction and diminished liver transduction. A second-generation variant achieved up to 128-fold improved transduction over AAV9 in human myotubes and therapeutic benefit in murine muscular-dystrophy models at 10- to 250-fold lower doses (Tabebordbar et al., 2021).

Model choice strongly affects selection outcomes. Deverman *et al* evolved AAV-PHP.B for efficient blood–brain-barrier crossing using a Cre-dependent selection in mice (Deverman et al., 2016), but subsequent studies showed that its neurotropism was strain-restricted and not reproduced in primates (Hordeaux et al., 2018). In contrast, AAV-Cap.Mac, evolved directly in macaques, achieved robust neuronal and glial transduction *in vivo* (Chuapoco et al., 2023). To generate cross-species variants, Rosales *et al* cycled capsid libraries through murine, porcine, and primate kidney models as well as human kidney organoids, identifying AAV.k13 and AAV.k20 variants with superior transduction efficiency over AAV9 across species (Rosales et al., 2025).

Directed evolution has also targeted immune evasion. Havlik *et al* combined cryo-EM mapping of AAV8 antigenic footprints with saturation mutagenesis and selection in human hepatocytes, generating AAVhum.8, which demonstrated 10-fold higher hepatocyte transduction in a humanised mouse model and reduced antibody neutralisation (Havlik et al., 2020). This study exemplifies integration of structural and evolutionary strategies.

High-throughput sequencing has accelerated these approaches. DNA barcoding enables multiplexed assessment of variants using short read next-generation sequencing (Pekrun et al., 2019), but recombination during packaging appears to introduce significant noise (Lalanne et al., 2025).

6.4.4 Computationally Guided Design

As data from high-throughput selection accumulate, computational approaches have emerged to analyse sequence–function relationships and design improved variants *in silico*. Early computational frameworks such as SCHEMA (Voigt et al., 2002) identified recombination breakpoints that preserve structural integrity during capsid shuffling. Applying this algorithm to six AAV serotypes yielded a chimeric library with a theoretical diversity of 1.6 million variants and led to discovery of SCH9, which displayed 24-fold greater neural-stem-cell transduction and increased antibody resistance (Ojala et al., 2018).

Ancestral sequence reconstruction represents another computational strategy. Santiago-Ortiz *et al* generated a phylogenetically inferred ancestral AAV library, identifying broadly infectious capsids with enhanced thermostability and independence from canonical glycan receptors (Santiago-Ortiz et al., 2015). Zinn *et al* used a similar approach to produce Anc80L65, which exhibited high stability and efficient transduction in mouse muscle, liver, and retina (Zinn et al., 2015).

Machine-learning methods have since broadened the scope of computational capsid design. Ogden *et al* created a comprehensive AAV2 fitness landscape by testing all possible single-codon substitutions, deletions, and insertions, linking sequence data to phenotypes such as packaging, thermostability, and antibody neutralization. Training supervised models on these datasets enabled prediction of novel functional sequences. The same study identified an unannotated open reading frame, the membrane-associated accessory protein (MAAP), that regulates viral egress (Ogden et al., 2019).

While still nascent, computational approaches promise to vastly increase the efficiency of AAV capsid discovery by maximising the functional diversity and viability of library variants.

6.5 B Cell Biology

6.5.1 Central Role of B Cells in Immunity

B cells play central roles in adaptive immunity through the generation of antibodies, the formation of immune memory, and antigen presentation to T cells. Naïve B cells express clonally unique B cell receptors (BCRs) formed by V(D)J recombination of immunoglobulin gene segments, enabling recognition of diverse antigens. Upon antigen encounter, B cells internalise antigen–BCR complexes and present processed peptides via MHC-II to T helper cells, which help to shape the immune response. Following activation, some B cells enter germinal centres where they undergo affinity maturation and class switching before differentiating into long-lived plasma cells or memory B cells. Beyond antibody production, B cells also regulate immune homeostasis through cytokine secretion and the activity of regulatory B cell subsets. Collectively, these processes provide the basis for long-lived and adaptable humoral immunity.

6.5.2 B Cell Development

B cell ontogeny begins in the bone marrow from haematopoietic stem cells that commit to the B lineage under the influence of transcription factors such as E2A, EBF1, and PAX5 (Nutt et al., 1999). Early pro-B cells initiate immunoglobulin heavy chain (IgH) locus recombination by joining D and J segments, followed by V-DJ rearrangement. Successful expression of a μ heavy chain with the surrogate light chain (VpreB/λ5) forms the pre-B cell receptor. Pre-B cells then proliferate and undergo Ig light-chain (VL–JL) rearrangement, producing surface IgM that marks the immature B cell stage. Surface

immunoglobulin, in combination with signalling proteins, forms the B cell receptor (BCR), coupling the specificity of antigen recognition to antibody production.

Immature B cells are subjected to central tolerance via receptor editing, clonal deletion, or anergy upon self-antigen recognition in the bone marrow. Surviving cells migrate to secondary lymphoid organs as transitional B cells, completing maturation into naïve B cells expressing both IgM and IgD. This developmental program is driven by sequential transcriptional modules including E2A and EBF1 that initiate B-lineage genes and PAX5 that fixes B cell identity and represses alternative fates (Busslinger, 2004).

Upon antigen exposure in secondary lymphoid tissue, naïve B cells interact with T cells at the T-B border. Depending on B cell receptor signal strength, co-stimulatory cues, and cytokines (IL-4, IL-21), they either differentiate into short-lived plasmablasts or enter the germinal centre (GC) pathway for affinity maturation (De Silva and Klein, 2015).

6.5.3 The Germinal Centre Reaction and B Cell Fates

GCs first arise approximately 4 days after immunisation when BCL-6–expressing B cells accumulate in follicles. Within GCs, a dark zone (DZ) is populated by rapidly proliferating B cells (centroblasts) undergoing somatic hypermutation (SHM) which randomly mutates the immunoglobulin sequence, and a light zone (LZ) where slowly dividing B cells (centrocytes) compete for antigen on follicular dendritic cells (FDCs) and for T follicular helper (TFH) signals (De Silva and Klein, 2015).

Within GCs, B cells cycle between DZ and LZ via chemokine gradients. After SHM in the DZ, cells migrate to the LZ where high-affinity mutants capture more antigen and present higher density MHC–peptide complexes to TFH cells. TFH cells then deliver CD40L, IL-4, and IL-21 signals to selected B cells and drive their re-entry into the DZ for further mutation and proliferation. This iterative cycle results in affinity maturation and the production of high-affinity memory and plasma cells (De Silva and Klein, 2015).

The plasma cell or memory cell fate decision is driven by the magnitude and duration of T cell help and intrinsic survival signals. Strong TFH-derived CD40 and cytokine signals favour plasma cell differentiation, while low T cell help promotes the memory fate. GC-independent MBCs also arise early at the T-B border. These often retain unswitched IgM isotypes and low-affinity BCRs, providing broad specificity to variant antigens. By contrast, GC-derived MBCs are class-switched and affinity-matured, ensuring rapid high-affinity responses upon re-exposure (De Silva and Klein, 2015).

Upon secondary challenge, MBCs rapidly differentiate into antibody-secreting cells or re-enter secondary GCs to refine affinity. Tissue-resident memory B cells (Brm) and persistent TFH cells in draining lymph nodes contribute to localised recall responses. The speed and strength of this recall response depend on the prior BCR affinity and the surrounding cytokine environment (Syeda et al., 2024).

6.5.4 Antibody effector functions

Antibodies, the soluble effectors of humoral immunity, and their membrane-bound counterparts on B cells (BCRs), link antigen recognition to diverse mechanisms that eliminate pathogens and maintain immune homeostasis. Each antibody molecule comprises two antigen-binding Fab regions and one constant Fc region. While Fab domains confer specificity through recognition of unique epitopes, the Fc region determines how that recognition is translated into effector activity by engaging complement components, Fc receptors, and other immune mediators.

The most direct function of antibodies is neutralization, which blocks the interaction between pathogens or toxins and host cell receptors. By coating viral or bacterial surfaces, antibodies prevent attachment, fusion, or entry, thereby halting infection at an early stage. Neutralisation often depends on multivalent interactions that compensate for the typically low monovalent affinity of antibody–antigen binding. IgM pentamers and dimeric IgA can cross-link viral particles or toxins into large immune aggregates that are readily cleared by phagocytic immune cells. In mucosal tissues, secretory IgA contributes an additional barrier by entrapping pathogens in mucus and preventing colonisation without provoking inflammation (Lu et al., 2018).

A second major effector pathway is complement activation through the classical pathway. When multiple Fc domains are clustered on a pathogen surface, the complement component C1q binds and activates the proteases C1r and C1s, leading to formation of the C4b2a C3 convertase. This cascade deposits C3b on microbial surfaces, which acts as an opsonin for complement receptors on phagocytic cells. Continuation of

the cascade leads to the assembly of the membrane attack complex which lyses pathogen cells. Pentameric IgM is particularly efficient at initiating complement, whereas IgG requires hexameric organisation of Fc regions to engage C1q effectively (Lu et al., 2018).

The third principal mechanism involves Fc receptor (FcR) engagement, which connects antibodies to the effector functions of innate immune cells. Fcγ receptors (FcγRs) on macrophages, neutrophils, and dendritic cells recognize the Fc portions of IgG bound to antigen. Crosslinking of activating FcγRs triggers immunoreceptor tyrosine-based activation motif (ITAM)-dependent signalling leading to phagocytosis (antibody-dependent cellular phagocytosis; ADCP) and cytokine secretion. On natural killer cells, FcγRIIIa (CD16a) mediates antibody-dependent cellular cytotoxicity (ADCC), releasing perforin and granzymes to kill antibody-coated targets (Lu et al., 2018).

6.5.5 B Cell Receptor Signalling

The B cell receptor (BCR) consists of a membrane-bound immunoglobulin non-covalently associated with the Igα (CD79A) and Igβ (CD79B) heterodimer, which provides intracellular signalling functions through immunoreceptor tyrosine-based activation motifs (ITAMs). Each B cell expresses a unique BCR derived from V(D)J recombination at the immunoglobulin heavy and light chain loci, producing an antigen-binding Fab region linked to a transmembrane domain anchored in the plasma membrane. Engagement of the BCR by multivalent antigen induces receptor clustering within lipid microdomains and activates Src-family kinases such as Lyn, which phosphorylate the ITAMs of Igα/Igβ. Phosphorylated ITAMs recruit the Syk tyrosine kinase, triggering a cascade involving BLNK (SLP-65) and PLCγ2. This generates diacylglycerol and inositol-1,4,5-trisphosphate, leading to calcium mobilisation and activation of PKCβ and NF-κB pathways. In parallel, Syk activates phosphoinositide 3-kinase (PI3K), producing phosphatidylinositol-3,4,5-trisphosphate that recruits and activates Akt, linking BCR signalling to cell survival and metabolic regulation (Gomez-Manriquez et al., 2025). Signal strength and duration are tightly modulated by negative regulators such as SHP-1 phosphatase and the FcγRIIB co-receptor, whose immunoreceptor tyrosine-based inhibitory motif (ITIM) attenuates BCR signalling.

6.5.6 Organization and Rearrangement of the Immunoglobulin Loci

Generation of antibody diversity begins with the somatic assembly of variable-region exons from discrete gene segments located within the immunoglobulin loci. In developing B cells, the recombination-activating gene products RAG1 and RAG2 mediate this process by introducing site-specific DNA double-strand breaks at recombination signal sequences (RSSs) that flank the variable (V), diversity (D), and joining (J) gene segments. Each RSS contains a conserved heptamer–nonamer pair separated by either a 12- or 23-base-pair spacer. The RAG complex cleaves one 12-spacer and one 23-spacer RSS in accordance with the 12/23 rule, ensuring ordered and compatible joining of gene segments. The resulting coding and signal ends are ligated by the nonhomologous end-joining machinery to create a contiguous V(D)J exon capable of encoding the immunoglobulin variable domain (Jung et al., 2006).

The human immunoglobulin heavy-chain (IGH) locus is situated on chromosome 14 and spans roughly 1.2 Mb. It contains a linear array of 40-50 functional VH genes upstream of a compact cluster of 27 DH and 6 JH segments, which together generate the variable domain of the heavy chain. Downstream of JH lie the constant-region genes (Cμ, Cδ, Cγ3, Cγ1, Cα1, Cγ2, Cγ4, Cε, Cα2) extending over 300 kb. Enhancers and regulatory elements, including the intronic Eμ and 3' regulatory region, orchestrate long-range chromatin interactions that bring distant VH promoters into proximity with recombination centres during B cell development. Allelic variation within V and D clusters, as well as structural polymorphisms in regulatory domains, contribute to inter-individual differences in antibody repertoires while maintaining the conserved architecture required for ordered recombination.

The locus organization establishes several hierarchical layers of diversity. Germline differences among VH segments determine complementarity-determining regions 1 and 2, while the junction between VH, DH, and JH segments generates complementarity-determining region 3 (CDR3), the principal determinant of antigen specificity. Productive VDJ joints yield a μ heavy chain that pairs with the surrogate light chain, forming the pre-B-cell receptor providing a checkpoint signal to progress with B cell development (Jung et al., 2006).

Rearrangement follows a strict temporal hierarchy. Each IgH allele first undergoes DH-to-JH recombination then one allele completes VH-to-DJH joining, enforcing allelic exclusion. Expression of a functional μ chain suppresses further rearrangements on the second allele and initiates light-chain recombination. The general architecture and regulatory logic of the mouse locus are conserved in humans, though the number and spacing of V, D, and J segments differ (Jung et al. 2006).

6.6 Engineering B Cells to Express Therapeutic Antibodies

Conventional monoclonal antibody (mAb) therapies provide potent and specific treatments across many diseases but require repeat infusions because serum antibodies decay within weeks (Keizer et al., 2010). B cells naturally generate long-lived plasma and memory populations capable of sustained, antigen-responsive antibody secretion. Engineering B cells to express defined antibody genes therefore offers a strategy for durable monoclonal antibody production *in vivo*.

Early work aimed to randomly integrate antibody genes using lentiviruses. However, the transduction of primary B cells using VSV-G pseudotyped lentivirus is inefficient (Serafini et al., 2004). The first demonstration of reprogramming human primary B cells to express a defined antibody used VSV-G pseudotyped lentivirus to deliver anti-HIV antibodies to hematopoietic stem cells which were then differentiated into plasma cells (Luo et al., 2009). Concurrently, alternative envelope proteins were tested to enable gene delivery in mature B cells using lentiviruses including measles (Frecha et al., 2009) and, later, baboon (Levy et al., 2016) envelopes. Fusil *et al* used a Baboon-envelope-pseudotyped lentiviral vector carrying immunoglobulin regulatory elements that reproduced the natural switch between membrane and secreted antibody forms, demonstrating that integrated constructs can harness normal B-cell transcriptional control to regulate antibody production (Fusil et al., 2015).

The targeted integration of antibody sequences into the endogenous immunoglobulin loci, using CRISPR/CRISPR-associated protein 9 (Cas9) technologies, enables the use of

contextual transcriptional regulatory sequences to more accurately mimic natural antibody production. Additionally, for the engineering of B cells to express antibodies against pathogens, it raises the possibility of enabling affinity maturation and class switching to dynamically adapt the antibody responses. The groundwork was laid in 2018 to show that electroporation of CRISPR/Cas9 ribonucleoprotein complexes into dividing primary B cells with AAV6-encoded template DNA sequences could enable the precise knock-in of defined sequences by homology-directed repair (HDR) (Hung et al., 2018, Johnson et al., 2018).

In 2019, multiple groups published work demonstrating the knock in of antibody sequences in the immunoglobulin loci of primary mouse and human B cells (Greiner et al., 2019, Hartweger et al., 2019, Moffett et al., 2019, Voss et al., 2019). From these a common approach emerged in which a coding sequence containing a full immunoglobulin light chain followed by a heavy chain VDJ region, delivered using AAV6, was inserted between the most 3' J region and the intronic Eμ enhancer. A splice donor sequence was placed at the 3' end of the inserted heavy chain VDJ sequence to enable splicing onto the endogenous downstream constant regions. The light chain and heavy chain were either separated by a P2A site or physically tethered by a flexible peptide linker. Insertion into the heavy chain locus disrupted the endogenous heavy chain expression, however the endogenous light chain was still expressed. Therefore, the use of a peptide linker between the two chains minimised the possibility that the engineered heavy chain could associate with the endogenous light chain (Moffett et al., 2019), which could generate a pathogenic antibody.

The durability of the antibody production using the methods established in 2019 was short lived in mouse models (Moffett et al., 2019). The culture conditions used mimic T cell-dependent activation of B cells and produces germinal centre-like cells and plasma cells with a lower memory and survival potential. Therefore, toll-like receptor agonists were used to activate B cells through T-cell independent pathways and this approach was shown to enable durable immunological memory (Nahmad et al., 2020).

In vivo engineering of mouse B cells using dual AAV delivery of *Staphylococcus aureus* Cas9 and the broadly neutralizing anti-HIV antibody 3BNC117 has since been demonstrated using the engineered AAV-DJ capsid. The approach achieved durable B cell editing with sustained bNAb secretion at neutralizing serum levels and minimal off-target activity (Nahmad et al., 2022). *In vivo* editing vastly simplifies the engineering process and may democratise B cell therapies, however AAV-DJ is inefficient in primary human B cells (Hung et al., 2018) and AAV6 may not be able to support *in vivo* editing due to low transduction efficiency of human B cells (Ahmadi et al., 2021) and due to neutralisation by pre-existing antibodies (Chhabra et al., 2024).

6.7 Cancer and the application of monoclonal antibodies

6.7.1 Cancer

Cancer remains a leading global health challenge and in 2020 alone, approximately 19.3 million new cases and 9.96 million deaths were recorded (Sung et al., 2021). Conventional therapies, including surgery, cytotoxic chemotherapy, and radiotherapy,

remain indispensable yet inherently limited. Surgery requires a defined resectable mass and chemotherapy is constrained by toxicity and resistance. These shortcomings have stimulated the development of targeted biologics with improved selectivity. Among them, monoclonal antibodies (mAbs) have redefined cancer treatment through their ability to recognise specific molecular targets with high affinity and to recruit immune effector mechanisms (Beck et al., 2017, Hudis, 2007b).

6.7.2 Overview of Monoclonal Antibodies

The concept of monoclonal antibodies originated with the 1975 description of the hybridoma technique, which enabled production of antibodies with predetermined specificity (Köhler and Milstein, 1975). The first therapeutic breakthrough came two decades later with the approval of rituximab, a chimeric IgG1 antibody against CD20, for relapsed non-Hodgkin lymphoma (Grillo-López, 2000). This marked the beginning of antibody-based oncology treatments, and by 2024 nearly 100 antibody drugs had gained regulatory approval worldwide for the treatment of cancer (Toledo-Stuardo et al., 2024).

Early murine antibodies elicited immune responses in humans, prompting the development of chimeric, humanised, and later fully human antibodies through recombinant and transgenic technologies. These advances minimised immunogenicity and improved pharmacokinetics. Unlike small molecules that often act promiscuously, antibodies typically bind discrete epitopes with high specificity, allowing precise engagement of tumour-associated targets and improved tolerability.

6.7.3 Mechanisms of Action in Cancer Treatment

Monoclonal antibodies eliminate cancer cells through complementary biological and immune mechanisms. One principal effect is the inhibition of oncogenic signalling. Antibodies such as trastuzumab, cetuximab, and bevacizumab block receptor-ligand interactions essential for cell proliferation and angiogenesis, sensitising tumours to conventional therapy (Hudis, 2007b).

A second mechanism involves immune-effector recruitment. The Fc region of therapeutic IgG binds Fc receptors on natural killer cells and macrophages, mediating antibody-dependent cellular cytotoxicity (ADCC) and phagocytosis (ADCP). Complement activation triggers complement-dependent cytotoxicity (CDC), exemplified by ofatumumab and alemtuzumab (Pawluczkowycz et al., 2009, Zent et al., 2008).

A third strategy, immune checkpoint blockade, acts indirectly by reactivating the anti-tumour immune response. Tumours suppress immunity through inhibitory receptors such as PD-1 and CTLA-4. Antibodies that block these checkpoints, such as ipilimumab, nivolumab, and pembrolizumab, restore cytotoxic T-cell activity and have achieved durable responses across multiple cancers (Sharma et al., 2023).

Finally, antibody–drug conjugates (ADCs) combine an antibody with a potent cytotoxin, delivering the drug selectively to tumour cells while limiting systemic exposure (Beck et al., 2017). Trastuzumab emtansine (T-DM1) exemplifies this design and has demonstrated a significant survival benefit in HER2-positive breast cancer (Verma et al., 2012).

6.7.4 Clinical Impact

The addition of monoclonal antibodies to standard regimens has transformed outcomes in several cancers. Rituximab, combined with CHOP chemotherapy, significantly improved response rates and overall survival in diffuse large B-cell lymphoma, establishing a new standard of care (Coiffier et al., 2002). In HER2-positive breast cancer, trastuzumab plus chemotherapy extended median overall survival from about 20 to 25 months in metastatic disease and reduced recurrence rates in early-stage disease (Slamon et al., 2001a). Bevacizumab added to standard chemotherapy improved outcomes in metastatic colorectal cancer (Hurwitz et al., 2004).

Nevertheless, resistance and access barriers remain. Tumours may escape through antigen loss or alternative signalling pathways, and high production costs limit global availability.

6.8 Thesis Overview

Adeno-associated virus (AAV) has emerged as the gold standard for safe and efficient gene delivery in humans. Nevertheless, its clinical utility is constrained by limited tissue tropism, pre-existing immunity, and the poor transduction efficiency of several primary human cell types. B lymphocytes, as key effectors of humoral immunity and long-lived producers of antibodies, represent an attractive platform for therapeutic gene delivery. Recent advances combining CRISPR/Cas9 genome editing with AAV-mediated donor delivery have enabled the reprogramming of B cells to express defined monoclonal antibodies, offering a promising new avenue for durable in vivo immunotherapy. However, existing AAV serotypes transduce primary human B cells inefficiently, restricting current approaches to inefficient and costly *ex vivo* manipulation. Moreover, substantial molecular and vector design challenges remain in adapting these strategies to generate B cells capable of producing anti-cancer monoclonal antibodies. This thesis was therefore driven by the dual objective of improving AAV vector performance in human B cells and exploring their engineering for the expression of therapeutic anti-cancer antibodies.

The overarching objective of this thesis was to develop a platform for the engineering of anti-cancer B cell therapies, encompassing both vector optimisation and antibody cassette design. To this end, we aimed to develop novel AAV library production systems that preserve the genetic linkage between capsids variants and their encoding genomes, to apply these systems for the discovery of B cell-optimised capsid variants, and to

demonstrate the targeted knock-in of full-length anti-cancer antibodies into the immunoglobulin heavy-chain (IgH) locus of primary human B cells.

The first results chapter presents the development of an inducible two-step AAV capsid library platform designed to prevent cross-packaging and mosaicism that may introduce noise into libraries produced using existing methods. Inducible gene-control systems were evaluated to suppress genome-encoded capsid expression during the production of AAV2 shuttle vectors and to activate it only during repackaging into the genome-encoded capsid. Two alternative mechanisms were explored including a protein transactivator system and a Cre recombinase system. Although both provided stringent repression under plasmid conditions, neither achieved the required level of control during AAV replication, revealing key complexities in inducible regulation systems within replicating AAV genomes.

To enable screening, the next chapter develops capsid libraries optimised for screening in primary human B cells using the established one-step library approach. Quantitative assays showed that B-cell resistance to AAV6 does not appear to be due to impaired binding or entry, guiding the design of libraries focused on post-entry determinants of transduction. Two complementary library types were created: a cross-serotype DNA-shuffled library and a targeted mutagenesis library of the VP1 N-terminus region, implicated in endosomal escape and uncoating. Systematic optimisation of helper plasmid design, genome input, and AAP complementation defined conditions was performed with the aim to maximise library complexity while minimising cross-packaging noise. Finally, the libraries were screened using *in vitro* primary human B cell cultures.

The final chapter describes the development of a strategy for inserting full-length monoclonal antibody genes into the IgH locus of primary human B cells. Homology-directed repair templates were designed to reproduce the natural competition between secretory and membrane-bound immunoglobulin isoforms. To test these constructs, anti-cancer antibodies were expressed in membrane-bound and secreted forms in HEK293 cells, and secreted trastuzumab retained antigen specificity and the ability to mediate antibody-depended cellular phagocytosis. CRISPR/Cas9 ribonucleoprotein complexes targeting the IgHJ6 intron were combined with AAV6 vectors to deliver donor templates, achieving on-target integration verified by junction PCR. A B cell–specific VH1-69 promoter supported expression only following integration. Overall, the work demonstrated the simultaneous expression of membrane-bound and secreted anti-cancer antibodies from homology-directed repair templated designed for the knock-in of sequences at the human IgH locus, though further work is required to validate functionality in primary cells.

6.9 Overall Aims

1. Develop a novel two-step AAV capsid library production platform using AAV2 shuttle vectors carrying inducible capsid variant sequences.
2. Design and produce AAV capsid libraries tailored towards overcoming bottlenecks in the transduction of primary human B cells.
3. Develop a strategy for the engineering of primary human B cells expressing functional full-length anti-cancer monoclonal antibodies.

7 Materials and Methods

7.1 Molecular Biology

7.1.1 Bacterial Transformation

Transformation by heat shock

For AAV plasmids containing ITRs, New England Biolabs Stable (NEB, USA) Competent *E. coli* was used to minimise recombination during growth of transformed bacteria, for all other plasmids NEB 10β cells were used. For both strains of bacteria, 2 µL of plasmid DNA was incubated with 48 µL bacteria on ice for 20 minutes. The cells were then heat-shocked at 42 °C for 30 seconds and rested on ice for 2 minutes, before addition of 450 µL SOC medium. For plasmids encoding a bacterial kanamycin resistance marker, transformed E. coli were then incubated at 37 °C for 1 hour with shaking at 210 RPM. 100 µL of bacteria was streaked onto Luria Bertani (LB; Sigma-Aldrich, USA)-agar plates containing 50 µg/mL of kanamycin sulfate (Merck, Germany) or 50 µg/mL ampicillin sodium salt (Sigma-Aldrich, USA). Individual colonies were picked into LB broth containing 50 µg/mL of kanamycin sulfate or 50 µg/mL ampicillin sodium salt and incubated overnight at 37 °C with shaking at 210 RPM.

Transformation by electroporation

Electroporation was used for the transformation of high complexity plasmid libraries to maximise the number of transformants. 2 µL of purified ligation reaction was added to 18 µL MegaX DH10B T1 Electrocomp Cells (Invitrogen, USA) on ice and mixed gently by pipetting 2-3 times. The cells were then pipetted into a chilled 0.1 cm cuvette and

electroporated using a MicroPulser Electroporator (BioRad, USA) set to 2.0 kV. The electroporated cells were mixed with 1mL Recovery Medium (Invitrogen, USA) before shaking at 210 RPM for 2 hours at 30 °C. 1 µL of sample was diluted 1/10,000 and spread onto LB-agar plates containing 50 µg/mL of kanamycin sulfate and incubated overnight at 37 °C. The reminder of the outgrowth sample was used to inoculate 500 mL of LB containing 50 µg/mL of Kanamycin Sulfate, which was incubated overnight at 30 °C with shaking at 210 RPM.

7.1.2 DNA Purification

1.5 mL (mini scale), 35 mL (midi scale) or 500 mL (mega scale) overnight cultures of transformed *E. coli* were processed using NEB Monarch® Plasmid Miniprep Kit (NEB, USA), QIAGEN Plasmid Plus Midi Kit (QIAGEN, UK) or QIAGEN Plasmid Plus Mega Kit (QIAGEN, UK) according to the manufacturer's protocols.

7.1.3 Polymerase Chain Reaction (PCR)

Oligonucleotide primers were synthesised by Integrated DNA Technologies (IDT, USA). DNA fragments were ordered as gblocks from IDT.

7.1.3.1 Colony PCR

To verify correct assembly during plasmid production, colony PCR was performed using bacteria picked directly from individual bacterial colonies grown on agar-plates into the following PCR reaction mix:

Table 1. Colony PCR reaction mix components.

Component	Volume (µL)
OneTaq Quick-Load 2X Master Mix (NEB, USA)	12.5
Forward Primer (10 µM)	0.5
Reverse Primer (10 µM)	0.5
Nuclease-Free Water (Ambion, USA)	11.5

The PCR was performed using the following thermocycler parameters:

Table 2. Colony PCR thermocycler program.

Step	Temperature (°C)	Time (s)	Cycles
Lysis	94	300	1
Denaturation	94	15	30
Annealing	*	15	
Extension	68	60 per kb	
Final Extension	68	300	1
Hold	10	Infinite	1

* The annealing temperature was calculated based on primer sequence using the NEB melting temperature calculator (tmcalculator.neb.com)

7.1.3.2 High-fidelity PCR

For cloning or for the production of amplicons for sequencing, high fidelity PCR was performed using 10 ng of template DNA in the following reaction mix:

Table 3. High-fidelity PCR reaction mix components.

Component	**Volume (µL)**
Q5 High-Fidelity 2X Master Mix (NEB, USA)	25
Forward Primer (10 µM)	2.5
Reverse Primer (10 µM)	2.5
Nuclease-Free Water (Ambion, USA)	To 25
Template DNA	Variable

The PCR was performed using the following thermocycler parameters:

Table 4. High-fidelity PCR thermocycler program.

Step	**Temperature (°C)**	**Time (s)**	**Cycles**
Initial Denaturation	98	30	1
Denaturation	98	10	30-35
Annealing	*	15	
Extension	68	20 per kb	
Final Extension	68	120	1
Hold	10	Infinite	1

* The annealing temperature was calculated based on primer sequence using the New England Biolabs melting temperature calculator (tmcalculator.neb.com)

7.1.4 Restriction Enzyme Digestion

Restriction digestion was used to validate the correct assembly of plasmid DNA isolated from bacterial clones following plasmid assembly, to produce compatible DNA fragments for ligation cloning, to linearise plasmid DNA for transfection, or to deplete

bacterial template DNA from PCR reactions. Restriction enzymes were sourced from New England Biolabs (NEB) and restriction digests were performed in the supplied buffer. For routine digestions 10 U of restriction enzyme was used per µg of DNA and the digestion reaction was performed for 1 hour at the temperature specified by the manufacturer for the enzyme. For the depletion of bacterial template DNA from PCR reactions, 1 µL of DpnI (NEB, USA) was added directly to the completed PCR reaction and incubated at 37 °C for 1 hour.

7.1.5 Agarose Gel Electrophoresis

To separate DNA based on fragment size, DNA samples were mixed with 1/5 volume of 6X Gel Loading Dye (NEB, USA) and loaded onto a 1% (w/v) agarose gel pre-stained with a 1/10,000 dilution of GelRed Nucleic Acid Gel Stain (Biotium, USA). 1 kb Plus DNA Ladder (NEB, USA) was used as DNA reference ladders to determine DNA fragment sizes. Gel electrophoresis was performed at 4 V/cm for 1 hour in 1X Tris-Acetate-Ethylenediaminetetraacetic acid (EDTA) (TAE) buffer (40 mM Tris, 50 mM EDTA, 20 mM acetic acid). The separated DNA was visualised using a ChemiDoc Imaging System (Bio-Rad, USA).

7.1.6 PCR amplicon purification

To purify DNA directly from PCR reactions, the Monarch® Spin PCR & DNA Cleanup Kit (NEB, USA) was used. Purification of DNA fragments from agarose gel following gel electrophoresis was performed using the Monarch Spin DNA Gel Extraction Kit (NEB, USA). Both kits were used according to the manufacturer protocols.

7.1.7 Nucleic Acid Quantification

Purified nucleic acids were quantified using a Nanodrop spectrophotometer (Thermo Fisher Scientific, USA) using nuclease-free water, or kit-specific elution buffers to blank the device.

7.1.8 Illumina amplicon Sequencing

Amplicon sequencing was performed by Genewiz (UK) using the Amplicon-EZ service.

7.1.9 Sanger Sequencing

Sanger sequencing was carried out by Genewiz (UK) using custom primers. Sanger traces were aligned to reference DNA using SnapGene (USA) to confirm sequence identity.

7.1.10 Whole plasmid nanopore sequencing

Whole plasmid nanopore sequencing was carried out by Plasmidsaurus (USA). Consensus DNA sequences were aligned to reference DNA using SnapGene (USA) to confirm sequence identity.

7.1.11 HiFi Assembly

The routine assembly of plasmid DNA was performed using NEBuilder HiFi Assembly Master Mix (NEB, USA). For assemblies involving ITR containing plasmids, the ITR-containing backbone fragment was generated by restriction digest. All other DNA

fragments were amplified by PCR from plasmid DNA or synthesised DNA fragments. HiFi assemblies were designed so that each fragment was flanked by 20 bp overlap regions identical to the adjacent fragment. For simple 2-3 fragment assemblies, a vector:insert ratio of 1:2 was used with a total of 3-200 fmol DNA fragments. For 4-6 fragment assemblies, a 1:1 vector-to-insert ratio was used with a total of 200 fmol DNA fragments. DNA fragments were assembled in a 20 µL reaction with 10 µL NEBuilder HiFi Assembly Master Mix.

7.1.12 Restriction-ligation cloning

For the assembly of plasmid AAV capsid libraries, restriction-ligation cloning was used.

Table 5. DNA ligation reaction mix

Component	**Volume (µL)**
10X T4 DNA Ligase Buffer (NEB, USA)	20
Digested library vector DNA fragment (576 ng)	Variable
Digested library capsid DNA fragment pool (788.2 ng)	Variable
T4 DNA Ligase (NEB, USA)	10
Nuclease-Free Water (Ambion, USA)	To 200

The ligation reaction mixture was incubated at 16 °C overnight and then purified using the PCR amplicon purification method. The ligated product was eluted in 10 µL Nuclease-Free Water (Ambion, USA) for electroporation into *E. coli*.

7.1.13 Site directed mutagenesis of plasmid DNA

For small 1-3 nucleotide modifications of plasmid DNA, 5' phosphorylated primers (IDT, USA), incorporating the modification, were designed to amplify the entire plasmid. Following PCR amplification and depletion of residual plasmid DNA by DpnI digestion, 30 ng of the linearised plasmid was re-circularised in a 20 µL ligation reaction with 2 µL T4 DNA Ligase (NEB, USA) and 1 µL T4 DNA Ligase Buffer (NEB, USA). The ligation reaction was incubated at 21 °C for 1 hour. Larger modifications were performed using HiFi Assembly.

7.1.14 Assembly of Plasmid DNA Library

Mutagenesis

Mutagenesis of the capsid N-terminal region between the VP1 and VP3 start codons was performed using the GeneMorph II Random Mutagenesis Kit (Agilent, USA) following the manufacturer's protocol. Error-prone PCR was performed using the RepCap6 helper plasmid as template DNA and primers listed in Table 6. To generate amplicons with a range of mutation rates, the amount of template DNA was varied from 9901 ng to 1.41 ng plasmid DNA. To further increase mutation rates, nested error prone PCRs were performed.

Table 6. Error-prone PCR primers

Primer	**Sequence**
VP12u Morph Fw	CTTTGAACAATAAATGATTTAAATCAGGTATG
VP12u Morph Rv	CCGCCTGAAGCCATTG

Capsid Shuffling

Capsid sequences from AAV1-9 were amplified by PCR and pooled in an equimolar ratio. 4 µg of pooled DNA was fragmented using 0.04U DNase I (NEB, USA) in a 45 µL reaction incubated for 10-15 minutes at 25 °C. The reaction mixture was immediately mixed with 5 µL 50 mM EDTA and incubated at 75 °C for 10 minutes to halt fragmentation. The DNA was separated by gel electrophoresis and fragments between 100 bp and 500 bp were purified by gel purification. The purified fragments were used as input for a primerless PCR reaction:

Table 7 – Primerless PCR reaction mix

Component	**Volume (µL)**
5X Phusion HF Buffer (NEB, USA)	10
10 mM dNTPs (NEB, USA)	1
Phusion Polymerase (NEB, USA)	0.5
DMSO (Sigma-Aldrich, USA)	1.5
Fragment pool (1 µg)	Variable
Nuclease-Free Water (Ambion, USA)	To 50

Table 8. Primerless PCR thermocyler program.

Step	**Temperature (°C)**	**Time (s)**	**Cycles**
Initial Denaturation	98	30	1
Denaturation	98	10	40
Annealing	42	30	
Extension	72	60	
Final Extension	72	600	1

Full length capsid sequences were amplified from the primerless PCR reaction using high-fidelity PCR.

Table 9. Full length capsid rescue PCR primers.

Primer	Sequence
Primerless Rescue Fw	CCACTCGACACACCC
Primerless Rescue Rv	TGCGGCTACCGTCTC

Analysis of capsid shuffling was carried out using ShuffleAnalyzer (Schweiggert et al., 2024). Plasmid DNA was isolated from randomly selected colonies of *E. coli* transformed with the shuffled library plasmid and sequenced by sanger sequencing.

7.2 Cell Culture

7.2.1 Cell Line Maintenance

293-AAV (AAV-100, Cell BioLabs, USA), HEK293 (ATCC), MDA-MB-231 and BT474 (kindly supplied by Prof. Susan Brooks, Oxford Brookes University) cells were cultured in Dulbecco's Modified Eagle Medium (Sigma-Aldrich, USA) containing 10% (v/v) heat-inactivated foetal bovine serum (FBS; Merck, USA). RAMOS cells (Cell Line Services, Germany) were cultured in RPMI-1640 (Sigma-Aldrich, USA) supplemented with 10% FBS. All cells were maintained at 37 °C, 5% CO2 at 95% humidity. Adherent cells were passaged at 80-90% confluency and RAMOS cells were maintained at a cell density of 0.2-1 x 10^6 cells/mL. To passage adherent cells, the cells were gently washed with room-temperature phosphate buffered saline (PBS; Thermo Fisher Scientific, USA) and incubated with trypsin-EDTA (Sigma-Aldrich, USA) for 5 minutes at 37 °C. The detached cells were diluted in a 10-fold excess of warm culture medium, and a 1/10 volume of cells was added to a tissue culture-treated flask with fresh culture medium. Cell lines were

tested for mycoplasma contamination using a MycoAlert Mycoplasma Detection Kit (Lonza, Switzerland).

7.2.2 Peripheral Blood Mononuclear Cell (PBMC) Isolation

Non-clinical issue leukocyte cones (leukocyte reduction system cones) were purchased from the National Health Service Blood and Transplant service at the John Radcliffe Hospital.

The cone was wiped with 70% industrial methylated spirit and placed over an open 50 mL centrifuge tube. The top and bottom tubing was cut to allow the sample to flow into the 50 mL tube. The sample was topped up to 30 mL with room temperature PBS containing 2% FBS. PBMCs were isolated using STEMCELL SepMate PBMC Isolation Tubes (STEMCELL, UK) according to the manufacturer's protocol. The PBMC sample was centrifuged at 800 x g for 10 minutes to pellet. The sample was resuspended in 10 mL Puregene RBC Lysis Solution (QIAGEN, UK) and incubated for 10 minutes at 4 °C. The sample was topped up to 50 mL with PBS 2% FBS, centrifuged at 800 x g for 10 minutes and then washed again with 50 mL PBS 2% FBS.

7.2.3 Primary B Cell Isolation and Culture

Primary human B cells were isolated from PBMCs by negative selection using the EasySep Human B Cells Isolation Kit (STEMCELL, UK) following the manufacturer's instructions.

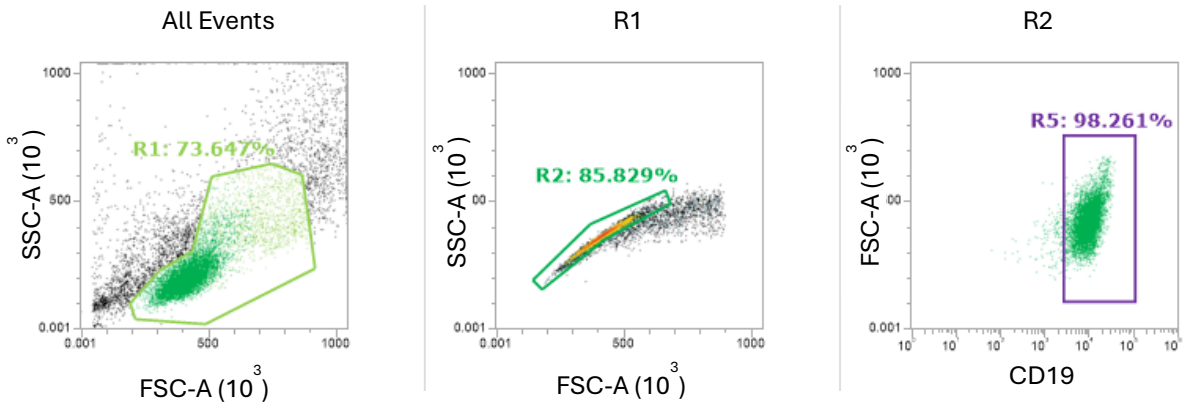


Figure 2 – Representative B cell purity. Frequency of CD19+ cells measured by flow cytometry

To stimulate, purified B cells were cultured at 1-2 x 10^6 cells/mL for 72 hours in B cell medium (RPMI-1640, supplemented with 10% heat-inactivated FBS, 10 mM HEPES, penicillin/streptomycin, and 53 µM 2-mercaptoethanol) containing 50 ng/mL IL-2 (Peprotech, UK), 50 ng/mL IL-10 (Peprotech, UK), 10 ng/mL IL-15 (Peprotech, UK), 10 µg/mL ODN 2006 (Miltenyi Biotec, Germany) and 100 ng/mL MEGACD40L (Enzo Life Sciences, USA).

7.2.4 Primary Monocyte Isolation and Culture

Following the isolation of PBMCs from LRS blood products (section 7.2.2), PBMCs were resuspended in 20 mL RPMI-1640 containing 10% (v/v) FBS. 5 mL PBMC suspension was overlayed onto 6 mL room temperature 46% isotonic Percoll PLUS (Cytivia, UK) in RPMI 10% FBS (285 mOsm). The Percoll PLUS density gradient was centrifuged using an Allegra X-15R Centrifuge (Beckman Coulter, USA) at 950 x g (acceleration = 5, deceleration = 0) for 30 minutes. The monocyte interphase formed during self-assembly of the Percoll

gradient was added to a new tube. The monocytes were then washed three times with PBS and then resuspended in RPMI 10% FBS.

To culture human PBMC-derived macrophages, the cells were resuspended in X-VIVO®-10 serum-free medium (Lonza, Switzerland) supplemented with 2% human serum. 80 x 10^6 cells were cultured in a T-175 flask for 72 hours at 37 °C, 5% CO2 at 95% humidity.

7.2.5 Cryopreservation and recovery of cells

All cell lines were cryopreserved in the culturing medium containing 10% dimethyl sulfoxide (DMSO; Sigma-Aldrich, USA). Purified primary human B cells were cryopreserved in CryoStor® CS10 (STEMCELL, UK). The freezing rate was controlled using CellCool containers (Corning, USA) and frozen cells were kept at -80 °C.

To thaw cryopreserved cells, the cells were rapidly thawed using constant agitation in a 37 °C water bath. The thawed cells were added to 20 mL of pre-warmed culture medium. The culture medium was exchanged to remove residual DMSO by centrifugation at 400 x g for 5 minutes for cell lines and 400 x g for 10 minutes for primary human B cells. Primary B cells were rested for 1 hour at 37 °C in B cell medium before stimulation with cytokines.

7.2.6 DNA Transfection

HEK293 and 293-AAV cells were seeded at 6.3 x 10^4 cells/cm2 in tissue culture-treated plates (Corning, USA) and cultured in DMEM 10% FBS for 16-20 hours, or until 80-90% confluent. For each transfection, a DNA mix (Tube A) was prepared using 312.5 ng/cm2

plasmid DNA diluted in Opti-MEM (Thermo Fisher Scientific, USA). A second mix (Tube B) was prepared containing 4 µg polyethyleneimine hydrochloride (PEI) MAX (Kyfora Bio, USA) per 1 µg of plasmid DNA in Tube A, diluted in Opti-MEM. Tube A and Tube B were mixed at a 1:1 ratio, by transferring the contents of Tube B to Tube A, and incubated at room temperature for 10 minutes. The DNA/PEI mixture was then added dropwise to the cells and mixed with the cell culture medium by gentle rocking. The cells were cultured in DMEM 10% FBS for 48-72 hours.

For shRNA knock-down experiments and for the transfection of Cre expression plasmids, cells were transfected at the time of seeding.

Table 10. shRNA sequences.

shRNA	**Sequence**
Universal	ATTAGTTCATAGCCCATATAT
AAV2 capsid	CATGGGTGGATTCGGACTTAA
AAV5 capsid	CCAGAGATCCAGTACACAAAC
AAV8 capsid	TGTGGACTTTGCTGTTAATAC

7.2.7 DNA Purification from mammalian cells

Plasmid DNA was purified from cultured cells using the DNeasy Blood and Tissue kit (QIAGEN, UK). Genomic cellular DNA and AAV DNA was isolated using the Monarch Spin gDNA Extraction Kit (NEB, USA). DNA purification was performed using the manufacturer's protocols. For the purification of AAV DNA, the proteinase K step of the Monarch Spin gDNA Extraction Kit was extended to 45 minutes for AAV capsid lysis.

7.2.8 Production and purification of soluble Trastuzumab

To produce purified Trastuzumab for antibody-dependent cell-mediated phagocytosis (ADCP) assays, a T-175 flask of 293-AAV cells was transfected with 50 µg of SFFV-Trastuzumab-minimal intron plasmid and 72 hours post-transfection the cell culture medium was removed and clarified by centrifugation at 1000 x g for 5 minutes at room temperature. The Trastuzumab was purified using a Strep-Tactin®XT 4Flow® column (IBA Lifesciences, Germany) following the manufacturer's protocol. All steps were performed at 4 °C in a walk-in cold room. Briefly, the column was equilibrated with 10 mL wash buffer (100 mM Tris/HCl pH 8.0, 150 mM NaCl, 1 mM EDTA). The cell culture medium was applied to the column and allowed to flow through. The column was washed 5 times with 5 mL wash buffer. To elute the bound Trastuzumab, 3 mL 1X BXT buffer (IBA Lifesciences, Germany) was applied to the column, and the eluate was collected. The elution step was repeated with 8 mL, and then 4 mL, of 1X BXT buffer. Following an assessment of cellular protein contamination using SYPRO Ruby Total Protein staining (Figure 3), the second and third eluates (fractions 9 and 10) containing purified Trastuzumab were combined and stored at -20 °C.

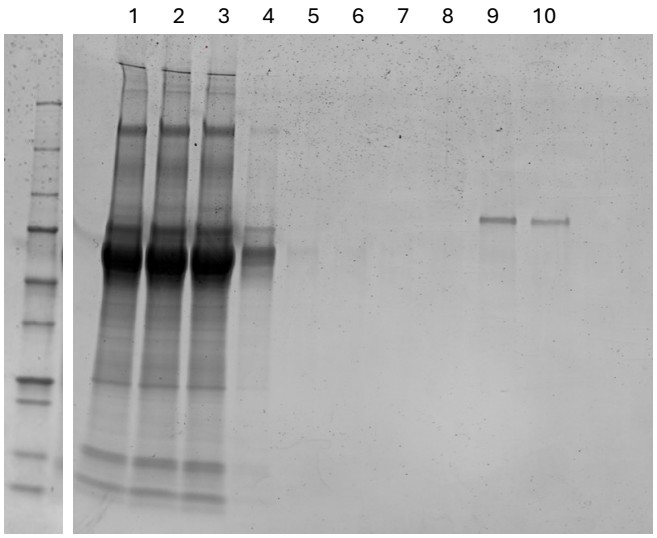


Figure 3. Total protein (SYPRO Ruby) staining to assess Trastuzumab protein contamination levels following column purification.

7.3 Gene Editing

7.3.1 CRISPR RNA (crRNA) Design

The human IgH locus RefSeq sequence (NG_001019.6) was downloaded from the National Center for Biotechnology Information (NCBI) National Library of Medicine (NLM) website (www.ncbi.nlm.nih.gov). CRISPR RNAs were designed to target the IgHJ6-Eμ region sequence (positions 1017150-1017974) using the UC Santa Cruz CRISPOR (Concordet and Haeussler, 2018) design tool (crispor.gi.ucsc.edu). The GRCh38/hg38 + SNPs: dbSNP148, Kavier database was used for off-target scoring, and the Streptococcus pyogenes Cas9 (Sp. Cas9) 20bp-NGG protospacer adjacent motif (PAM) spacer configuration was used for crRNA design. The generated sequences were filtered using an MIT specificity score and ’16 Doench score cutoff of 70. CRISPR RNAs were selected so that 300 bp homology arms contained intronic sequences only and did not

include the Eμ enhancer element to prevent the risk of off-target gene activation in the event of off-target insertion of a homology-directed repair template.

7.3.2 Editing of Primary Human B Cells

Primary human B cells were stimulated (section 7.2.3) for 72 hours prior to editing. To prepare the guide RNA (gRNA) complex, 5 µL 100 µM Alt-R CRISPR-Cas9 crRNA (Table 11; synthesised by IDT, USA) was mixed with 5 µL 100 µM Alt-R CRISPR-Cas9 tracrRNA (IDT, USA). The tracrRNA/crRNA mixture was incubated at 95 °C for 5 minutes and allowed to cool to room temperature to form the annealed gRNA complex.

Table 11. CRISPR RNA (crRNA) sequences

crRNA	**Sequence**
466	CCCTGCTCTCATCAAGACCG
483	TCAGCCATCACTAAGACCCC

To prepare ribonucleoprotein (RNP) complexes, 2.4 µL of annealed gRNA was mixed with 1.6 µL Alt-R SpCas9 nuclease (IDT, USA). The mixture was vortexed to mix and incubated for 10 minutes at room temperature. 2.27 µL Alt-R Cas9 Electroporation Enhancer (IDT, USA) was added to the RNP mixture.

1×10^6 stimulated primary human B cells were nucleofected using the P3 Primary Cell 4D-Nucleofector X Kit S (Lonza, Switzerland) following the manufacturer's protocol using the EO-117 program. Following nucleofection, the cells were plated in 100 µL B cell stimulation medium (see section 7.2.3) and cultured for 72 hours. For knock-in

experiments using AAV, the indicated amount of AAV was mixed with the cell culture medium immediately after plating.

7.3.3 Quantification of Double-Strand Break Efficiency

To measure the insertion-deletion (InDel) frequency at the IgHJ6 intron editing site, genomic DNA was isolated from B cells edited without AAV homology-directed repair templates, and from mock nucleofected B cells. The IGHJ6 intronic region containing the cut site was amplified by high-fidelity PCR. The amplicon was column purified and sequenced by sanger sequencing. To estimate the InDel frequency, the sanger sequencing traces were analysed using the Inference of CRISPR Edits tool (Conant et al., 2022) (ice.synthego.com).

7.4 AAV Production and Characterisation

7.4.1 Small-scale adherent AAV productions

Small-scale AAV production was performed using either the helper-free triple-transfection method or using TESSA. To produce AAV using the triple-transfection method, HEK293 cells were co-transfected with AAV genome plasmid, RepCap plasmid and adenoviral helper plasmid at an equimolar ratio. To produce AAV using TESSA, the cells were transfected with AAV genome plasmid and transduced with 25 tissue culture infectious dose 50% (TCID50) per cell of TESSA. 72 hours post-transfection, the cells and culture medium were harvested by pipetting. Cells were lysed using three freeze/thaw

rounds and the sample was clarified by centrifugation at 3200 x g for 20 minutes to pellet cell debris.

7.4.2 Production of nanobody-display AAVs

The production of nanobody-displaying AAVs was identical to the small-scale AAV production procedure except the RepCap helper plasmid is replaced with a VP1-nanobody expression plasmid and a VP2/3 expression plasmid. The four plasmids were co-transfected at an equimolar ratio.

7.4.3 Large scale AAV production

AAV used for gene editing and capsid library screening was produced using the helper-free triple-transfection method at large scale using 10 x 15 cm diameter dishes per production. The AAV genome plasmid, RepCap helper plasmid (Rep-MAAP-AAP for library production) and adenoviral helper plasmid were co-transfected at an equimolar ratio using a total amount of 35 µg DNA per dish. The AAV was processed using an adapted version of a previously published protocol (Kimura et al., 2019). 72 hours post-transfection, the cells were harvested using a cell scraper and pelleted by centrifugation at 750 x g for 15 minutes at 4 °C. The cell culture medium was set aside, and cells were washed with PBS. The cells were incubated with 2.5 mL 110 mM citrate buffer pH 4.2 (55 mM citric acid, 55 mM sodium citrate, 800 mM NaCl) for 10 minutes to lyse the cells. The lysate was centrifuged at 5251 x g for 30 minutes at 4 °C and the supernatant was collected and neutralised using a 1/5 volume of 2M HEPES pH 8.0. The cell culture medium and cell lysate was combined and a ¼ volume of 5X polyethylene glycol

(PEG)/NaCl solution (40% w/v PEG 8000, 2.5 M NaCl) was added. The AAV/PEG solution was incubated at 4 °C overnight and then centrifuged at 2500 x g for 1 hour at 4 °C to pellet the precipitated AAV. The supernatant was discarded and the AAV pellet was resuspended in TNE buffer (100mM Tris-Cl pH 8.0, 150 mM NaCl, 20 mM EDTA).

To reduce lipid and cellular protein contamination, the AAV sample was mixed with an equal volume of chloroform (Sigma-Aldrich, USA) and vortexed for 2 minutes. The sample was centrifuged at 1000 x g for 5 minutes and the top AAV-containing aqueous phase was retained. Residual chloroform was allowed to evaporate for 30 minutes and then the sample was subjected to buffer exchange using Amicon filters (see section 7.4.5).

7.4.4 Iodixanol Purification of AAV

For gene editing and AAV capsid library screening experiments, AAV produced using the large-scale production protocol was purified over an iodixanol gradient. The following iodixanol solutions were underlaid in ascending concentration in a 32.5 mL OptiSeal ultracentrifuge tube (Beckman Coulter, USA):

Table 12. Iodixanol Solutions. *PBS-MK (1mM MgCl2, 2.5 mM KCl, in PBS).*

OptiPrep Concentration (%)	**Volume (mL)**	**Phenol Red (%)**	**Diluent**
15	8	-	1M NaCl/PBS-MK
25	5	0.001	PBS-MK
40	5	-	PBS-MK
60	3	0.002	-

5 mL of chloroform-purified AAV was overlaid onto the iodixanol gradient and topped up with PBS. The gradient was centrifuged at 350000 x g for 90 minutes at 10 °C in a T70i fixed

angle rotor (Beckman Coulter, USA) using maximum acceleration and deceleration setting 9. To recover the purified AAV, the 40% fraction was isolated using an 18G needle.

7.4.5 Concentration of AAV

AAV samples were concentrated using 100 KDa cutoff Amicon filters (Millipore, USA). The filters were washed with TNE containing 0.01% Pluronic F68 (Gibco, USA) and centrifuged at 3500 x g for 10 minutes at room temperature. The AAV samples were added and centrifuged at 3500 x g at room temperature until a volume of <50 µL was reached. The sample was topped up with PBS 0.01% Pluronic F68, mixed and centrifuged again. This was repeated three times to remove residual iodixanol or chloroform. The concentrated virus was then aliquoted and stored at -80 °C.

7.4.6 Measurement of AAV Physical Titre by qPCR

For the titration of crude AAV produced during the development of the two-step AAV capsid library approach, a GFP-specific TaqMan qPCR assay was used. Due to the wide range of AAV genome designs, common ITR primers were used for the remaining work.

7.4.6.1 Titration of crude AAV using GFP-specific primers

2 µL of crude AAV was incubated for 2 hours at 37 °C with 2 µL TURBO DNase (Thermo Fisher Scientific, USA) in a total volume of 20 µL, followed by a 15 minute incubation at 75 °C. 0.5 µL proteinase K (QIAGEN, UK) was added and the reaction was incubated at 55 °C for 1 hour, followed by 95 °C for 20 minutes.

The treated sample was used in the following reaction mix:

Table 13 – TaqMan AAV DNA qPCR reaction mix.

Component	**Volume (µL)**
TaqMan Fast Advanced Master Mix (Applied Biosciences, USA)	10
20X eGFP-specific primer/probe mix	1
AAV Sample (1/100)	5
Nuclease-Free Water (Ambion, USA)	4

Table 14. TaqMan AAV DNA qPCR eGFP primer/probe sequences

Primer	**Forward sequence**	**Reverse sequence**	**Probe sequence**	**Dye/Quencher**
eGFP	GAACCGCATCGAG CTGAA	TGCTTGTCGGCCAT GATATAG	ATCGACTTCAAGGA GGACGGCAAC	FAM/ZEN/3IABkFQ

An 8-point serial dilution (10^8-10^1 copies/reaction) of a DNA fragment containing the eGFP amplicon sequences (see Appendix section 13.1) was used as a standard curve. The absolute DNA copies was measured using a StepOne Plus qPCR system (Applied Biosciences, USA) using the following run parameters:

Table 15. TaqMan AAV DNA qPCR run parameters.

Step	**Temperature (°C)**	**Time (s)**	**Cycles**
Initial Denaturation	95	600	1
Denaturation	95	1	40
Annealing/Extension	60	20	

The genome copies per mL value was calculated as:

$$Titre\ \left(\frac{gc}{mL}\right) = copies\ per\ reaction \times DF \times (2 \times 10^5)$$

where DF is the sample preparation dilution factor.

7.4.6.2 Titration of purified AAV using ITR-specific primers

For the quantification of purified AAV samples, a published protocol was used (Kimura et al., 2019) to quantify ITR copies. A detailed protocol is provided in the paper's, supplementary information. However, the DNase digestion step was replaced with a 1 hour incubation at 37 °C with 2 µL TURBO DNase (Thermo Fisher Scientific, USA) in a total volume of 20 µL.

7.4.7 Assessment of AAV cross-packaging

In the context of AAV capsid libraries, cross-packaging is the erroneous packaging of AAV genomes encoding capsid sequences into the incorrect capsid protein. This results in the loss of genome-capsid correlation required for the recovery of favourable capsids during library screening.

To assess cross-packaging under library production conditions, AAV was produced using the large scale AAV production methods but using one 15 cm dish per production. A mixture of CLV WTCap6 plasmid and CLV Cap6-FS plasmid was used as the AAV genome at a 95:5 ratio. The virus was iodixanol purified, concentrated and buffer exchanged. The AAV was then incubated with 1 µL TURBO DNase (Thermo Fisher Scientific, USA) in a total volume of 20 µL for 1 hour at 37 °C, and then immediately heated to 95 °C for 10 minutes to release AAV genomes. High fidelity PCR was performed with 1 µL treated AAV sample

or 10 ng AAV genome plasmid mixture primers, using designed to flank the frame-shift region in CLV Cap6-FS.

Table 16. Cross-packaging PCR primers.

Primer	Sequence
Frame-shift Fw	ACTGCTATGGCCTCACAC
Frame-shift Rv	GGCCCAAATAGGACCCTG

The amplicons were purified using spin columns and sent for amplicon sequencing. Sequencing read quality control, alignment and variant analysis was performed by Genewiz (UK).

7.5 Cell-based Assays

7.5.1 Flow cytometry

Adherent cells were harvested by trypsinisation. Prior to staining for surface markers human B cells and PBMC-derived macrophages were incubated with FcR Blocking Reagent (Miltenyi Biotec, Germany), used at a 1/100 dilution in PBS 2 % FBS, for 20 minutes at 4 °C. For surface staining, cells were resuspended in an antibody cocktail containing relevant antibodies from Table 17 and a 1/5000 dilution of LIVE/DEAD Fixable Dead Cell Stain Kit (Thermo Fisher Scientific, USA), and incubated for 20 minutes at 4 °C. Cells were then fixed in 10% formalin (Sigma-Aldrich, USA) for 10 minutes in the dark at room temperature, washed with PBS 2% FBS. The cells were then resuspended in PBS 2% FBS and analysed using the NxT Attune (Thermo Fisher Scientific, USA).

Table 17. Antibodies used for flow cytometry.

Antibody	Conjugate	Clone	Dilution	Supplier
Anti-CD19	APC	HIB19	1/200	BioLegend (USA)
Anti-CD11b	APC	ICRF44	1/200	BioLegend (USA)
Anti-Strep Tag II	None	5A9F9	1/1000	GenScript (USA)
Anti-mouse IgG	PE	RMG1-1	1/1000	BioLegend (USA)
Anti-human IgG	PE-Cy7	M1310G05	1/200	BioLegend (USA)
Anti-HER2	FITC	2G11	1/100	Invitrogen (USA)

7.5.2 Microscopy

Images were taken using an EVOS M5000 imaging system (Invitrogen, USA) or a Nikon DS-U2 camera (Nikon, Japan) mounted to a Zeiss Axiovert 25 inverted light/fluorescence microscope (Carl Zeiss AG, Germany).

7.5.3 Fluorescent Reporter AAV Transduction Assay

To measure the functional titre of AAV preparations, AAV samples were serially diluted in DMEM 2% FBS to generate an 8-point dilution series. HEK293 cells were seeded in a 96 well plate and, once 90% confluent, transduced with 100 µL of each dilution. The cells were incubated with the virus for 72 hours at 37 °C and the frequency of eGFP-positive cells was determined by flow cytometry. The following formula was used to calculate the functional titre:

$$Functional\ Titre\ \left(\frac{TU}{mL}\right) = \frac{N\ \times f\ \times\ D}{V}$$

Where, N is the number of cells at transduction, f is the frequency of eGFP-positive cells, D is the dilution factor relative to the original AAV stock, and V is the volume diluted virus added to the cells (mL).

A dilution giving $f < 0.1$ (10%) was chosen so that, under the Poisson distribution, $f \approx m$, where m is the mean number of viral particles per cell. This ensures that most infected cells receive only one viral genome.

7.5.4 AAV Binding/Internalisation Assay

1 x 10^5 primary human B cells, or 293-AAV cells in suspension, were plated in RMPI 2% FBS in a V-bottom tissue culture plate and incubated at 4 °C for 30 minutes. The cells were centrifuged at 800 x g for 5 minutes and then incubated in 50 µL RPMI 2% FBS containing iodixanol purified SFFV-eGFP AAV6 (MOI of 1 x 10^5) for 1 hour at 4 °C to allow binding. For binding-only samples, the cells were washed three times with cold PBS and immediately processed to isolate cellular and AAV genomic DNA with an extended proteinase K treatment (see section 7.2.7). For internalisation samples, the cells were incubated at 37 °C for 1 hour and then washed with room temperature PBS before incubation with trypsin-EDTA (Sigma-Aldrich, USA) at 37 °C for 10 minutes to remove bound virus. The cells were then washed with RPMI 10% FBS, followed by two washes with room temperature PBS, and then processed as above to isolate genomic cellular and AAV DNA.

To measure the concentration of eGFP and GAPDH gene copies for each sample, the purified DNA sample was used in the following reaction mix:

Table 18. TaqMan binding/internalisation assay qPCR reaction mix.

Component	**Volume (µL)**
TaqMan Fast Advanced Master Mix (Applied Biosciences, USA)	5
20X eGFP-specific primer/probe mix	0.5
20X GAPDH-specific primer/probe mix	0.5
DNA sample	2
Nuclease-Free Water (Ambion, USA)	2

Table 19. TaqMan binding/internalisation assay qPCR primer/probe sequences

Primer	**Forward sequence**	**Reverse sequence**	**Probe sequence**	**Dye/Quencher**
eGFP	GAACCGCATCGAG CTGAA	TGCTTGTCGGCCAT GATATAG	ATCGACTTCAAGGA GGACGGCAAC	FAM/ZEN/3IABkFQ
GAPDH	CCCTTCATACCCTC ACGTATTC	ATGACAAGCTTCCC GTTCTC	CAAATTCCATGGCA CCGTCAAGGC	

For absolute quantification, a DNA fragment containing the eGFP and GAPDH amplicon sequences (See appendix section 13.2) was used to generate an 8-point standard curve using 1/10 serial dilutions from 8.59×10^9 copies/µL. The DNA copy number concentration was measured using a StepOne Plus qPCR system (Applied Biosciences, USA) using the following run parameters:

Table 20. TaqMan AAV DNA qPCR run parameters.

Step	Temperature (°C)	Time (s)	Cycles
Initial Denaturation	95	600	1
Denaturation	95	1	40
Annealing/Extension	60	20	

The ratio of eGFP copies/GAPDH copies was calculated to normalise by cell number.

7.5.5 Antibody-dependent Cell-mediated Phagocytosis (ADCP) Assay

40 nM Anti-CD47 (Invitrogen, USA; clone: B6H12) or purified Trastuzumab were incubated with or without soluble 100 nM HER2 in PBS 2% FBS for 1 hour at 4 °C. BT474 target cells were stained with 1 µL CFSE per 1 x 10^6 cells in RPMI 2% FBS for 20 minutes at 37 °C. The CFSE was then quenched with 15 mL RPMI 10% FBS for 5 minutes at 37 °C. Target cells were resuspended in Anti-CD47 or Trastuzumab mixes and incubated for 30 minutes at 37 °C. and then resuspended in RPMI 2% FBS. PBMC-derived macrophages were detached using 20 mM EDTA and resuspended in RPMI 2% FBS. 3.5 x 10^4 macrophages were mixed with 3.5 x 10^4 antibody-treated BT474 cells in RPMI 2% FBS and incubated for 3 hours. The cell mixtures were then analysed by flow cytometry using anti-CD11b antibodies to stain for macrophages.

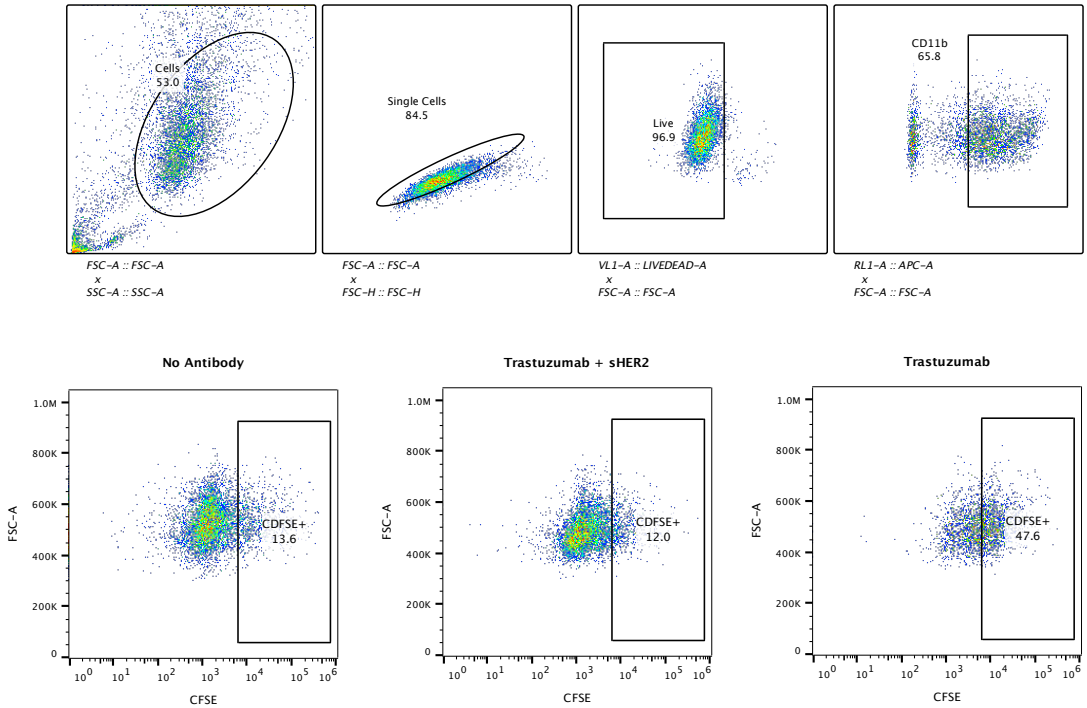


Figure 4. Representative ADCP assay flow cytometry analysis. *Top row shows gating strategy. Bottom row shows representative examples of CFSE+ populations within the CD11b+ population. CD11b+CFSE+ cells represent macrophages that have phagocytosed CFSE-stained BT474 target cells. Gate positions were set using unstained cells.*

7.6 AAV Library Screening

7.6.1 Variant Selection in Primary Human B Cells

5 x 10^5 stimulated primary human B cells were incubated in 250 µL of stimulating B cell culture medium with iodixanol purified AAV capsid library virus at an MOI of 1 x 10^4 for 2 hours. 250 µL of stimulating B cell culture medium was then added and the cells were cultured for 72 hours.

7.6.2 Recovery of Selected Capsid Sequences

For samples treated with splicing inhibitor, cells were incubated with 100 nM Pladienolide-B (Santa Cruz Biotechnology, USA) for 4 hours before RNA isolation. Total cellular RNA was isolated from transduced B cells using the Monarch Spin RNA Isolation Kit (NEB, USA) according to the manufacturer's protocol. To remove contaminating DNA to undetectable levels, purified RNA was treated with 1 µL TURBO DNase (Thermo Fisher Scientific, USA) in a 50 µL reaction for 20 minutes at 37 °C. The RNA was purified using the Monarch Spin RNA Isolation Kit following the RNA cleanup protocol. Capsid RNA was converted to cDNA using the LunaScript RT Master Mix Kit (Primer-free) (NEB, USA):

Table 21. Capsid cDNA synthesis reaction mix.

Component	**Volume (µL)**
LunaScript RT Master Mix (Primer-free) (5X)	4
Capsid cDNA Synthesis primer (10uM)	1
RNA (500 ng)	Variable
Nuclease-Free Water (Ambion, USA)	To 20

The cDNA reaction was incubated at 55 °C for 10 minutes and then 95 °C for 1 minute.

Table 22. Capsid cDNA synthesis primer

Primer	**Sequence**
Capsid cDNA Synthesis	TGTTTATTGCAGCTTATAATGGTTACAAATAAAGC

Pre-mRNA or spliced capsid sequences were then amplified by high-fidelity PCR using the following primers:

Table 23. Library screen capsid recovery primers

Primer	**Sequence**
Pre-mRNA Capsid Fw (intron-binding)	TGTGGATTTGGATGACTGC
Spliced Capsid Fw (5' UTR-binding)	GCGCAGCCATCGACG
Capsid Rv	TTTGTCCAAACTCATCAATGTATCTTATC

7.7 Common Methods

7.7.1 Polyacrylamide Gel Electrophoresis (PAGE)

To prepare protein samples from cultured cells for PAGE, cells were washed with PBS and harvested by trypsinisation. The cells were then resuspended in RIPA Lysis and Extraction Buffer (Thermo Fisher Scientific, USA) supplemented with Halt Protease Inhibitor Cocktail (Thermo Fisher Scientific, USA). The lysate was incubated at 4 °C for 20 minutes with constant rotation. The lysate was then clarified by centrifugation at 15600 x g for 20 minutes at 4 °C. To measure total protein concentration in clarified lysates, the Pierce BCA Protein Assay Kit (Thermo Fisher Scientific, USA) was used according to the manufacturer's protocol.

Protein samples were diluted in 6X Laemmli SDS Reducing Sample Buffer (Thermo Fisher Scientific, USA) and heated to 100 °C for 10 minutes. Protein samples were then separated, alongside a PageRuler pre-stained protein ladder (Thermo Fisher Scientific,

USA), at 200 V using precast 4-20% gradient SDS-PAGE gels (Bio-Rad, USA) with Tris/Glycine/SDS running buffer (25 mM Tris, 192 mM glycine, 0.1% SDS).

7.7.2 Immunoblotting

7.7.2.1 Automated Western Blot

Automated western blots were performed using the Wes automated western blot system (ProteinSimple, USA) according to the manufacturer's protocol.

7.7.2.2 Western Blot

Following SDS-PAGE, separated proteins were transferred to nitrocellulose membranes using the Trans-Blot Turbo transfer system (Bio-Rad, USA). The membrane was then blocked using EveryBlot Blocking Buffer (Bio-Rad, USA) at room temperature for 5 minutes, and then incubated with primary antibodies in blocking buffer overnight at 4 °C. The membrane was washed three times using PBST (PBS 0.1% Tween-20) for 5 minutes at room temperature and then incubated with horseradish peroxidase-conjugated secondary antibodies in blocking buffer for 1 hour at room temperature. The membrane was washed again three times and then incubated for 1 minute at room temperature with SuperSignal West Dura Extended Duration Substrate (Thermo Fisher Scientific, USA). Chemiluminescent signal was detected using the ChemiDoc Imaging System (Bio-Rad, USA).

Table 24. Antibodies used for western blotting.

Antibody	Clone	Dilution	Supplier
Anti-Strep Tag II	5A9F9	1/5000	GenScript (USA)
Anti-Capsid	B1	1/200	PROGEN (Germany)
Anti-Cofilin	D59	1/1000	Cell Signaling Technology (USA)
Anti-β-actin	13E5	1/1000	Cell Signaling Technology (USA)
Anti-Cre	D3U7F	1/200	Cell Signaling Technology (USA)
Anti-Mouse IgG HRP- linked	7076	1/10000	Cell Signaling Technology (USA)
Anti-Rabbit IgG HRP- linked	7074	1/10000	Cell Signaling Technology (USA)

7.7.3 SYPRO Ruby staining

To assess the level of contaminating endogenous proteins following the purification Trastuzumab, 10 µL of each sample collected for during the wash and elution steps was separated by SDS-PAGE. Total protein staining was performed using SYPRO Ruby Protein Gel Stain (Thermo Fisher Scientific, USA) according to the manufacturer's protocol. Stained gels were imaged using the ChemiDoc Imaging System (Bio-Rad, USA).

7.7.4 Statistics

Statistical analyses were performed using GraphPad Prism (GraphPad Software, San Diego, CA). Plots show experimental replicates as single points with the median displayed as a horizontal line.

Where stated in figure legends, data spanning multiple orders of magnitude were log-transformed prior to statistical analysis to improve normality and variance homogeneity. Assumptions of parametric testing, including approximate normality and homogeneity of variance, were assessed prior to analysis.

Comparisons between two groups were performed using unpaired two-tailed Student's t-tests. Comparisons between multiple groups were performed using one-way or two-way analysis of variance (ANOVA) as appropriate. Post-hoc multiple comparison testing was performed using Tukey's or Sidak's correction as indicated to control the Type I error rate.

Comparisons were performed as stated in each figure legend, but only significant pairwise comparisons are shown.

8 Development of a Two-Step AAV Capsid Library Platform

8.1 Introduction

During AAV capsid library screening, capsid performance is evaluated using DNA sequencing to measure enrichment of capsid variant sequences under defined selection conditions. For most effect use, each member of the viral library must encode the capsid variant used to encapsulate it. Loss of this genome-capsid correlation during library production would result in the failure to correctly identify the most effective capsids.

Early in the development of complex AAV capsid libraries, Müller *et al* acknowledged that the transfection of the amounts of AAV genome plasmid DNA similar to those used for standard AAV production could risk the cross-packaging of genomes into the incorrect capsid and the formation of mosaic capsids (Müller et al., 2003). To mitigate this risk, Müller *et al* employed a two-step library production method in which HEK293 cells were co-transfected with an AAV2-based peptide insertion plasmid library and a wild-type AAV2 capsid expression plasmid at a 1:1 ratio, along with helper plasmids, to produce chimeric AAV2 shuttle vectors. The shuttle vectors were then used to transduce producer cells at a low multiplicity of infection (MOI), significantly reducing the probability of delivering multiple AAV genomes per cell, and the AAV genomes were re-packaged into the genome encoded capsid. An alternative early approach involved the transfection of very low amounts of plasmid library and was used successfully to identify novel capsids

(Maheshri et al., 2006). A direct comparison between the two methods revealed that, for the production of random peptide display libraries with the goal of retargeting AAV2, the use of a two-step library skews the amino-acid contribution of the insertion towards sequences that exhibit increased heparin affinity (Körbelin et al., 2017), reflecting the use of heparan sulfate proteoglycan (HSPG) as a primary attachment factor for AAV2 (Summerford and Samulski, 1998a). Therefore, the use of chimeric shuttle vectors biased the library based on producer cell transduction efficiency before selection in the target cells.

The sequential discovery of more natural serotypes with divergent transduction properties led to library designs based on serotypes other than AAV2. The use of such libraries with chimeric AAV2-based shuttle vectors would likely lead to significant bias towards sequences from serotypes that confer high transduction efficiency in HEK293 cells. Due to the risk of bias using a chimeric shuttle vector approach to library production, the low plasmid mass transfection method (one-step approach) has been widely adopted. However, the extent to which this method ensures faithful genotype–phenotype linkage in large complex libraries remains unclear. Using simple systems, the genome-capsid correlation and degree of mosaicism has been evaluated for different plasmid library configurations (Nonnenmacher et al., 2015a, Schmit et al., 2020). Cross-packaging and mosaicism appeared to rapidly decrease upon plasmid dilution when assessing productions using two serotypes. However, the *in vivo* screening of a small library of natural serotypes did not fully recapitulate the enrichment of serotypes using a mock library of individually produced vectors, although libraries produced with low plasmid input performed better than high plasmid input libraries (Schmit et al., 2020).

Therefore, although the one-step approach has been used extensively and simple production models suggest cross-packaging and mosaicism is limited, significant noise may be present in large complex libraries from which rare variants must be detected.

To overcome these limitations and enable the use of a two-step strategy with modern library designs, a new shuttle system in which capsid expression is tightly supressed during shuttle vector production, and inducible during shuttle vector re-packaging, is required (Figure 5). Several classes of inducible regulatory systems are available to achieve such control, including drug-inducible transcriptional repressors/activators, CRISPR/Cas9 approaches, recombinase systems and post-transcriptional strategies such as riboswitches and RNA interference. There are several crucial requirements that a gene control system must satisfy for this purpose. Capsid leakage must be negligible to prevent expression during shuttle vector production. The system must also be fully reversible and avoid crosstalk with inducible AAV helper systems. Finally, the trigger must be suitable for a scalable and efficient production platform. The requirement for very low capsid leakage favours inducible systems, over repression systems, as amplification of the AAV genome during production may lead to saturation of the cellular levels of repressor.

An additional consideration in the design of modern AAV libraries is the choice of helper system used to support vector production. Standard triple transfection approaches, employing plasmid-based adenoviral helpers, remain widely used but are limited in scalability and in the yield and quality of the AAV produced. To address these limitations, Tetracycline-Enabled Self-Silencing Adenovirus (TESSA) was developed to provide the

full complement of adenoviral helper functions while restricting adenoviral replication through self-silencing of the adenovirus major late promoter. This platform has been shown to increase rAAV yields and improve particle quality (Su et al., 2022b), and therefore represents an ideal helper system for the implementation of two-step library production strategies. Incorporation of TESSA also raises the possibility of a plasmid-free repackaging step during library production, simplifying library production.

The aim of this chapter was to develop and characterise a two-step AAV capsid library production strategy that preserves genome–capsid linkage while avoiding the biases associated with earlier shuttle vector systems. We hypothesised that placing the AAV genome-encoded capsid under the control of a tightly regulated inducible gene control system would enable shuttle vectors to be propagated without capsid production in producer cells, and subsequently re-packaged into their genome-encoded capsids following induction. This approach was expected to reduce noise during library screening associated with cross-packaging and mosaicism relative to standard one-step methods, while maintaining compatibility with modern high-complexity library designs

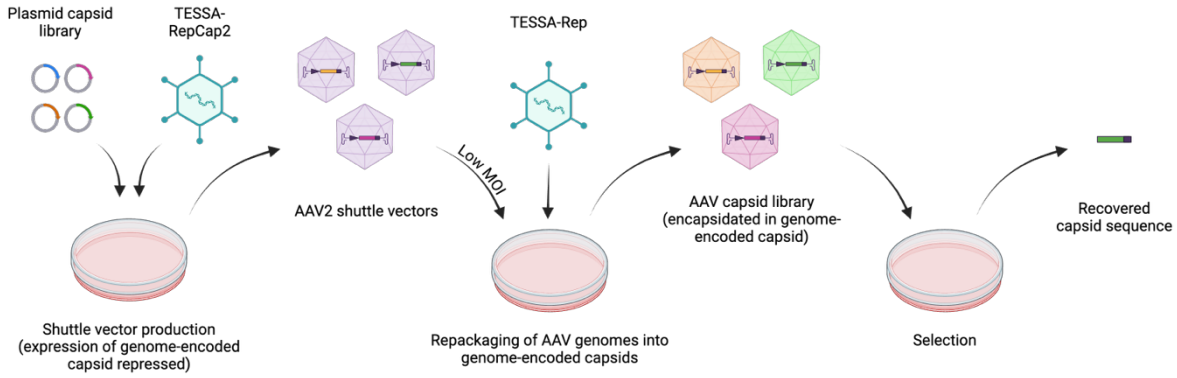


Figure 5 – Schematic representation of an AAV capsid library production using the inducible two-step library. *In the first step of the AAV capsid library production, the plasmid library is transfected into producer cells in the presence of TESSA-RepCap2, which supplies adenoviral helper functions along with AAV Rep and AAV2 capsid proteins from a single adenoviral vector. During this step, expression of the genome-encoded is restrained and the capsid library is packaged into AAV2 capsid. The AAV2 shuttle vectors are used to transduce producer cells in the presence of TESSA-Rep at a low MOI to reduce the liklihood of cells receiving multiple AAV genomes. During this second step, the genome-encoded capsid is expressed and, in the absence of trans-supplied capsid, the capsid library is re-packaged into the genome-encoded capsid. Created with BioRender.com.*

8.2 Aims

1. Evaluate TESSA-compatible transcriptional switches for their capacity to induce genome-encoded capsid expression while maintaining strict repression in the absence of induction.
2. Characterise the performance of inducible shuttle vectors during repackaging into their genome-encoded capsid variants.

8.3 Results

8.3.1 Characterisation of a CymR-VP16-inducible System.

The two-step library approach presented here uses a genetic switch to allow a diverse library encoding many different AAV capsid sequences to be packaged within identical shuttle vectors comprising AAV2 capsids. This design would enable uniform genome delivery to packaging cells and decrease the multiplicity of genomes per cell, thereby minimising the formation of mosaic capsids and the mis-packaging of genomes into inappropriate capsid proteins. To facilitate this, a system is required in which capsid expression can be tightly controlled so that, in the "off state", genome-encoded capsid is not expressed during packaging into AAV2 shuttle vectors. The resulting shuttle vectors are then used to transduce producer cells at a low multiplicity-of-infection (MOI), in the absence of trans capsid, and re-packaged into the genome-encoded capsid by switching on capsid expression (on-state).

The initial design utilised a CymR-VP16-inducible system in which bacterial repressor CymR, fused to the transactivation domain of VP16, binds cumate operator sequences (CuO) to activate expression from a minimal CMV promoter (Mullick et al., 2006). In this design, capsid expression is controlled by the CymR-VP16-inducible minimal CMV promoter and an eGFP reporter gene is driven by a constitutive promoter (p567) (Figure 6a).

To test for leakage of capsid gene expression in the absence of CymR-VP16, and for adequate capsid expression in the presence of CymR-VP16, HEK293 cells were

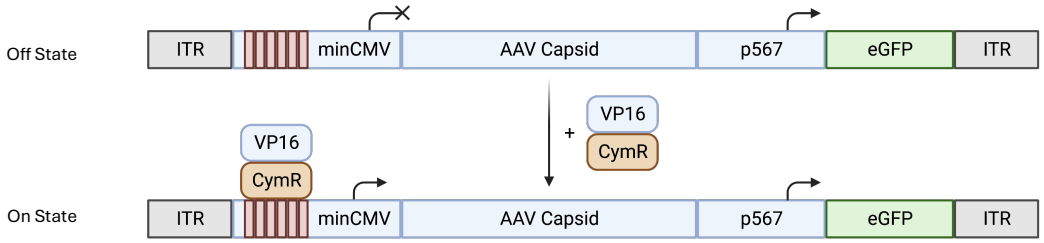
transiently transfected with the AAV capsid library constructs encoding AAV2, AAV5 and AAV8 capsids, with or without a CymR-VP16 expression plasmid. In the absence of CymR-VP16, capsid protein was undetectable by western blot for all three serotypes, while robust capsid expression was detected in presence of CymR-VP16 (Figure 6b). These results confirmed that, in a transient transfection setting, the system enables robust capsid expression when induced and without detectable leakage in the absence of CymR-VP16.

The system was then tested during AAV2 shuttle vector production. AAV2 shuttle vectors were produced using TESSA-RepCap2 to provide AAV Rep and trans AAV2 capsid in the absence of CymR-VP16. Shuttle vector titres, measured by qPCR detection of genome copy numbers, were within the expected range for crude AAV2 (Figure 6c). To assess whether the genome-encoded capsid sequences were inappropriately expressed during shuttle vector production, HEK293 cells were separately transfected with CuOx6-Cap2, CuOx6-Cap5 and CuOx6-Cap8 plasmids in the absence of CymR-VP16 and transduced with TESSA-Rep, which supports AAV genome replication without providing trans capsid. Under these conditions, any detected capsid protein would be derived exclusively from the genome-encoded capsid. Western blot analysis revealed clear expression of AAV2, AAV5, and AAV8 capsids (Figure 6d), suggesting that repression of genome-encoded capsid expression was insufficient to restrain expression from the amplified AAV genome.

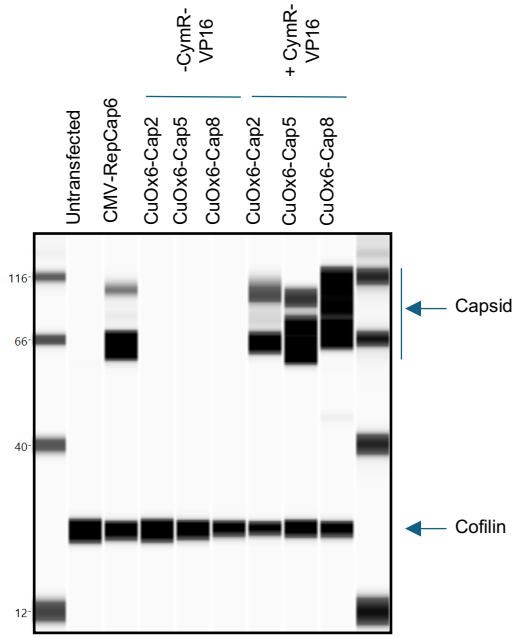
The transduction efficiency of shuttle vectors produced using the inducible constructs was then evaluated using an eGFP reporter-based assay. AAV2 shuttle vectors generated using the CuOx6-Cap2 construct were compared to control AAV2 vectors expressing

eGFP from either the p567 or CMV promoter. The CuOx6-Cap2 shuttle vectors exhibited significantly lower apparent transduction efficiency compared to both control vectors, reflected in elevated GC/TU ratios (Figure 6e). The p567 is a weaker promoter than CMV, accounting for the difference in apparent infectivity between the p567-eGFP and CMV-eGFP control viruses (Figure 7). As the AAV2 shuttle vector functional assay used an AAV2 genome-encoded capsid, the reduction in infectivity could not be explained by mosaicism between different serotypes. Furthermore, the eGFP signal from cells transfected with the p567-eGFP and CuOx6-Cap5 library plasmid was indistinguishable, confirming that the low GC/TU values are not a result of poor eGFP expression at the plasmid level (Figure 7).

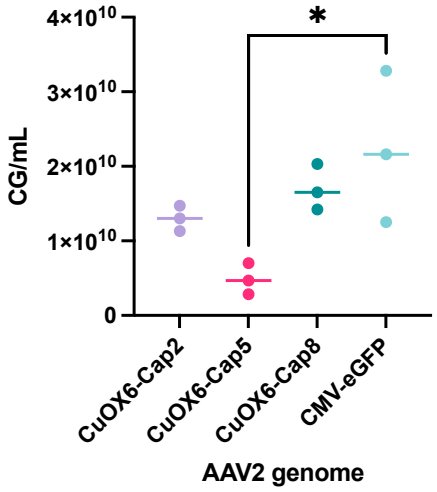
a)


Off State
ITR
minCMV
AAV Capsid
p567
eGFP
ITR
VP16
CymR
On State
ITR
minCMV
AAV Capsid
p567
eGFP
ITR

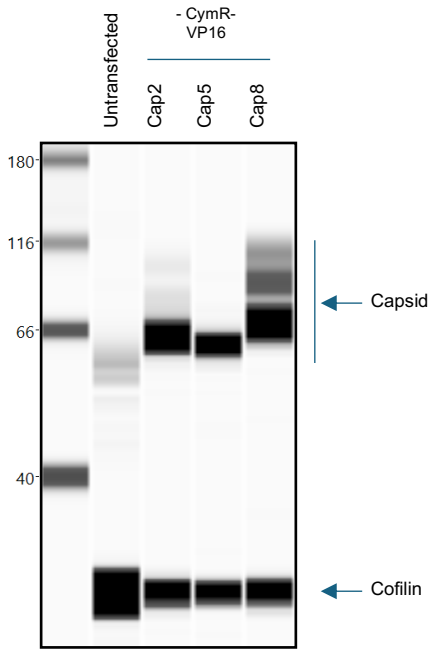
b)


-CymR-VP16
+ CymR-VP16
Untransfected
CMV-RepCap6
CuOx6-Cap2
CuOx6-Cap5
CuOx6-Cap8
CuOx6-Cap2
CuOx6-Cap5
CuOx6-Cap8
116
66
40
12
Capsid
Cofilin

c)


4×10^10
3×10^10
2×10^10
1×10^10
0
CG/mL
*
CuOX6-Cap2
CuOX6-Cap5
CuOX6-Cap8
CMV-eGFP
AAV2 genome

d)


- CymR-VP16
Untransfected
Cap2
Cap5
Cap8
180
116
66
40
Capsid
Cofilin

e)

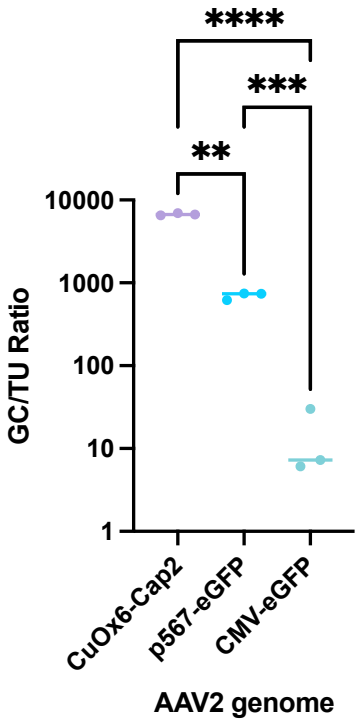
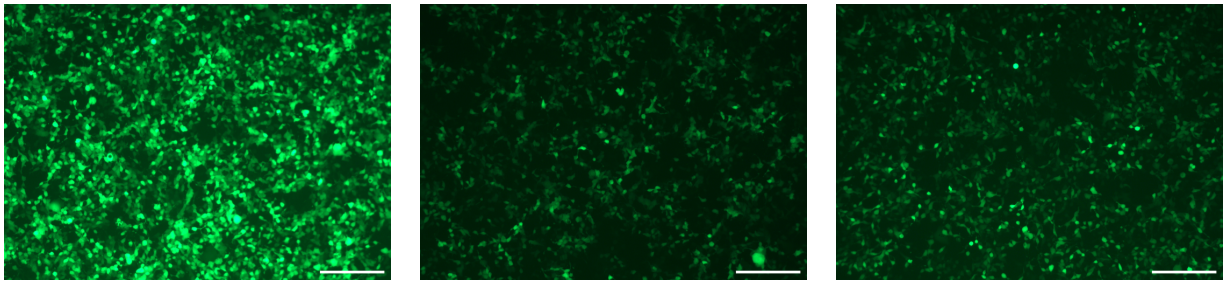

****
***
**
10000
1000
100
10
1
GC/TU Ratio
CuOx6-Cap2
p567-eGFP
CMV-eGFP
AAV2 genome

Figure 6 – Characterisation of CymR-VP16-inducible capsid expression and shuttle vector performance. *A) schematic of CymR-VP16-inducibe AAV capsid library system. miniCMV (minimal CMV promoter), enhanced GFP (eGFP). B) Detection of AAV capsid protein in HEK293 cells transiently transfected with AAV2, AAV5 or AAV8 encoding CymR-VP16-inducible library construct plasmids, with or without a CymR-VP16 expression plasmid, 72 hours post-transfection by western blot. C) CymRVP16-inducible AAV2 shuttle vector genome titres (genome copies/mL) measured by GFP-specific qPCR. D) Detection of AAV capsid protein in HEK293 cells transduced with TESSA-Rep and transiently transfected with AAV2, AAV5 or AAV8 encoding CymR-VP16-inducible library construct plasmids 72 hours post-transfection by western blot. E) Genome copy/transducing units (GC/TU) ratio of AAV2 shuttle vectors, compared to p567-eGFP and CMV-eGFP AAV2 control virus, 72 hours post-transduction in HEK293 cells. Each data point represents an independent experiment. Statistical analysis performed using one-way ANOVA with correction for multiple post-hoc comparisons using Tukey's test. For data shown in plot E, the data was log-transformed before statistical testing. ****P ≤ 0.000.1, **P ≤ 0.01, *P ≤ 0.05*



CMV-eGFP P567-eGFP CuOx6-Cap5

expression levels from CymR-VP16 inducible library plasmid, compared to p567-eGFP and CMV-eGFP, controls following transient transfection in HEK293 cells. *HEK293 cells were transiently transfected with equimolar amounts of CMV-eGFP, p567-eGFP or CuOx6-Cap5-p567-eGFP plasmids and imaged after 72 hours in culture. Scale bar, 1000 µm.*

While mosaicism cannot account for the poor performance of the shuttle vectors, it is possible that excess capsid protein reduces packaging efficiency, leading to elevated empty capsid levels that interfere with cell binding and saturate trafficking processes during transduction. To assess whether leakage of the genome-encoded capsid could account for the reduced transduction efficiency of the shuttle vectors, shRNAs were designed to knock-down AAV5 and AAV8 capsid proteins. However, to use this strategy within a two-step library platform, which incorporates many diverse sequences, the shRNA would need to target the genome-encoded capsid variants independently of the capsid sequence. Therefore, a universal shRNA was also designed targeting the 3' UTR of the genome-encoded capsid that would spare the trans capsid supplied by TESSA. HEK293 cells were co-transfected separately with the CuOx6-Cap2, CuOx6-Cap5 or CuOx6-Cap8 plasmids, and a stuffer plasmid or capsid-specific shRNA expression plasmid. The cells were also transduced with TESSA-Rep to initiate AAV genome replication as before. Efficient knock-down by the serotype-specific shRNAs was observed by western blot for the CuOx6-Cap5 and CuOx6-Cap8 plasmids (Figure 8c), however no rescue of apparent transduction efficiency was observed for shuttle vectors produced in the presence of serotype-specific shRNAs compared to the stuffer DNA control (Figure 8a). Robust knock-down of the CuOx6-Cap2 and CuOx6-Cap5 plasmid-encoded capsid was observed using the 3'UTR-specific universal shRNA, and moderate knock-down was observed for the CuOx6-Cap8 plasmid (Figure 8d). However, no improvement was seen in the apparent transduction efficiency of shuttle vectors produced with concurrent knock-down of the genome-encoded capsid using universal shRNAs (Figure 8b). These findings suggest that excess capsid in the shuttle vector preparations does not account for the low observed transduction efficiency.

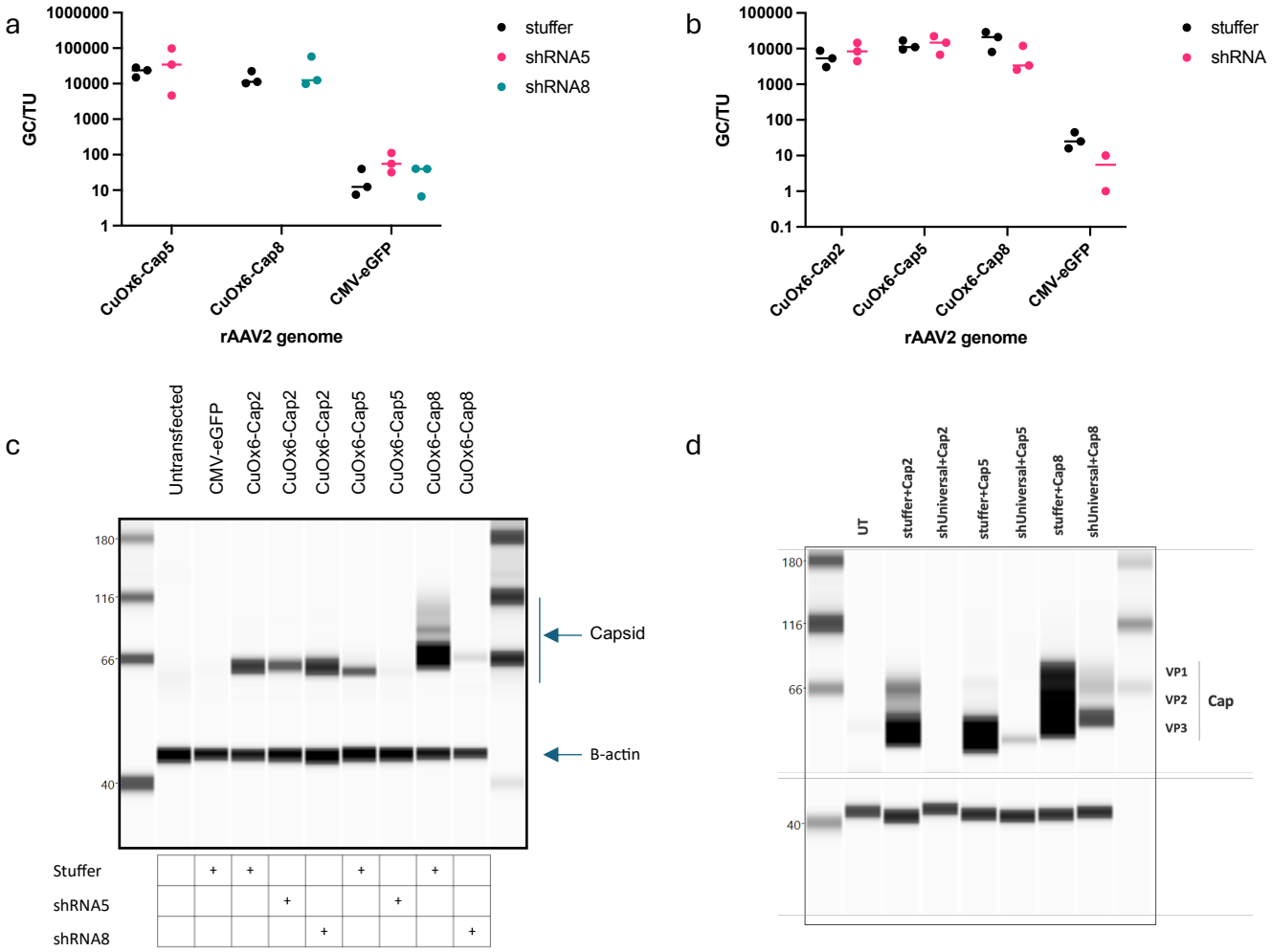


Figure 8 – Selective shRNA mediated knockdown of AAV genome encoded capsid proteins and the effect on shuttle vector performance. *A) GC/TU ratio of CymR-VP16 inducible shuttle vectors, in HEK293 cells, produced in the presence of serotype-specific shRNAs or stuffer DNA. B) GC/TU ratio cells of CymR-VP16 inducible shuttle vectors, in HEK293 cells, produced in the presence of stuffer DNA or an shRNA targeting the 3'UTR of the genome-encoded capsid. C) Capsid-specific western blot of a). d) Capsid-specific western blot of b). Each data point represents an independent experiment. Statistical analysis performed using two-way ANOVA with correction for multiple comparisons using Tukey's test. Comparisons were limited to within AAV genome groups only, in line with the null hypothesis that shRNA treatment does not affect GC/TU value for a given AAV2 genome.*

While shRNA-mediated knockdown of genome-encoded capsid transcripts did not improve shuttle vector performance, the presence of leakage remains a critical concern, since uncontrolled capsid expression during library production could cause mosaicism and introduce sequence bias during repackaging, as documented in earlier implementations of two-step methods (Körbelin et al., 2017). The AAV ITRs are known to have innate promoter activity (Earley et al., 2020) which could account for the leakage of capsid expression during shuttle vector production. To test whether the AAV ITRs could drive capsid expression using our system, the CuOx6-minimal CMV promoter was deleted from the CuOx6-Cap2 plasmid. HEK293 cells were transfected with CMV-eGFP, CuOx6-Cap2 or the CuOx6-deleted plasmid, in the presence of TESSARep. After 72 hours, the cell lysates were added to HEK293 cells and GFP expression was assessed 72 hours later. Low level GFP signal was detected in cells exposed to lysates from CuOx6-deleted plasmid transfections, supporting the possibility that the AAV ITRs alone can drive capsid expression (Figure 9a). This was not an artifact of DNA carryover as no GFP fluorescence signal was detected for the plasmid control lacking the cap2 sequence (CMV-eGFP) (Figure 9a). The expression of capsid protein in the absence of the CuOx6 promoter was confirmed by western blot of the cell lysates (Figure 9b). Therefore, the leakage can partially be explained by inherent ITR activity, however, considerably higher levels of capsid protein was detected in lysates from cells transfected with the cuOx6-Cap2 control compared to those from CauOx6-deleted plasmid transfections. This suggests that ITR promoter activity may contribute to capsid expression during shuttle vector production but is unlikely to be solely responsible.

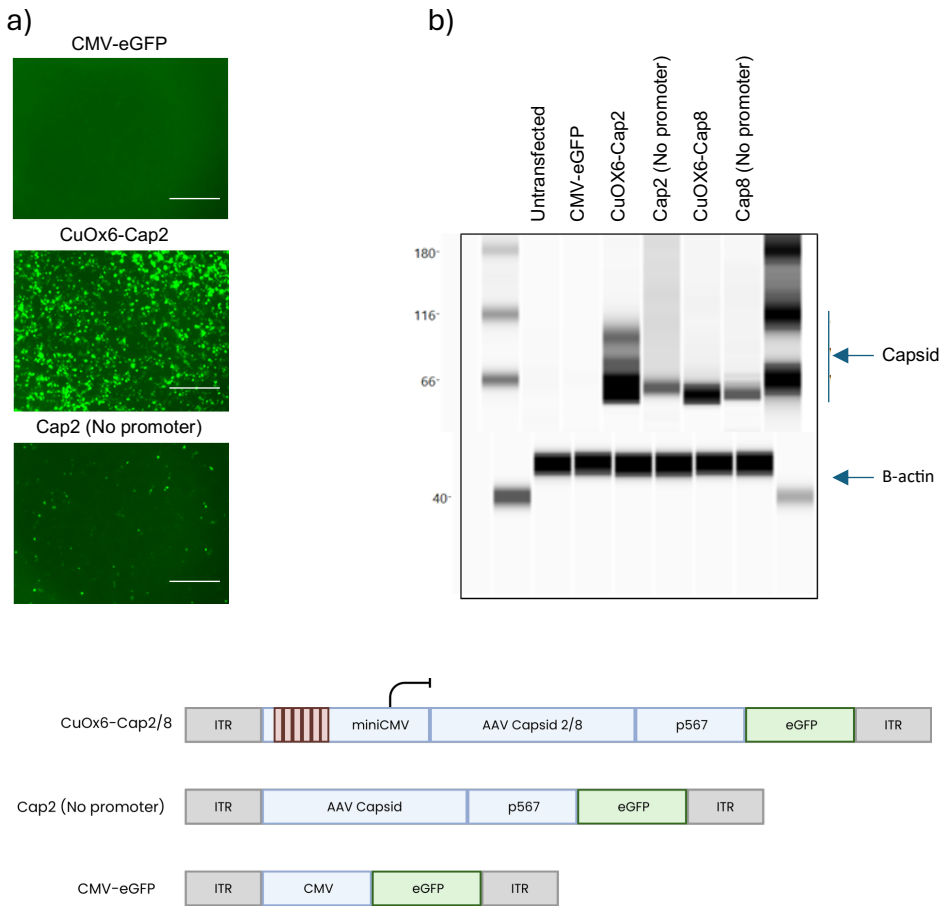


Figure 9 –Capsid expression driven by intrinsic ITR promoter activity using plasmids lacking the CuOx6 miniCMV promoter. *A) eGFP expression in HEK293 cells 72 hours after exposure to lysates from TESSA-Rep transduced cells transfected with the CymR-VP16 library construct, with or without the CuOx6 promoter. B) Capsid-specific western blot of lysates used in a). Schematic depicts plasmid variants used for transfections. Scale bar, 750 µm.*

To examine the effect of individual construct elements on the apparent transduction efficiency of AAV2 shuttle vectors, variants were generated in which the capsid sequence and CuOx6 promoter were removed individually or in combination. (Figure 10a). Removal of the CuOx6 promoter alone did not improve transduction efficiency compared to the p567-eGFP control AAV2 virus, however removal of the capsid sequence rescued apparent transduction efficiency to p567-eGFP control levels. To probe whether the presence of capsid DNA sequences affected transduction efficiency, the capsid

sequence was replaced with firefly luciferase (FLuc). Similar GC/TU values were obtained using AAV2 shuttle vectors produced using the FLuc construct compared to the CuOx6-Cap2 construct (Figure 10b). Therefore, the inclusion of a coding sequence upstream of the p567 promoter appears to limit apparent shuttle vectors transduction efficiency.

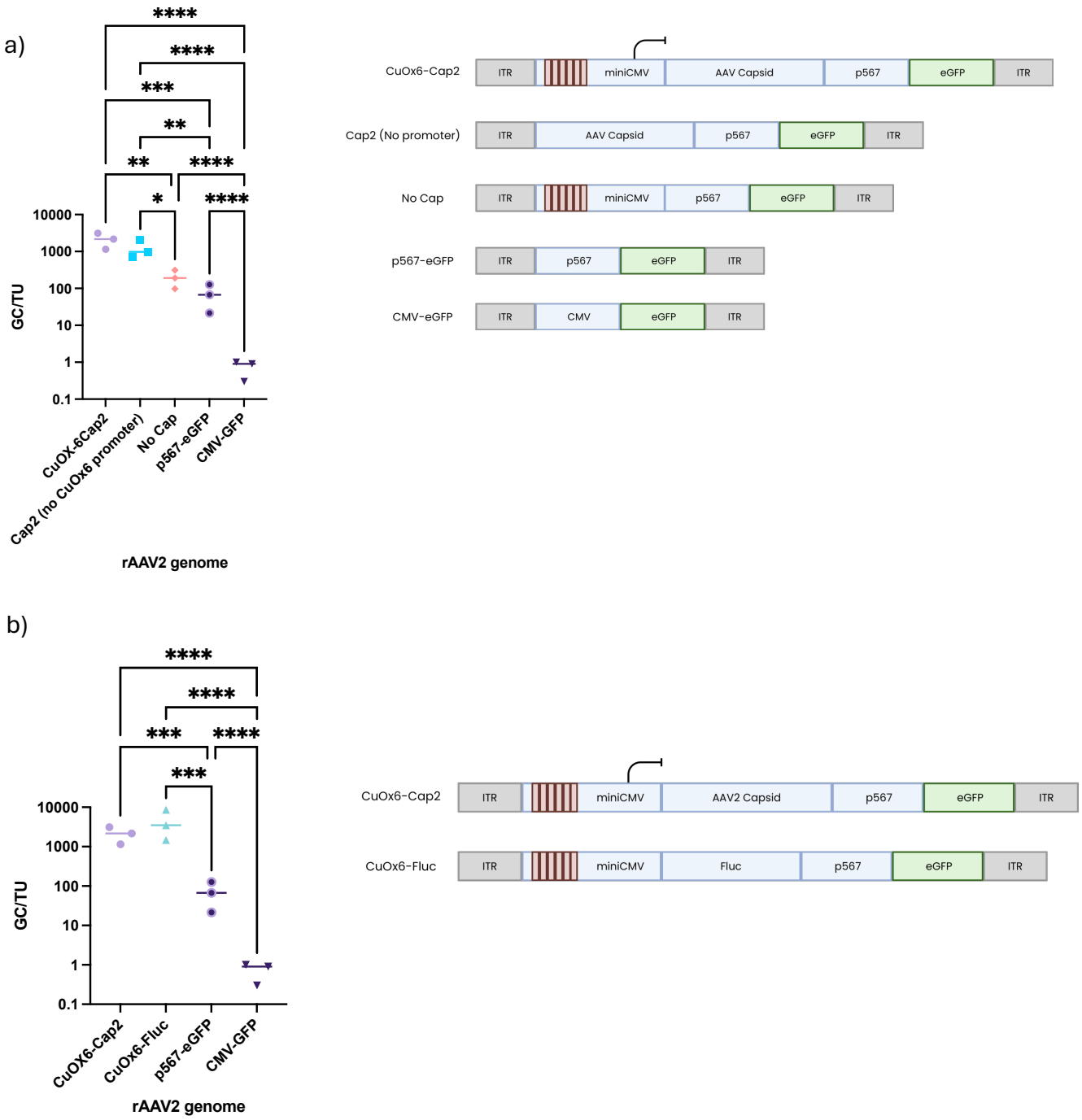


Figure 10 – Removal of CymR-inducible library construct elements to probe promoter and capsid sequence effect on apparent shuttle vector transduction efficiency. *A) GC/TU ratio of CymR-CP16 inducible shuttle vector variants, in HEK293 cells. Schematic depicts variants used to produce shuttle vectors. B) GC/TU ratio of CymR-CP16 inducible shuttle vectors encoding AAV2 capsid or FLuc, in HEK293 cells. Each data point represents an independent experiment. Statistical analysis performed on log-transformed data using one-way ANOVA with correction for multiple comparisons using Tukey's test (a,b). **P ≤ 0.01, *P ≤ 0.05.*

Due to the limited packaging capacity of rAAV, a polyA termination sequence between the capsid sequence and the GFP gene was omitted during the project design phase. One possible explanation for the low apparent transduction efficiency of vectors containing a capsid or FLuc sequence upstream of eGFP could be transcriptional repression of GFP expression due to readthrough from the upstream gene. To test this, the bovine growth hormone (bGH) polyA sequence, or a small synthetic polyA sequence, was inserted immediately after the capsid sequence in the CuOx6-Cap5 plasmid. In addition, an outward configuration plasmid was produced in which the capsid and eGFP genes are transcribed in opposite directions to exclude the possibility of transcriptional read-through. However, neither the addition of polyA sequences or the reversal of the capsid gene orientation improved the apparent transduction efficiency of AAV2 shuttle vectors (Figure 11a).

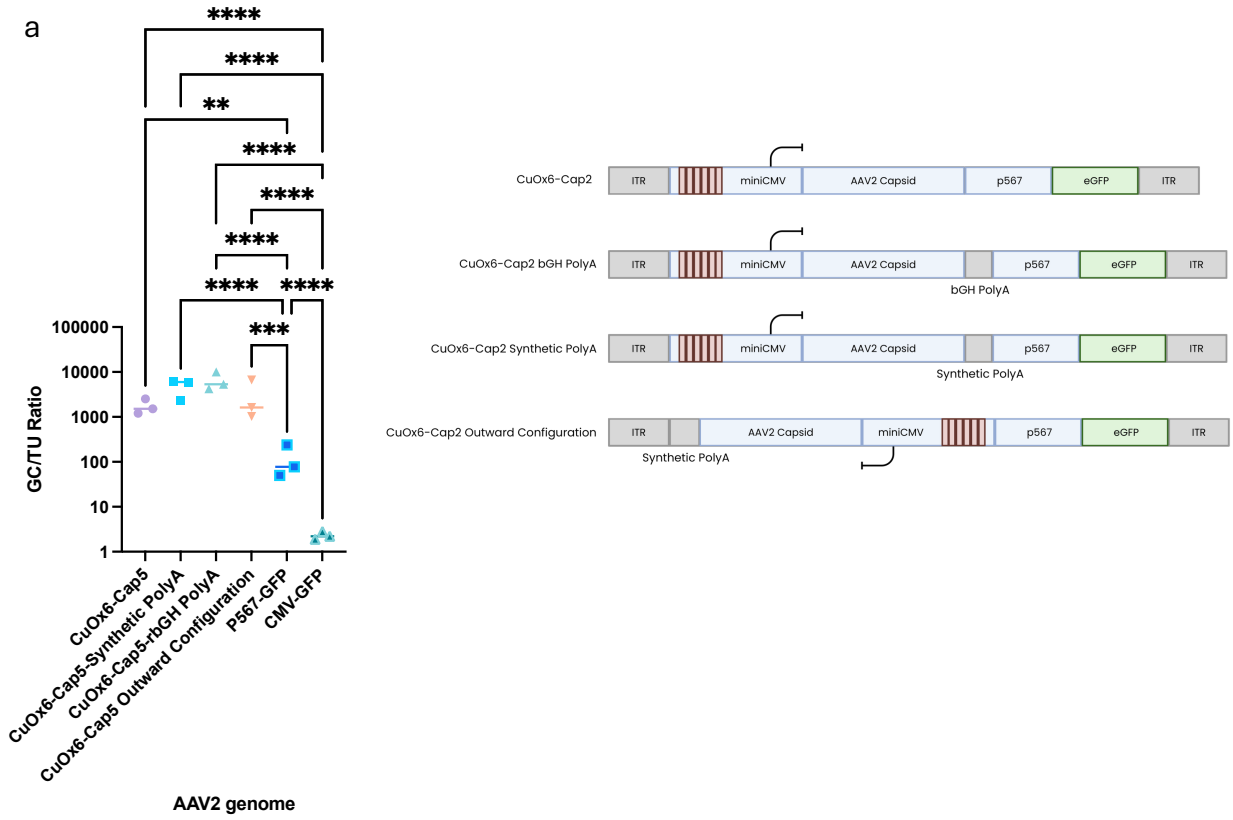


Figure 11 – Introduction of polyA sequences and reversal of the capsid gene in the CymR-VP16-inducible capsid construct to investigate the role of transcriptional read-through on apparent transduction efficiency. *A) GC/TU ratio of CymR-CP16 inducible shuttle vector variants, in HEK293 cells. Schematic depicts variants used to produce shuttle vectors. Each data point represents an independent experiment. Statistical analysis performed on log-transformed data using one-way ANOVA with correction for multiple comparisons using Tukey's test (a,b). *P ≤ 0.05.*

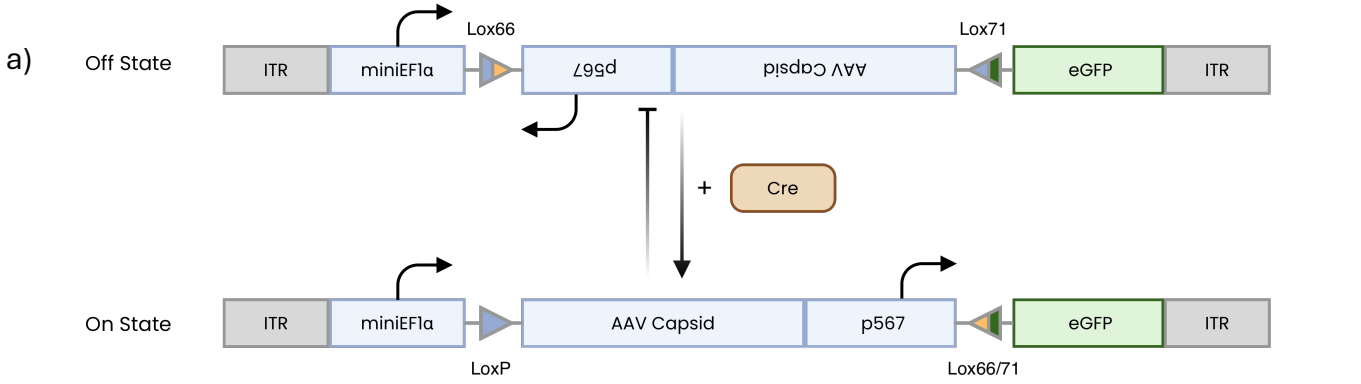
Taken together, these experiments demonstrated that although the CymR-VP16 system functioned as a reliable inducible switch in transient transfection experiments, it did not provide sufficiently tight repression during shuttle vector production. Genome-encoded capsid proteins were expressed in the absence of CymR-VP16, and the inclusion of capsid or other large coding sequences impaired shuttle vectors performance. It was therefore concluded that the CymR-VP16 system does not meet the requirements for unbiased two-step library production and that alternative inducible systems would need to be explored.

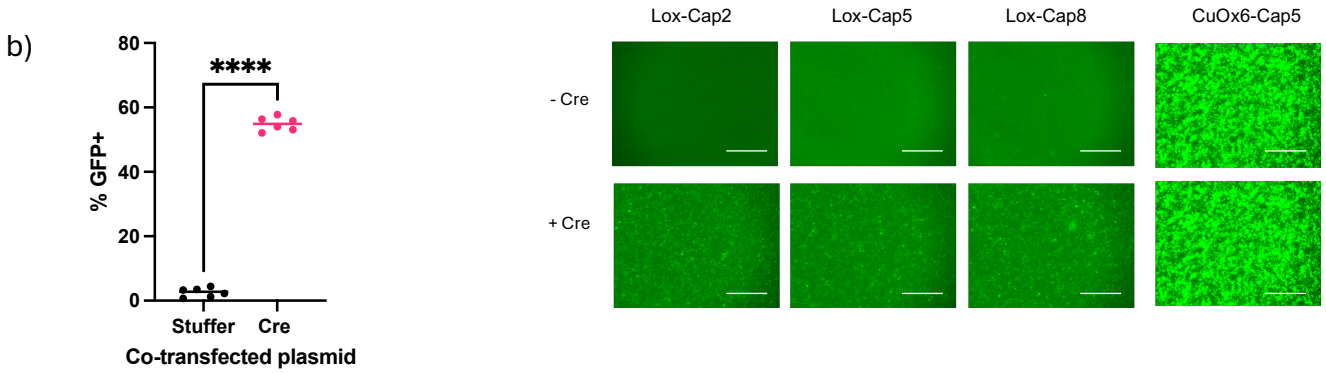
8.3.2 Characterisation of a Cre-inducible System.

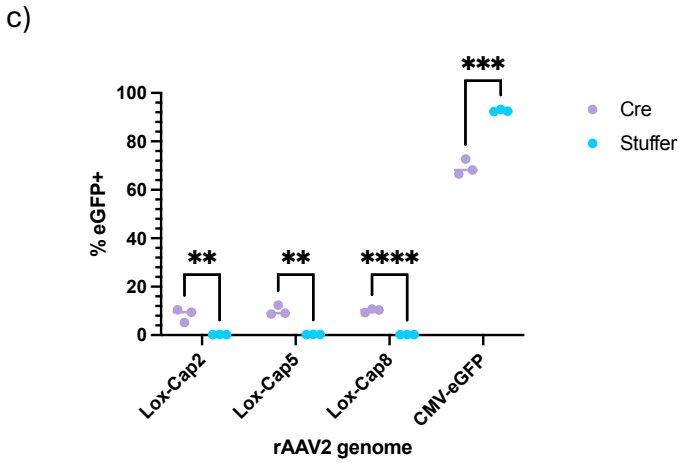
The leakage of genome-encoded capsid during shuttle vector production observed during the characterisation of the CymR-inducible system (section 4.3.1) introduces a risk of mosaic capsid formation. Such mosaic particles could bias the library during the repackaging step by favouring capsid sequences that form functional hybrids with AAV2 capsid and retain the ability to transduce HEK293 producer cells. As shown in section 4.3.1, this may be improved using shRNAs against the 3' UTR of the genome-encoded capsid; however, knock-down using the universal shRNA was not equally effective across different capsids (Figure 8b), potentially introducing an additional source of bias during library production. Therefore, a more robust binary expression system was designed using the Cre-Lox system (Figure 12a). In this system, the capsid and p567 promoter sequences are flipped in the "off state", preventing eGFP or capsid expression. In the presence of Cre recombinase, the capsid and p567 promoter sequences are flipped into the "on state" in which capsid expression is driven by a minimal EF1α promoter and eGFP is driven by the p567 promoter. This design allows recombination during library production to be monitored by eGFP expression. Furthermore, following repackaging of shuttle vectors into the genome-encoded capsids, the final library contains AAV genomes in the "on state" allowing for FACS sorting during screening. Heterotypic mutant loxP sites were used to limit "flip-flopping" between the two states due to the low binding affinity of Cre recombinase to the lox66/71 (lox72) site in the "on state" configuration (Albert et al., 1995).

To confirm that the design was capable of recombination, an AAV2 capsid-encoding Cre-inducible shuttle vector plasmid was co-transfected into HEK293 cells with a Cre recombinase expression plasmid, or stuffer DNA, and the frequency of eGFP expressing cells was assessed by flow cytometry. Encouragingly, eGFP expression was observed in approximately 60% of Cre-expressing cells, indicating successful recombination, while very few cells expressed eGFP in cells that lacked Cre-recombinase (Figure 12b). However, the intensity of the GFP signal in GFP-positive cells appeared relatively low (Figure 12b images).

The Cre-inducible AAV genomes were packaged into AAV2 capsid using TESSA-RepCap2 and used to transduce HEK293 cells in the presence or absence of Cre recombinase or stuffer DNA. Low frequencies of eGFP positive cells were observed in cells expressing Cre recombinase, demonstrating that AAV2 shuttle vector genomes are capable of recombination (Figure 12c). However, during amplification of the Cre-inducible constructs in cells transduced with TESSARep, and co-transfected with a Cre-expressing construct or stuffer DNA, no capsid was detected by western blot in Cre-expressing cells (Figure 12d).


a)
Off State
ITR
miniEF1α
Lox66
p567
AAV Capsid
Lox71
eGFP
ITR
+
Cre
On State
ITR
miniEF1α
LoxP
AAV Capsid
p567
Lox66/71
eGFP
ITR


b)
% GFP+
80
60
40
20
0
****
Stuffer
Cre
Co-transfected plasmid
Lox-Cap2
Lox-Cap5
Lox-Cap8
CuOx6-Cap5
- Cre
+ Cre


c)
% eGFP+
100
80
60
40
20
0
**
**
****
***
Lox-Cap2
Lox-Cap5
Lox-Cap8
CMV-eGFP
rAAV2 genome
Cre
Stuffer

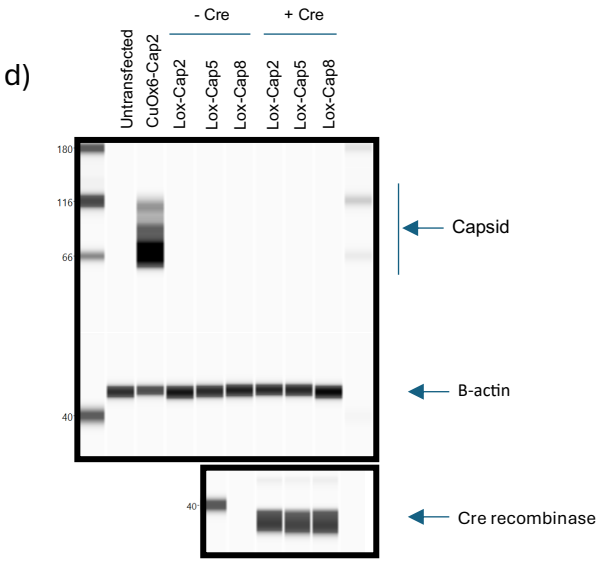

d)
- Cre
+ Cre
Untransfected
CuOx6-Cap2
Lox-Cap2
Lox-Cap5
Lox-Cap8
Lox-Cap2
Lox-Cap5
Lox-Cap8
180
116
66
40
40
Capsid
B-actin
Cre recombinase

Figure 12 – Characterisation of a Cre-inducible AAV capsid library system. *A) Schematic of Cre inducible library system. miniEF1α minimal Eukaryotic Translation Elongation Factor 1α promoter. B) Frequency of eGFP positive HEK2983 cells following co-transfection with a AAV5 capsid encoding Cre inducible library plasmid and a Cre expression plasmid, or stuffer DNA. Representative images of HEK293 cells transfected with Cre-inducible library plasmids encoding AAV2, AAV5 or AAV8 capsids (Lox-Cap2/5/8) or CymR-VP16-inducible library plasmid encoding AAV5 capsid in the presence or absence of a Cre expression plasmid. Scale bar, 750 µm C) Frequency of eGFP positive HEK293 cells transduced with AAV2 shuttle vectors (MOI = 5000) produced with Lox-Cap2/5/8 plasmids and transfected with Cre expression plasmid of stuffer DNA. D) Detection, by western blot, of AAV capsid protein in HEK293 cells transfected with Cre-inducible library plasmids, with or without a Cre expression plasmid, and transduced with TESSA-Rep. Each data point represents an independent experiment. Statistical analysis performed using an unpaired t-test (b) or by one-way ANOVA with correction for multiple comparisons using Tukey's test (c). ****P ≤ 0.0001 , ***P ≤ 0.001, **P ≤ 0.01, *P ≤ 0.05.*

The low recombination efficiency observed in cells transduced with Cre-inducible shuttle vectors, the weak eGFP signal intensity following transient transfection, and the lack of capsid expression from amplifying Cre-inducible AAV genomes in the presence of Cre could be explained by aberrant recombination events, resulting in species other than the "on state". Therefore, the resolution of loxP sites following recombination was assessed by sequencing clonal plasmid DNA recovered from HEK293 cells co-transfected with the Cre-inducible library plasmid and Cre-expression plasmid. Firstly, to exclude the possibility of a difference in recombination efficiency between circular plasmid DNA and a linear DNA fragment representing an AAV genome following second-strand synthesis and prior to concatermarisation, which would suggest differences in recombination resolution, the Cre-inducible plasmid was linearised by restriction digestion outside of the AAV ITRs and gel purified. The linearised plasmid was co-transfected into HEK293 cells in the presence of absence of Cre recombinase. GFP expression levels were similar between linear and circular DNA species in the presence of Cre (Figure 13a). Additionally, linearisation of the plasmid DNA did not rescue capsid

expression during AAV replication in Cre expressing cells (Figure 13b). Next, clonal plasmid DNA isolated from transiently transfected cells expressing Cre recombinase was sequenced by sanger sequenced to check for the incorrect resolution of LoxP junctions. The left and right junctions from all 6 sequenced plasmids were identical to the predicted sequence (Figure 13c). Together these findings suggest that resolution of the LoxP sites does not generate sequences that would be expected to interfere with capsid or GFP expression.

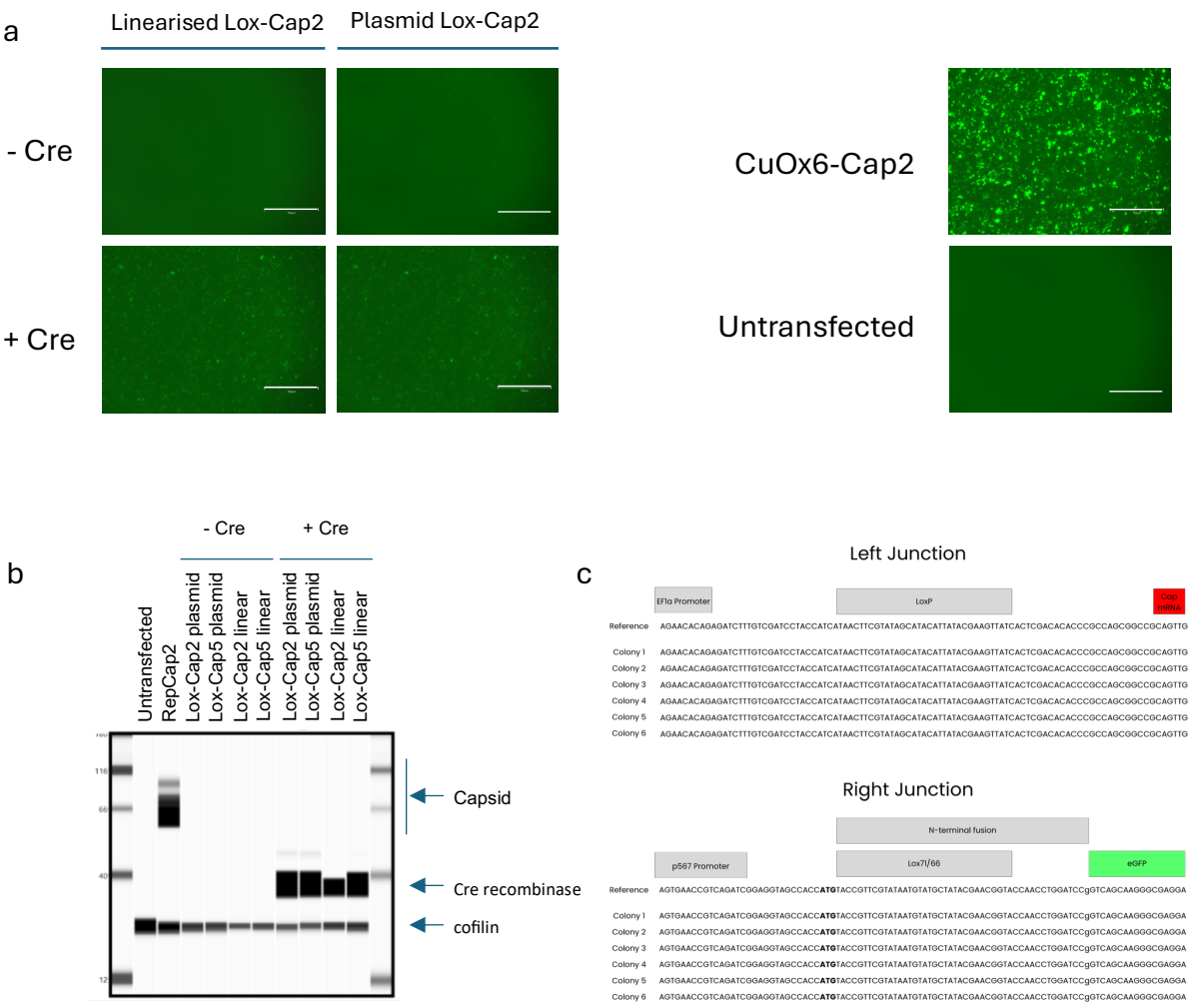


Figure 13 – GFP and Capsid expression following recombination of linearised Cre-inducible library plasmid to test for aberrant recombination species. *A) representative images of HEK293 cells transfected with linearised or intact AAV2 capsid-encoding Cre-inducible library plasmid DNA, in the presence or absence of a Cre expression plasmid. Scale bar, 750 µm B) Detection, by western blot, of AAV capsid protein in HEK293 cells transfected with linearised or intact Cre-inducible library plasmid DNA, with or without a Cre expression plasmid, and transduced with TESSA-Rep. RepCap2 is a trans AAV capsid expression plasmid using for helper-free AAV production used as a positive control for capsid detection. C) Alignment of sanger sequencing reads from plasmid colonies, isolated from HEK293 cells co-transfected with Cre-inducible library plasmid and Cre expression plasmid, to the predicted left and right junctions following Cre recombination.*

To investigate more directly the low eGFP expression and lack of capsid expression in the presence of Cre, a plasmid was produced with the Cre-inducible system in the “on state” (Figure 12a). Capsid expression was detected from the transiently transfected “on state” plasmid by western blot, but interestingly was undetectable in the presence of Cre (Figure 14). Similarly, during AAV replication in the presence of TESSA-Rep, capsid expression from the “on state” plasmid was present in the absence of Cre but lost in cells expressing Cre (Figure 14c). Taken together, these observations suggest that Cre recombinase activity or Cre binding significantly interferes with capsid expression from the recombined library plasmid.

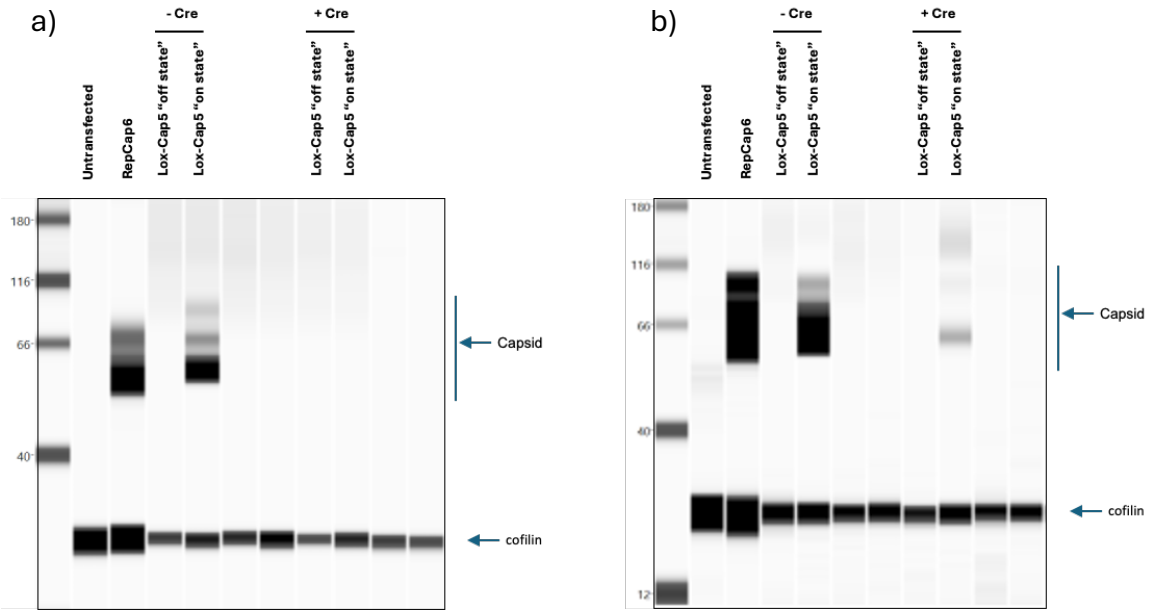


Figure 14 – GFP and Capsid expression from an “on state” Cre-inducible library plasmid in the presence and absence of Cre recombinase. *Detection, by western blot, of AAV capsid protein in HEK293 cells transfected with Cre-inducible library plasmid DNA in the “on state” or “off state” configuration, with or without a Cre expression plasmid, without (b) or with (c) transduction with TESSA-Rep.*

To examine the mechanism behind the repression of capsid expression in the presence of Cre, a mutant version of Cre (K201A) was used that is capable of binding to LoxP sites but lacks recombinase activity (Ghosh et al., 2007). HEK293 cells were co-transfected with the Lox-Cap5 plasmid in the “on state” or “off state”, with either stuffer DNA or a plasmid encoding the wild-type or mutant Cre. As expected, in the presence of stuffer DNA, very few cells were GFP positive in cells transfected with the “off state” plasmid and up to 80% of cells were GFP positive in cells transfected with the “on state” plasmid. In the presence of wild-type Cre, the frequency of GFP positive cells increased to approximately 55% in cells transfected with the “off state” plasmid due to recombination to the “on state”. GFP frequency remained very low in cells expressing the mutant Cre and transfected with the “off state” plasmid, confirming that the mutant Cre is unable to catalyse recombination. The frequency of GFP positive cells transfected with the “on state” plasmid was unaffected by the presence of wild-type or mutant Cre (Figure 15a). Capsid expression in TESSA-Rep transduced cells, co-transfected with Lox-Cap5 “on state” plasmid and stuffer DNA, was detectable by western blot but lost in the presence of either wild-type or mutant Cre (Figure 15b). These results suggest that that catalytic activity of Cre recombinase does not cause the observed repression of capsid expression. Toxicity related to Cre expression is largely dependent on the recombinase activity of Cre (Loonstra et al., 2001), suggesting that Cre related toxicity is unlikely to explain reduced transgene expression. As the K201A mutant Cre retains the ability to bind the high affinity LoxP site, it is possible that bound Cre sterically inhibits transcription of the capsid.

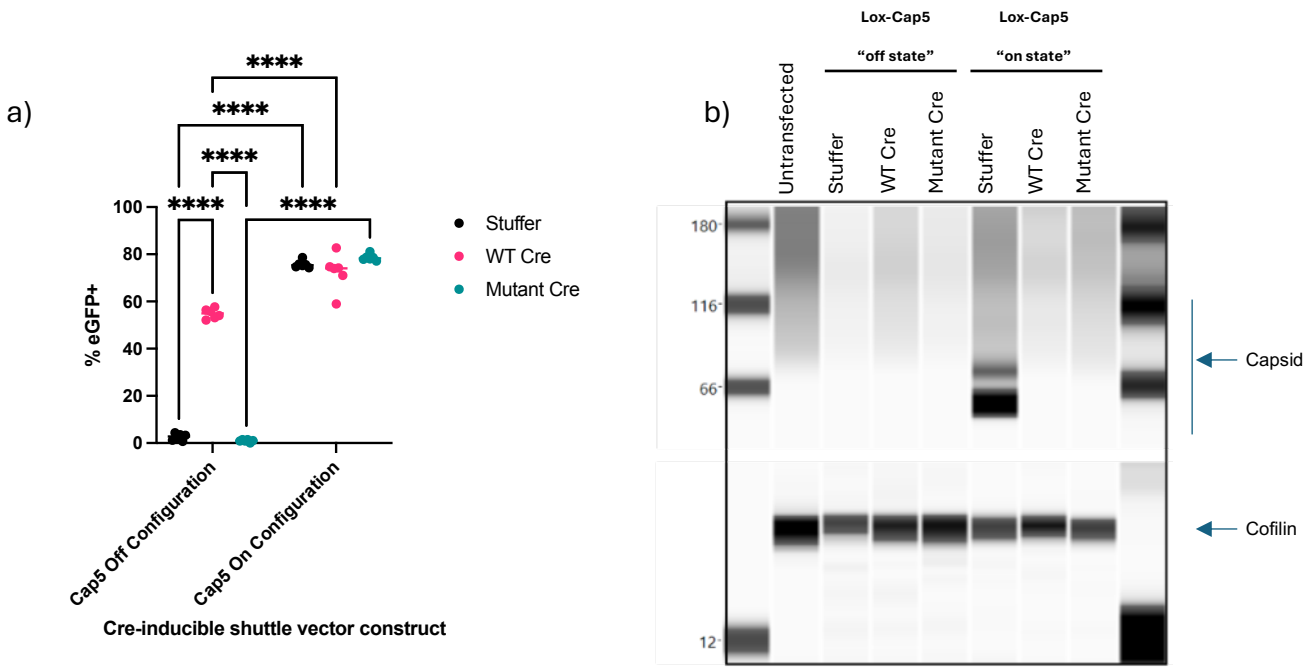


Figure 15 – GFP and Capsid expression from a "on state" and "off state" Cre-inducible library plasmids in the presence of wild-type or K201A mutant Cre recombinase. *A) Frequency of eGFP positive HEK293 cells transfected with AAV5 capsid-encoding Cre-inducible library plasmid DNA in the "on state" or "off state" configuration, in the presence of wild-type or K201A mutant Cre expression plasmid, or stuffer DNA. B) Detection, by western blot, of AAV capsid protein in TESSA-Rep transduced HEK293 cells transfected with Cre-inducible library plasmid DNA, with wild-type or mutant Cre expression plasmid. Each data point represents an independent experiment. Statistical analysis performed using a two-way ANOVA with correction for multiple comparisons using Tukey's test. Comparisons were limited to within shuttle vector construct group or within Cre condition contrasts. ****P ≤ 0.0001 , ***P ≤ 0.001, **P ≤ 0.01, *P ≤ 0.05.*

Taken together, these experiments demonstrate the that the Cre–lox recombination system configuration enables recombination between states that can be monitored using GFP reporter expression. The system provides stringent repression of capsid expression in the absence of Cre, but the presence of Cre during transient transfection experiments, or during shuttle vector production, inhibits expression of the genome-

encoded capsid independently of Cre recombinase activity. Shuttle vectors produced using this system are capable of low efficiency recombination in HEK293 cells transfected with a Cre-expression plasmid.

8.4 Discussion

This chapter set out to develop a two-step AAV library production system in which genome–capsid linkage is preserved without introducing the biases reported for earlier shuttle vector systems (Müller et al., 2003). The central approach was to place genome-encoded capsid expression under inducible control, such that capsid expression remained repressed during shuttle vector production and became active only during repackaging. We found that genome-encoded capsid was expressed at high levels during shuttle vector production in the absence of the transactivator protein. Furthermore, we observed that the performance of the shuttle vectors was impaired compared to control virus and that this deficit appeared to be dependent on the inclusion of a coding sequence in addition to the GFP gene. This effect was not due to transcriptional readthrough from the GFP-upstream gene.

The CymR-VP16-inducible system was selected on account of its reported low basal activity and its use of regulatory components distinct from those of the tetracycline-inducible system employed in TESSA, thereby minimising the likelihood of crosstalk. In agreement with the low basal expression of the CymR-VP16 system described during the original characterisation (Mullick et al., 2006), we were unable to detect leakage of capsid expression in cells transiently transfected with the library plasmid in the absence of CymR-VP16. However, Mullick *et al* detected very low level β-galactosidase activity in the absence of transactivator, indicating that the minimal CMV promoter exhibits low but measurable basal activity (Mullick et al., 2006). Our results suggest that the basal expression from the minimal CMV promoter is sufficient to drive the accumulation of

significant levels of genome-encoded capsid protein during AAV production. However, we cannot exclude the possibility that adenoviral helper genes contribute to the increase of capsid proteins to detectable levels, for example through enhanced cytoplasmic accumulation of capsid mRNA supported by adenoviral E4orf6 (Samulski and Shenk, 1988). We demonstrated that intrinsic ITR promoter activity alone was sufficient to drive capsid expression and generate productive AAV particles in the absence of an upstream promoter. However, during the production of shuttle vectors, the level of protein capsid was lower for the promoter-deleted vectors, suggesting that the basal activity of the inducible promoter has an additive effect of capsid expression. The contribution of ITR-driven transcription remains an important consideration for alternative inducible library designs. The use of alternative ITRs with lower intrinsic promoter activity, such as those from AAV1 and AAV6 (Earley et al., 2020), may help to minimise this effect. The inclusion of a repressor gene, such as CymR, within the AAV genome may limit capsid expression during production by coupling repressor abundance to genome amplification, reducing the risk of repressor saturation. However, this would exceed the packaging limit for AAV given our desire to include a fluorescent reporter to aid with library screening.

AAV2 shuttle vectors produced using the CymR-VP16 system also displayed poor apparent transduction efficiency measured using a GFP-reporter assay. We initially suspected that the poor shuttle vector performance may be related to the inappropriate expression of the genome-encoded capsid. However, shRNAs designed to specifically knock-down the genome-encoded capsid failed to rescue vector performance. Notably, shuttle vectors encoding AAV5 or AAV8 capsids exhibited transduction efficiencies comparable to those encoding AAV2. As AAV5, and to a lesser extent AAV8, transduce

HEK293 cells inefficiently (Ellis et al., 2013), substantial expression of these capsids during shuttle vector production would be expected to diminish the relative proportion of highly efficient AAV2 particles, thereby impairing shuttle vector performance for these serotypes. Together, this suggests that the leakage of the genome-encoded capsid and the poor shuttle vector performance are distinct challenges for the CymR-VP16-inducible system. We next sought to systematically delete sequences from the library construct to isolate the cause of the low shuttle vector transduction efficiency. Removal of the capsid sequence, but not the CuOx6-minimal CMV promoter, rescued the apparent transduction efficiency. We found that this effect was independent of the capsid sequence, as replacement with a FLuc reporter produced vectors with similarly poor performance. Improved transduction efficiency following removal of the large capsid sequence could conceivably reflect more efficient packaging of full-length genomes using the smaller construct, reducing the representation of GFP genes. However, we measured the viral titre using GFP primers and found no difference in copy number between vectors produced with the CuOx6-Cap2 and CMV-eGFP genomes (Figure 6c), and our 4.5 kb CymR-VP16-inducible construct falls well within the expected limit for efficient packaging of full-length AAV genomes (Kosaka et al., 2025).

One possible explanation for the reduced apparent transduction efficiency is transcriptional interference (TI) arising from an upstream gene lacking a transcriptional termination sequence, thereby suppressing expression of the GFP gene. TI describes the cis-acting repression of one promoter by the transcriptional activity of another proximal promoter. TI from tandem promoters has been studied in detail in *E. coli* (Chauhan et al., 2024) and has been reported in *Drosophila melanogaster* (Petruk et al., 2006) and

Saccharomyces cerevisiae (Martens et al., 2004). In mammalian cells, direct evidence is limited, although it has been reported for synthetic constructs where termination sequences relieve transcriptional repression from tandem promoters (Proudfoot, 1986). The mechanism in eukaryotic cells remains undefined and is likely context dependent. Notably, enhanced HIV-1 gene expression has been observed when the integrated genome is orientated in the same direction as a neighbouring gene (Han et al., 2008), possibly reflecting cooperative effects of trailing polymerases in releasing stalled polymerases (Saeki and Svejstrup, 2009). The inclusion of polyA sequences upstream of the GFP promoter, or reversal of the capsid gene orientation, did not rescue AAV2 shuttle vector performance, suggesting that TI is unlikely to account for the low apparent transduction efficiency. Furthermore, it is improbable that the very low basal activity of the minimal CMV promoter could support sufficient transcription in transduced cells to significantly hinder transcription from the p567 promoter.

Since this work, additional studies have examined the transcriptional activities of the AAV ITRs that may help explain these findings. As noted, the promoter activity of the AAV ITRs is well established (Flotte et al., 1993, Earley et al., 2020). Reports of enhancer activity also exist, although these observations were made in the context of Rep-mediated transactivation of AAV promoters in the wild-type AAV genome (Mccarty et al., 1991, Beaton et al., 1989b). Recently, evidence for the enhancer activity of the AAV ITRs in the context of recombinant AAV has emerged through the development of micro-promoters which lack intrinsic enhancer activity but support impressive transgene expression when placed near the AAV ITR. Interestingly, gene expression from these micro-promoters was equivalent to the chicken beta-actin (CAG) promoter strength when delivered via AAV, but

was inferior in plasmid form (Chai et al., 2023). In the light of these findings, the increase in the apparent transduction efficiency of our shuttle vectors when large sequences between the left ITR and p567 promoter were deleted may be due to the closer proximity between the ITR enhancer elements and the p567 promoter, boosting GFP expression and therefore apparent transduction efficiency. In this case, the poor shuttle vector performance may be due to the inherently weak promoter activity of p567, and the performance of the p567-eGFP control is misleading due to the enhanced GFP expression beyond the inherent capability of the p567 promoter. The difference in enhancer activity between AAV genomic DNA and AAV plasmid DNA observed by Chai *et al* may explain why we did not see a difference in GFP expression between the p567-eGFP control and CuOx6-Cap5 plasmid in transient transfection experiments (Figure 7). Furthermore, the enhancer activity of the ITR may contribute significantly to the robust capsid expression observed during shuttle vector production in the absence of cymR-VP16. The inclusion of insulator sequences inside of each AAV ITR may help to block ITR enhancer activity, an approach that has been explored to limit the ITR-driven expression of bacterial proteins from the reverse packaging of AAV plasmids (Taylor et al., 2024).

To address the limitations of the CymR-VP16-inducible system, we designed and tested a Cre-inducible system. This system was capable of strict repression of capsid expression during shuttle vector production but failed to express capsid protein in the presence of Cre due to an unresolved mechanism independent of the catalytic activity of Cre. Furthermore, we found that the AAV2 shuttle vector recombination efficiency appeared low in HEK293 cells expressing high levels of Cre recombinase.

During initial characterisation experiments using the Cre-inducible system, we noted that the viral shuttle vector recombination efficiency appeared low when measuring the frequency of GFP-positive transduced HEK293 cells. The mutant lox66/71 are known to recombine less efficiently than wild-type loxP sequences (Oberdoerffer et al., 2003). The suboptimal recombination efficiency of the mutant lox sites, together with the relatively weak activity of the p567 promoter, likely contributed to the low frequency of transduced GFP-positive cells. However, because the two-step library design relies on repackaging at a low genome copy number per cell, we anticipated that limited recombination efficiency would not constitute a major barrier. In principle, the amount of virus used during packaging could be titrated to balance the number of "on-state" genomes per cell with overall viral yield. Therefore, we focussed on the failure to express capsids following recombination.

The lack of capsid expression from both plasmid DNA in the presence of Cre recombinase suggested an inherent flaw in the vector design that enabled low level GFP expression but blocked capsid expression. The absence of capsid expression in the presence of TESSARep is indicative of a complete block to expression rather than a reduction in efficiency, as even minimally expressed capsids would be expected to accumulate during AAV genome amplification. Therefore, we suspected that recombination resulted in unintended sequences that prevented capsid expression. For instance, if the lox sites were inadvertently arranged in tandem orientation, recombination would excise the capsid gene, leaving the GFP reporter under the control of the minimal EF1α promoter, which may be weaker than the intended p567 promoter. We isolated DNA from cells co-transfected with the Cre-inducible shuttle vector plasmid

and a Cre-expression plasmid, and sequenced clonal plasmids purified from transformed *E. coli*. The junctions for all sequenced plasmids matched the predicted sequences, suggesting that aberrant recombination is unlikely to account for our observations.

In light of the sequencing results, we confirmed that the predicted recombined sequence could express capsid in the absence of Cre. Interestingly, we observed that capsid expression was abolished in the presence of wild-type Cre or a catalytically dead mutant, suggesting that the presence of Cre interferes with capsid expression independently of recombinase activity. Cre recombinase is widely used to generate conditional mouse models and in lineage-tracing studies, typically using Cre mouse lines that express Cre from low copy genes integrated into genomic sites (Liu et al., 2020, Kim et al., 2018). Cre recombinase mouse lines have been extensively used in neuroscience with AAV vectors encoding Flip-excision (FLEX) designs to enable Cre-dependent tissue-specific transgene expression (Matsushita et al., 2023). The FLEX system uses a configuration similar to our Cre-inducible construct but uses two pairs of heterotypic lox sites either side of the flipped sequence. Sequential recombination events result in a reoriented sequence flanked by incompatible lox2272 and loxP sites. The widespread adoption of these vectors illustrates that Cre-mediated inversion is a valid strategy for controlling gene expression from AAV vectors. However, the level of cellular Cre is likely much lower in *in vivo* applications than in the transient transfection setting used in this chapter and it is possible that very high levels of Cre interfere with transgene expression as our data suggest. The toxic effect of Cre recombinase appear to depend on the catalytic activity of Cre (Loonstra et al., 2001) and the repression of capsid expression in the presence of

the catalytically dead Cre suggests that this effect was not due to toxicity, however, we cannot exclude that especially high levels of catalytically dead Cre are toxic. Furthermore, Cre expression in HEK293 cells appeared to reduce the transduction efficiency of a CMV-eGFP control vector (Figure 12c), suggesting that high levels of Cre may interfere with AAV transduction, possibly through elevated DNA damage responses (Loonstra et al., 2001, Schwartz et al., 2007). It is possible that the use of a single copy number Cre-expressing HEK293 cell line would support sufficient recombination rates for library production without capsid repression caused by excess Cre.

Taken together, our results demonstrate the challenges in restraining capsid expression during viral production and recent publications support the notion that inherent features of the AAV ITRs may enhance basal expression from a minimal promoter. We also showed that a Cre recombinase-based system strictly restrains capsid expression during shuttle vector production but is susceptible to gene repression in the in the presence of Cre through an unknown mechanism independent of recombinase activity. Although the one-step library production method has been widely adopted, recent comparisons between the one-step and two-step approaches (Stamataki et al., 2025) highlights a continued interest in the two-step approach. The use of a low copy number Cre-expressing AAV producer cell line, in combination with a strong promoter to drive GFP expression, may overcome the major challenges posed by the Cre-inducible library design. A comparison between inducible two-step and conventional one-step libraries in recovering rare sequences from a mock library could clarify the degree to which the two-step design reduces noise during screening.

9 Production and Screening of AAV Capsid Libraries Tailored Towards Primary human B Cells

9.1 Introduction

The previous chapter described the potential improvements in cross-packaging and mosaicism of a two-step strategy over typical one-step AAV capsid library production methods. However, the two-step approach requires significant development given the challenge of restraining the expression of an AAV genome-encoded capsid. To enable the identification of novel capsids capable of efficient B cell transduction, this chapter focusses on the development of an AAV capsid library using the established one-step production approach.

The rationale behind AAV capsid library design should be guided by the goal of the capsid screen. For example, if selectivity in a mixture of permissive cell types is the goal, then the capsid library design may focus on the retargeting of residues involved in cell binding. For cell types that are challenging to transduce, such as B cells, an understanding of the bottlenecks in transduction should be used to guide capsid library design.

Natural serotypes such as AAV2 and AAV6 can transduce primary human B cells, but overall transduction efficiencies remain low and are highly dependent on activation state (Theiss et al., 2003, Ahmadi et al., 2021). In a comparison of natural serotypes 1-7, clade A (AAV1 and AAV6) and clade B (AAV2) serotypes performed significantly better than other

serotypes in EBV infected B-cells (Ahmadi et al., 2021), and AAV6 was found to be the most efficient serotype in stimulated primary human B cells (Hung et al., 2018). However, a single-stranded AAV (ssAAV) dose escalation study in Raji cells, an EBV-positive Burkitt's lymphoma cell line, showed that a multiplicity-of-infection between $1x10^5$-$1x10^6$ was required to achieve moderate (25-60 %) transduction (Rogers et al., 2021).

The basis for the low permissiveness of EBV-negative primary human B cells to transduction by AAV is poorly understood and the transduction pathway could be restricted at any of several critical steps. One significant observation is that B cell activation status appears to have a striking effect on AAV transduction efficiency. While transduction of unstimulated primary human B cells at a multiplicity-of-infection (MOI) of $1x10^5$ appeared negligible (Ahmadi et al., 2021), Hung *et al* achieved up to 40 % transduction efficiency using an MOI of $2.4x10^4$ following a 48 hour pre-stimulation period using CD40L, CpG oligonucleotides and a cytokine cocktail (Hung et al., 2018), both using GFP-encoding self-complementary AAVs (scAAVs). However, in the context of gene editing in stimulated and electroporated primary human B cells using ssAAVs, MOIs of $1x10^5$-$1x10^6$ are typically required (Hartweger et al., 2019, Moffett et al., 2019). The difference in transduction efficiency between scAAV and ssAAV suggests that second-strand synthesis contributes to the inefficiency of ssAAV in B cells. However, as second-strand synthesis is broadly limiting even in permissive cell lines (McCarty et al., 2001, Ferrari et al., 1996), it is unlikely to be the principal factor accounting for the markedly lower permissiveness of primary human B cells. Instead, other binding or post-entry barriers in B cells are more probable explanations. Since stimulation is required for both

the *in vitro* culture of primary B cells and to support homology-directed repair (Lin et al., 2014), this chapter focusses on ssAAV transduction in stimulated primary human B cells.

Endosomal escape is another critical step in AAV transduction that could be impaired in primary human B cells. The failure of AAV to accumulate in the Golgi of non-permissive cell lines including NIH3T3 cells (Nonnenmacher et al., 2015c) suggests endosomal trafficking can be a major bottleneck in some cell types. Professional antigen-presenting cells, including B cells, have specialised endosomal environments that may present unique challenges for AAV. However, the effect of *in vitro* stimulating conditions on the endosomal environment in B cells is poorly resolved. B cells internalise BCR-bound antigen through clathrin-mediated endocytosis and traffic the antigen to MHC class II peptide-loading compartments that are enriched in proteolytic enzymes (Adler et al., 2017). Interestingly, it has been shown that CD40L-stimulated human B cells are capable of presenting HBV core peptide antigen (Wu et al., 2010). This suggests that CD40L stimulation can support trafficking of soluble proteins into specialised endosomal compartments without BCR stimulation. If this process is efficient, this may in part explain the limited transduction efficiency of AAV in *in vitro* stimulated B cells.

Finally, nuclear import, uncoating and second-strand synthesis may be impaired. In dendritic cells (DCs), post-entry steps in AAV transduction appear to be compromised, however a peptide display library yielded a capsid with improved transduction efficiency. Increased levels of concatemerised AAV genome was detected using the isolated peptide insertion variant, suggesting that vector uncoating limits AAV transduction in

DCs (Rossi et al., 2019). It is possible that B cells display a similar bottleneck to antigen-presenting DCs.

Although identifying biological restrictions is important for guiding library design, library production also poses technical challenges that are independent of the target cell type and may influence the likelihood of identifying novel capsids. Typical library designs create enormous theoretical sequence spaces. For example, a relatively simple 7-mer peptide insertion library has 20^7 (1.28×10^9) amino acid combinations. Therefore, diversity should be prioritised to maximise the sequence space that is sampled by the library. During the production of the viral library, a large proportion of variants are lost because they fail to assemble (Bryant et al., 2021). Because the capsid sequence encodes non-functional proteins that are important for AAV assembly and egress, it is possible that loss-of-function mutations occur in accessory proteins in sequences that are otherwise functional. The assembly-activating protein (AAP) is encoded within an out-of-frame coding sequence in the capsid gene and acts as a chaperone that shuttles VP proteins to the nucleus and nucleates capsid assembly. The requirement for AAP varies by serotype and it appears to have evolved to allow more flexibility in capsid sequence by using AAP to compensate for reduce stability (Maurer et al., 2018). Trans-complementation assays in HEK293 cells show that supplying wild-type AAP can restore capsid production from an AAP-knockout cap gene (Sonntag et al., 2011). It is possible that providing AAP in trans during AAV library production may rescue sequences that would otherwise fail to assemble stable particles.

In this chapter we aim to establish whether cell entry is the major bottleneck in AAV transduction in human B cells. Using this knowledge, we aim to design and produce an appropriate diverse AAV capsid library, which will then be used to screen for novel AAV capsids in primary human B cells. Additionally, we aim to investigate the effect of trans-complementation of AAP2 (AAV2 AAP) on the production of complex AAV libraries and to assess the extent of cross-packaging using the chosen production parameters. Finally, we aim to perform a screen in stimulated primary human B cells with the objective of identifying capsid variants with superior transduction efficiency.

9.2 Aims

1. Investigate whether cell binding represents a major barrier to transduction using AAV6.
2. Design and construct AAV capsid libraries tailored towards our understanding of B cell transduction bottlenecks.
3. Evaluate the effect of AAV helper plasmid design on the production of our viral AAV capsid libraries.
4. Measure the potential for cross-packaging using the AAV library plasmid designs and production parameters.
5. Screen the viral capsid libraries using stimulated primary human B cells to identify novel capsid variants with improved transduction efficiency.

9.3 Results

9.3.1 Investigation into the bottlenecks impairing B Cell transduction by AAV.

To aid in the design of an AAV capsid library tailored towards primary human B cells, initial experiments sought to understand the degree to which cell binding is a major barrier to transduction.

The dynamics of AAV6 binding and internalisation were assessed in stimulated primary human B cells and 293-AAV cells, a HEK293 derived cell line with improved AAV production properties, using a qPCR-based binding/internalisation assay (Figure 16a). HEK293, and derived cell lines, are permissive to AAV6 transduction and act here as a reference for qualifying receptor expression and entry pathways. In this assay, the cells were pre-cooled to 4ºC to arrest endocytosis, allowing virus binding but not internalisation. Cells were incubated at 4ºC for 1 hour with AAV6 virus encoding eGFP driven by a spleen focus-forming virus (SFFV) promoter. Unbound AAV was removed through washes with PBS. Binding-only samples were immediately processed to isolate AAV and cellular DNA using an extended protease treatment to degrade AAV capsids. For internalisation samples, the cells with bound virus were incubated at 37ºC for 1 hour to allow cell entry. Bound, but uninternalized, virus was then removed by treatment with trypsin and further washes with PBS. AAV and cellular DNA was isolated as described for the binding-only samples. The number of AAV particles bound or internalised in each condition was measured by quantification of GFP gene copy numbers in the DNA sample,

normalised to GAPDH gene copy numbers. Importantly, 293-AAV and B cells differ considerably in size and therefore direct comparisons in absolute binding or internalisation values should not be made between the two cell types because the surface area is likely a confounding variable in AAV binding. Furthermore, HEK293 cells possess abnormal and unstable karyotypes and GAPDH copies per cell will likely differ between 293-AAV cells and primary human B cells. However, the comparison of the ratio of binding to internalisation between cell types is still valid. As expected, a significantly smaller number of AAV genomes were internalised compared with the number bound to the cell for both cell types (Figure 16b), confirming that only a subset of bound particles undergo entry. The ratio of bound to internalised AAV was smaller in the primary B cells (1.50) compared to the 293-AAV sample (2.78; Figure 16c), suggesting that the efficiency of internalisation is lower in B cells, or that internalised genomes are more rapidly degraded in B cells compared to 293-AAV cells. Although this may in part explain why transduction efficiency in primary human B cells is lower than in permissive cell lines, the less than two-fold reduction in entry efficiency is unable to account for the vast difference in MOIs required to achieve detectable levels of transduction in stimulated primary human B cells compared to HEK293 cells.

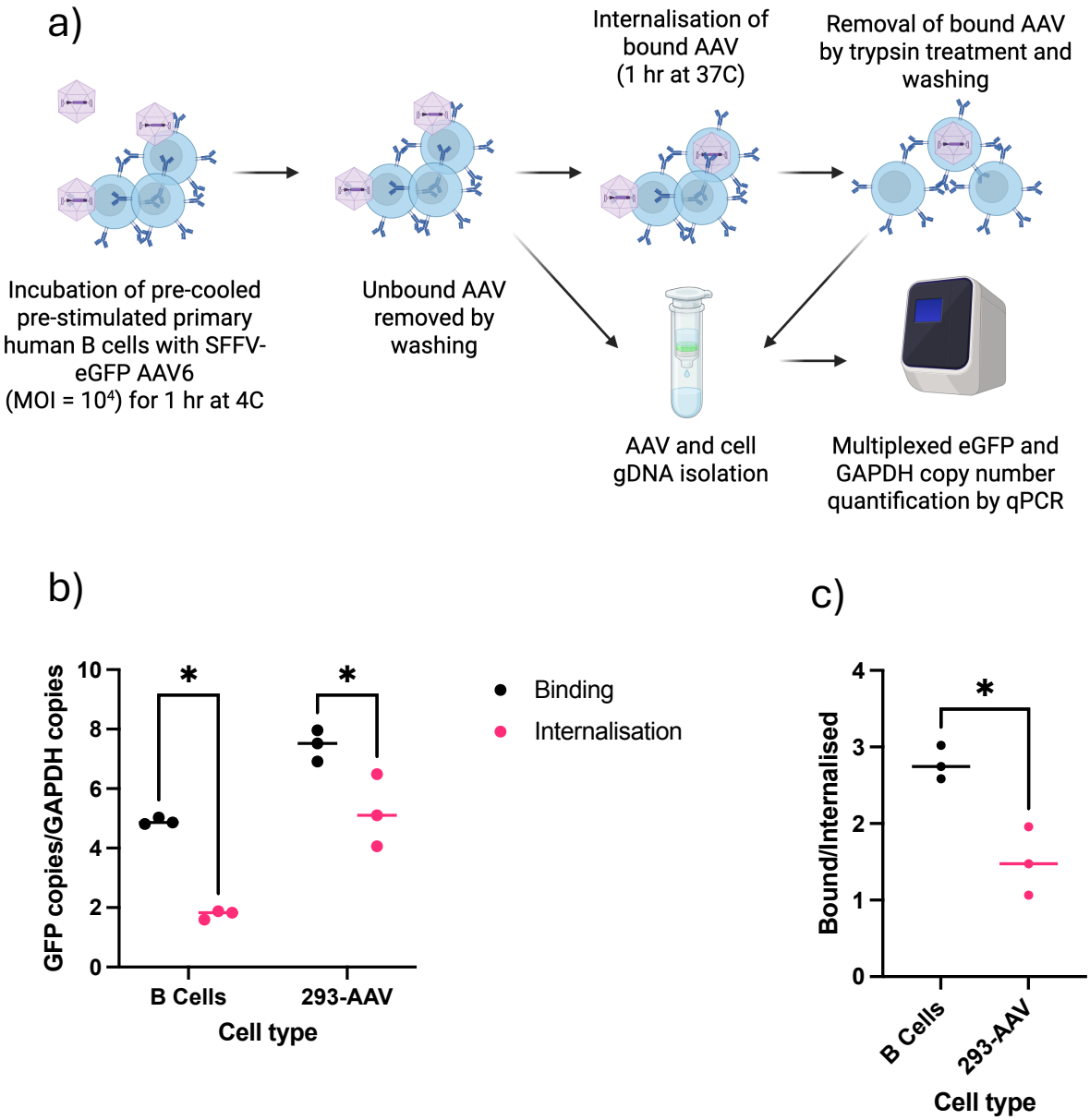
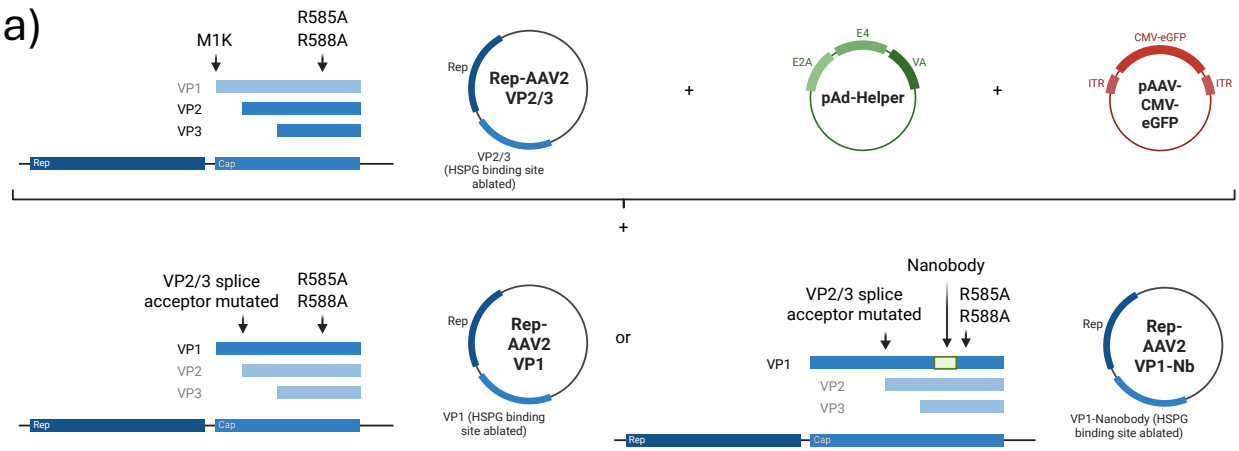
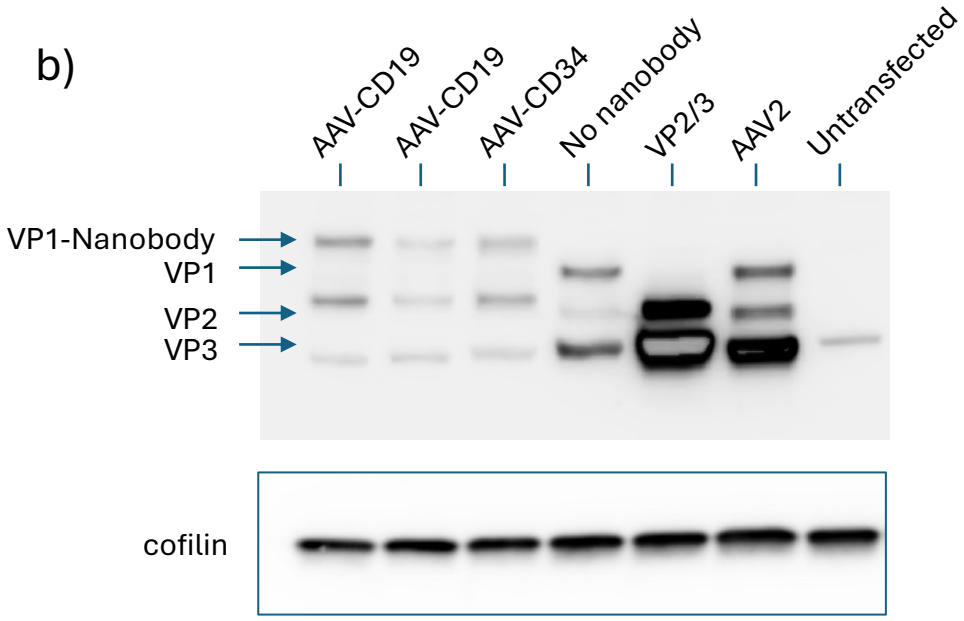


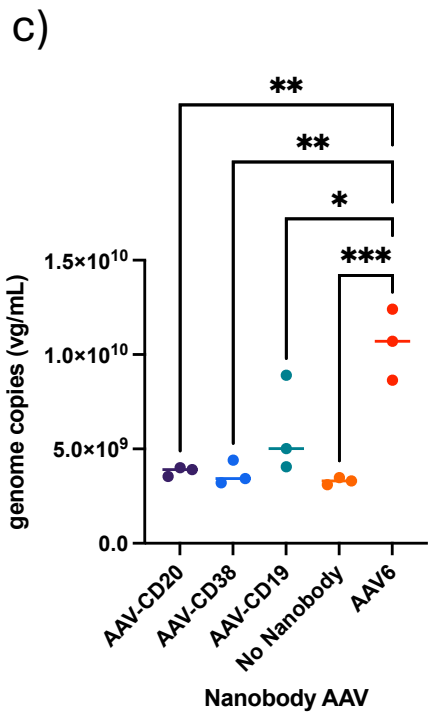
Figure 16 – Binding and internalisation kinetics of AAV6 in stimulated primary human B cells, from a single donor, and HEK293-derived cells. *A) Schematic of Binding/Internalisation assay methodology. Created with BioRender.com B) The binding/internalisation dynamics of SFFV-eGFP AAV6 was assessed using the methodology depicted in A) for stimulated primary human B cells and 293-AAV cells. The absolute number of GFP gene copies, normalised to GAPDH gene copies, is plotted in B) and the ratio of normalised GFP copies in the binding samples to the internalisation samples is plotted in C). Each condition was performed with 3 replicates from a single donor. Statistical significance was assessed using multiple unpaired t-tests corrected for multiple comparisons using Holm-Sidak with the alpha threshold set to 0.05 (B) or by unpaired t-test (C). *P ≤ 0.05.*

The binding/internalisation assay offered some insight into the entry kinetics of AAV6 in primary human B cells but, due to the experimental design, it was unable to resolve deficiencies in AAV binding to B cells. To explore this, a previously described AAV-nanobody display system (Eichhoff et al., 2019) was used to target AAV to B cell receptors. To achieve this, the four plasmid AAV-nanobody production system (Figure 17a) was reproduced by mutating the start codon of VP1, in an AAV2 Rep-Cap helper plasmid, to produce a VP2/3 expression plasmid. Additionally, a VP1-only expression plasmid was produced by mutating the VP2/3 splice acceptor so that all capsid transcripts from this plasmid use the VP1 splice acceptor, preventing the expression of VP2 and VP3. In all plasmids the heparan sulfate proteoglycan (HSPG) binding site was mutated (R585A and R588A) to blind the AAV2 capsid to the native cell attachment factor used for cell binding. Anti-CD19, anti-CD20 and anti-CD38 nanobodies sequences, flanked by $(GGGS)_5$ sequences, were inserted in the variable region IV of the AAV2 capsid by replacing residues 453-459. This places the nanobody in an accessible position at the apex of the 3-fold spike. By placing the nanobody into VP1 only, the number of nanobodies per particle is limited, due to the stoichiometry of capsid assembly, to 5 nanobodies per AAV particle, assuming normal assembly stoichiometry is maintained. To produce the AAV-nanobody particles, HEK293 cells were transfected with equimolar amounts of VP1 expression plasmid (with or without nanobody insertion), VP2/3 expression plasmid, an adenoviral helper plasmid and a CMV-eGFP AAV genome plasmid. The nanobody integration was confirmed by anti-capsid western blot of 293-AAV cell lysates following transfection with nanobody display production system plasmids (Figure 17b) and the titre was measured by absolute quantification of genome copies by GFP-specific qPCR. Nanobody integration did not impair AAV production when

using the four-plasmid system (Figure 17c). To see if re-targeting of AAV2-derived nanobody-displaying AAV to human B cells enhances transduction, RAMOS cells, an EBV-negative human B cell line derived from Burkitt's lymphoma patient cells, were transduced with 10^3 nanobody AAVs, or control virus (Figure 17d). 72 hours post-transduction, no eGFP expression was observed by flow cytometry for the no-nanobody control virus or the AAV particles displaying anti-CD20 or anti-CD19 nanobodies. However, display of the anti-CD38 nanobody enabled the transduction of RAMOS cells to AAV6 control levels. As the transduction efficiency of the anti-CD38 nanobody displaying AAV did not exceed the transduction efficiency of AAV6, these results suggest that the low transduction efficiency achieved by AAV6 is limited by post-binding steps in cell transduction.


a)
M1K
R585A
R588A
VP1
VP2
VP3
Rep
Cap
Rep-AAV2 VP2/3
VP2/3 (HSPG binding site ablated)
E4
E2A
VA
pAd-Helper
CMV-eGFP
ITR
pAAV-CMV-eGFP
VP2/3 splice acceptor mutated
Rep-AAV2 VP1
VP1 (HSPG binding site ablated)
or
Nanobody
Rep-AAV2 VP1-Nb
VP1-Nanobody (HSPG binding site ablated)


b)
AAV-CD19
AAV-CD19
AAV-CD34
No nanobody
VP2/3
AAV2
Untransfected
VP1-Nanobody
VP1
VP2
VP3
cofilin


c)
genome copies (vg/mL)
1.5×10^10
1.0×10^10
5.0×10^9
0.0
AAV-CD20
AAV-CD38
AAV-CD19
No Nanobody
AAV6
Nanobody AAV

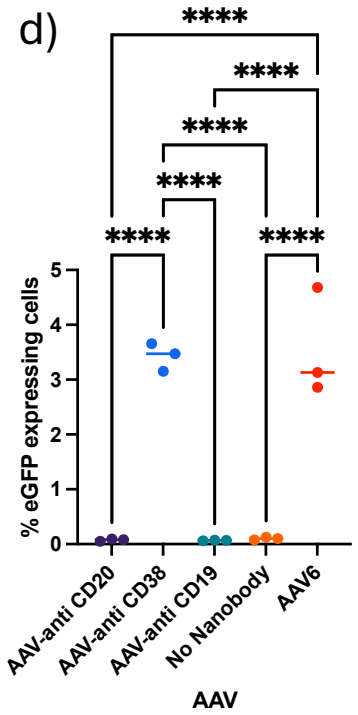

d)
% eGFP expressing cells
AAV-anti CD20
AAV-anti CD38
AAV-anti CD19
No Nanobody
AAV6
AAV

Figure 17 – Anti-CD38 nanobody-displaying AAV2 enables transduction in RAMOS cells to AAV6 levels. *A) Schematic representation of four plasmid AAV nanobody display production system. The AAV nanobody display production system was used to produce AAVs displaying anti-CD19, anti-CD20 or anti-CD38 nanobodies. As a control for transduction using the best available natural serotype, AAV6 was also produced using a typical three plasmid production system. B) Anti-capsid western blot on HEK293 lysates following transfection with AAV nanobody display production plasmids. C) Genome titres (vg/mL) for crude nanobody displaying AAVs measured by qPCR using GFP primers. D) Frequency of eGFP-expressing RAMOS cells transduced with nanobody displaying AAVs at an MOI of 10^3 72 hours post-transduction, measured by flow cytometry. Each point represents a separate AAV production (C) or a biological replicate (d). Statistical significance was assessed using one-way ANOVAs with correction for multiple post-hoc comparisons using Tukey's test. ****P ≤ 0.0001 , ***P ≤ 0.001, **P ≤ 0.01, *P ≤ 0.05.*

9.3.2 Cloning of a shuffled AAV capsid library.

The binding/internalisation and nanobody data together suggest that binding is not a major barrier to human B cell transduction and that, although entry appears impaired compared to permissive HEK293 cells, reduced entry kinetics alone is unlikely to account for the need for extremely high MOIs to achieve transduction in human B cells. Given this data, a peptide display library design, which usually focusses on viral attachment, is unlikely to yield a capsid with improved transduction efficiency over natural serotypes. Therefore, strategies that create diversity across the entire capsid length, and therefore are more likely to change the characteristics of the viral variants during post-attachment steps, are more suitable.

The sequence diversity of capsid libraries is restrained by practical limitations during plasmid library production. Therefore, it is important that the library design maximises functional diversity. Shuffled capsid libraries use a homology-driven shuffling process to

create chimeric capsid sequences from multiple natural serotypes. This process creates novel capsid sequences in which different serotype characteristics are grafted together into new combinations. Although this approach appears limiting due to the use of naturally occurring sequences, the shuffling of sequences with divergent characteristics is more likely to yield high functional diversity compared random mutagenesis across the entire capsid sequence, which would yield many variants that are functionally identical.

To produce a shuffled capsid library, a capsid library vector (CLV) plasmid was designed (Figure 18a) to allow for constitutive capsid expression in primary human B cells during screening to enable the rescue of capsid variant sequences from RNA transcripts. This design avoids the recovery of capsid sequences from AAV variants that fail to complete the transduction pathway, such as variants that bind but fail to enter the cell, or variants that fail endosomal escape and nuclear entry. The capsid variants are expressed from an SFFV promoter, chosen for its durable expression in hematopoietic lineage cells, and terminated by an SV40 polyA sequence. NotI (5') and NsiI (3') sites were placed either side of the capsid sequence for convenient cloning of capsids into the plasmid. For the shuffling of capsid sequences (Figure 18b), natural serotypes from AAV1-9 were subcloned into the CLV plasmid to add 3' NsiI sites, and amplified by PCR using Subclone Fw and Subclone Rv primers, adding 70-90 bp buffer regions to ensure restriction site integrity is maintained during capsid shuffling. The amplified capsid sequences were pooled at an equimolar ratio and randomly fragmented using a controlled DNase I digestion using a suboptimal temperature and limited incubation time to generate 100-500 bp fragments (Figure 18c). The fragments were gel purified and used as an input for a primerless PCR reaction in which regions of sequence similarity enables priming between serotypes followed by DNA extension by the Taq polymerase. Finally, full length

shuffled capsid sequences were amplified from the primerless PCR reaction using primers binding common regions flanking the capsid coding sequence, and cloned into the CLV plasmid using restriction-ligation cloning via the NotI and NsiI sites. The assembled library was electroporated into *E.coli* cells for plasmid amplification achieving a transformant count of 1.323×10^7 colony forming units (CFU).

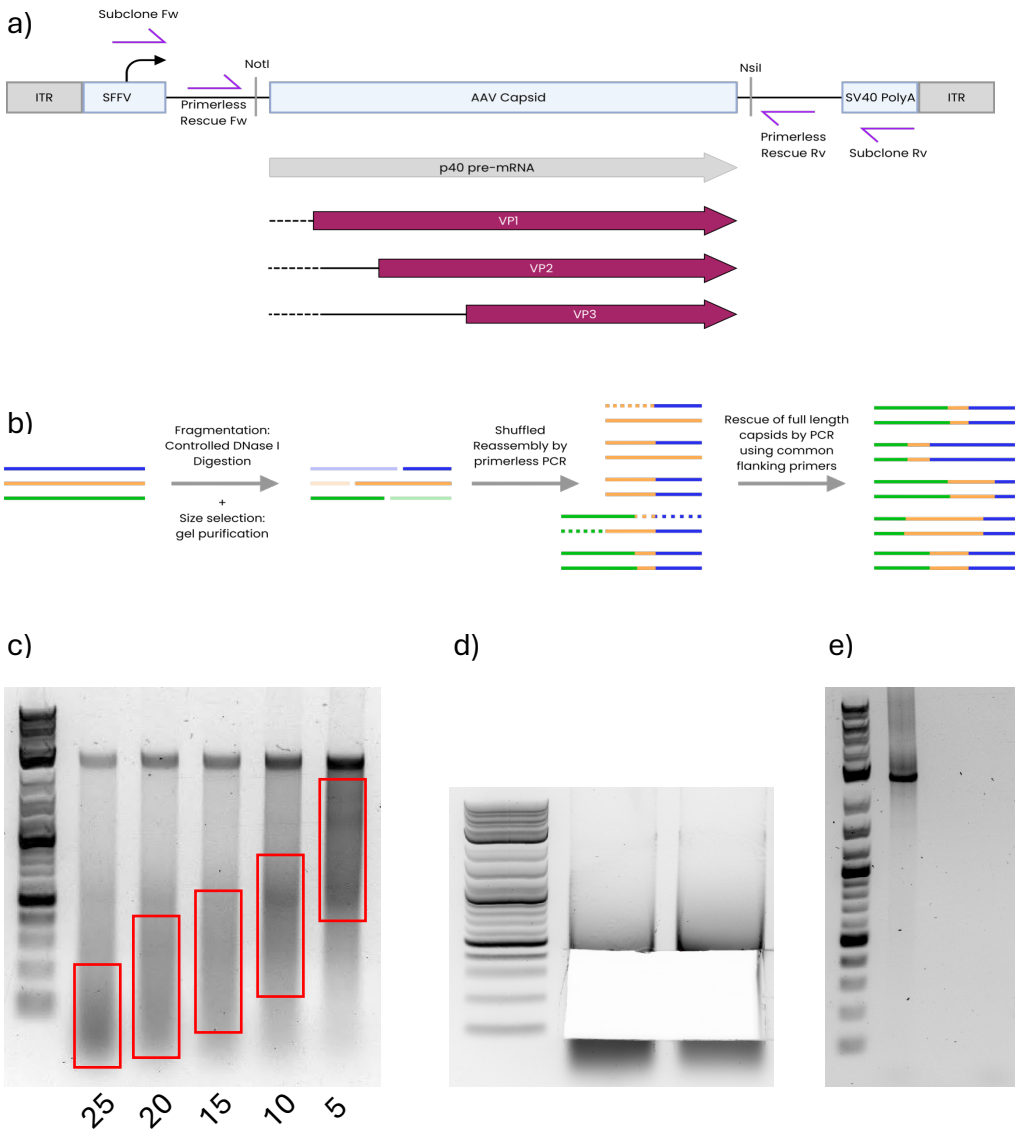
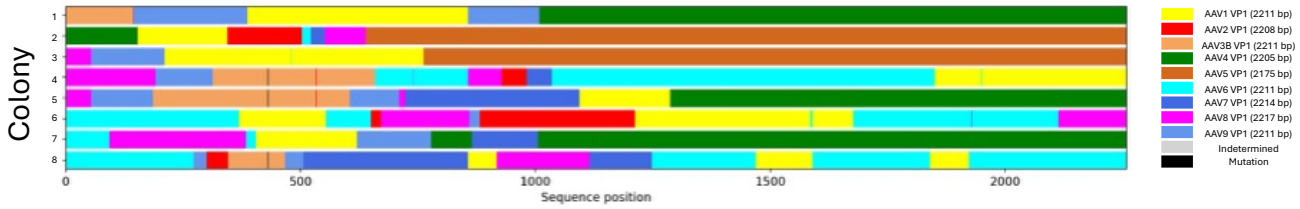


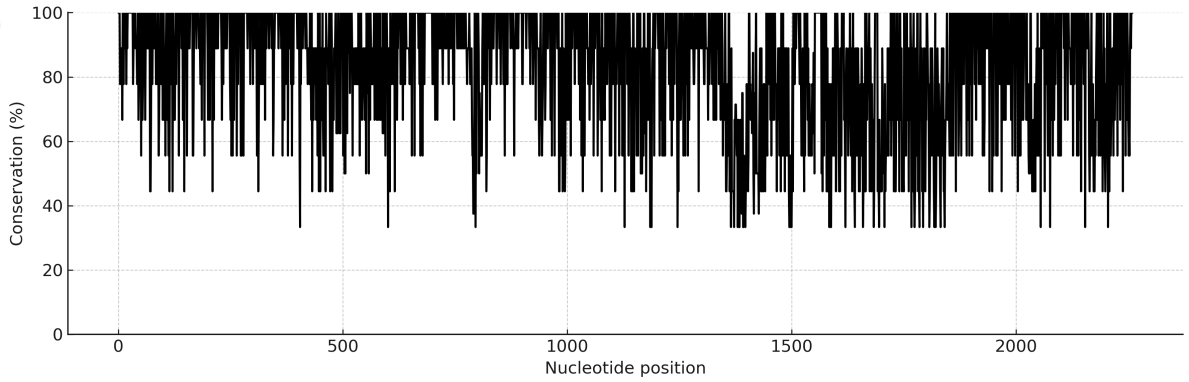
Figure 18 – Generation of shuffled AAV capsids. *A) AAV capsid library vector (CLV) schematic showing NotI and NsiI restriction sites used for capsid insertion, selected because they are noncutters for any of the AAV1-9 capsid sequences, and primers used for the shuffling process. AAV1-9 were subcloned into the capsid library construct via NotI and NsiI restriction sites to add the restriction sites and to allow amplification of 70-90bp buffer regions to ensure integrity of the restriction sites following shuffling. B) Schematic representation of the capsid shuffling process. AAV capsid amplicons generated using Subclone Fw and Subclone Rv primers were pooled at an equimolar ratio. 100-500bp fragments were created through controlled digestions using DNase I at 25C over a range of short incubation periods (c; red boxes highlight the range of digested fragments and numbers indicate incubation times), followed by gel extraction (d). The purified capsid fragments were used as input for a primerless PCR in which regions between serotypes with sufficient sequence identity act as primers for extension across serotypes. Through random fragmentation and homology-driven extension, a pool of shuffled capsids with different crossover junctions are generated. Full-length capsids were PCR amplified from the primerless PCR reaction using Primerless Rescue Fw and Primerless Rescue Rv primers (e). Finally, the pool of full-length capsids were cloned into the capsid library construct using the NotI and NsiI restriction sites.*

Following transformation of the shuffled plasmid library into *E.coli* for plasmid amplification, plasmids were isolated from eight single colonies and sequenced by whole plasmid nanopore sequencing. The mean crossover number between capsid serotypes was 9.75 across the eight sequences and generally cultured towards the first half of the sequence, although three of the eight sequences featured multiple crossovers in the 3' half of the sequence (Figure 19a). As the shuffling process requires regions of sequence similarity, it is expected that crossovers will occur more frequently in the 5' half of the sequence due to higher average conservation between AAV1-9 VP1 sequences (Figure 19b). Most segment lengths assigned by the analysis software were below 500 bp (Figure 19c), which is consistent with the size selection used during fragmentation. The 3' half of the capsid sequences were dominated by regions assigned to AAV6, AAV5, AAV4 and AAV1 parental sequences, while the 5' half included a more varied selection of parental sequences including AAV1, AAV2, AAV3B, AAV6, AAV8 and AAV9, with less contribution from AAV4 and AAV6 (Figure 19d). In terms of overall sequence similarity between variants, no two variants were identical and of the twenty-eight unique comparisons, twenty-one had an average sequence identity below 85 % (Figure 19e).

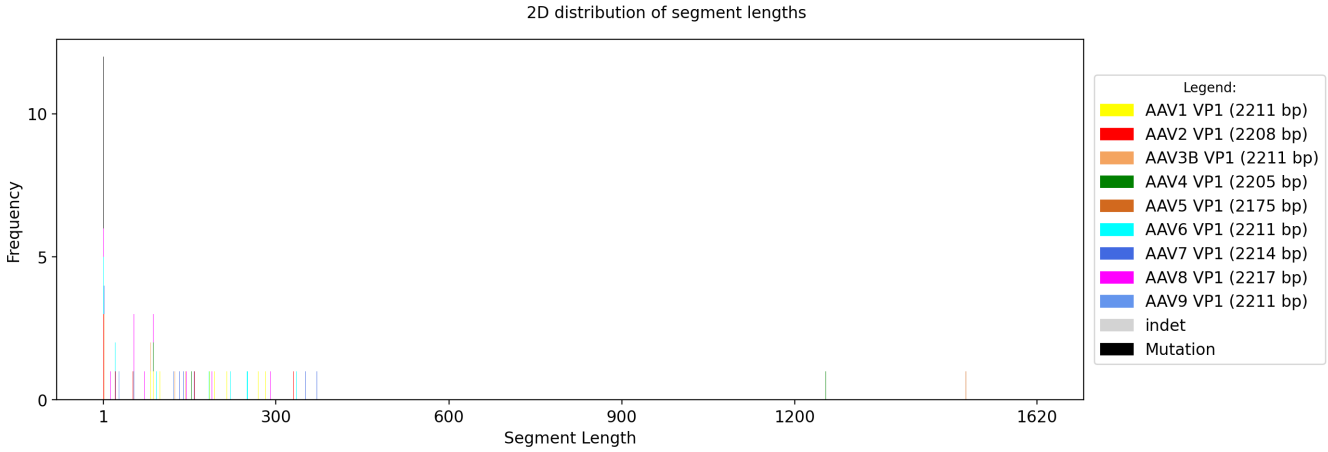
a


Colony
Sequence position
AAV1 VP1 (2211 bp)
AAV2 VP1 (2208 bp)
AAV3B VP1 (2211 bp)
AAV4 VP1 (2205 bp)
AAV5 VP1 (2175 bp)
AAV6 VP1 (2211 bp)
AAV7 VP1 (2214 bp)
AAV8 VP1 (2217 bp)
AAV9 VP1 (2211 bp)
Indetermined
Mutation

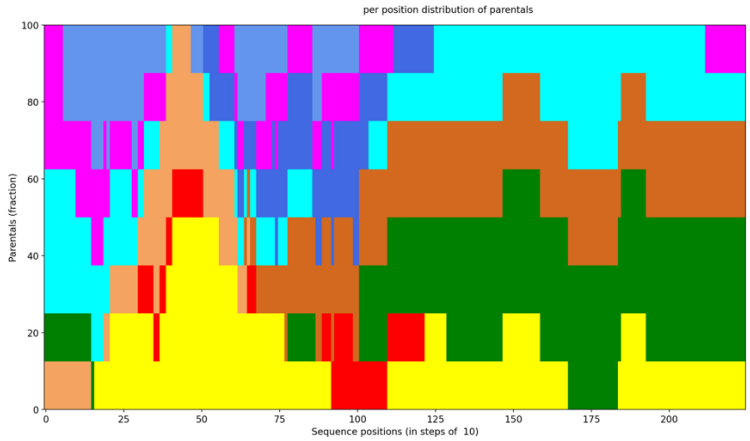
b


Conservation (%)
Nucleotide position

c


2D distribution of segment lengths
Frequency
Segment Length
Legend:
AAV1 VP1 (2211 bp)
AAV2 VP1 (2208 bp)
AAV3B VP1 (2211 bp)
AAV4 VP1 (2205 bp)
AAV5 VP1 (2175 bp)
AAV6 VP1 (2211 bp)
AAV7 VP1 (2214 bp)
AAV8 VP1 (2217 bp)
AAV9 VP1 (2211 bp)
indet
Mutation

d


per position distribution of parentals
Parentals (fraction)
Sequence positions (in steps of 10)

e

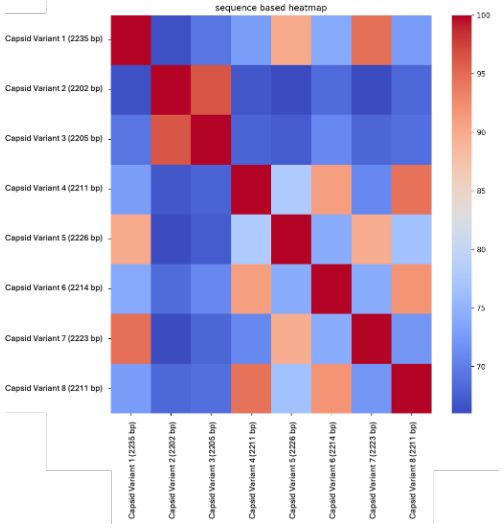
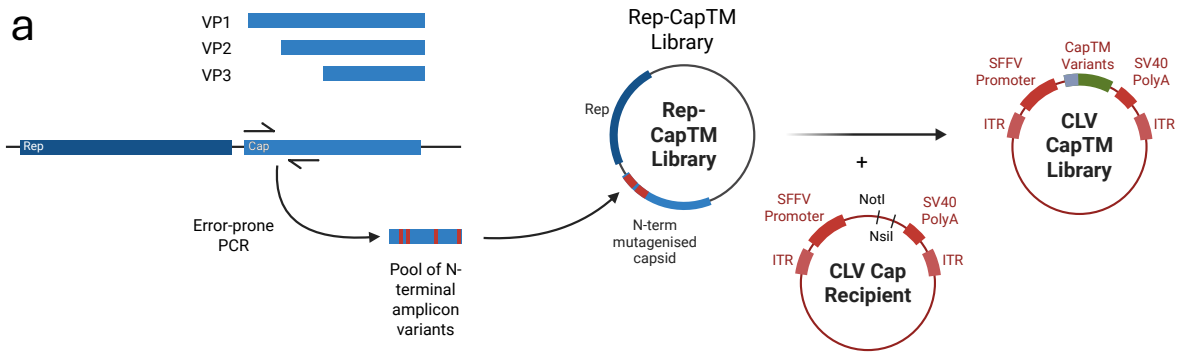

sequence based heatmap
Capsid Variant 1 (2235 bp)
Capsid Variant 2 (2202 bp)
Capsid Variant 3 (2205 bp)
Capsid Variant 4 (2211 bp)
Capsid Variant 5 (2226 bp)
Capsid Variant 6 (2214 bp)
Capsid Variant 7 (2223 bp)
Capsid Variant 8 (2211 bp)

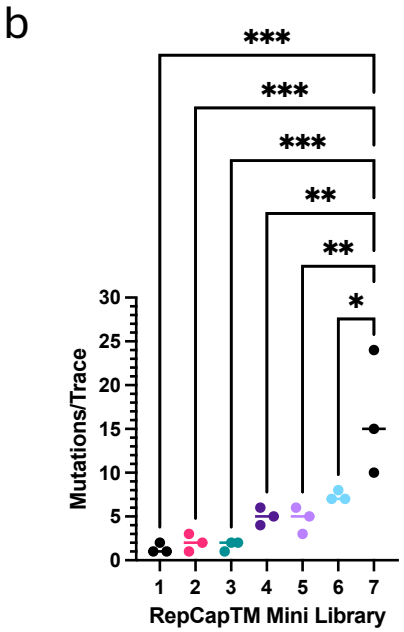
Figure 19 - Characterisation of capsids shuffling. *Following transformation of the shuffled capsid library into E.coli for plasmid production, plasmid DNA was isolated from randomly picked colonies and sequenced by full plasmid nanopore sequencing. The capsid sequences were analysed using ShuffleAnalyzer. A) Horizontal bar plot showing the assignment of parental capsid sequences to the isolated shuffled sequences. Mean crossover value = 9.75 (range: 4-15), indicating the average number of recombination junctions per capsid variant. B) Average per-nucleotide conservation between AAV1-9 serotype VP1 sequences C) Bar chart showing the fragment length distribution of fragments shown in (A). D) Frequency per position diagram showing the percentage of each parental sequence assigned to the 10bp window at the given position. E) Heatmap showing overall sequence similarity between all shuffled capsid variants.*

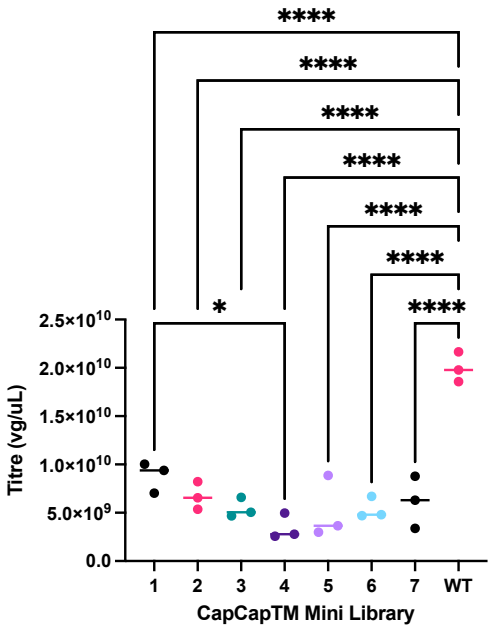
9.3.3 Cloning of a targeted random mutagenesis AAV capsid library.

As described in section 5.3.2, the limited sequence diversity available during plasmid library production necessitates careful consideration of library design to maximise functional diversity. An alternative approach to using a shuffled capsid library is to perform hypothesis-driven targeted mutagenesis on a limited region of the capsid sequence known to be involved with a post-entry step in transduction. As professional antigen-presenting cells, stimulated B cells possess specialised endosomal pathways that are tailored towards antigen preservation and presentation. As pH-dependent endosomal escape is a critical step in successful AAV transduction, we hypothesised that the altered endosomal environment in stimulated primary human B cells is suboptimal for AAV phospholipase activation and subsequent viral escape. To generate a complex capsid library containing variants with altered endosomal escape characteristics, the ~600 bp 5’ region of AAV6 VP1 sequence that encodes the VP1-unique (VP1u) and VP2/3 common region was randomly mutagenised using error-prone PCR, using various mutation rates, and cloned into a RepCap AAV helper plasmid (Figure

20a). The resulting RepCap-Targeted Mutatgenesis (TM) libraries was transformed into *E.coli* to create seven mini-libraries with transformant counts between 10^5 and 10^6 CFU, here referred to as RepCapTM 1-7 mini libraries with increasing mutagenesis rates. Three colonies from each transformation were sequenced by sanger sequencing to confirm that the mutation rate increases in a stepwise manner between mini libraries (Figure 20b). The RepCapTM mini libraries were then used to package a CMV-eGFP AAV genome to produce small scale crude viral libraries. All small-scale libraries produced AAV at a lower titre compared to a wild-type AAV6 control, however all mini libraries produced titres within a 2-fold range and the highest mutation rate (RepCapTM 7) produced titres in the 10^9-10^{10} range (Figure 20c). To confirm that increasing mutation rates in this region translated to functional differences between the libraries, HEK293 cells were transduced at an MOI of 500 and eGFP expression was assessed after 72 hours. As expected for a permissive cell type, a clear decrease in transduction efficiency was observed in cells transduced with higher mutation rate libraries compared to lower mutation rate libraries, or the wild-type AAV6 control virus (Figure 20d). These results confirmed that random mutagenesis of the 5' unresolved region of the AAV6 VP1 sequence produces libraries in which increased mutagenesis rates translate to increase functional differences between the library and wild-type AAV6 capsid. Furthermore, it shows that high mutation rates are tolerated during AAV production. To produce full scale capsid libraries, the full capsid sequences from the assembly of mutation rate 4 and mutation rate 7 (medium and high mutation rates, respectively) libraries were amplified and cloned into the CLV vector (Figure 20a). These CLV CapTM libraries were electroporated into *E.coli* with transformant counts of 1.815×10^7 and 2.376×10^7 CFU for the medium and high mutation rate libraries respectively.


a
VP1
VP2
VP3
Rep
Cap
Error-prone PCR
Pool of N-terminal amplicon variants
Rep-CapTM Library
Rep
Rep-CapTM Library
N-term mutagenised capsid
+
SFFV Promoter
NotI
NsiI
SV40 PolyA
ITR
ITR
CLV Cap Recipient
CapTM Variants
SFFV Promoter
SV40 PolyA
ITR
ITR
CLV CapTM Library


b
Mutations/Trace
RepCapTM Mini Library


c
Titre (vg/uL)
CapCapTM Mini Library

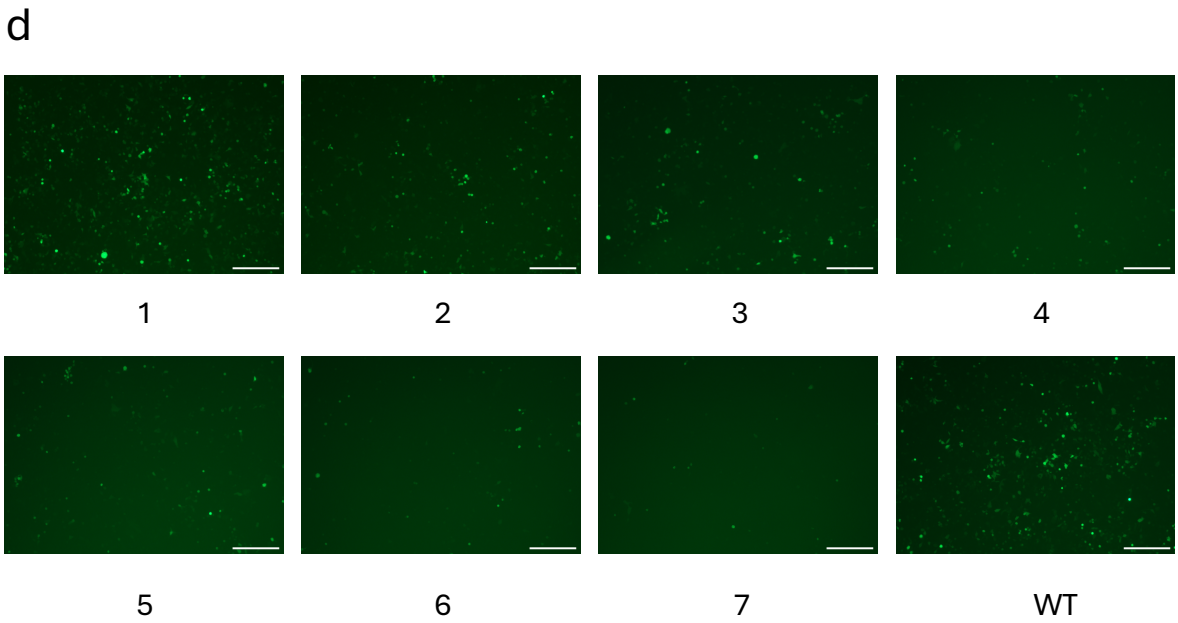

d
1
2
3
4
5
6
7
WT

Figure 20 – Targeted mutagenesis of the structurally unresolved N-terminal domain of VP1. *Schematic outlining the mutagenesis and cloning of the RepCap-targeted mutagenesis (TM) and capsid library vector (CLV) CapTM libraries. Error-prone PCR was used to randomly mutagenise the N-terminal region between the VP1and VP3 start codons. The amount of input DNA used in the error-prone PCR was titrated to create a range of mutation rates (mutation rate 1-6). To achieve a higher mutation rate, a nested error prone PCRs was performed using the mutation rate 5 amplicons as input (mutation rate 7). The mutagenised fragments were cloned into a recipient RepCap6 AAV helper plasmid, in which this region was replaced with a single ScaI blunt end restriction site, using HiFi assembly to create the RepCapTM library plasmids. To clone the full capsid sequences containing the mutagenised region into the CLV plasmid*, the capsid sequences were PCR amplified directly from the HiFi assembly reaction using primers containing the NotI and NsiI restriction sites, which were then used for restriction-ligation cloning into the CLV NotI-NsI sites. B) Following transformation into *E.coli*, three transformants were randomly selected for each RepCapTM library and the mutagenised region was sequenced by sanger sequencing of the isolated plasmids. C) and D) The RepCapTM plasmid libraries, along with a wild-type (WT) Cap6 control, were used to package a CMV-eGFP AAV genome into small scale AAV libraries. The virus produced was titrated by GFP-specific qPCR (C) and used to transduce HEK293 cells at an MOI of 500. 72 hours post-transduction, eGFP expression was assessed using a fluorescent microscope. Scale bar, 1000 µm. *Each point represents a separate sequenced colony (B) or a separate AAV production (C). Statistical significance was assessed using one-way ANOVAs with correction for multiple post-hoc comparisons using Tukey's test. ****P ≤ 0.0001 , ***P ≤ 0.001, **P ≤ 0.01, *P ≤ 0.05.*

9.3.4 Characterisation of viral AAV capsid library production

The capsid library vector plasmids encode the AAV capsid variants, but Rep must be supplied in trans by a helper plasmid. Furthermore, the library designs used for this work are likely to result in chimeric or mutagenised AAP sequences and it is not possible to predict whether AAP variants will negatively impact the production of capsid variants with improved transduction efficiency in primary human B cells. It is unclear therefore whether supplying AAP in trans would improve the diversity of the viral library by avoiding the loss of capsids with unfavourable AAP sequences. To generate a Rep expression plasmid that encodes AAP within its natural genomic context, VP1 R245 of a RepCap2 plasmid was mutated to a stop codon so that the mutation is silent with regard to the out-

of-frame AAP coding sequence. In addition, a no-AAP Rep expression plasmid was produced by adding an additional mutation within the AAP coding sequence to replace P71 with a stop codon (Figure 21a). To confirm that the Rep helper plasmid was able to support AAV production only in the presence of trans capsid, no-AAP plasmid was co-transfected into HEK293 cells along with an adenoviral helper plasmid and a CMV-eGFP AAV genome, a CMV-Cap6 AAV genome or a CMV-Cap6-Frameshift (Cap6-FS) AAV genome plasmid. A RepCap6 plasmid was used to package the CMV-eGFP genome into AAV6 capsid as a positive control for virus production and transduction. The Cap6-FS capsid sequence harbours a 26 bp stuffer sequence that introduces a frameshift in the capsid coding sequence, rendering the capsid non-functional. Crude lysates were harvested and AAV genome copy numbers were determine by ITR-specific qPCR (Figure 21b). The Rep-MAAP plasmid was able to support the packaging of the CMV-Cap6 AAV genome into the genome encoded capsid to levels that exceeded the RepCap6 plasmid positive control. In the absence of trans-supplied capsid, using either the CMV-eGFP or CMV-Cap6-FS AAV genome plasmids, the ITR genome copy number was in the 10^8 – 10^9 vg/mL range using the Rep-MAAP helper plasmid. To confirm that the copy number values calculated for these samples represent the background for the qPCR assay under these conditions, 10 µL of the crude lysate, from cells transfected with the CMV-eGFP AAV genome and either the Rep-MAAP or RepCap6 plasmid, was added directly to 293-AAV cells in a 96 well plate format. The frequency of eGFP positive cells was measured 72 hours later by flow cytometry (Figure 21c). Approximately 93 % of cells exposed to lysates from cells transfected with CMV-eGFP and RepCap6 plasmids were eGFP positive. No eGFP positive cells were detected in cells exposed to lysates from cells transfected with

CMV-eGFP and Rep-MAAP plasmids, confirming that all AAV particles in the CMV-Cap6 + Rep-MAAP samples were packaged into the AAV genome-encoded capsid sequence.

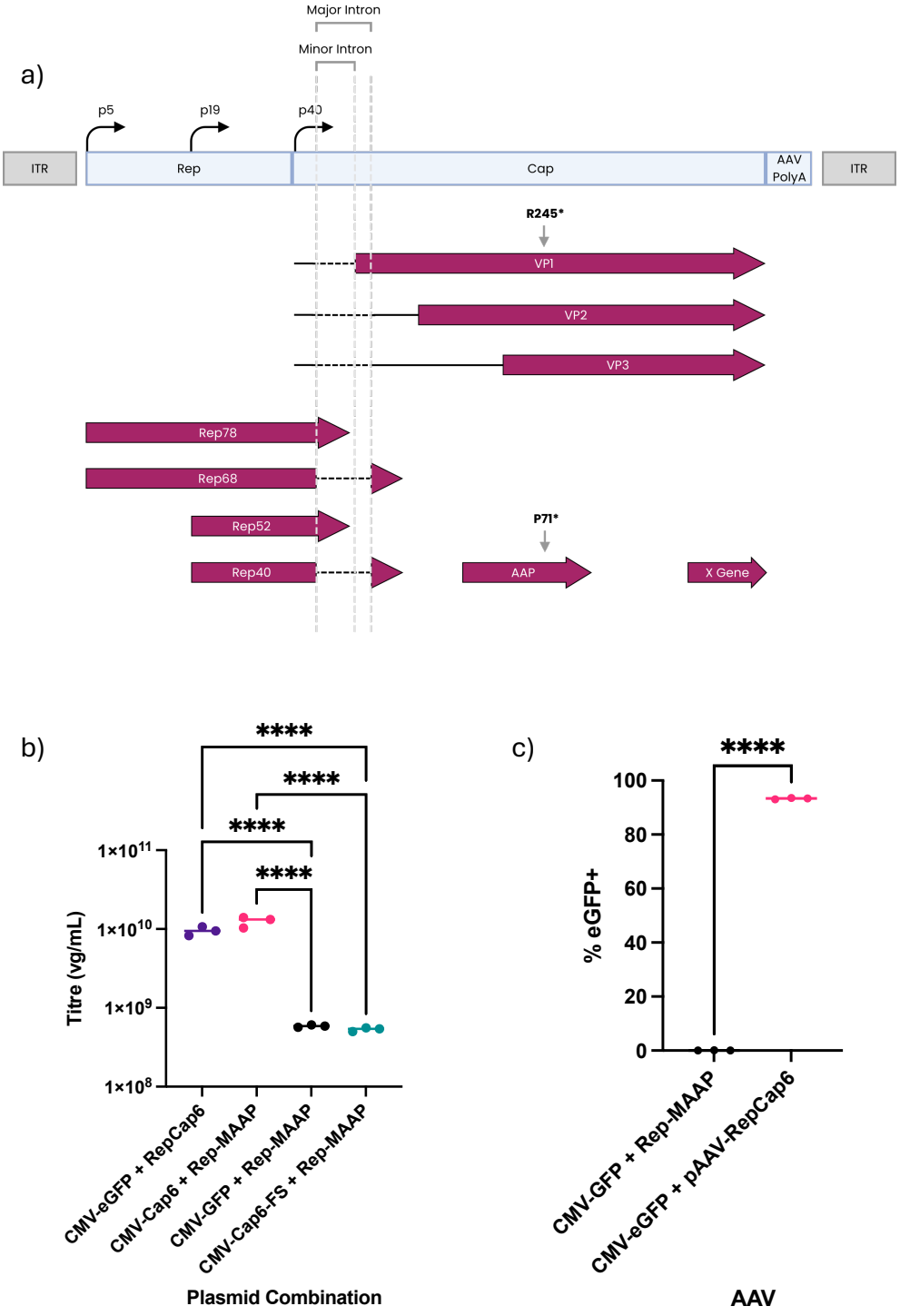


Figure 21 – Validation of Rep-MAAP-AAP and Rep-MAAP AAV capsid library production helper plasmids. *A) Schematic representation of the AAV genome and mutations used to create helper plasmids. To prevent the expression of the structural capsid proteins (VP1, VP2 and VP3), a single nucleotide change was made to mutate Arg 245 of VP1 to a stop codon. This mutation is silent with regard to the out-of-frame AAP coding sequence. To disrupt AAP expression in addition to the structural proteins, a further stop codon was placed at Pro 71 of AAP. B) HEK293 cells were transfected with an adenoviral helper plasmid and a combination of RepCap6 or Rep-MAAP helper plasmid with CMV-eGFP, CMV-Cap6 or CMV-Cap6-Frameshift (FS) AAV genome plasmid. AAV genome copies within crude lysates were determined by ITR-specific qPCR. C) 10 µL of crude lysate from cells transfected with CMV-eGFP AAV genome in B) was added to 293-AAV cells in a 96 well plate and the frequency of eGFP expressing cells was measured 72 hours later by flow cytometry. Each point represents a separate AAV production (B) or a biological replicate (C). Statistical significance was assessed using one-way ANOVAs on log-transformed data with correction for multiple post-hoc comparisons using Tukey's test (B) or using an unpaired t-test (C). ****P ≤ 0.0001, *P ≤ 0.05.*

Before full-scale production of the AAV capsid libraries for screening, the productivity of the different capsid libraries (Figure 22a) in the presence or absence of AAP was assessed. 293-AAV cells were transfected with equimolar ratios of capsid library plasmid, adenoviral helper plasmid and Rep-MAAP (Figure 22b) or Rep-MAAP-AAP (Figure 22c) plasmid. The productivity was determined by titration of genome copy numbers in crude lysates using ITR-specific qPCR. In the presence of AAP, the two targeted mutagenesis libraries (CLV TM4 and CLV TM7) titres were equivalent and produced titres greater than the shuffled library. The shuffled library and CLV Cap6 control tires were not statistically different. However, in the absence of AAP, the genome copy number of viruses produced using the shuffled library plasmid was lower than the CLV Cap6 control and CLV TM4 libraries. Under these conditions, the CLV TM7 library titre was 26.3% lower than the CLV TM4 library and statistically equivalent to the CLV shuffled library. By directly comparing the titre in the presence or absence of AAP (Figure 22c), the

presence or absence of AAP did not appear to affect the production of the CLV Cap6 control, the CLV shuffled library or the CLV TM4 library. However, the productivity of the CLV TM7 library was modestly impaired by the presence of trans-supplied AAP.

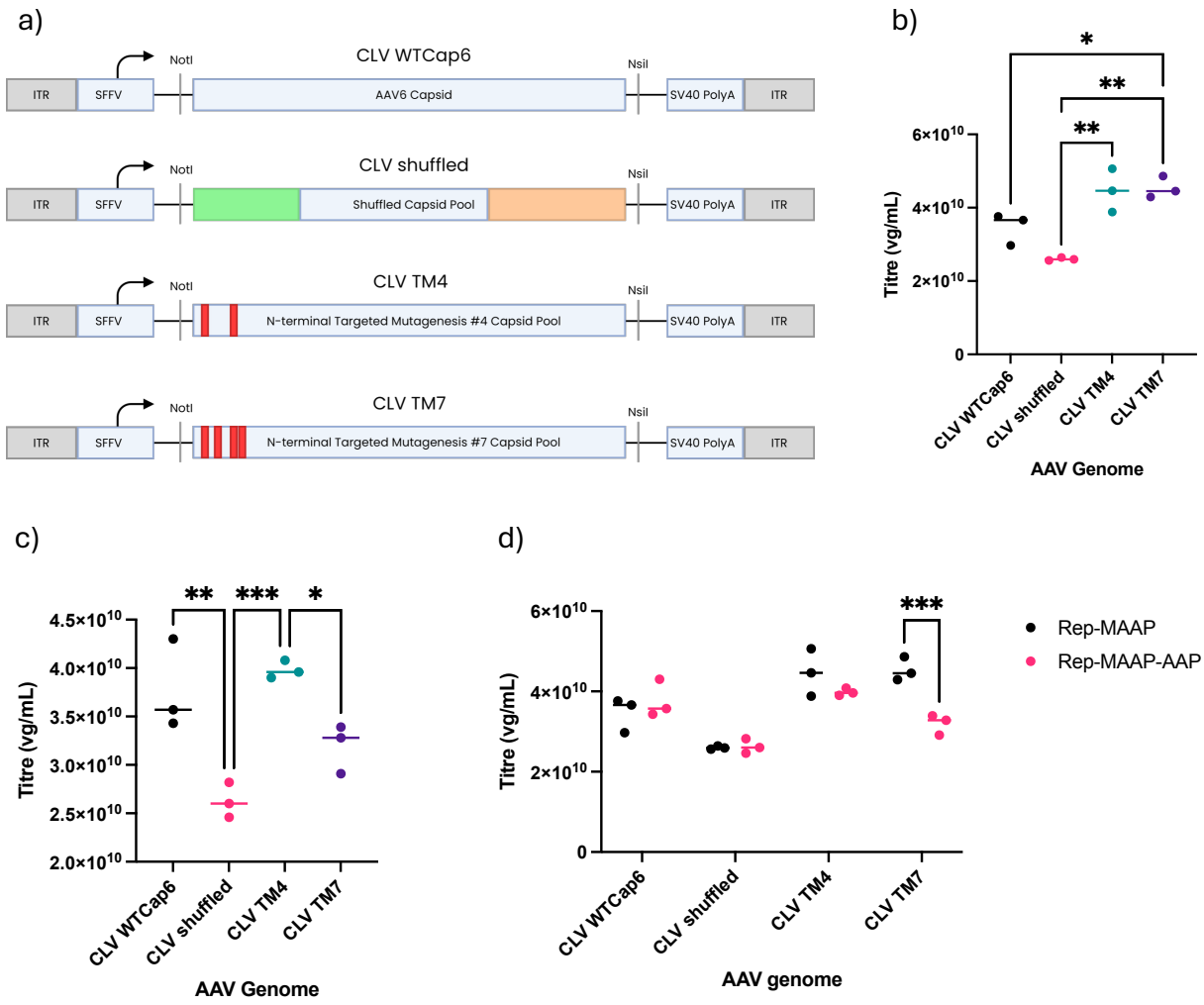


Figure 22 – AAV capsid library production efficiency in the presence or absence of AAP. *A) Schematic overview of high diversity AAV capsid plasmid libraries, and a capsid library vector encoding WT AAV 6 capsid (WTCap6) as a control plasmid for virus productions. Small scale viral libraries were produced by triple-transfection of an adenoviral helper plasmid, 65.86 ng/cm² AAV capsid library plasmid, or WTCap6-encoding capsid library plasmid, and either Rep-MAAP (B) or Rep-MAAP-AAP (C) plasmids. AAV genome copies/mL was determined by ITR-specific qPCR on crude AAV preparations. B) and C) compare across capsid library genome plasmids. D) displays the effect of AAP on each capsid library. Each point represents a separate AAV production. Statistical significance was assessed using one-way ANOVAs with correction for multiple post-hoc comparisons using Tukey's test (B) and (C) or using a two-way ANOVA with post-hoc analyses limited to within AAV genome groups only and corrected for multiple comparisons using the Sidak method (D). ***P ≤ 0.001, **P ≤ 0.01, *P ≤ 0.05.*

To limit the likelihood of cross-packaging and mosaic capsid formation during one-step capsid library production, the genome mass used is kept low. To see how the productivity of the different capsid libraries compared under low genome input conditions, 293-cells were transfected with an amount equivalent to 100 ng per 15-cm dish (0.6897 ng/cm^2) of each capsid library, compared to 65.86 ng/cm^2 used in the previous experiments. As before, productivity was assessed using qPCR measurement of ITR-specific genome copy number. In contrast to the high genome input experiment in shown in Figure 22, CLV TM7 production is reduced compared to CLV TM4 and CLV Cap6 samples in both the absence (Figure 23a) or presence (Figure 23b) of AAP. The productivity of the CLV shuffled library remains lower compared to the CLV Cap6 control both in the presence or absence of AAP. Unlike the high genome input experiment, CLV TM4 viral library titres are modestly reduced compared to the CLV Cap6 control in the presence of AAP. Here, the direct comparison of productions with, or without, AAP within samples using the same AAV genomes shows that at low genome input, the productivity of the CLV TM4 library is reduced by 23.2% in the presence of AAP (Figure 23c).

In summary, under normal AAV production conditions (equimolar plasmid amounts), the presence of AAP negatively affects the productivity of the high mutation rate targeted mutagenesis library (CLV TM7) alone. In contrast, under low genome input conditions, the medium rate targeted mutagenesis library (CLV TM4) titre is marginally reduced in the presence of AAP compared to samples where trans AAP is withdrawn. Under low genome conditions, the CLV TM7 library shows reduced productivity compared to the CLV TM4 and control productions regardless of the presence, or absence, of trans AAP. One explanation for these findings could be that, as AAP is known to act as a homodimer, the

mutation of AAP in the genome-encoded capsid creates AAP proteins that are incompatible with wild-type AAP2 (WT AAP), resulting in the disruption of capsid assembly upon heterodimerisation. This explanation alone appears inconsistent with the finding that the effect is observed in the lower mutation rate library under low genome input conditions, but not in the high genome input conditions. Under this hypothesis, it would be expected that a higher dose of mutated AAP (high genome input) would lead to impaired capsid assembly at lower mutation rates, and that higher mutation rates would be required to observe the effect at lower genome inputs. Given the low sample size used for the data presented here, caution should be taken not to overinterpret. However, if these differences represent a true effect, this contradiction could be explained by a model (Figure 24) in which a low dose of mutant AAP early in the production does not provide an advantage to WT AAP due to the low availability of capsid protein to assemble, and fast turnover of WT AAP inhibits accumulation in order to outcompete mutant AAP later in the production. In this model, the use of a high genome input provides a longer window of time during the production in which capsid is available for the suboptimal mixture of homodimeric and heterodimeric AAP to assemble viral particles. This could explain why a very high mutation rate is required is required to detect a decrease in AAV production when a high amount of genome is used, while a decrease in titre can be detected using a lower mutation rate during a production with a low amount of input genome.

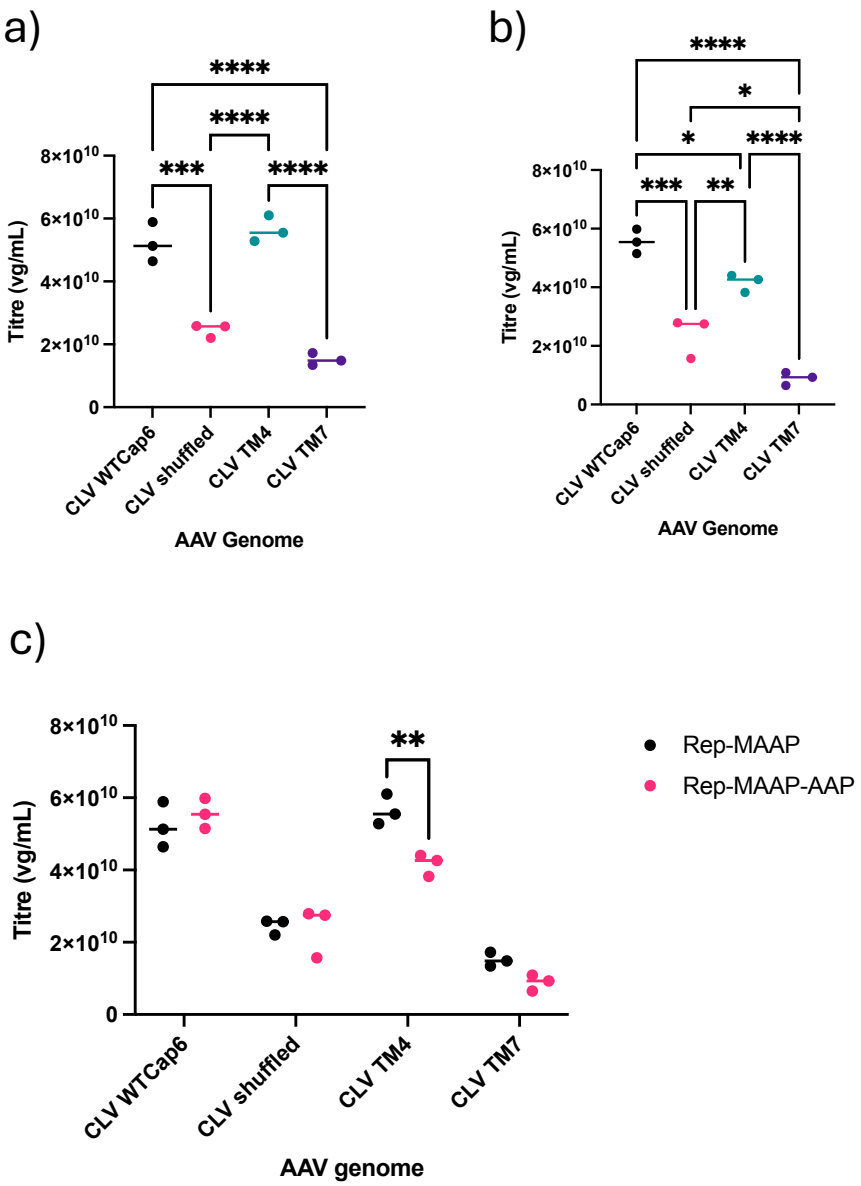


Figure 23 - AAV capsid library production efficiency, using low genome input, in the presence or absence of AAP. *Small scale viral libraries were produced by triple-transfection of an adenoviral helper plasmid, 0.6897 ng/cm² of AAV capsid library plasmid, or WTCap6-encoding capsid library plasmid, and either Rep-MAAP (A) or Rep-MAAP-AAP (B) plasmids. AAV genome copies/mL was determined by ITR-specific qPCR on crude AAV preparations. A) and B) compare across capsid library genome plasmids. C) displays the effect of AAP on each capsid library. Each point represents a separate AAV production. Statistical significance was assessed using one-way ANOVAs with correction for multiple post-hoc comparisons using Tukey's test (A) and (B) or using a two-way ANOVA with post-hoc analyses limited to within AAV genome groups only and corrected for multiple comparisons using the Sidak method (C). ****P ≤ 0.0001 , ***P ≤ 0.001, **P ≤ 0.01, *P ≤ 0.05.*

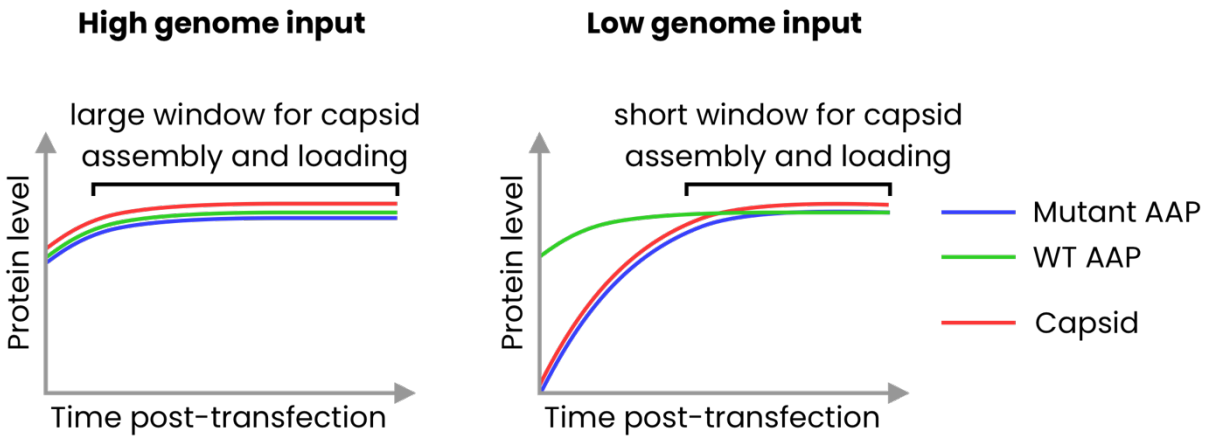


Figure 24 – Model for the effect of mutant genome-encoded AAP on the production of AAV libraries at high versus low plasmid genome inputs. Following transfection of equimolar ratios of Rep-MAAP-AAP plasmid, supplying WT AAP, and the capsid library genome, supplying mutant AAP, capsid protein levels quickly accumulate creating a large time window for capsid assembly and genome loading, even with suboptimal mixtures of heterodimeric and homodimeric AAP. When a very low amount of genome plasmid is supplied, it takes longer for the capsid protein levels to accumulate. Additionally, if the turnover of WT AAP is fast, the WT AAP does not accumulate to outcompete the mutant AAP. This results in a short time window for capsid assembly where the ratio of WT and mutant AAP is similar to that found in the high genome input condition.

9.3.5 Measurement of AAV genome cross-packaging during viral library production.

The work presented in chapter 4 sought to produce a system that could reduce the likelihood of cross-packaging to negligible levels. The more typical one-step library production design in this chapter necessitates some assessment of this risk. The average cross-packaging frequency during one-step library production has been investigated in detail by other researchers (Nonnenmacher et al., 2015a), interestingly however there is a difference in capsid-genome correlation behaviour between wild-type and recombinant AAV. This suggests that capsid library genome, and Rep helper plasmid, configuration has an impact on cross-packaging potential. To assess the potential for cross-packaging using the plasmids produced for the libraries presented here, the CLV

WTCap6 plasmids was mixed at 95:5 molar ratio with a CLV plasmid encoding Cap6-Framsheft (FS) (Figure 25a), a version of WTCap6 harbouring a 26 bp stuffer sequence that renders the capsid non-functional (Figure 21). 8.77 µg (equimolar genome mix:adenoviral-helper:Rep-MAAP-AAP ratio) or 100 ng of the WTCap6:Cap6-FS mixture was co-transfected with an adenoviral helper plasmid and Rep-MAAP-AAP plasmid into 293-AAV cells in a 15 cm dish. The resulting AAV was iodixanol purified and DNase I treated to remove residual unpackaged AAV genomes. The Cap6-FS region containing the 26 bp stuffer DNA was amplified by PCR and sequenced by Illumina sequencing. Using this experimental design, the percentage cross-packaging was quantified by calculating the frequency of frameshift reads in the viral preparation relative to the frequency of frameshift reads in the plasmid input sample (Figure 25b). Therefore, a cross-packaging percentage of 100 % reflects a situation in which the loading a AAV genomes is completely random, while a percentage less than 100 % reflects bias towards the loading of WTCap6 sequences. The frequency of frameshift reads in the plasmid input sample was 5.15 %, approximately as expected given the 95:5 ratio of WTCap6 to Cap6-FS in the plasmid input mixture. Using the high genome input, the cross-packaging percentage was 4.78-5.78%. Reduction of the genome input mass to 100 ng significantly reduced cross-packaging to 0.14-0.23%, corresponding with just 5-8 reads amongst 65000-68000 total reads. Together, this data shows that, even using a high genome input, cross-packaging occurs at a much lower rate than would be expected by random assignment of AAV genomes to assembled capsids. Never-the-less, 5% cross-packaging is likely to introduce significant noise into complex libraries. The reduction of AAV genome input to levels typically used for AAV library production significantly reduces cross-packaging to negligible levels.

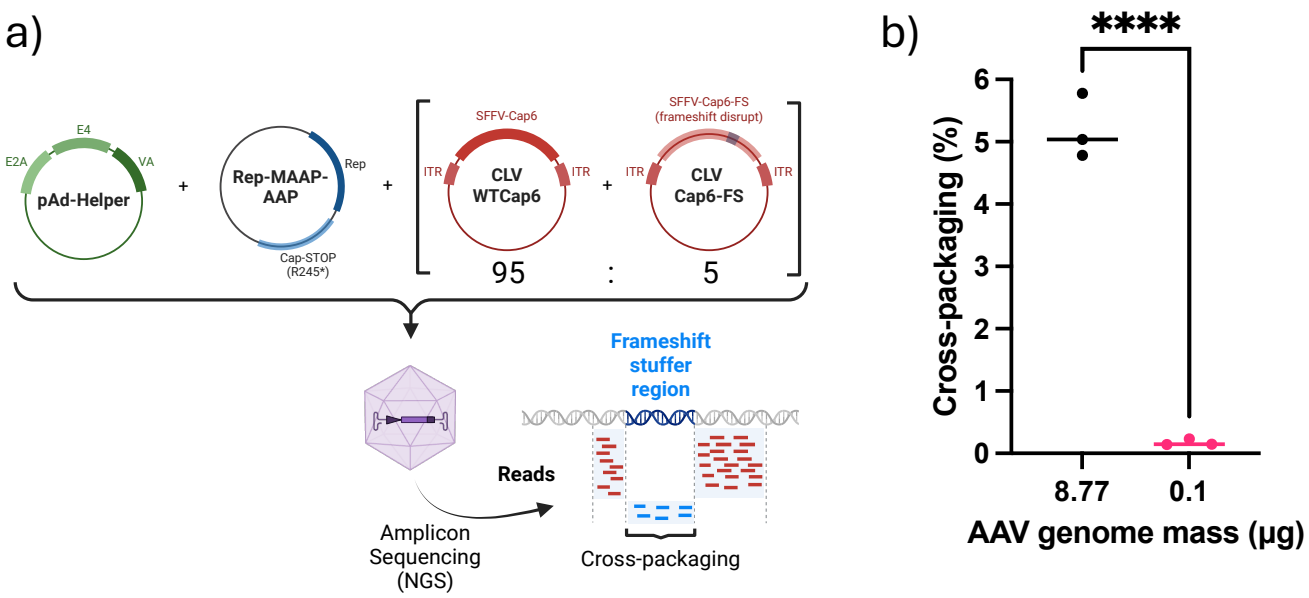


Figure 25 – Assessment of cross-packaging of a capsid library plasmid encoding a disrupted AAV6 capsid sequence into intact AAV6 particles. *A) Schematic of experiment methodology. Capsid library vector (CLV) plasmid encoding wild-type AAV6 capsid (WTCap6) was mixed with a CLV plasmid encoding a WTCap6 sequence featuring a stuffer sequence that introduces a frameshift, preventing capsid expression, at a 95:5 ratio. 8.77 µg (high genome mass sample) or 100 ng (low genome mass sample) of the mixture was co-transfected with 10.46 µg adenoviral helper plasmid and 15.77 µg Rep-MAAP-AAP plasmid. The frameshift region was PCR amplified using flanking primers from iodixanol purified AAV and sequenced by Illumina sequencing. B) Percentage cross-packaging, calculated as the frequency of frameshift reads in the viral preparation relative to the frequency of frameshift reads in the plasmid input sample (5.15 %), expressed as a percentage. Each point represents a separate AAV production. Statistical significance was assessed using an unpaired t-test. ****P ≤ 0.0001.*

9.3.6 Screening in primary human B Cells.

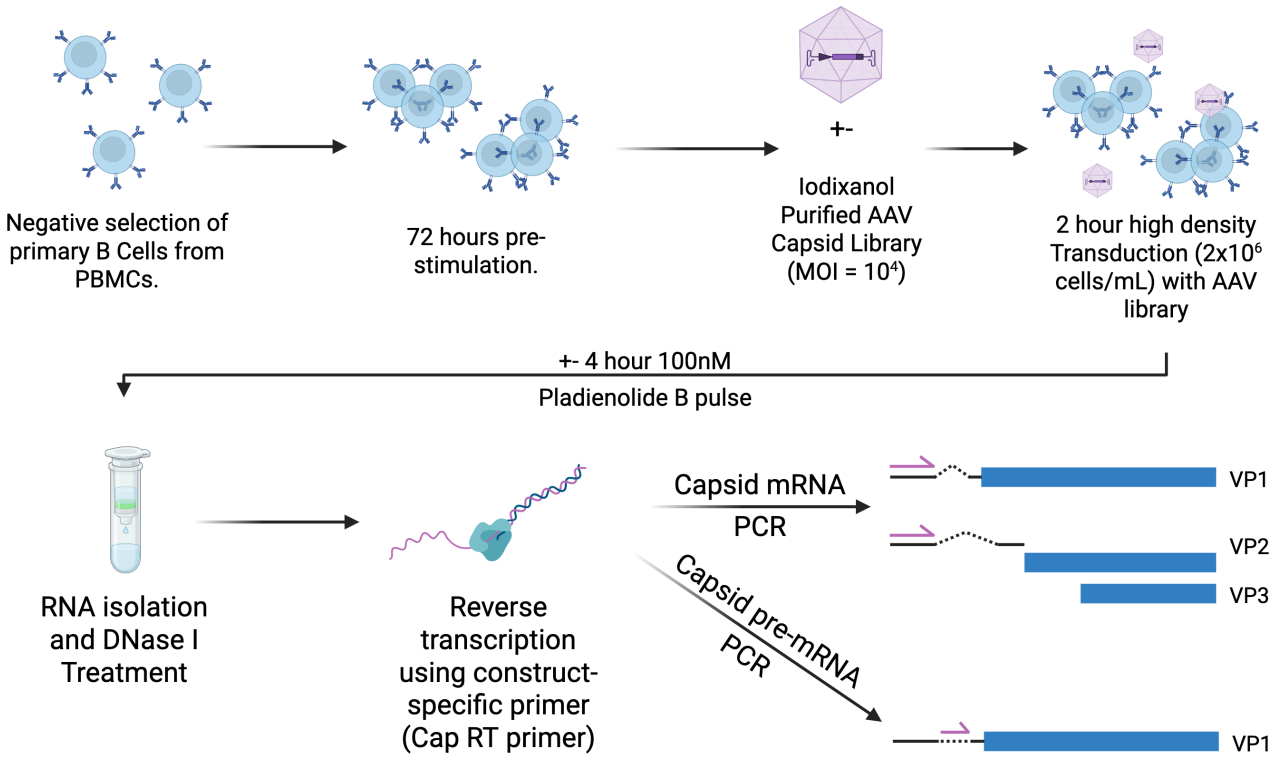
CLV shuffled, CLV TM4 and CLV TM7 libraries were produced at large scale (10 x 15 cm diameter dishes per production) using 100 ng of plasmid library per 15 cm dish to minimise the risk of cross-packaging and mosaic capsid formation, while enabling adequate titres for screening.

To screen for AAV capsids with improved transduction efficiency in stimulated primary human B cells, $5x10^5$ purified B cells pre-stimulated for 72 hours were exposed to CLV shuffled, CLV TM4 or CLV TM7 viral capsid libraries at an MOI of 10^4 (Figure 26a). The cells were transduced at a high density ($2x10^6$ cells/mL) for 2 hours to increase the chance of encounters between the cells and viral particles. The cells were then cultured for 72 hours at $1x10^6$ cells/mL to allow time for capsid RNA expression (Figure 26b).

The CLV plasmid design lacks unique barcode identifiers and therefore the capsid sequence recovered from the RNA isolated during screening must be the full length VP1 sequence. The AAV capsid subunits are translated from two mRNA transcripts formed through alternative splicing. Both transcripts share a single splice donor located within the Rep coding sequence, however the VP1 transcript uses a splice acceptor at -2 nucleotides (nt) from the start of the VP1 coding sequence (minor intron) and the VP2/3 transcript uses a splice acceptor shared with Rep 68 and Rep 40 at +25 nt (major intron) (Figure 21a). As a result, recovery of AAV capsid mRNA from cells exposed to the capsid libraries would yield a mixture of transcripts containing full and partial VP1 sequences. To circumvent this issue, primers were designed to bind to the capsid minor intron to selectively amplify only full-length pre-mRNA sequences from cDNA produced during the

screening method. As the splicing efficiency, and therefore abundance of pre-mRNA transcripts, was unknown, a condition was also included utilising and 4-hour pulse with 100 nM Pladienolide B (Plad-B), an RNA splicing inhibitor. If effective, this would enrich pre-mRNA and allow the recovery of full length VP1 sequences using the pre-mRNA specific forward primer, or a 5' UTR exon-binding forward primer. Following the 72-hour cell culture, with or without Plad-B pulse, RNA was isolated and DNase I treated to remove any contaminating AAV genomic DNA. Capsid cDNA was synthesised using a construct-specific primer that binds within the SV40 PolyA sequence, upstream from the polyA cleavage site.

a


Negative selection of primary B Cells from PBMCs.
72 hours pre-stimulation.
+-
Iodixanol Purified AAV Capsid Library (MOI = 10⁴)
2 hour high density Transduction (2x10⁶ cells/mL) with AAV library
+- 4 hour 100nM Pladienolide B pulse
RNA isolation and DNase I Treatment
Reverse transcription using construct-specific primer (Cap RT primer)
Capsid mRNA PCR
Capsid pre-mRNA PCR
VP1
VP2
VP3
VP1

b

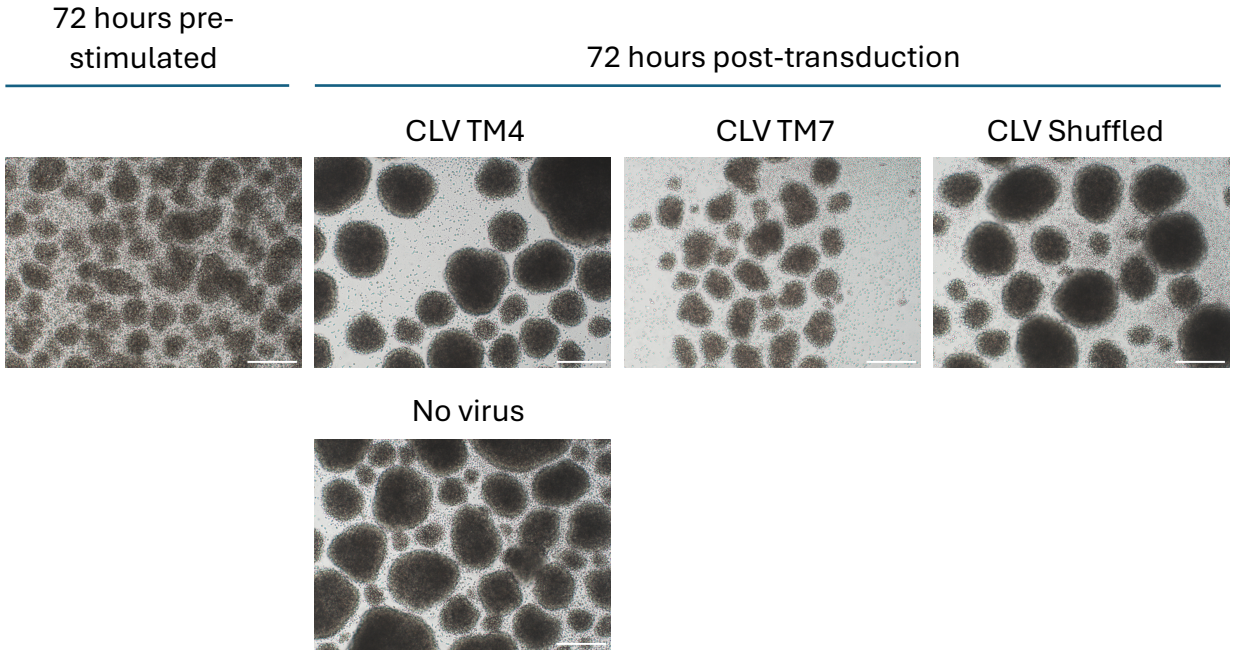

72 hours pre-stimulated
72 hours post-transduction
CLV TM4
CLV TM7
CLV Shuffled
No virus

Figure 26 – Process for screening of AAV capsid libraries in stimulated primary human B cells. *A) Schematic of screening methodology. Created with BioRender.com B) For each screen, 5x10^5 Primary human B cells were stimulated for 72 hours and then viral capsid libraries were added an MOI of 10^4 at high cell density (2x10^6 cells/mL) to minimise the cell culture volume and therefore increase the likelihood of encounters between cells and virus. After 2 hours, the cell culture volume was increased, and cells were cultured for a further 72 hours. Cells receiving Pladienolide B were pulsed for 4 hours with 100 nM prior to RNA isolation. Scale bar, 1000 µm.*

To test for the presence of recovered capsid sequences, and to validate that AAV genomic DNA was efficiently removed, total transcribed capsid sequences were amplified by PCR using a common reverse primer and a 5' UTR exon-binding capsid-specific forward primer, and visualised on an agarose gel (Figure 27a). For the no Plad-B condition, 2 distinct amplicons were observed, for the cDNA synthesised from cellular RNA isolated from cells exposed to the CLV shuffled and CLV TM4 libraries. No amplification of capsids rescued from the CLV TM7 exposed cells could be visualised. For the samples subject to the Plad-B pulse, two distinct bands were present for the CLV shuffled library and a single band, representing the full-length capsid sequence, was visualised for the CLV TM4 library. Importantly, no amplicons were detectable in the RNA samples, confirming full removal of contaminating AAV genomes. To specifically recover the full-length capsid pre-mRNA for the CLV shuffled and CLV TM4 samples, a capsid minor intron-binding forward primer was used (Figure 27b). Single amplicons were present for all samples. This data shows, in addition to the successful recovery of full-length capsid sequences from the CLV Shuffled and CLV TM4 libraries, that Plad-B may be used to selectively recover full length capsid sequences for non-barcoded AAV capsid libraries. However, the dose and length of the Plad-B pulse may need to be optimised to ensure consistency

across samples. The reduced signal intensity for the visualised amplicons from the Plad-B samples suggests that the Plad-B pulse reduced the overall RNA content of the sample, possibly due to toxicity related apoptosis. For the purpose of this work, the use of a pre-mRNA specific primer enabled full length sequence recovery without the need for Plad-B.

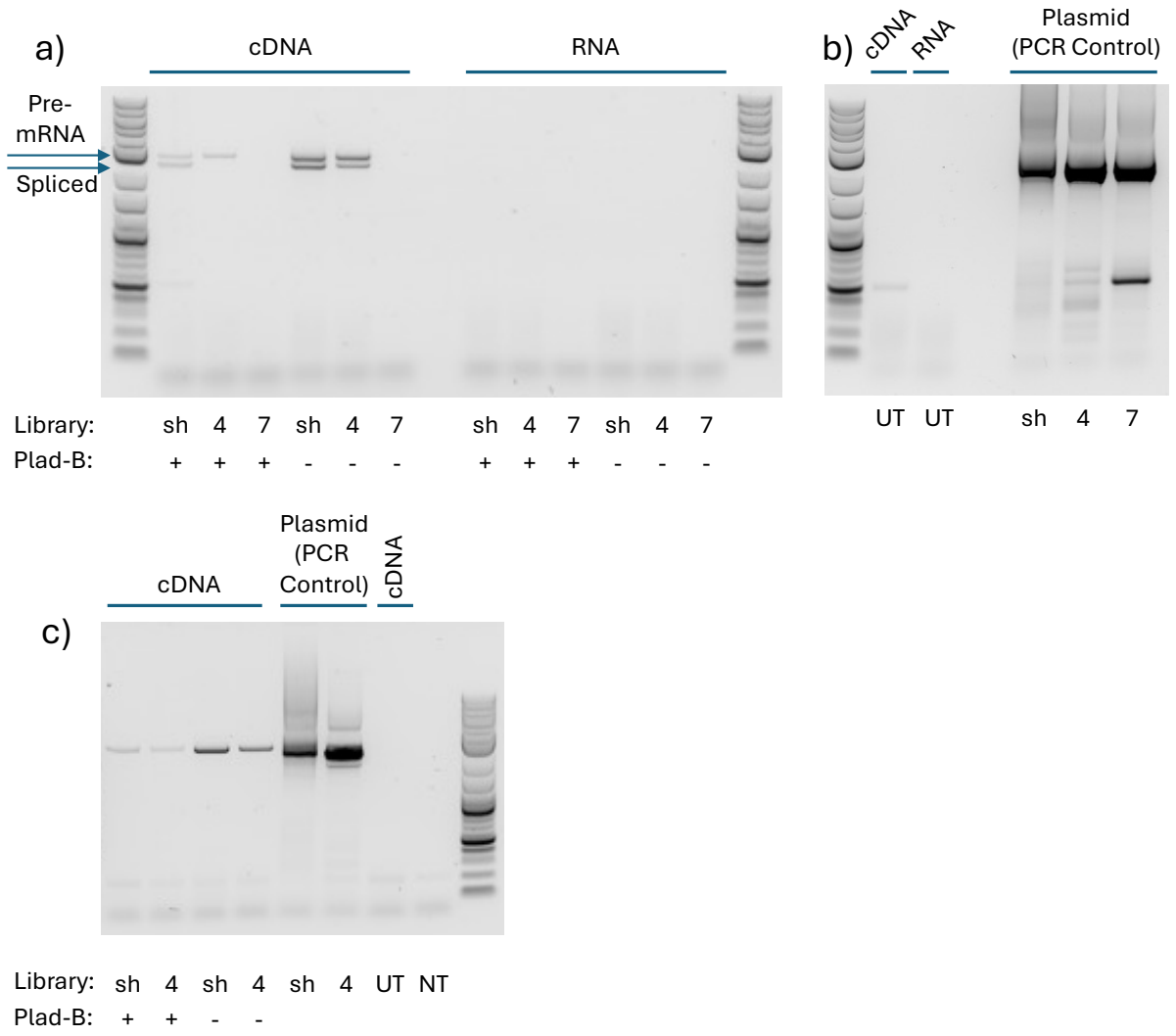


Figure 27 – Rescue of full-length transcript-level capsid variant cDNA from screening of capsid libraries in stimulated primary human B cells. *Following RNA isolation and DNase I treatment, cDNA synthesis was performed using a construct-specific primer. Capsid variant DNA was amplified from cDNA using a common 3' UTR reverse primer and forward primers that bind upstream of the capsid variant coding sequences to amplify either total capsid mRNA (5' untranslated exon binding) or pre-mRNA (5' intron binding). Amplification of pre-mRNA recovers the full VP1 sequence while amplification of total capsid mRNA recovers a mixture of VP1 and VP2/3 transcripts. The total capsid mRNA PCR products, along with no-reverse transcription controls (A) and pre-mRNA products (B) were visualised on an agarose gel.*

Prior full-scale sequencing of recovered capsid sequences, the shuffled capsid library variants, amplified using pre-mRNA-specific primers, were cloned into plasmid DNA and used for bacterial transformation. Plasmid DNA was isolated from a small sample of colonies and sent for nanopore sequencing (Figure 28).

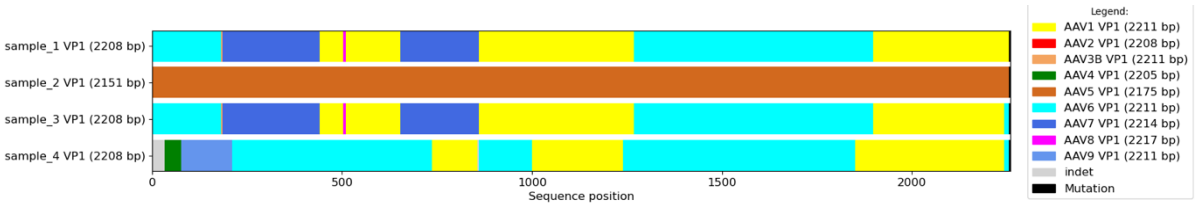


Figure 28 - Parental sequence contribution to recovered shuffled library capsid sequences recovered following screening in primary human B cells. *Capsid sequences recovered from the shuffled library capsid screen, using pre-mRNA-specific primers, were cloned into plasmid DNA, which was then used to transform E. coli. Plasmid DNA was isolated from randomly picked colonies and sequenced by full plasmid nanopore sequencing. The capsid sequences were analysed using ShuffleAnalyzer.*

9.4 Discussion

In this chapter we assessed the level of impaired binding and entry of AAV6 in human B cells. The internalisation appeared somewhat less efficient in primary B cells compared to permissive 293-AAV cells, but not sufficiently to explain the orders of magnitude difference in MOIs required to transduce the two cell types in practice. Using nanobody-displaying AAVs we saw that targeting AAV2-based capsids to CD38 expressed on B cells achieved similar transduction efficiencies to AAV6, suggesting that post-entry bottlenecks limit transduction efficiency. Guided by these experiments, we produced shuffled and targeted mutagenesis AAV capsid libraries. The role of trans AAP in library production yield was assessed, and cross-packaging frequency was determined using next-generation sequencing. Lastly, we screened the library in stimulated primary human B cells and recovered capsid sequences.

To establish whether early cell binding or entry steps represent major barriers in stimulated primary human B cells, we performed a binding/internalisation assay. This assay was designed to compare the rate of entry, normalised by the amount of bound virus, compared to the permissive 293-AAV cell line. We saw a 1.85-fold decrease in entry efficiency compared to 293-AAV cells. Differences in the expression or availability of AAVR and GPR108 between primary human B cells and permissive 293-derived cells may contribute to the reduced entry efficiency observed in B cells, such that while attachment via glycans occurs in both cell types, lower levels of critical entry factors could limit the proportion of internalised virions. Public transcriptomic resources detect low AAVR expression across immune populations and readily detected GPR108 expression in B-

cell subsets (Alam, 2021). To our knowledge. AAVR and GPR108a protein levels have not been measured directly in primary human B cells. Therefore, low AAVR levels may contribute to the reduced entry efficiency observed using the binding/internalisation assay. However, the modest decrease in entry efficiency is unlikely to account for the remarkably high MOIs required to achieve transduction in primary B cells. We used the same donor for the binding/internalisation assay as for the screen, therefore although we cannot generalise these findings across biological donors, these findings are valid for guiding the design of the library design in this donor. The performance of capsids recovered following screening will need to be validated across donors to confirm that the improved transduction efficiency is not limited to a subset of donors.

To investigate whether AAV binding is impaired in primary human B cells due to low glycan availability, we generated AAV2-based AAV particles incorporating nanobodies against CD20, CD19 and CD38. RAMOS cells express moderate surface levels of CD19 and high levels of CD20 and CD38 (Flavell et al., 2006, Kang et al., 2018). We saw that AAVs displaying anti-CD20 or anti-CD19 nanobodies failed to transduce RAMOS cells at an MOI of 10^3. In contrast, incorporation of an anti-CD38 nanobody rescued the ability of the no nanobody AAV2 (HSPG binding site-ablated) control to transduce RAMOS cells to wild-type AAV6 levels. Since re-targeting of AAVs to a surface molecule with high expression did not improve transduction efficiency, it is improbable that low glycan abundance could explain the poor transduction efficiency in RAMOS cells. RAMOS cells are an EBV-negative B cells line and may not accurately reflect the availability of attachment factors on stimulated primary B cells. However, together with reports that stimulation of primary human B cells through CD40 induces strong expression of HSPGs

(van der Voort et al., 2000) and that α2-6-linked sialylated glycans are abundant on human tonsillar B cells (Macauley et al., 2015), the attachment of AAV6 to primary human B cells is unlikely to be a major bottleneck in AAV transduction. It is not clear why the anti-CD19 and anti-CD20 nanobodies failed to rescue AAV transduction. Without performing a binding assay to confirm binding, it cannot be confirmed that these nanobodies were functional once incorporated into AAV particles.

The binding/internalisation assay and nanobody experiment data together suggest that the impairment in B cell transduction efficiency is due to bottlenecks in endosomal trafficking and escape, nuclear entry or uncoating, or increased proteasomal degradation and epigenetic silencing. AAV6.2FF (AAVTM), an AAV6 variant rationally designed for reduced proteasomal degradation, transduced EBV-negative and EBV-positive B cell lines with similar efficiency to wild-type AAV6 (Ahmadi et al., 2021) suggesting that proteasomal degradation is unlikely to severely limit B cell transduction. A study into the bottlenecks in dendritic cell transduction revealed that AAV appears to traffic to the nucleus but is limited by impaired uncoating (Rossi et al., 2019). The transduction efficiency of an AAV variant with reduce stability and improved uncoating improved transduction efficiency in monocyte-derived immature and mature DCs, but overall transduction efficiency remained very low. However, when DCs were matured after incubation with AAV, transduction efficiency was drastically improved. As speculated by Rossi *et al*, this may be explained by changes in the endosomal environment that accompany DC maturation and are more favourable for AAV transduction, while mature DCs may have a reduced capacity for particle uptake. Recent biophysical studies point to a requirement for unfolding of the VP1 N-terminal region, normally associated with pH

changes involved in endosomal escape and exposure of nuclear-localisation signals (Penzes et al., 2021, Venkatakrishnan et al., 2013), to facilitate genome release (Yamaguchi et al., 2025, Gliwa et al., 2025). This may provide some mechanistic explanation for the observation by Rossi *et al* that uncoating appeared defective in DCs. We suspect that as antigen presenting cells, the endosomal environment is suboptimal for AAV transduction in B cells and, like DCs, may prevent conformational changes in the AAV capsid which are critical for endosomal escape, nuclear localisation and possibly uncoating.

Guided by our investigation into the bottlenecks in B cell transduction by AAV6, we produced two complementary AAV capsid libraries: a full-capsid shuffling library spanning AAV1–AAV9 and a focussed mutagenesis library (CLV TM) targeting the VP1u-VP2/3 regions. Similar shuffling strategies have been used successfully in other contexts to generate high-diversity capsid pools with improved tropism or immune evasion compared to parental sequences (Li et al., 2008, Grimm et al., 2008, Lisowski et al., 2014). To our knowledge, the random mutagenesis library presented here is novel and has not been reported before. However, supporting our rationale, one study grafted combinations of VP1u and VP2/3 regions from AAV4, -5 ,-11 and 12 onto the AAV6 sequence (Viney et al., 2021). AAV6 variants with VP12u-VP2/3 regions from AAV4, -11 and -12 enabled significantly higher transduction efficiency in CD8+ T-cells and CD34+ hematopoietic stem cells. The increased transduction efficiency came at the cost of reduced viral titres, which could be partially restored by changing an 8-residue stretch of amino acids to match AAV6 AAP (AAP6) or by supplying AAP6 in trans. While restoration of the AAP sequence in the variant capsid sequence reduced transduction efficiency

approximately 10-fold, providing AAP in trans largely rescued transduction efficiency. This study supports our reasoning that mutagenesis of the N-terminal region of the AAV6 capsid may yield more efficient AAV capsids.

To produce a shuffled capsid library, we used a classic DNA shuffling process (Stemmer, 1994) that has been adapted for the production of shuffled AAV capsid libraries (Grimm et al., 2008). The capsid shuffling process yielded sequences with similar crossover rates and crossover distributions as reported by others (Koerber et al., 2008a, Pekrun et al., 2019). A limitation of this method is that it relies on stretches of DNA sequence with high sequence similarity between natural serotypes to create crossover points. This is illustrated by the tendency of crossover points to occur in the 5' half of the capsid DNA sequence (Figure 19a). This means that features in the 3' half of the DNA sequence are more likely to be inherited together, reducing functional diversity. We found that the viral yield for the shuffled library was lower than wild-type AAV6 production. This is expected given that random shuffling is likely to result in chimeras with impaired assembly and stability. The desire to maximise functional diversity within capsid libraries with retrained sequence diversity has led to the development of machine-learning (ML)-guided library designs. These use ML to build predictive models, based on experimental capsid stability or tropism data, that can be used to enrich libraries with sequences that are likely to be functional (Zhu et al., 2024). However, this field is still in its infancy and requires significant investment to generate high quality and appropriate training data.

To produce the targeted mutagenesis library, mutagenesis was first performed for AAV6 capsid sequences encoded in a RepCap helper plasmid. This allowed us to produce eGFP-encoding AAV libraries to assess the effect of mutagenesis on AAV transduction. We anticipate that the optimal mutation rate for screening in primary B cells must strike a balance between generating enough diversity to capture beneficial mutations and preserving a sufficient proportion of functional variants within the viral library. To guide mutation rate selection, we produced a gradient a mutation rates and assessed the average impact on functionality in the viral library in HEK293 cells. We chose to produce two full scale libraries using average mutation rates of approximately 5 (TM 4) and 15 (TM 7) mutations per clone.

The work by Viney *et al* also highlights our motivation for characterising the effect of suppling AAP in trans during library production. Interestingly, Viney *et al* found that the capsid sequence conferring enhanced transduction efficiency overlapped with the out-of-frame AAP sequence, as restoring this region to the AAP6 sequence yielded virus with wild-type AAV6 performance. This suggests that, in a high-diversity pooled library containing mutated or chimeric AAP sequences, variants with superior transduction potential could be underrepresented due to poor assembly. To explore the effect of trans AAP on the viral yield using our libraires, we produced and validated AAV helper plasmids in which the capsid sequences has been disrupted without affecting the AAP coding sequence. All shuffled variants sequences by sanger sequencing contained at least one cross-over point within the AAP sequence (sequence position 530 – 1141 using AAV2 VP1 numbering). Interestingly, we saw no difference in shuffled library viral titre between AAV produced in the presence or absence of trans AAV2 AAP (AAP2). This was true for both

high genome input and low genome input conditions that replicate typical full-scale library production conditions. AAP2 was chosen because it improves virus production across a wide range of AAP-deficient wild-type capsid sequences (Grosse et al., 2017). Our findings are consistent with studies showing that supplementation of AAV productions using randomly selected chimeric capsid sequences with a pool of wild-type AAPs appeared to have no effect on AAV production for 18 of the 19 functional chimeras (Herrmann et al., 2019). Interestingly, the yield or transduction efficiency of one variant was markedly improved through trans supplementation with wild-type AAPs. This suggests that while there is no effect on virus production for most chimeric variants, supplementation with wild-type AAP sequences may rescue a small number of defective sequences. In contrast to the effect of trans AAP on the shuffled capsid library yield, we observed a small negative effect on the production of our targeted mutagenesis libraries in the presence of AAP that was dependent on the genome input mass during production. When equimolar amounts of the CLV TM 7 library, REP-MAAP(-AAP) and adenoviral helper plasmids were used, there was a modest decrease in viral library yield. Under these conditions the yield for the lower mutation rate library, CLV TM 4, was unaffected. When low genome input conditions were used CLV TM 4 exhibited a similar decrease in vector yield in the presence of trans AAP2. The negative effect on yield of adding wild-type AAP could suggest that mutant AAP and wild-type AAP heterodimers tend to be more deleterious than mutant or wild-type AAP homodimers, perhaps due to a requirement for a symmetrical interface. The relatively small effect size and low statistical power of these experiments mean that more data is required to be confident that the effect observed represents a true result. With these limitations in mind, we cautiously proposed a model which relies on the fast turnover of wild-type AAP, limiting its accumulation during the

AAV production. According to this model, the delayed availability of capsid protein and mutant AAP due to a low AAV genome input may not afford wild-type AAP an advantage over defective mutant AAPs. As a result, the amount of genome input would not be expected to affect the ratio of wild-type AAP to mutant AAP and the only variable between the high and low genome input conditions becomes the time available for particle assembly. Therefore, when high genome input is used, the amount of capsid could accumulate quickly resulting in a large time window for assembly. Consequently, higher mutation rates may be required to detect the deleterious effect of the mutated AAPs. When low genome input is used, we expect that capsid proteins accumulate more slowly resulting in a short window for assembly. Therefore, lower mutation rates may be required to detect a reduced viral yield. The turnover rate of AAP in mammalian cells is not well defined. AAP appears to be turned over relatively quickly in insect cells, which cautiously points to possible inherent instability (Grosse et al., 2017). Measurements of AAV RNA and protein levels during AAV production in HEK293 cells suggest that AAP accumulates between the 24-hour and 72-hour timepoints tested (Lu et al., 2024). However, capsid proteins also accumulated during the same period and may mutually stabilise AAP. Measurements of AAP protein levels in cells transfected with Rep-MAAP-AAP at multiple time points following transfection may support this model if AAP levels rapidly reach a steady state. If the steady state is reached before AAV genomic DNA and capsid protein accumulation during a CLV TM4 library production under the low genome input conditions, this data would add confidence to our model.

Prior to full scale production of our viral libraries, we measured cross-packaging using an NGS-based assay which aims to replicate the library production conditions. Other

researchers have attempted to quantify cross-packaging during AAV production using qPCR-based measurements of genome identity in following physical separation based capsid-antibody binding (Nonnenmacher et al., 2015a), or detection of GFP reporter sequences cross-packaged into a single capsid (Schmit et al., 2020). Our finding that cross-packaging appears to be surprisingly low in viral preparations produced using a normal amount of genome input is consistent with Nonnenmacher *et al* when using a similar plasmid configuration. The significant decrease in cross-packaging when using low amounts of AAV genome plasmid suggests that cross-packaging is unlikely to add significant noise to AAV capsid library screening results. Mosaic capsid analysis on a simple mixture of AAV capsids using differential scanning fluorimetry also adds confidence that reduced genome inputs during capsid library production limits mosaicism.

Lastly, we produced large scale viral libraries and performed screens in pre-stimulated primary human B cells. Selection during AAV capsid library screens is typically performed using relatively high MOIs between 10^2-10^5 (Hüttermann et al., 2025, Giacomoni et al., 2024, Schröder et al., 2025, Pekrun et al., 2019). Given that B cells appear particularly resistant to transduction by AAV, we chose a high MOI to achieve high coverage of each AAV variant during screening to provide the best chance of recovering rare variants. Assuming that each transformant during library production contains a unique variant, the quantities of virus and cells used would produce a coverage of 210-378 instances of each viral variant per screen, depending on the library. The diversity is likely far below the transformant count due to the dropout of non-functional variants during viral library production, and due to duplicate variants in the plasmid library. The duplication of

variants is more probable for the shuffled library where the semi-random homology driven assembly process is more likely to yield duplicates within a large library than through a fully random mutagenesis process. Therefore, the coverage of each viral variant is likely higher than the numbers based on the transformant counts.

Due to alternative splicing of the capsid transcripts, we tested two alternative methods to selectively recover full length transcripts for sequencing. Plad-B, a potent RNA splicing inhibitor, was added to the primary B cell cultures prior to RNA extraction to enrich unspliced capsid transcripts. We found that this a viable approach but requires optimisation. Using our capsid library plasmid configuration, the abundance of unspliced transcripts proved to be sufficient for rescue by PCR using intron-specific primers without the need for a spliceosome inhibitor. We were able to successfully rescue transcribed capsid variant sequences from libraries screened in primary human B cells for the shuffled and CLV TM 4 libraries. We suspect that the mutation rate used for the CLV TM 7 library was too deleterious to transduction to recover capsids containing potentially beneficial mutations.

As an initial validation step, we isolated capsid sequences through re-cloning and bacterial transformation, then analysed using nanopore sequencing. Unexpectedly, the wild-type AAV5 capsid sequence was recovered. AAV5 represents the most divergent of the natural serotypes, with few regions of homology sufficient to support priming between other serotypes during capsid shuffling. Therefore, it is possible that the full-length AAV5 sequence is over-represented in the viral library. Although AAV5 exhibits very poor transduction efficiency across most cell types (Ellis et al., 2013), it appears capable

of low-level transduction in primary human B cells (Hung et al., 2018). Consequently, while wild-type AAV5 sequences may drop out during library screening for other cell types, recovery of the sequence in B cells appears possible. If this small sample is representative, it appears that wild-type AAV5 may constitute a substantial proportion of recovered sequences. NGS analysis of the viral library and recovered sequences will be necessary to determine the extent to which overrepresentation in the input library contributes to the recovery of this sequence following screening. The closely related AAV1 and AAV6 (clade A) sequences are well represented in this small sample of recovered sequences, consistent with wild-type AAV6 being the most efficient natural capsid in B cells. Both recovered variants containing AAV1- and AAV6-derived sequences also incorporated non-clade A regions in their N-terminal domains. If these variants support more efficient transduction than wild-type AAV6, their chimeric N-terminal regions may confer more favourable endosomal escape or nuclear import characteristics, given the established role of the N-terminal domain in these steps.

Due to time constraints, the following experiments were not completed within the scope of this thesis but represent logical next steps for further investigation. Following long-read next-generation sequencing of recovered capsid variants, the distribution of variant enrichment will be analysed to determine whether additional rounds of selection are required. The transduction efficiency in primary human B cells, and production characteristics, of the top performing capsid variants will then be compared directly with wild-type AAV6. If capsid variants demonstrating improved performance relative to wild-type AAV6 are identified, future work will focus on investigating the mechanisms underlying enhanced transduction efficiency, with the aim of defining key post-entry

bottlenecks in B-cell transduction and capsid–host interactions. In addition, although all three capsid libraries have been generated in a barcoded format, these barcoded libraries were not characterised within the timeframe of this study. Each capsid variant is uniquely labelled by a DNA barcode located in the 3' untranslated region, enabling capsid performance to be tracked at the RNA level using short-read next-generation sequencing platforms. This approach will support future complex screening experiments in novel fully human model systems.

10 Development of a B cell Engineering Strategy for Anti-cancer Antibody Expression

10.1 Introduction

The previous chapter described the design, development and screening of AAV capsid libraries designed to identify novel AAV capsids with improved transduction efficiency in primary human B cells. This was done with the aim to facilitate more efficient editing of B cells, with the aspirational aim to enable *in vivo* editing in humans. Given the recent progress in the editing of primary B cells to express therapeutic antibodies towards infectious diseases, we saw potential in adapting this work to enable the expression of antibodies used in the treatment of non-infectious diseases, in particular cancer. This chapter describes progress towards addressing some of the challenges posed by this goal.

Current work engineering primary B cells to express antibodies for the treatment of infectious disease uses a knock-in strategy that places a V_H region splice-donor immediately downstream of the knocked-in antibody V_H coding sequence. The knocked-in antibody sequence splices onto endogenous downstream constant regions enabling the B cell to select immunoglobulin (Ig) isotypes in response to infection through class-switching (Nahmad et al., 2020, Hartweger et al., 2019, Moffett et al., 2019). In the context of anti-cancer monoclonal therapeutics, the antibody isotype is carefully selected for therapeutic effect and constant region sequences are often manipulated to

improve stability or function. Therefore, to adapt current B cell engineering strategies for the knock-in of anti-cancer monoclonal antibodies, the monoclonal antibody constant region must be included in the homology-directed repair template (HDRT).

Naturally, the immunoglobulin (Ig) molecule acts as both a sensing membrane-bound receptor and as the secreted effector molecule for B cells. In engineering B cells to express anti-cancer monoclonal antibodies, the primary goal is for edited B cells to express the secreted antibody, however the membrane-bound receptor is critical to support B cell survival through tonic signaling (Lam et al., 1997, Kraus et al., 2004). Therefore, to enable long term survival of edited B cells, the knocked-in antibody sequence must contain the sequences required for the regulated expression of both the membrane-bound and secreted Ig forms. Furthermore, this technology could be used for the treatment of diseases where the level of antigen waxes and wanes during the course of disease development. Replication of the natural B cells ability to sense antigen and produce secreted antibody in response may allow engineered B cells to tune the amount of secreted antibody to the disease status.

The membrane-bound and secreted Ig forms are produced through competitive 3' end processing of a single IgH pre-mRNA transcript. In naïve B cells, which function primarily as antigen sensors, mRNA processing favours the membrane-bound Ig isoform; following activation and plasma cell differentiation, the balance shifts toward production of the secreted form. The constant region genes for each isotype encode exons that are common to both the membrane-bound and secreted forms of the Ig (Figure 29). The last of these common exons ends with an internal splice donor site (5'ss) followed

immediately by the coding sequence specific to the secreted form (C_S). The M1 and M2 exons of the membrane-bound Ig are located downstream and are separated from the C_S exon by a large intron (approximately 2kb for IgM) containing a weak polyA sequence (pA_S) that terminates the secretory mRNA, while M1 possess a 5' splice acceptor site (3'ss). The balance of secreted to membrane-bound mRNA is controlled by the competition between splicing and the weak polyA sequence (pA_S). When termination using pA_S dominates, polyadenylation removes the M1/M2 sequences from the mature mRNA, resulting in a secretory mRNA sequence. When splicing dominates, the C_S and pA_S sequences are removed to produce the membrane-bound Ig mRNA sequence (Peterson, 2011). The model for how the balance between the secreted and membrane-bound forms of Ig is poorly defined and constantly evolving. Several proteins involved in polyadenylation, splicing and transcription elongation have been implicated highlighting that the regulation involves multiple layers of control and involve global changes in core proteins associated with transcription and end processing. Key proteins include transcriptional elongation factor eleven nineteen lysine-rich leukaemia gene (ELL2) (Martincic et al., 2009), cleavage stimulation factor 64 KDa (CstF64K) (Takagaki et al., 1996), Heterogeneous nuclear ribonucleoprotein L-like (hnRNPLL) (Benson et al., 2012). To enable engineered B cells to express both membrane-bound and secreted antibody, this competitive end processing mechanism must be incorporated into the HDRT design. This is particularly challenging given the large size of the C_S-M1 intron and the limited packaging capacity of AAV vectors.

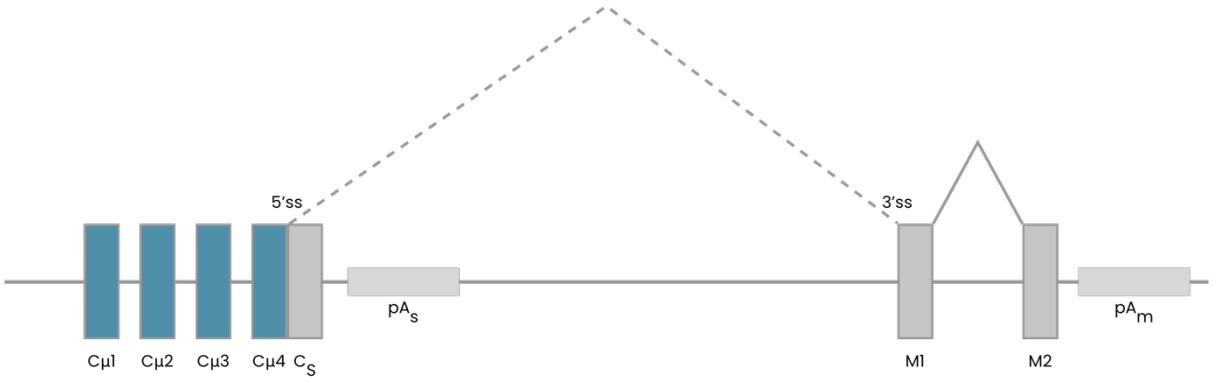


Figure 29 - Schematic of IgHM gene 3' end. *Abbreviations: Cμ1-4, μ constant exon 1-4; C_S, secretory sequence; ss, splice site; pA_s, secretory form polyA sequence; M1-2, transmembrane region exon 1-2; pA_m, membrane form polyA sequence. Solid line represents RNA splicing. Dotted line represents alternative RNA splicing.*

Beyond the need to encode both membrane-bound and secreted isoforms, an additional consideration is the choice of promoter used to drive antibody expression. The use of a cell type-specific promoter reduces the risk of antibody expression due to off-target integration. Expression of the natural antibody sequence is driven by IgH variable region (VH) promoters placed upstream of each variable segment in the IgH locus. VH promoters appear to be preferentially active in B cells based on plasmid transfection studies (Mason et al., 1985). The use of a VH promoter to drive expression of a monoclonal antibody in the genomic context may improve tissue specificity and drive the correct level of Ig expression based on B cell state.

In this chapter we aim to show that full length anti-cancer monoclonal antibodies can be expressed in both membrane-bound and secreted forms using the native 3' end processing mechanism of B cells, as well as a minimal version suitable for AAV vectorisation. Furthermore, we aim to show that monoclonal antibodies produced in this

way retain functionality. Finally, we aim to knock the validated antibody sequences into primary human B cells and validate the expression of surface Ig on activated B cells and the secretion of functional antibody upon differentiation into plasma cells.

10.2 Aims

1. Develop a knock-in strategy for full-length anti-cancer monoclonal antibodies at the IgH locus of primary human B cells that enables expression of both membrane-bound and secreted Ig.
2. Validate HDR template design in human cell lines, confirming production of both isoforms and demonstrating that performance is independent of antibody sequence.
3. Establish antibody functionality by testing HDRT-derived antibodies in appropriate cell-based assays for the anti-cancer antibody.
4. Demonstrate knock-in in primary human B cells and verify that the relative expression of membrane-bound and secreted isoforms reflects the stage of B cell activation and differentiation

10.3 Results

10.3.1 Validation of HDRT designs for the engineering of anti-cancer monoclonal-expressing primary human B cells.

To enable the knock-in of anti-cancer monoclonal antibodies into primary B cells, HDRTs were designed that included the heavy chain antibody constant region engineered to allow alternative splicing between membrane-bound and secreted forms by including the intron between the IgG secretory coding sequence (Cs) and IgG M1 exon (C_s-C_m) (Figure 30).

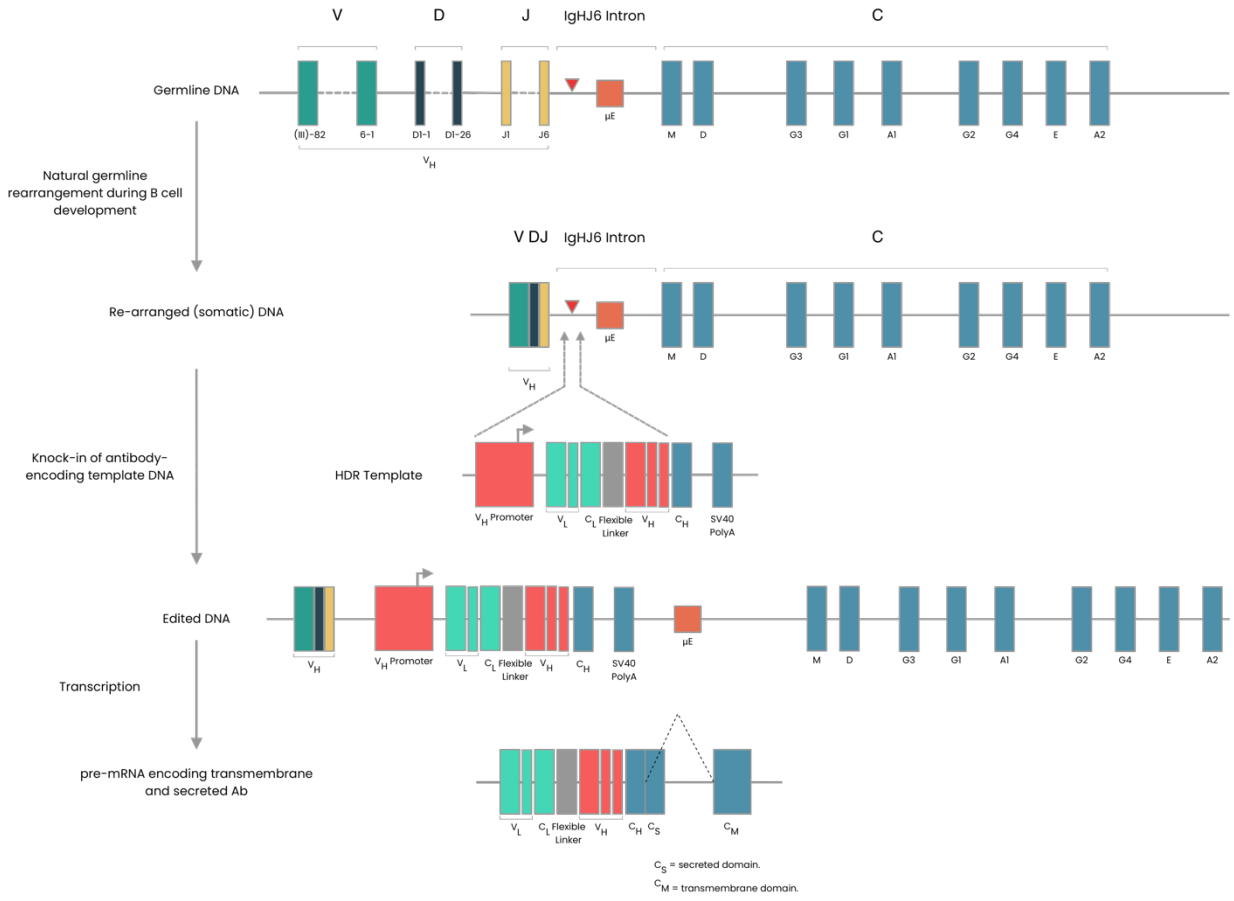


Figure 30 – Schematic representation of the human IgH locus and the gene editing strategy designed to knock-in an anti-cancer monoclonal antibody enabling membrane-bound and secreted antibody expression in vivo. *Abbreviations: V, variable region; D, diversity region; J, joining region; $V_{H,}$ heavy chain variable domain; μE, μ-enhancer; C, constant region; constant region exons for IgM (M), IgD (D), IgG3 (G3), IgG1 (G1), IgA1 (A1), IgG2 (G2), IgG4 (G4), IgE (E), IgA2 (A2); V_L, light chain variable domain; C_L, light chain constant domain; SV40, simian virus 40.*

To validate that both membrane-bound and secreted antibody could be expressed from the HDRT, the anti-PD-1 antibody Nivolumab was used as a proof-of-principle antibody sequence. Three different configurations were produced, each driving Nivolumab expression using a constitutively active spleen focus-forming virus promoter (Figure 31). A version containing the full C_s-C_m intron (Nivo-full C_s-C_m intron) was produced to replicate the natural splicing conditions, with the transmembrane exons (M1 and M2) fused into a single coding sequence. Due to the length of the full C_s-C_m intronic sequence, the version was approximately 5.2 kb in length which would exceed the packaging capacity of AAV. Therefore, a minimal intron version was created in which all DNA sequences were removed from the intronic sequence except for the sequences known to be involved in polyadenylation of the secretory coding sequence and DNA splicing between the secretory and membrane-bound immunoglobulin forms (AAUAAA, GU-rich downstream element, branch site, pyrimidine-rich tract). Finally, a variant encoding the membrane-bound form of Nivolumab was produced as a positive control for the detection of surface Nivolumab. This variant also included a P2A-linked eGFP reporter to allow to confirm transcription in the absence of detectable surface Nivolumab.



Figure 31 - Schematic representation of HDR template designs for antibody expression validation experiments. *ITR-ITR DNA length shown in brackets. Abbreviations: RHA, right homology arm; SFFV, spleen focus-forming virus promoter; Nivo, Nivolumab; VK, variable kappa chain; CK, constant kappa chain; VH, variable heavy chain; CH, constant heavy chain; C_m, transmembrane constant region exons; C_s, secretory coding sequence, FCS, furin cleavage site; P2A, porcine teschovirus-1 2A sequence; eGFP, enhanced green fluorescent protein; SV40, simian vacuolating virus 40; polyA, polyadenylation sequence; ITR, AAV inverted terminal repeat. Green and blue circles represent splice donor and splice acceptor sites respectively. The strep tag II linker consists of A Gly/Ser/Ala rich linker containing 3 strep tag II tags.*

To test for secreted and membrane-bound Nivolumab, 293-AAV cells were transfected with plasmid DNA encoding the three HDRT variants. 72 hours later, the frequency of cells expressing surface Nivolumab (Figure 32a), and eGFP-positive cells (Figure 32b), was assessed by flow cytometry. The frequency of Nivolumab-expressing cells was similar to the frequency of eGFP-positive cells, confirming efficient expression and trafficking of Nivolumab to the cell surface within transfected cells. 5-10 % of cells transfected with the minimal intron or full intron variants were positive for surface Nivolumab expression and there was no difference between the two intron variants. The flow cytometry data confirmed that the minimal and full intron variants could produce membrane-bound Nivolumab. To confirm that these HDRTs could produce both membrane-bound and secreted Nivolumab, a western blot was performed on cell lysates and cell media 72 hours post-transfection to detect the presence of Nivolumab (Figure 32c), using an

antibody that binds the tandem strep tag II tags within the flexible peptide linker. A version encoding the secreted version of Nivolumab was used as a positive control for Nivolumab detection in the cell media. Nivolumab was detected in both the cell lysate and cell media for the full intron and minimal intron variants, as well as the variant encoding secreted Nivolumab. The membrane-bound version was only detected in the cell lysate and the protein size difference between the membrane-bound and secreted version represents the additional protein sequence added by the M1-M2 exons. The membrane-bound form of nivolumab was not detected by western blot for the full intron or minimal intron variants in the cell lysates of transfected cells. Together these data show that both secreted and membrane-bound Nivolumab can be expressed from a single Nivolumab sequence using the HDRT designs developed here. In 293-AAV cells, the amount of the membrane-bound form appeared to be low compared to the amount of secreted Nivolumab, given that surface Nivolumab could be detected by flow cytometry but not by western blot. This is consistent with preferential use of the secretory polyA site over the use of the weak M1 splice acceptor in 293-AAV cells.

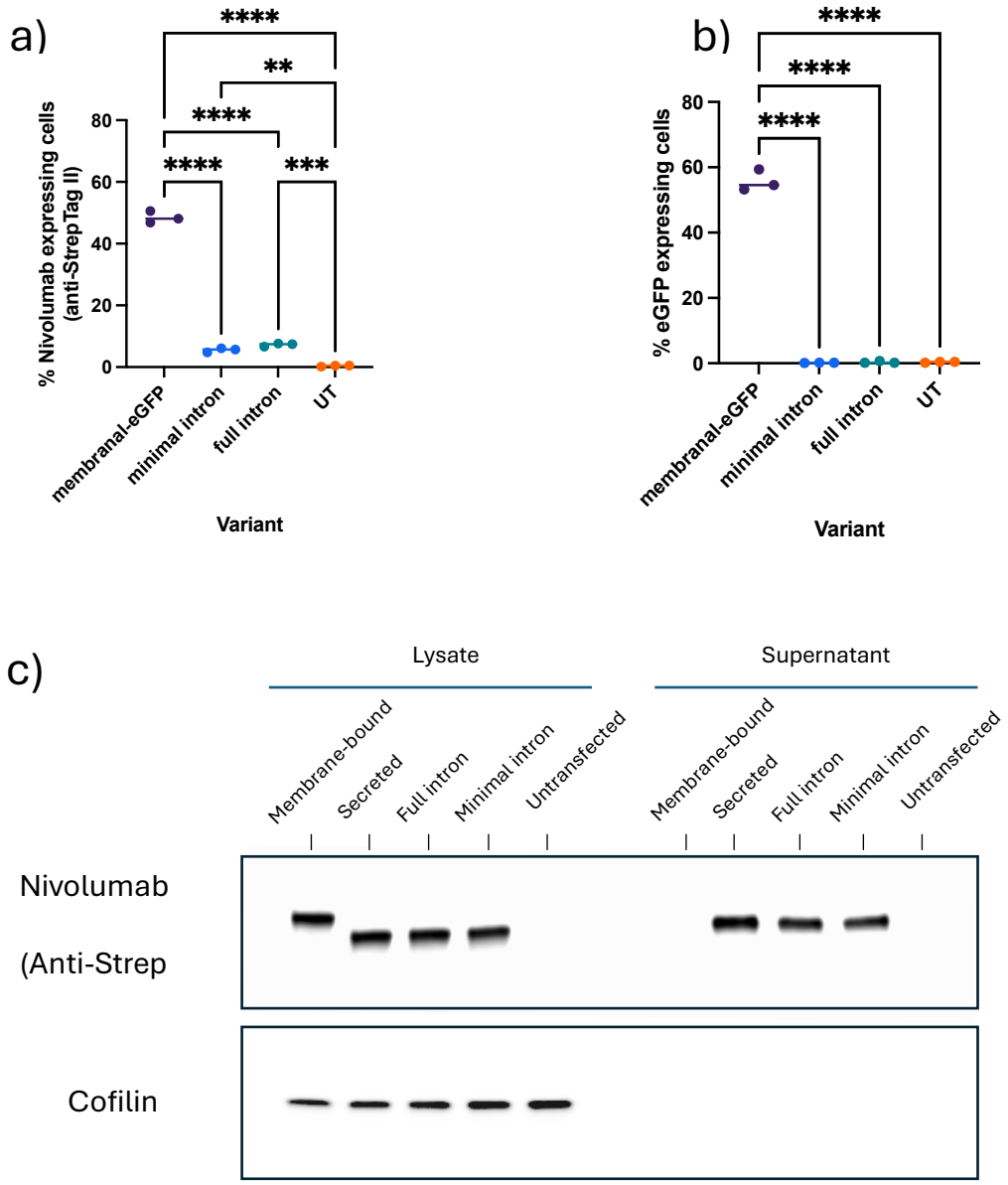


Figure 32 - Validation of HDR Template episomal antibody expression in 293-AAV cells. *293-AAV cells were transfected with plasmid DNA encoding Nivolumab homology-directed repair templates (HDRTs; Figure 31) and the frequency of surface Nivolumab-positive cells (a), as determined by surface Strep-Tage II detection, and eGFP-positive cells (b) was determined by flow cytometry 72 hours post-transfection. Cell lysates and cell culture media was also analysed, by western blot, for the presence of nivolumab. Cofilin detection was used as a loading control (c). Untransfected (UT). Each point represents a single biological replicate. Statistical significance was assessed using one-way ANOVAs with correction for multiple post-hoc comparisons using Tukey's test. ****P ≤ 0.0001 , ***P ≤ 0.001, **P ≤ 0.01.*

Trastuzumab, an anti-HER2 antibody commonly used in the treatment of HER2+ breast cancer (Slamon et al., 2001b), was chosen as an antibody sequence for proof-of-principle experiments due to its multiple complementary effector functions, enabling the use of functional assays to test both cell-mediated and direct antibody functions. To validate that the HDRT design could be adapted for the expression of Trastuzumab, the Nivolumab antibody sequences were exchanged for Trastuzumab and the IgG4 sequences (M1-M2) were exchanged for IgG1 to ensure compatibility with the IgG1 Trastuzumab sequence. 293-AAV cells were transfected with the trastuzumab minimal intron variant, trastuzumab full intron variant or nivolumab minimal intron variant and the presence of antibody in the cell lysate of cell media was assessed 72 hours later by western blot (Figure 33). Trastuzumab was detected in both the lysate and cell media using the minimal intron or full intron variants. The level of antibody appeared higher for the trastuzumab minimal intron variant compared to the nivolumab minimal intron or trastuzumab full intron variants.

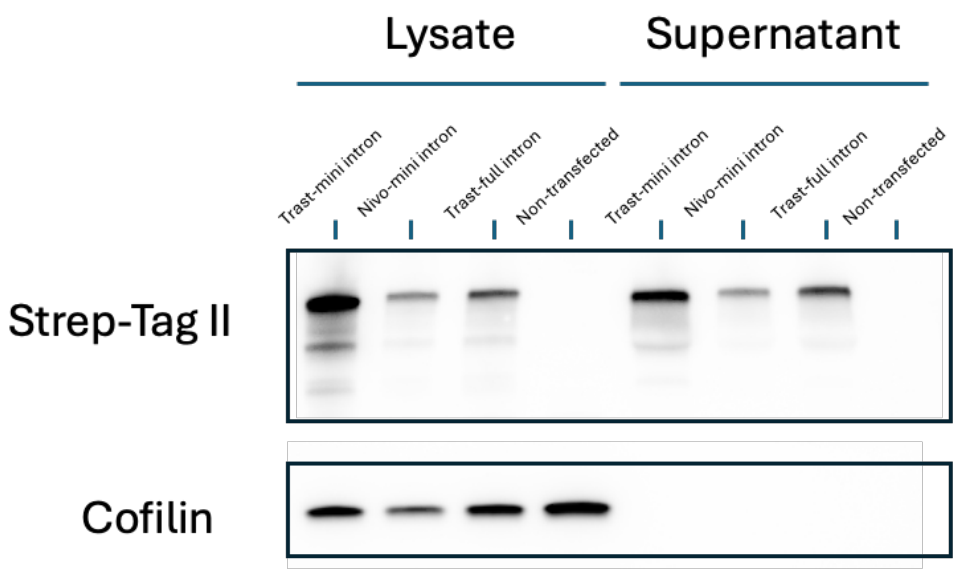


Figure 33- Validation of HDR template episomal Trastuzumab expression in HEK293 cells. *293-AAV cells were transfected with plasmid DNA encoding Trastuzumab or Nivolumab homology-directed repair templates. Cell lysates and cell culture media was analysed, by western blot, for the presence of trastuzumab or nivolumab. Cofilin detection was used as a loading control.*

To validate that the secreted trastuzumab was functional, trastuzumab-containing cell culture media from 293-AAV cells transfected with the trastuzumab minimal intron HDRT was collected 72 hours post-transfection. The trastuzumab was purified by gravity flow affinity chromatography via capture of the strep tag II tags within the flexible linker. BT474 cells, a HER2-amplified human breast cancer cell line with high HER2 expression levels (Figure 34b), were used to confirm that the purified trastuzumab selectively bound HER2 present on the surface of target cells. To do this, trastuzumab was pre-incubated with different concentrations of soluble HER2 (sHER2) to block the HER2 antigen-binding site. BT474 cells were then incubated with 40 nM purified trastuzumab and detected using an anti-strep tag II antibody (Figure 34c). Titration of sHER2 showed that at 80 nM or higher

concentrations of sHER2, trastuzumab binding was indistinguishable from the no trastuzumab or unstained controls, reflecting blocking of both HER2 binding sites on each trastuzumab molecule. This confirmed that trastuzumab binding was HER2-specific. Next the cell-mediated functionality of trastuzumab was assessed using a flow cytometry-based antibody-dependent cell-mediated phagocytosis assay (ADCP) using human peripheral blood mononuclear cell-derived macrophages (Figure 34a). In this assay, purified trastuzumab, or commercial anti-CD47 as a positive control was pre-incubated with or without 80 nM sHER2. CD47 is a ubiquitously expressed transmembrane protein that acts to inhibit the phagocytosis of healthy cells. The antibodies were then added to CFSE-stained BT474 cells to allowing binding. The BT474 cells were then mixed with human macrophages at a 1:1 ratio and incubated for 3 hours at 37 ºC. Finally, phagocytosis was determined using flow cytometry to measure the frequency of CFSE-positive macrophages (Figure 34d). As expected, the BT474 cells incubated with anti-CD47 antibody in the absence of sHER2 were readily phagocytosed compared to the no antibody no sHER2 control. The presence of sHER2 did not affect ADCP of BT474 cells in the anti-CD47 or no antibody samples. The phagocytosis of BT474 cells incubated with trastuzumab in the absence of sHER2 was indistinguishable from the anti-CD47 controls. However, trastuzumab-mediated ADCP was abrogated in the presence of sHER2 to levels equivalent to the no antibody controls, confirming that the ADCP effect for the purified trastuzumab was mediated through its interaction with surface HER2 on the BT474 cells.

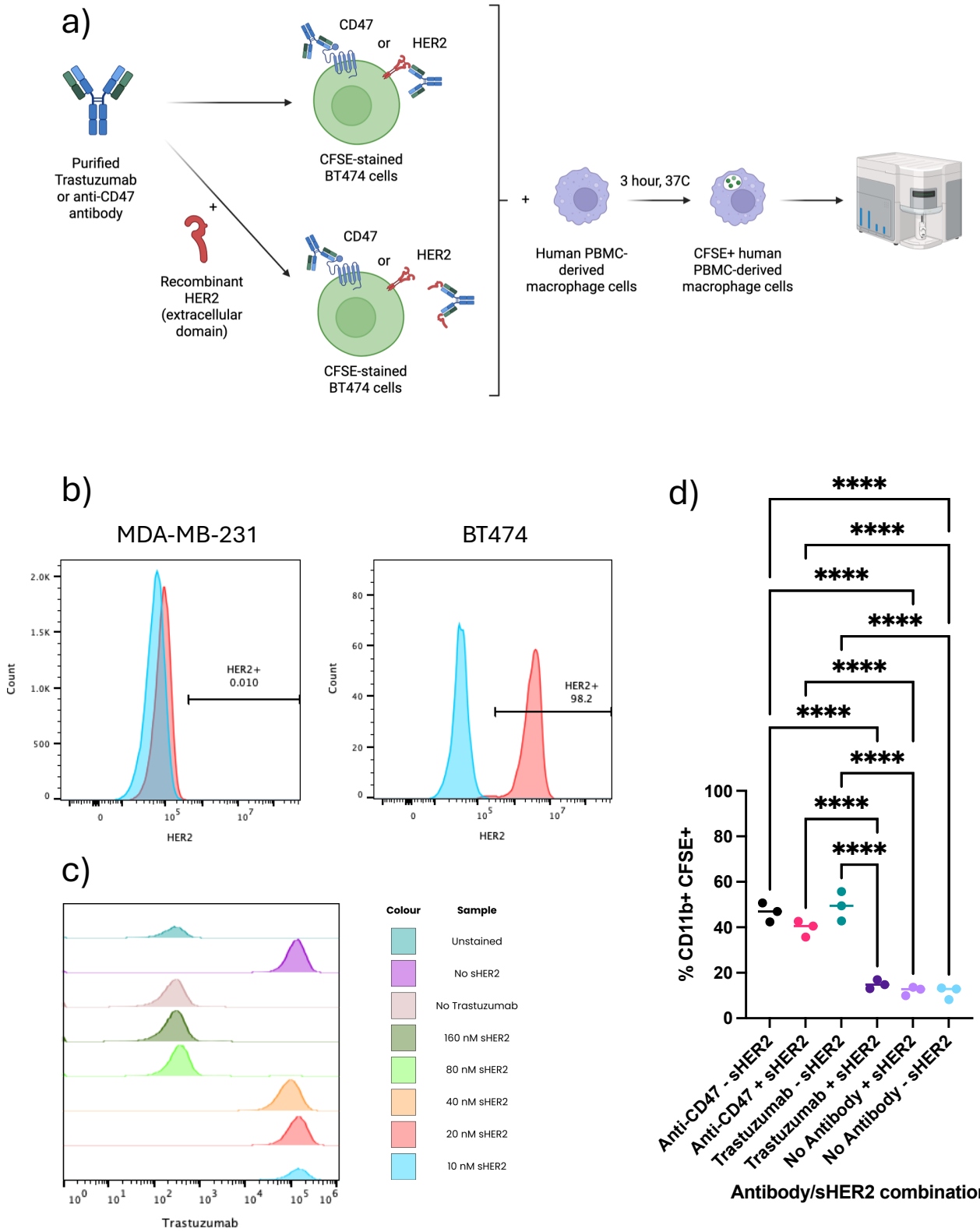

a)
CD47
or
HER2
Purified Trastuzumab or anti-CD47 antibody
CFSE-stained BT474 cells
+
Recombinant HER2 (extracellular domain)
CD47
or
HER2
CFSE-stained BT474 cells
+
Human PBMC-derived macrophage cells
3 hour, 37C
CFSE+ human PBMC-derived macrophage cells
b)
MDA-MB-231
BT474
Count
HER2+ 0.010
HER2+ 98.2
HER2
c)
Trastuzumab
Colour
Sample
Unstained
No sHER2
No Trastuzumab
160 nM sHER2
80 nM sHER2
40 nM sHER2
20 nM sHER2
10 nM sHER2
d)
% CD11b+ CFSE+
****
Anti-CD47 - sHER2
Anti-CD47 + sHER2
Trastuzumab - sHER2
Trastuzumab + sHER2
No Antibody + sHER2
No Antibody - sHER2
Antibody/sHER2 combination

Figure 34 – Functional validation of purified Trastuzumab produced using SFFV-Trastuzumab HDRT. *A) Schematic representation of antibody-dependent cell-mediated phagocytosis (ADCP) assay to functionally validate Trastuzumab purified from 293-AAV cells transfected with SFFV-Trastuzumab-minimal intron HDRT plasmid DNA. Purified Trastuzumab, or commercial anti-CD47, was incubated with or without soluble HER2. CFSE-stained BT474 cells were then incubated with the antibodies before mixing 1:1 with PBMC-derived human macrophages for 3 hours at 37 ºC. The frequency of CFSE-positive macrophages (CD11b+CFSE+) was determined by flow cytometry. Created with BioRender.com B) validation of HER2 expression on BT474 cells by flow cytometry. Blue = unstained cells, red = cells stained for surface HER2 expression. C) Dose dependent interference, of 40 nM (3.22 µg/mL) Trastuzumab binding to surface HER2 on BT474 cells, using different concentrations of soluble HER2, measured using flow cytometry. D) The frequency of CD11b+CFSE+ cells following the ADCP methodology in (A), using different combinations of antibody and soluble HER2 (sHER2). Each point represents the mean of three technical replicates from a single donor, 3 individual donors in total. Statistical significance was assessed using one-way ANOVAs with correction for multiple post-hoc comparisons using Tukey's test. ****P ≤ 0.0001.*

10.3.2 Knock-in of DNA sequences into the IgHJ6 intron of primary human B cells

Efficient generation of double-strand breaks is a prerequisite for supporting HDR-mediated DNA knock-in. To achieve this in primary human B cells, guide RNA sequences were designed to target the IgHJ6 intron (Figure 35b) and were synthesised as CRISPR RNAs (crRNAs) for assembly into Cas9 ribonucleoprotein (RNP) complexes. Pre-stimulated primary human B cells were nucleofected with RNP complexes made using 2 different crRNA sequences (Figure 35a). After 72 hours, genomic DNA was isolated from the B cells and insertion-deletion (InDel) frequency was assessed using Inference of CRISPR Edits (ICE; Conant et al. (2022); Figure 35c). The RNP complex formed with crRNA 466 was able to support up to a 68 % editing frequency as inferred by ICE and performed significantly better than the 483 crRNA.

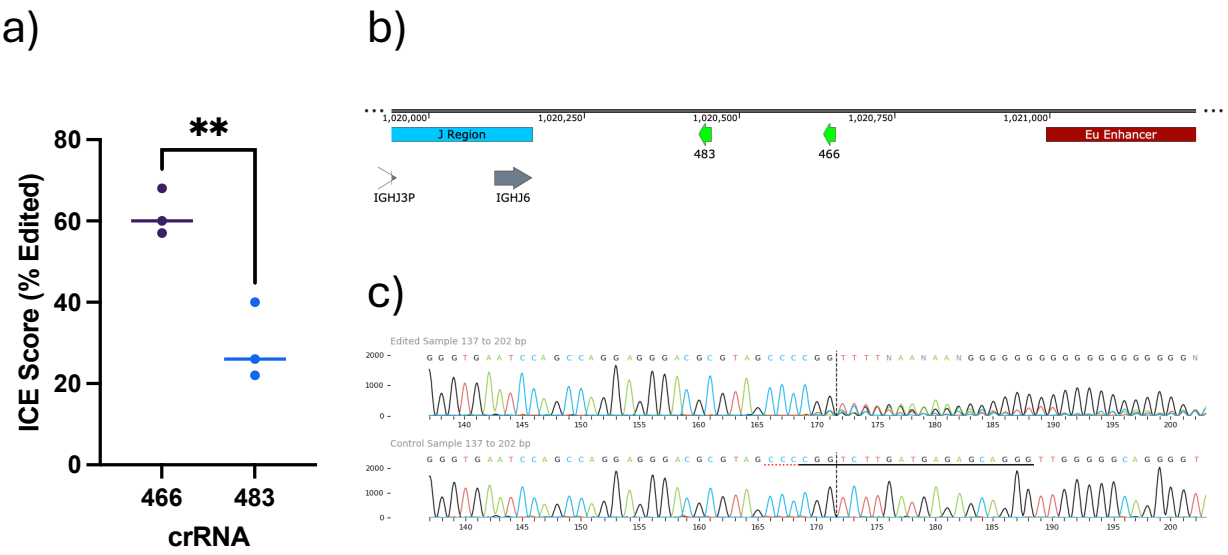


Figure 35 - Validation of IgHJ6-targeting guide RNAs by nucleofection of SpCas9 RNP in primary human B cells. *Pre-stimulated primary human B cells were nucleofected with SpCas9 RNP using 2 different crRNAs (b) and the InDel frequency was determined by ICE (a). C) a representative trace output using ICE for crRNA 466; showing the cut site (dashed vertical line), spacer sequence (black solid line) and protospacer adjacent motif (GGG, red dashed line). The crRNA sequence is oriented in the anti-sense direction. Each data point represents a separate nucleofection experiments using a single donor. Statistical significance was assessed using an unpaired t-test. **P ≤ 0.01.*

To validate targeted knock-in at the IgHJ6 intron in primary human B cells using the validated crRNA and the HDRT homology arms, an HDRT encoding eGFP driven by an SFFV promoter was cloned into an AAV genome plasmid. The use of a constitutive promoter in this HDRT means that unedited transduced cells will express eGFP. However, the IgH enhancer elements near the genomic insertion site should significantly boost eGFP expression levels in edited cells. The HDRT was packaged into high-titre AAV6 and purified by 2-phase partitioning using chloroform or by chloroform purification followed by purification over an iodixanol gradient. Purification using chloroform is frequently used for *in vitro* experiments using AAV, however purification over an iodixanol gradient

produces AAV preparations with exceptionally high purity (Kimura et al., 2019), which may support more efficient knock-in when using AAV to deliver HDRT sequences. Pre-stimulated primary human B cells were nucleofected in the presence or absence or SpCas9 protein complexed with crRNA 466 (RNP:466) and immediately transduced with $1x10^5$ vg/cell of chloroform or iodixanol purified SFFV-eGFP AAV6 containing homology arms for the crRNA 466 cut site. The homology arms were designed so that insertion of the DNA sequence would split the crRNA recognition sequence and remove the protospacer adjacent motif (PAM), preventing editing by residual RNP:466 following HDR. 72 hours post-editing, the cells were imaged to assess eGFP expression levels (Figure 36). Weak eGFP expression was observed in cells transduced with iodixanol purified SFFV-eGFP AAV6 in the absence of RNP:466. The signal from cells nucleofected in the presence of RNP:466 was markedly stronger suggesting successful integration into the IgHJ6 intron. Notably, the iodixanol purified virus appeared to support higher knock-in efficiency compared to the chloroform purified virus. However, the samples receiving iodixanol purified virus showed more robust cell clustering under activating conditions and therefore the increased eGFP signal could be due to greater proliferation of edited cells. Importantly, both the chloroform and iodixanol purified samples were buffer exchanged and therefore the difference in performance was unlikely to be due to the toxicity of residual chloroform in the chloroform purified sample.

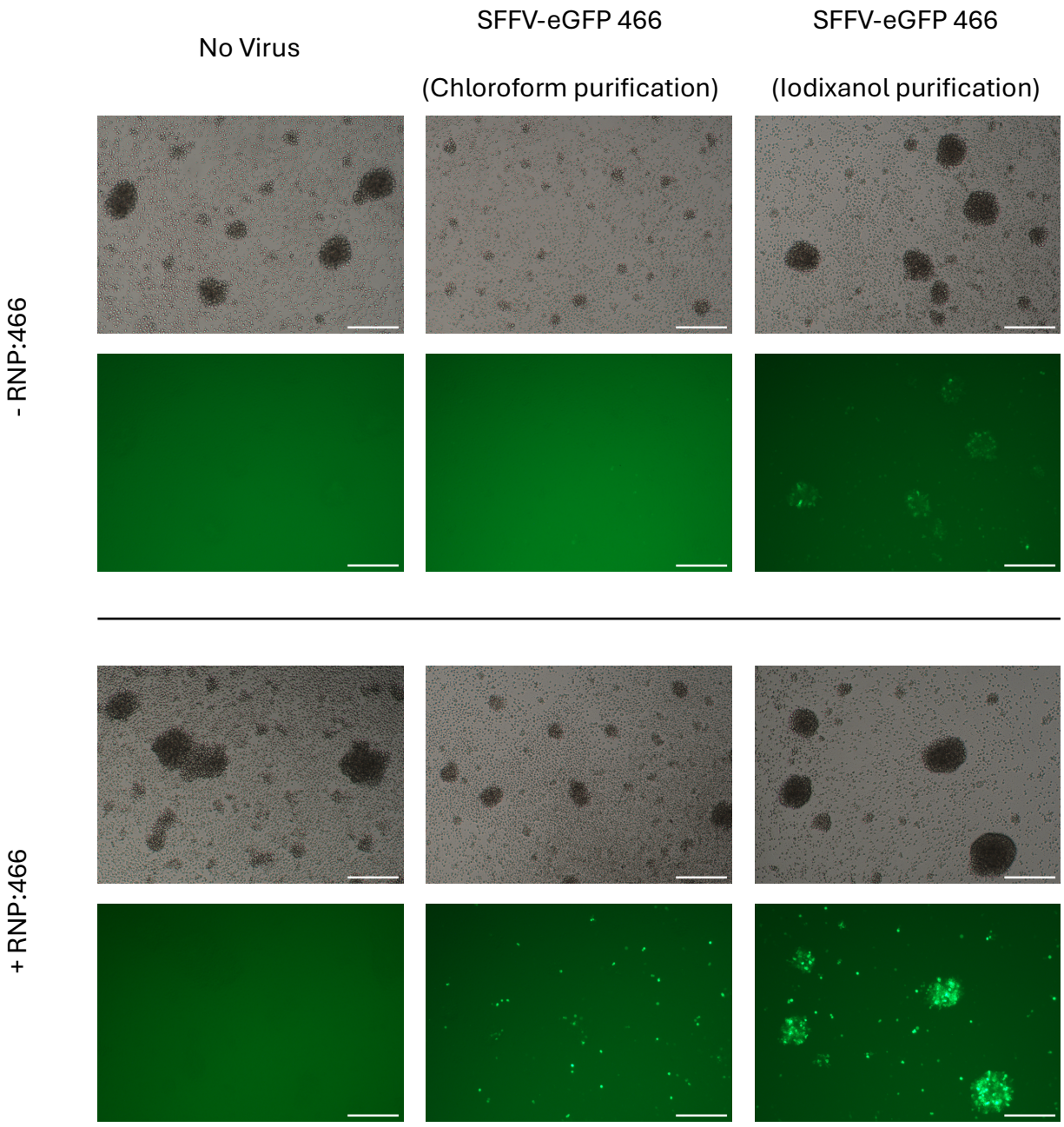
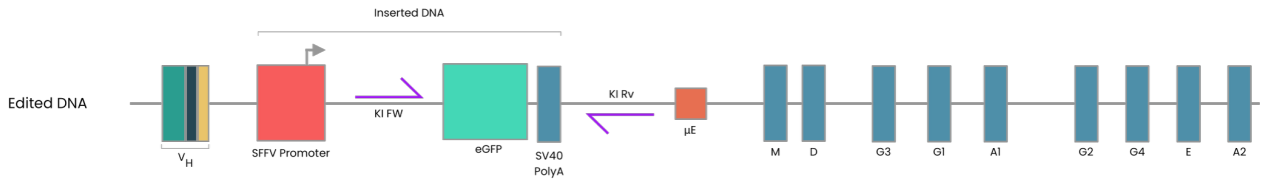


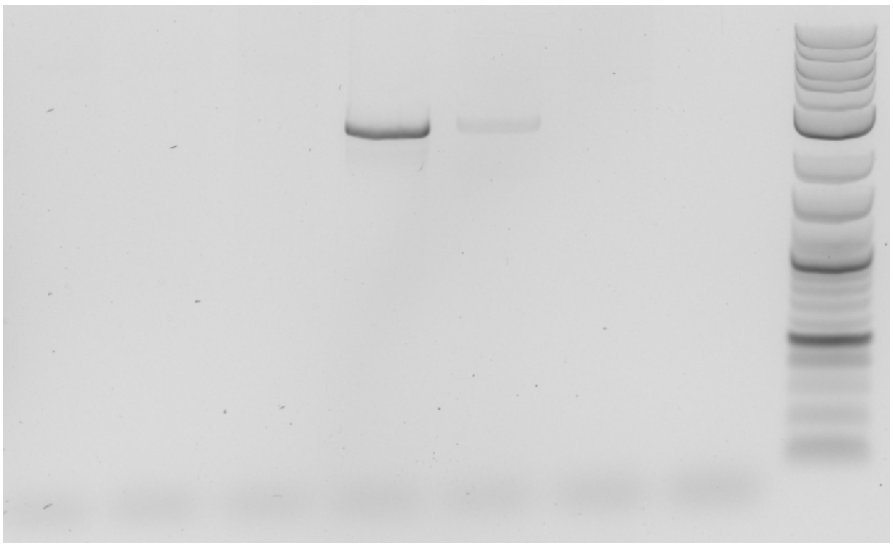
Figure 36 - Knock-in of SFFV-eGFP into pre-stimulated primary human B cells using nucleofection of SpCas9 RNP and AAV6-mediated HDRT delivery. *Pre-stimulated primary human B cells were nucleofected with or without SpCas9 RNP:466 and immediately transduced with 1x10^5 vg/cell of AAV6 encoding an SFFV-eGFP HDRT sequence. The AAV6 was either purified by 2-phase partitioning using chloroform, or by chloroform purification followed by purification over an iodixanol gradient. 72 hours post-transduction, the cells were imaged to assess eGFP expression. Scale bar 1000 µm.*

To provide more conclusive evidence for the on-target integration of the SFFV-eGFP sequence, genomic DNA was isolated from the edited B cells and a PCR was performed using primers that flank the right-hand junction between the HDRT sequence and the genomic sequence (Figure 37a). Using gel electrophoresis to visualise the PCR products, no amplicons were detectable for any of the samples nucleofected without RNP:466 as expected. Amplicons were detected for both samples nucleofected with RNP:466 and transduced with SFFV-eGFP AAV6. In line with the images in Figure 36, there was a stronger amplicon signal for the cells transduced with iodixanol purified AAV6 compared to the chloroform purified AAV6. The PCR data confirms that RNP:466 and the HDRT homology arms enable knock-in of DNA sequences into the IgHJ6 intron in pre-stimulated primary human B cells. Furthermore, it confirms that the weak eGFP expression observed in the cells transduced with iodixanol purified SFFV-eGFP AAV6 was due to expression from episomal AAV genomes and not on-target integration.

a)



b)



RNP:	-	-	-	+	+	+	-
AAV6:	+	+	-	+	+	-	-
Purification:	I	C	-	I	C	-	-

Figure 37 – Confirmation of on-target SFFV-eGFP knock-in in primary human B cells by In-Out PCR. *A) schematic representation of the edited IgH locus following the on-target knock-in of SFFV-eGFP, showing primers used for the validation of successful knock-in in purple. B) Following nucleofection and transduction with SFFV-eGFP HDRT AAV6, genomic DNA was isolated and the right-hand junction between the HDRT sequence and genomic sequence was amplified by PCR. The PCR product was visualised by gel electrophoresis, along with a 1 kb DNA ladder. Iodixanol purified (I), chloroform purified (C).*

To prevent the expression of the inserted monoclonal antibody sequence in the event of off-target integration and align the expression with natural b cell receptor and secreted antibody expression levels, a heavy chain variable region promoter (VH1-69) was selected due to its dependency on the proximal μE enhancer. To test that VH1-69 promoter activity required integration, the SFFV promoter of the SFFV-eGFP HDRT AAV genome plasmid was replaced with the full VH1-69 promoter sequence (**Figure 38**a). Iodixanol purified VH1-69-eGFP AAV6 was produced and used to transduce pre-stimulated primary human B cells following nucleofection with or without RNP:466. The cells were imaged 72 hours post-editing to assess eGFP expression (**Figure 38**b). Robust eGFP expression was clearly detected in the sample transduced with VH1-69-eGFP AAV6 and nucelofected in the presence of RNP:466. No eGFP expression was observed in the B cells transduced with VH1-69-eGFP and nucleofected in the absence of RNP:466. This confirmed that the VH1-69 promoter was unable to drive transgene expression from episomal AAV genomes and once integrated into the B cell genome, it yielded strong transgene expression.

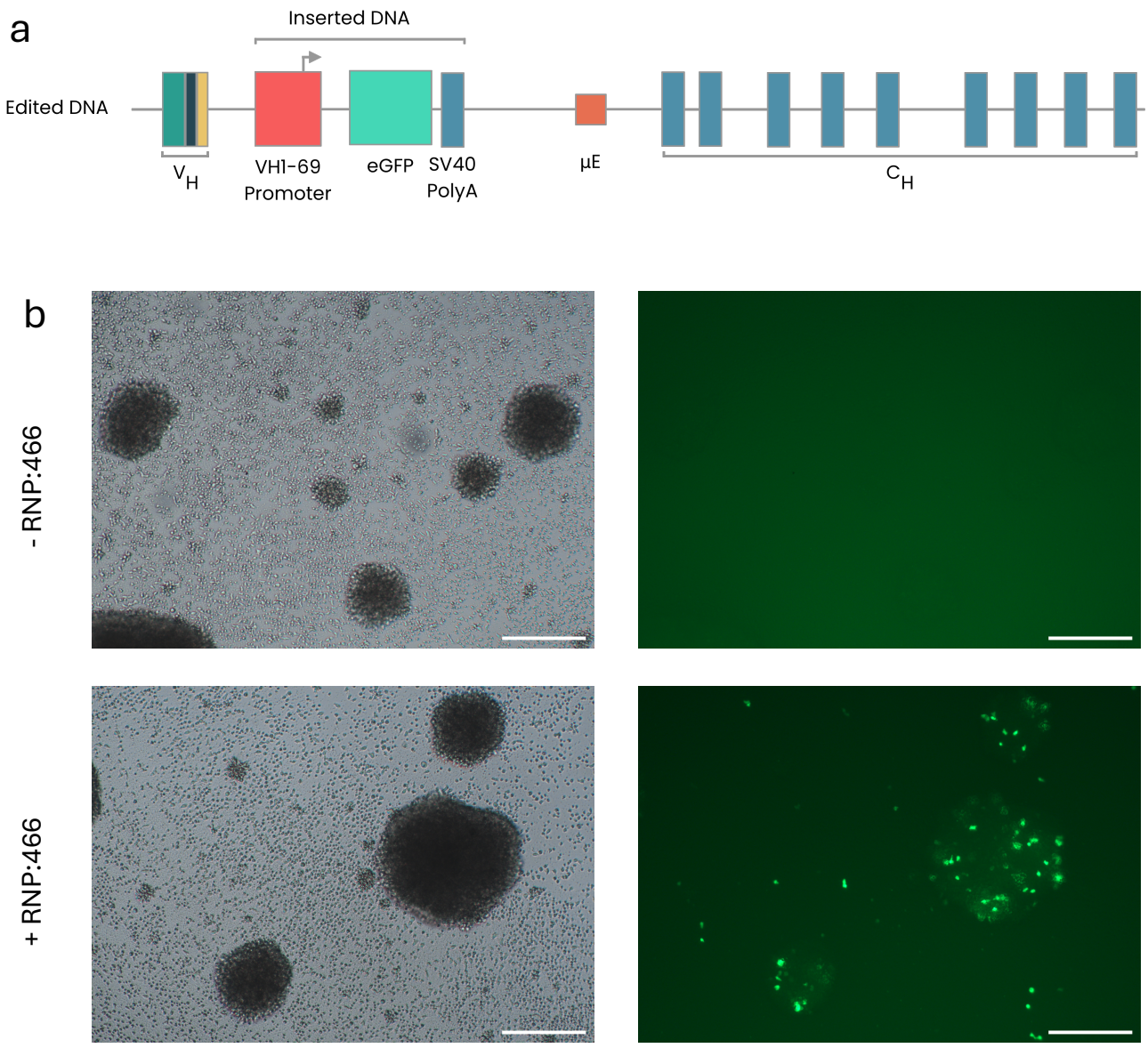


Figure 38 - Knock-in of VH1-69-eGFP into pre-stimulated primary human B cells using nucleofection of SpCas9 RNP and AAV6-mediated HDRT delivery. *Pre-stimulated primary human B cells were nucleofected with or without SpCas9:466 RNP and immediately transduced with 5x10^4 vg/cell of Iodixanol purified AAV6 encoding a VH1-69-eGFP HDRT sequence. 72 hours post-transduction, the cells were imaged to assess eGFP expression. Scale bar, 1000 μm.*

10.4 Discussion

In this chapter we developed a strategy for the knock-in of full-length anti-cancer antibodies into the IgH locus of primary human B cells. We demonstrated expression of the membrane-bound and secreted Ig forms of two independent antibody sequences, from a single HDRT, in HEK293 cells. Furthermore, we showed that HDRT-encoded trastuzumab was specific and functional using a cell-based ADCP assay. Finally, we laid the groundwork for the knock-in of anti-cancer antibodies into primary human B cells using Cas9 RNP and AAV6, and validated the integration-specific expression of a reporter under the control of the B cell-specific VH1-69 promoter.

Prior to the development of CRISPR/Cas9-mediated knock-in of antibody sequences into primary human B cells (Hartweger et al., 2019, Moffett et al., 2019), Fusil *et al* demonstrated that a neutralising anti-hepatitis C virus antibody could be expressed in both membrane-bound and secreted forms in primary B cells and B cell lines (Fusil et al., 2015). A lentiviral vector, pseudotyped with a baboon endogenous virus glycoprotein, was used to deliver the light and heavy chain sequences of the antibody separated by a F2A site. Earlier studies used artificial 2A (Yu et al., 2012) or splicing systems (Horlick et al., 2013) to force the simultaneous production of both membrane-bound and secreted antibody. These systems failed to replicate the regulation of Ig production based on B cell maturation. Constitutive expression of membrane-bound and secreted antibody may lead to suboptimal expression of both forms. Reduced surface Ig expression in mature B cells may lead to reduced tonic signaling and B cell survival, while suboptimal expression of secreted antibodies in plasma cells may reduce therapeutic efficacy. Fusil *et al* used

a minimal CH3-M1 intron, similar to the sequence designed for this thesis, to demonstrate preferential expression of the membrane-bound or secreted form based on the B cell maturation status. However, in contrast to random integration using lentiviral vectors, targeted integration into the IgH locus disrupts the endogenous heavy chain allele preventing inappropriate production of endogenous antibodies in engineered cells (Moffett et al., 2019). Additionally, targeted integration places the antibody in the natural genomic context which is likely to result in more robust secreted antibody production and correctly tuned surface Ig levels. The differentiation of mature B cells into plasma cells is accompanied by a drastic increase in Ig transcription (Kassambara et al., 2021). This effect appears to depend on the intronic 3' regulatory element (3'RR) located in the IgH constant region (Vincent-Fabert et al., 2010). Therefore, targeted integration harnesses the regulatory elements around the IgH locus which may boost monoclonal antibody production in plasma cells.

To validate the expression of secreted and membrane-bound forms of a monoclonal antibody, an HDRT was designed using the anti-PD-1 antibody nivolumab as a proof-of-principle sequence. In 293AAV cells the level of surface nivolumab, as measured by flow cytometry, was low for the HDRT versions engineered to splice between the membrane-bound and secreted Ig forms. This is likely a consequence of low CD79A/B expression in HEK293 cells and high 3' end processing due to HEK293 cell programming. CD79A and CD79B, which are not expressed in HEK293 cells (Indraccolo et al., 2002), are required in B cells for surface Ig expression (Huse et al., 2022). It is unexpected therefore that we observed surface nivolumab-positive cells following transfection of the membrane-only variant. In this construct, nivolumab is under the control of a strong SFFV promoter, so

high constitutive expression of nivolumab in the membrane configuration appears to overcome the requirement for CD79A/B. As previously discussed, the regulation of splicing between the membrane-bound and secreted forms of antibodies is tightly regulated in B cells, apparently through global changes in 3' end processing proteins that enable more efficient usage of the weak secretory polyA sequence. For example, it was shown that overexpression of CstF-64k was sufficient to switch the IgM transcript from the membrane-bound to secretory form (Takagaki et al., 1996). However, no clear model has emerged, and regulation likely involves many layers control. The high levels of secreted nivolumab and trastuzumab expression in HEK293 cells suggest that the HEK293 cell program supports efficient polyadenylation from the weak intronic polyA site. Interestingly, immunoglobulin secretion has been reported for several cancer cell lines (Zhao et al., 2021) illustrating that non-lymphoid cells can utilise the secretory polyA site.

Once we confirmed that the HDRT design was capable of producing both secreted and membrane-bound Ig, the functionality of trastuzumab was tested using a cell-based ADCP assay. Trastuzumab is a humanized IgG1 monoclonal antibody targeting the HER2 receptor and is widely used in the treatment of HER2-positive breast and gastric cancers (Hudis, 2007a). Proposed mechanisms of action for trastuzumab include direct inhibition of HER2 signaling (Yakes et al., 2002) and antibody-dependent cellular cytotoxicity (Barok et al., 2007). More recently ADCP has been proposed as a major mechanism of antitumour activity (Tsao et al., 2019). The data presented here confirms that the HDRT derived secreted trastuzumab bound specifically to HER2, using soluble HER2 to compete for trastuzumab antigen-binding sites. Furthermore, the Fc-mediated

functionality of the antibody appears intact. BT474 cells are sensitive to trastuzumab through direct inhibition of Her2 signalling, therefore, antibody-independent mechanisms such as efferocytosis of stressed or dying BT474 cells, rather than FcγR-mediated ADCP. No difference was observed in the frequency of BT474 cells staining positive using the cell membrane permeability dye (live/dead dye) between BT474 cells incubated with trastuzumab compared to no antibody control cells (Appendix Figure 39e). However, this is a marker of late apoptosis and clearance of apoptotic cells before the loss of membrane integrity occurs rapidly through the exposure of stress signals such as phosphatidylserine (PS) (Lauber et al., 2004). Others have observed no increase in PS exposure, as measured by Annexin-V staining, on BT474 cells treated with 10 ug/mL trastuzumab for 24 hours (Milella et al., 2004). Given the dose and exposure time used for the ADCP assay, it is unlikely that the trastuzumab-mediated phagocytosis observed in the ADCP assay was due to increased exposure of stress signals. To confirm this, the ADCP assay could be repeated with trastuzumab harbouring Fc mutations that eliminate Fcγ receptor binding (Wilkinson et al., 2021).

In preparation of CRISPR-mediated knock-in of DNA sequences into primary human B cells, we next designed and screened guide RNAs designed to initiate double-strand breaks within the IgHJ6 intron. Using guide 466, the indel frequency was similar to those achieved by others using Cas9 RNP to edit the IgHJ6 intron (Hartweger et al., 2019, Moffett et al., 2019).

Next, we produced AAV6 vectors encoding an eGFP reporter driven by the SFFV promoter, flanked by homology arms containing 340-380 bp of DNA sequence from either side of

the IgHJ6 cut site. This construct was designed as a positive control virus for transduction during editing experiments. However, following transduction of primary human B cells nucleofected with RNP:466 we observed a marked increase in GFP expression. This was likely due to the proximity of the B-cell-specific Eμ or 3'RR enhancer elements to the insertion site. We also noticed a pronounced difference in the GFP signal between AAV purified using 2-phase chloroform purification or AAV with an additional purification step over an iodixanol gradient. This difference was clear in eGFP expression using both the RNP:466 nucleofected and mock nucleofected samples, as well as by confirmation of the on-target knock-in by PCR. Chloroform purification provides a cheaper and much less labour-intensive method to produce high quality AAV that can perform as well as iodixanol purified virus (Davidsson et al., 2020). The difference in performance observed in the knock-in experiments is unlikely to be due to carry over chloroform due to the extensive buffer exchange step used. Davidsson *et al* measured a 5 % difference in the percentage of empty capsids, with chloroform purified AAV preparations containing more empty particles. Davidsson *et al* used the commonly utilised AAV RepCap helper plasmid configuration resembling the natural AAV genome in which the capsid gene is under the control of its native p40 promoter. Our helper plasmid uses a CMV promoter to drive capsid expression which can increase viral titres through elevated capsid expression (Vincent et al., 1997), but may come at the cost of increased levels of empty capsids if Rep protein and AAV genome levels limit genome loading into assembled capsids. Excess empty capsids in the chloroform-only purified sample could compete for attachment receptors or post-entry pathways.

Lastly, to validate that the VH1-69 promoter supports transgene expression following integration, primary B cells nucleofected with RNP:466 were transduced with iodixanol purified AAV6 encoding VH1-69-eGFP flanked by homology arms. Bright eGFP expression was observed in cells nucleofected with RNP:466, but no eGFP expression was detected in mock nucleofected cells. This confirms that integration is required for expression using the VH1-69 promoter, however, further experiments are required to show that the promoter is in active in non-B cells following both on-target and off-target integration. It has been shown that some VH promoters, including VH4-59 (Zhu et al., 2010) and VH6-1 (Wu et al., 2014), are active in non-B cell cancer cell lines. It must also be shown that VH1-69 is not active following off-target integration in B cells.

Since these experiments were conceived, work demonstrating the knock-in of anti-cancer antibodies using modern CRISPR/Cas9 approaches has been published (Ueda et al., 2024, Bracha et al., 2025, Boucher et al., 2025). Importantly, these studies use the same antibody configuration as earlier studies (Nahmad et al., 2020, Moffett et al., 2019, Hartweger et al., 2019), relying on endogenous constant domains and enabling class-switching. Bracha *et al* and Boucher *et al* performed complementary studies to demonstrate intact effector functions of anti-cancer antibodies produced by mouse (Bracha et al., 2025) and human (Boucher et al., 2025) engineered B cells. Both studies focussed on non-self antigens from human papillomavirus (HPV), with Boucher *et al* also using claudin-6 as an example oncofoetal membrane-bound antigen. Boucher *et al* demonstrated that BCRs engineered human B cells were responsive to antigen stimulation through detection of ERK phosphylation, and that engineered cells spontaneously secreted poly-isotypic antibodies. Using a similar ADCP assay to the

experiment presented in this chapter but using a monocyte cell line, secreted antibodies were shown to support ADCP. Furthermore, Boucher *et al* showed that the secreted antibody supported complement-dependent cytotoxicity and antibody-depended cell-mediated cytotoxicity (ADCC). Interestingly, the recombinant antibody was markedly more potent in the ADCC assay compared to the secreted antibody. It is not clear if this was due to differences in the Fc regions or the single chain configuration of the secreted antibody. Antigen-presenting B cells have been found to be co-enriched in the tumours of multiple cancer types with tumour-infiltrating T-cells and induce anti-tumour T-cell responses *in vitro* (Wennhold et al., 2021). Given the apparent role for B cells in antigen presentation to anti-tumour T -cells, Boucher *et al* assessed the ability of engineered B-cells to activate class-II-restricted engineered T-cells *in vitro* through BCR-mediated antigen-presentation. A significant increase in secreted IFNγ was measured using engineered B cells compared no mock-engineered B cells using soluble antigen or tumour cells. Bracha *et al* demonstrated similar results for IFNγ release from both CD4+ and CD8+ mouse T-cells. These studies demonstrate that secreted anti-cancer antibodies from engineered B cells appear to retain the ability to mediate a wide range of effector functions. However, these papers do not address critical questions that remain unanswered.

As durability is a key advantage of the transplantation of B cells expressing anti-cancer monoclonals over the passive transfusion of recombinant monoclonals, engineered B cells must be able to persist long term *in vivo*. In mice it has been shown that culture conditions that drive B cells towards plasma cell differentiation reduces the durability of transplanted B cells (Moffett et al., 2019). Ideally transplanted cells should resemble

naïve or memory b cells which have the potential to provide long-lasting protection and differentiate in vivo into antibody-secreting plasma cells (Nahmad et al., 2020). It is not clear which B cell phenotype provides engineered cells with the optimal durability and recall response *in vivo*. In mice memory B cell subpopulations behave differently upon activation with distinct T-cell help requirements and germinal-center involvement (Zuccarino-Catania et al., 2014). If T-cell help is required for the differentiation of engineered B cells *in vivo*, the degree to which natural T-cells can provide helper signals will depend on the antigen recognised by the B cell. Boucher *et al* and Bracha *et al* both demonstrated that engineered B cells can present antigen to CD4+ T-cells. However, both studies used T-cells artificially engineered to recognise the antibody-specific antigen. The co-enrichment of B cells and T-cells withing tertiary lymphoid tissues resembling germinal centres in tumours, and the correlation between these structures and patient prognosis (Cabrita et al., 2020), suggests that T-cell help can be provided to anti-cancer B-cells. This is likely highly dependent on the antibody and cancer type due to the requirement for antigens that are sufficiently different to self-antigens and for sufficient T-cell infiltration.

As previously discussed, the current approach to B cell engineering enables somatic hypermutation and class-switch recombination. For B cells engineered to express antibodies to treat infectious diseases, class-switch recombination may improve the therapeutic potential by allowing the immune system to select the appropriate isotype and achieve broad immunity across different tissues. However, for monoclonals designed for the treatment of cancers, the Fc domain is carefully selected and optimised. The use of endogenous constant domains makes the *in vivo* effects less predictable and

potentially less potent. Currently published work engineering B cells to express therapeutic antibodies do not address this issue and use the currently established knock-in strategy. Additionally, anti-cancer monoclonals are likely to be subject to somatic hypermutation (SHM), as has been observed for anti-HIV antibodies in engineered B cells (Nahmad et al., 2020, Huang et al., 2020). Further studies are required to determine whether SHM allows engineered B cells to prevent tumour cells acquiring resistance through antigen mutation.

In summary, this chapter aimed to adapt recent developments in the engineering of primary B cells to express monoclonal antibodies for the treatment of infectious disease towards the treatment of cancer. A primary limitation of current methods includes the delegation of isotype selection to engineered B cells, preventing the use of optimised monoclonal sequences and possibly leading to disappointing therapeutic effects. We sought to integrate the endogenous mechanism for splicing between the membrane0bound and secreted forms of Ig to enable the use of full-length antibody sequences. We demonstrated membrane-bound and secreted Ig could be produced from a single construct for two distinct anti-cancer monoclonal sequences, and that both specificity and Fc-mediated functionality was maintained. Finally, we laid the groundwork for testing the HDRT design in primary human B cells. Next we aim to validate that engineered B cells express membrane-bound trastuzumab and, following differentiation into plasma cells, functional secreted antibody.

11 Final Discussion

11.1 Rationale and Aims

The field of B cell engineering has developed rapidly since 2018 when precise editing of primary human B cells using CRISPR/Cas9 was demonstrated using electroporation of ribonucleoprotein (RNP) complexes. Since then, primary mouse and human B cells have been engineered to express predetermined antibodies, most of which were developed to treat infectious diseases. Seminal work in 2019 demonstrated that these engineered cells can participate in humoral immune responses in mouse models. The engineering approach combined modern genome editing technologies with design principles reminiscent of early 2000s strategies used to generate murine models for fundamental studies of B cell biology (Phan et al., 2003). New applications for engineered B cells are emerging and given the extensive clinical use of monoclonal antibodies for the treatment of cancer, the use of engineered B cells as *in vivo* factories for these biologics appears attractive. The field is still very much nascent and clinical translation will require a re-evaluation of the gene editing methods to make translation commercially viable and accessible. Delivery of the antibody genes represents a major challenge and the current approach using electroporation is inefficient and requires *ex vivo* manipulation. Therefore, the discovery of vectors capable of efficient gene delivery to primary human B cells would represent a significant step forward towards actualising *in vivo* antibody production using autologous human B cells. This thesis was driven by the dual objective of improving AAV vector performance in human B cells and exploring their engineering for the expression of therapeutic anti- cancer antibodies. We sought to develop AAV capsid

libraries optimised for the discovery of vectors capable of efficient B cell transduction. As part of this goal, we saw an opportunity to improve current AAV capsid library production approach using an inducible AAV capsid library system. This proved challenging, in part due to mechanisms inherent to AAV biology, and requires continued development. We developed complementary capsid libraries tailored to post-entry bottlenecks in AAV transduction based on our findings of AAV entry and binding dynamics. In parallel, we adapted existing B cell engineering frameworks to enable the knock-in of full-length anti-cancer monoclonal antibodies. Using a homology-directed repair template designed to produce both membrane-bound and secreted immunoglobulin forms, we demonstrated that a clinically relevant antibody could be expressed in a human cell line while retaining antigen specificity and Fc-mediated effector function.

11.2 Vectors to actualise human B cell therapies

AAV is an effective DNA donor delivery vector for T-cells and haematopoietic stem cells for gene editing applications (Wang et al., 2015, Wang et al., 2016) and has been adopted for B cell engineering (Johnson et al., 2018). Furthermore, it has a well-established track record for *in vivo* gene delivery in humans. Therefore, AAV is an attractive vector for further development with the goal of improving B cell engineering efficiency and potentially enabling *in vivo* editing in humans. Furthermore, AAV capsid engineering has evolved substantially, with capsid library screening firmly established as an effective strategy for discovering potent new vectors.

The limits of AAV capsid library diversity are largely dictated by bacterial transformation during library production and by the number of producer cells used during library production. The practical limit for the number of bacterial transformants is typically in the 10^7-10^8 range, while the theoretical diversity of large capsid libraries is usually orders of magnitude higher than this. Therefore, it is critical that the design of the library is carefully considered and tailored towards the goal of the screen. Our data suggested that the attachment and entry of AAV6 in stimulated primary human B cells appears largely intact and is unlikely to account for poor transduction efficiency. We postulated that, as an antigen-presenting cell, the endosomal environment of activated B cells may be suboptimal for exposure and activation of the capsid N-terminal domain, which is critical for endosomal escape and possibly uncoating (Yamaguchi et al., 2025).

AAV is a relatively simple virus in structure and as a result it relies heavily on host machinery throughout the transduction pathway. Although the virus itself is simple, the host-virus interactions are hugely complex and appear to differ significantly between serotypes and cell types (Herrmann and Grimm, 2018). Other viruses with more sophisticated entry apparatus may be able to circumvent the endosomal route in B cells. As discussed previously, lentiviruses have been tested with various fusogenic proteins including the measles and baboon endogenous retrovirus (BaEV) envelope proteins (Levy et al., 2016, Frecha et al., 2009). To integrate the effective template delivery of lentiviral vectors with the precise, site-specific genome editing capabilities of CRISPR/Cas9, BaEV-pseudotyped integrase-deficient lentiviruses (IDLVs) have been employed to deliver homology-directed repair templates, facilitating targeted insertion of an anti–PD-1 antibody cassette into the GAPDH locus of human B cells. The IDLVs were also used to

deliver the CRISPR/Cas9 machinery, removing the need for electroporation. This approach yielded a knock-in efficiency of approximately 20%, which is within the range observed using electroporation methods (Luo et al., 2020, Moffett et al., 2019). BaEV recognises the alanine-serine-cysteine (ASC) neutral amino acid transport system transporters ASCT1 and ASCT2, which are expressed in a wide range of tissues (Levy et al., 2016). Unlike VSV-G pseudotyped lentiviruses, BaEV lentiviruses appear to be resistant to neutralisation by human serum components (Girard-Gagnepain et al., 2014), making BaEV-IDLVs an interesting candidate for *in vivo* delivery to B cells. One advantage of *ex vivo* nucleofection with RNP is the transient delivery of CRISPR/Cas9 components, limiting toxicity and the risk of off-target editing. Viral delivery of CRISPR/Cas9 involves the packaging of Cas9 and guide-RNA genes into viral genomes resulting in more sustained delivery. To enable the transient delivery, BaEV pseudotyped "virus-like particles" have been used to package RNPs and, along with AAV6-delivered DNA templates, tested *in vitro* with primary human B cells for the ability to support CRISPR knock-in. The knock-in efficiency was relatively low compared to electroporation methods and required very high doses of AAV6, but nonetheless it shows potential for *in vivo* editing of B cells (Gutierrez-Guerrero et al., 2021).

Epstein-Barr virus (EBV) is a non-integrating oncogenic human γ-herpesvirus that naturally infects human B cells via CD21 and HLA class II to establish a persistent latent infection. EBV uses envelope fusion to escape endosomes following endocytosis (Munz, 2025), which may be more effective than the capsid-mediated mechanism used by AAV. The natural tropism for human B cells has made EBV an interesting candidate as a B cell gene therapy vector. However, the enormous genome size and complexity have made it

challenging to engineer a replication-incompetent EBV vector that can be manufactured at scale without helper contamination (Hettich et al., 2006). Furthermore, the oncogenic nature of the natural virus raises safety concerns for EBV as a gene therapy vector.

In summary, the current B cell engineering methodology has enabled the first proof-of-principle experiments in mouse models, but substantial refinement is needed for clinical translation. AAV has a strong track record for in vivo delivery and can be paired with electroporation for *ex vivo* editing, yet performance in B cells remains limited. AAV capsid engineering may raise editing rates and novel vectors that bypass AAV bottlenecks in B cells may prove more effective.

11.3 B Cell Engineering for cancer

A secondary aim of this thesis was to develop a strategy for the expression of full-length anti-cancer monoclonal antibodies, adapting current B cell engineering approaches. As the current work in the field has largely focused on treating infectious diseases, the established method uses endogenous immunoglobulin constant regions to enable class-switch recombination. This approach is in some ways simpler than encoding a full-length antibody because it results in an HDR template with a smaller footprint and only a simple splice donor at the 3' end of the heavy chain sequence. In contrast, a full-length antibody requires a system that replicates the natural 3' end processing decision that regulates the production of membrane-bound or secreted immunoglobulins. Recently, the knock-in of antibodies against classic cancer targets, anti-HER2 and anti-CD20, was demonstrated in primary B cells using the previously established strategy. Using similar

approaches to ours, the retention of antibody specificity and Fc-mediated functions were demonstrated. Somatic hypermutation and class-switch recombination was also validated (Ueda et al., 2024). Although this validated the use of anti-cancer antibody sequences for B cell engineering, the study made no effort to address the need for control of the antibody Fc region sequences in the context of cancer treatment. It remains to be demonstrated whether the knock-in of full-length monoclonal antibodies is possible given the small packaging capacity of AAV and the complexity of constant region 3' end processing.

Beyond the difficulty of inserting full-length antibody sequences, substantial barriers remain to adapting B cell engineering for cancer therapy. Many of the monoclonal antibodies used clinically target overexpressed self-antigens such as growth factor receptors and angiogenic signalling proteins that support tumour growth. The ability of engineered B cells to tune antibody secretion to the level of antigen is an attractive feature, however in the context of cancer, the sustained production of antibodies against self-antigens may lead to increased and persistent toxicity. Two recent papers have specifically targeted neoantigens and oncofoetal antigens to circumvent this challenge (Bracha et al., 2025, Boucher et al., 2025), but this approach significantly narrows the applicability of the approach. To target self-antigens an effective suicide switch may be required to enable clinicians to end the treatment. With the development of cell-based therapies, several suicide gene designs have been developed; however, achieving complete elimination of transplanted cell may prove challenging, particularly for cells that can clonally expand rapidly such as B cells. Furthermore, for engineered B cells to differentiate into antibody producing plasma cells, they require T cell help or pathogen-

derived secondary signals. Oncofoetal antigens and neoantigens may activate T cells, allowing them to provide help within tertiary lymphoid structures in the tumour. This may not be the case for self-antigens due to negative selection of autoreactive T cells. Furthermore, B cells engineered to express antibodies against self-antigens are likely to undergo deletion without T-cell help due to peripheral tolerance mechanisms. The need for a secondary signal may be bypassed using signalling domains, such as CD40 intracellular domains, engineered into the knocked-in immunoglobulin sequence. However, this raises the risk of dysregulated stimulation and potentially lymphomagenesis (Freidman et al., 2020). Antigen responsive engineered B cells are particularly attractive for prophylactic use but for cancer treatment it may be preferable to engineer long-lived plasma cells (LLPCs) instead that home to the bone marrow and can persist for years (Koike et al., 2023).

11.4 Future Directions

This thesis describes the development, production and screening of AAV capsid libraries tailored towards primary human B cells. Sequencing and analysis of the rescued capsid sequences are ongoing. The performance of recovered sequences will be tested against AAV6 for both B cell transduction efficiency and resistance to neutralising antibodies. It is possible that a broad distribution of capsid sequences will be recovered, in which case additional rounds of screening may be required to narrow down the pool of variants. If promising capsid sequences are identified, we will apply rational design approaches, such as incorporating mutations known to decrease proteasomal degradation, with the aim of improving potency.

For the B cell engineering arm of this thesis, we have established a knock-in template design. The knock-in of these sequences into primary human B cells and their characterisation is ongoing. We aim to demonstrate that these cells can secrete functional trastuzumab upon differentiation into plasma cells.

The field of B cell engineering has significant hurdles to overcome before B cell-based antibody therapies can be used to treat human disease. Immediate next steps involve the validation of the approach in non-human primates. A protocol for the *ex vivo* engineering of primary rhesus macaque B cells was recently published, so we suspect that this work is ongoing (Hartweger et al., 2023).

12 References

ADLER, L. N., JIANG, W., BHAMIDIPATI, K., MILLICAN, M., MACAUBAS, C., HUNG, S. C. & MELLINS, E. D. 2017. The Other Function: Class II-Restricted Antigen Presentation by B Cells. *Frontiers in Immunology,* 8.

AFIONE, S., DIMATTIA, M. A., HALDER, S., DI PASQUALE, G., AGBANDJE-MCKENNA, M. & CHIORINI, J. A. 2015. Identification and mutagenesis of the adeno-associated virus 5 sialic acid binding region. *J Virol,* 89**,** 1660-72.

AHMADI, E., RAVANSHAD, M., XIE, J., PANIGRAHI, R., JUBBAL, S. S., GURU, S. K., GUANGPING, G., ZIYAEYAN, M. & FINGEROTH, J. 2021. Serotype-dependent recombinant adeno-associated vector (AAV) infection of Epstein-Barr virus-positive B-cells, towards recombinant AAV-based therapy of focal EBV plus lymphoproliferative disorders. *Virology Journal,* 18.

ALAM, O. 2021. A single-cell-type transcriptomics map of human tissues. *Nature Genetics,* 53**,** 1275-1275.

ALBERT, H., DALE, E. C., LEE, E. & OW, D. W. 1995. Site-Specific Integration of DNA into Wild-Type and Mutant Lox Sites Placed in the Plant Genome. *Plant Journal,* 7**,** 649-659.

ATASOY, D., APONTE, Y., SU, H. H. & STERNSON, S. M. 2008. A FLEX switch targets Channelrhodopsin-2 to multiple cell types for selective optical stimulation. *Journal of Neuroscience,* 28**,** 7025-7030.

ATCHISON, R. W. 1970. The role of herpesviruses in adenovirus-associated virus replication in vitro. *Virology,* 42**,** 155-62.

ATCHISON, R. W., CASTO, B. C. & HAMMON, W. M. 1965. Adenovirus-Associated Defective Virus Particles. *Science,* 149**,** 754-6.

BAINBRIDGE, J. W., SMITH, A. J., BARKER, S. S., ROBBIE, S., HENDERSON, R., BALAGGAN, K., VISWANATHAN, A., HOLDER, G. E., STOCKMAN, A., TYLER, N., PETERSEN-JONES, S., BHATTACHARYA, S. S., THRASHER, A. J., FITZKE, F. W., CARTER, B. J., RUBIN, G. S., MOORE, A. T. & ALI, R. R. 2008. Effect of gene therapy on visual function in Leber's congenital amaurosis. *N Engl J Med,* 358**,** 2231-9.

BAROK, M., ISOLA, J., PáLYI-KREKK, Z., NAGY, P., JUHáSZ, I., VEREB, G., KAURANIEMI, P., KAPANEN, A., TANNER, M., VEREB, G. & SZöLLöSI, J. 2007. Trastuzumab causes antibody-dependent cellular cytotoxicity-mediated growth inhibition of submacroscopic JIMT-1 breast cancer xenografts despite intrinsic drug resistance. *Molecular Cancer Therapeutics,* 6**,** 2065-2072.

BEATON, A., PALUMBO, P. & BERNS, K. I. 1989a. Expression from the adeno-associated virus p5 and p19 promoters is negatively regulated in trans by the rep protein. *J Virol,* 63**,** 4450-4.

BEATON, A., PALUMBO, P. & BERNS, K. I. 1989b. Expression from the Adeno-Associated Virus-P5 and Virus-P19 Promoters Is Negatively Regulated in Trans by the Rep Protein. *Journal of Virology,* 63**,** 4450-4454.

BECK, A., GOETSCH, L., DUMONTET, C. & CORVAïA, N. 2017. The next generation of antibody–drug conjugates. *Nature Reviews Drug Discovery,* 16**,** 315-337.

BENSON, M. J., ÄIJö, T., CHANG, X., GAGNON, J., PAPE, U. J., ANANTHARAMAN, V., ARAVIND, L., PURSIHEIMO, J. P., OBERDOERFFER, S., LIU, X. S., LAHESMAA, R., LäHDESMäKI, H. & RAO, A. 2012. Heterogeneous nuclear ribonucleoprotein L-like (hnRNPLL) and elongation factor, RNA polymerase II, 2 (ELL2) are regulators of mRNA processing in plasma cells. *Proceedings of the National Academy of Sciences of the United States of America,* 109**,** 16252-16257.

BERNS, K. I. & KELLY, T. J., JR. 1974. Letter: Visualization of the inverted terminal repetition in adeno-associated virus DNA. *J Mol Biol,* 82**,** 267-71.

BERNS, K. I., PINKERTON, T. C., THOMAS, G. F. & HOGGAN, M. D. 1975. Detection of Adeno-Associated Virus (Aav)-Specific Nucleotide-Sequences in DNA Isolated from Latently Infected Detroit 6 Cells. *Virology,* 68**,** 556-560.

BERNS, K. I. & ROSE, J. A. 1970. Evidence for a single-stranded adenovirus-associated virus genome: isolation and separation of complementary single strands. *J Virol,* 5**,** 693-9.

BLACKLOW, N. R., HOGGAN, M. D., KAPIKIAN, A. Z., AUSTIN, J. B. & ROWE, W. P. 1968. Epidemiology of adenovirus-associated virus infection in a nursery population. *Am J Epidemiol,* 88**,** 368-78.

BLACKLOW, N. R., HOGGAN, M. D. & ROWE, W. P. 1967. Isolation of adenovirus-associated viruses from man. *Proc Natl Acad Sci U S A,* 58**,** 1410-5.

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**,** 475-80.

BLEKER, S., PAWLITA, M. & KLEINSCHMIDT, J. A. 2006. Impact of capsid conformation and Rep-capsid interactions on adeno-associated virus type 2 genome packaging. *J Virol,* 80**,** 810-20.

BOUCHER, A., ANDERSON, C., HINMAN, R., KINDSCHUH, M., FUNG, J., WANG, T. S., KLOOSTER, I., KIM, E., ROTH, C., VANDER OEVER, M., KHAN, B., ZELIKSON, N., VAGIMA, Y., SARIBASAK, H., SANTRY, L., KLAPPER, L. N., HESS, S., MOONEY, J., BUBLIK, D. R., LAKEN, H., BARZEL, A., BORDEN, P., PLEWA, C., CHADBOURNE, A. M., BRIDGEN, D. & NAHMAD, A. D. 2025. Engineered human B cells targeting tumor-associated antigens exhibit antigen presentation and antibody-mediated functions. *Frontiers in Immunology,* 16.

BRACHA, M. G., BIBER, G., ZELIKSON, N., SHAVIT, S., AVRAHAM, R., VAGIMA, Y., BUBLIK, D. R., KATZ, Y., BARZEL, A., KLAPPER, L. N., HESS, S. & NAHMAD, A. D. 2025. Mouse B cells engineered to express an anti-HPV antibody elicit anti-tumor T cell responses. *Frontiers in Immunology,* 16.

BRYANT, D. H., BASHIR, A., SINAI, S., JAIN, N. K., OGDEN, P. J., RILEY, P. F., CHURCH, G. M., COLWELL, L. J. & KELSIC, E. D. 2021. Deep diversification of an AAV capsid protein by machine learning. *Nature Biotechnology,* 39**,** 691-696.

BUSSLINGER, M. 2004. Transcriptional control of early B cell development. *Annual Review of Immunology,* 22**,** 55-79.

CABRITA, R., LAUSS, M., SANNA, A., DONIA, M., LARSEN, M. S., MITRA, S., JOHANSSON, I., PHUNG, B., HARBST, K., VALLON-CHRISTERSSON, J., VAN SCHOIACK, A., LOEVGREN, K., WARREN, S., JIRSTROEM, K., OLSSON, H., PIETRAS, K., INGVAR, C., ISAKSSON, K., SCHADENDORF, D., SCHMIDT, H., BASTHOLT, L., CARNEIRO,

A., WARGO, J. A., SVANE, I. M. & JONSSON, G. 2020. Tertiary lymphoid structures improve immunotherapy and survival in melanoma. *Nature,* 577**,** 561-+.

CALCEDO, R., VANDENBERGHE, L. H., GAO, G., LIN, J. & WILSON, J. M. 2009. Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses. *J Infect Dis,* 199**,** 381-90.

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**,** e104596.

CHAI, S., 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.

CHAUHAN, V., BAPTISTA, I. S. C., ARSH, A. M., JAGADEESAN, R., DASH, S. & RIBEIRO, A. S. 2024. Transcription Attenuation in Synthetic Promoters in Nonoverlapping Tandem Formation. *Biochemistry,* 63**,** 2009-2022.

CHEJANOVSKY, N. & CARTER, B. J. 1989. Mutagenesis of an AUG codon in the adeno-associated virus rep gene: effects on viral DNA replication. *Virology,* 173**,** 120-8.

CHHABRA, A., BASHIRIANS, G., PETROPOULOS, C. J., WRIN, T., PALIWAL, Y., HENSTOCK, P. V., SOMANATHAN, S., DA FONSECA PEREIRA, C., WINBURN, I. & RASKO, J. E. J. 2024. Global seroprevalence of neutralizing antibodies against adeno-associated virus serotypes used for human gene therapies. *Mol Ther Methods Clin Dev,* 32**,** 101273.

CHUAPOCO, M. R., FLYTZANIS, N. C., GOEDEN, N., OCTEAU, J. C., ROXAS, K. M., CHAN, K. Y., SCHERRER, J., WINCHESTER, J., BLACKBURN, R. J., CAMPOS, L. J., MAN, K. N. M., SUN, J. Q., CHEN, X. H., LEFEVRE, A., SINGH, V. P., AROKIARAJ, C. M., SHAY, T. F., VENDEMIATTI, J., JANG, M. J., MICH, J. K., BISHAW, Y., GORE, B. B., OMSTEAD, V., TASKIN, N., WEED, N., LEVI, B. P., TING, J. T., MILLER, C. T., DEVERMAN, B. E., PICKEL, J., TIAN, L., FOX, A. S. & GRADINARU, V. 2023. Adeno-associated viral vectors for functional intravenous gene transfer throughout the non-human primate brain. *Nature Nanotechnology,* 18**,** 1241-+.

CIDECIYAN, A. V., ALEMAN, T. S., BOYE, S. L., SCHWARTZ, S. B., KAUSHAL, S., ROMAN, A. J., PANG, J. J., SUMAROKA, A., WINDSOR, E. A., WILSON, J. M., FLOTTE, T. R., FISHMAN, G. A., HEON, E., STONE, E. M., BYRNE, B. J., JACOBSON, S. G. & HAUSWIRTH, W. W. 2008. Human gene therapy for RPE65 isomerase deficiency activates the retinoid cycle of vision but with slow rod kinetics. *Proc Natl Acad Sci U S A,* 105**,** 15112-7.

COIFFIER, B., LEPAGE, E., BRIèRE, J., HERBRECHT, R., TILLY, H., BOUABDALLAH, R., MOREL, P., VAN DEN NESTE, E., SALLES, G., GAULARD, P., GISSELBRECHT, C., BASTION, Y., GERLIER, J., RECHER, C., REYNAUD-BOUGNOUX, A. & VOSE, J. M. 2002. CHOP chemotherapy plus rituximab compared with CHOP alone in elderly patients with diffuse large-B-cell lymphoma. *New England Journal of Medicine,* 346**,** 235-242.

CONANT, D., HSIAU, T., ROSSI, N., OKI, J., MAURES, T., WAITE, K., YANG, J. Y., JOSHI, S., KELSO, R., HOLDEN, K., ENZMANN, B. L. & STONER, R. 2022. Inference of CRISPR Edits from Sanger Trace Data. *Crispr Journal,* 5**,** 123-130.

CONCORDET, J. P. & HAEUSSLER, M. 2018. CRISPOR: intuitive guide selection for CRISPR/Cas9 genome editing experiments and screens. *Nucleic Acids Research,* 46**,** W242-W245.

CONWAY, J. E., RHYS, C. M. J., ZOLOTUKHIN, I., MUZYCZKA, N., HAYWARD, G. S. & BYRNE, B. J. 1999. High-titer recombinant adeno-associated virus production utilizing a recombinant herpes simplex virus type I vector expressing AAV-2 Rep and Cap. *Gene Therapy,* 6**,** 986-993.

DALKARA, D., BYRNE, L. C., KLIMCZAK, R. R., VISEL, M., YIN, L., MERIGAN, W. H., FLANNERY, J. G. & SCHAFFER, D. V. 2013. Vivo–Directed Evolution of a New Adeno-Associated Virus for Therapeutic Outer Retinal Gene Delivery from the Vitreous. *Science Translational Medicine,* 5**,** 189ra76.

DAVIDSSON, M., NEGRINI, M., HAUSER, S., SVANBERGSSON, A., LOCKOWANDT, M., TOMASELLO, G., MANFREDSSON, F. P. & HEUER, A. 2020. A comparison of AAV-vector production methods for gene therapy and preclinical assessment. *Scientific Reports,* 10.

DE SILVA, N. S. & KLEIN, U. 2015. Dynamics of B cells in germinal centres. *Nature Reviews Immunology,* 15**,** 137-148.

DEVERMAN, B. E., PRAVDO, P. L., SIMPSON, B. P., KUMAR, S. R., CHAN, K. Y., BANERJEE, A., WU, W. L., YANG, B., HUBER, N. & PASCA, S. P. 2016. Cre-dependent selection yields AAV variants for widespread gene transfer to the adult brain. *Nature Biotechnology,* 34**,** 204-209.

DOMENGER, C. & GRIMM, D. 2019. Next-generation AAV vectors: smart transgene regulation and reduced immunogenicity. *Human Molecular Genetics,* 28**,** R3-R14.

DONSANTE, A., MILLER, D. G., LI, Y., VOGLER, C., BRUNT, E. M., RUSSELL, D. W. & SANDS, M. S. 2007. AAV vector integration sites in mouse hepatocellular carcinoma. *Science,* 317**,** 477.

DOSHI, J., COUTO, E., STAITI, J., VANDENBERGHE, L. H. & ZABALETA, N. 2024. E2A, VA RNA I, and L4-22k adenoviral helper genes are sufficient for AAV production in HEK293 cells. *Mol Ther Methods Clin Dev,* 32**,** 101376.

DUAN, D. 2023. Lethal immunotoxicity in high-dose systemic AAV therapy. *Mol Ther,* 31**,** 3123-3126.

DUDEK, A. M., PILLAY, S., PUSCHNIK, A. S., NAGAMINE, C. M., CHENG, F., QIU, J., CARETTE, J. E. & VANDENBERGHE, L. H. 2018. An Alternate Route for Adeno-associated Virus (AAV) Entry Independent of AAV Receptor. *J Virol,* 92.

DUDEK, A. M., ZABALETA, N., ZINN, E., PILLAY, S., ZENGEL, J., PORTER, C., FRANCESCHINI, J. S., ESTELIEN, R., CARETTE, J. E., ZHOU, G. L. & VANDENBERGHE, L. H. 2020. GPR108 Is a Highly Conserved AAV Entry Factor. *Mol Ther,* 28**,** 367-381.

EARLEY, L. F., CONATSER, L. M., LUE, V. T. M., DOBBINS, A. L., LI, C. W., HIRSCH, M. L. & SAMULSKI, R. J. 2020. Adeno-Associated Virus Serotype-Specific Inverted Terminal Repeat Sequence Role in Vector Transgene Expression. *Human Gene Therapy,* 31**,** 151-162.

EICHHOFF, A. M., BöRNER, K., ALBRECHT, B., SCHäFER, W., BAUM, N., HAAG, F., KöRBELIN, J., TREPEL, M., BRAREN, I., GRIMM, D., ADRIOUCH, S. & KOCH-NOLTE, F. 2019. Nanobody-Enhanced Targeting of AAV Gene Therapy Vectors. *Molecular Therapy-Methods & Clinical Development,* 15**,** 211-220.

ELLIS, B. L., HIRSCH, M. L., BARKER, J. C., CONNELLY, J. P., STEININGER, R. J. & PORTEUS, M. H. 2013. A survey of ex vivo/in vitro transduction efficiency of mammalian primary cells and cell lines with Nine natural adeno-associated virus (AAV1-9) and one engineered adeno-associated virus serotype. *Virology Journal,* 10**,** 74.

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. *Nat Commun,* 12**,** 6239.

ERIKSSON, R. A. E., LEINONEN, E. & HAURU, M. 2022. Optimized riboswitch-regulated AAV vector for VEGF-B gene therapy. *Frontiers in Medicine,* 9**,** 1052318.

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**,** 3227-3234.

FLAVELL, D. J., WARNES, S. L., BRYSON, C. J., FIELD, S. A., NOSS, A. L., PACKHAM, G. & FLAVELL, S. U. 2006. The anti-CD20 antibody rituximab augments the immunospecific therapeutic effectiveness of an anti-CD19 immunotoxin directed against human B-cell lymphoma. *British Journal of Haematology,* 134**,** 157-170.

FLOTTE, T. R., AFIONE, S. A., SOLOW, R., DRUMM, M. L., MARKAKIS, D., GUGGINO, W. B., ZEITLIN, P. L. & CARTER, B. J. 1993. Expression of the Cystic-Fibrosis Transmembrane Conductance Regulator from a Novel Adenoassociated Virus Promoter. *Journal of Biological Chemistry,* 268**,** 3781-3790.

FLOTTE, T. R., TRAPNELL, B. C., HUMPHRIES, M., CAREY, B., CALCEDO, R., ROUHANI, F., CAMPBELL-THOMPSON, M., YACHNIS, A. T., SANDHAUS, R. A. & MCELVANEY, N. G. 2011. Phase 2 clinical trial of a recombinant adeno-associated viral vector expressing α1-antitrypsin: interim results. *Human Gene Therapy,* 22**,** 1239-1247.

FRECHA, C., COSTA, C., LéVY, C., NèGRE, D., RUSSELL, S. J., MAISNER, A., SALLES, G., PENG, K. W., COSSET, F.-L. & VERHOEYEN, E. 2009. Efficient and stable transduction of resting B lymphocytes and primary chronic lymphocyte leukemia cells using measles virus gp displaying lentiviral vectors. *Blood,* 114**,** 3173-3180.

FREIDMAN, N., CHEN, I., WU, Q., BRIOT, C., HOLST, J., FONT, J., VANDENBERG, R. & RYAN, R. 2020. Amino Acid Transporters and Exchangers from the SLC1A Family: Structure, Mechanism and Roles in Physiology and Cancer. *Neurochem Res,* 45**,** 1268-1286.

FRIEDMANN, T. & ROBLIN, R. 1972. Gene therapy for human genetic disease? *Science,* 175**,** 949-55.

FUSIL, F., CALATTINI, S., AMIRACHE, F., MANCIP, J., COSTA, C., ROBBINS, J. B., DOUAM, F., LAVILLETTE, D., LAW, M., DEFRANCE, T., VERHOEYEN, E. & COSSET, F.-L. 2015. A Lentiviral Vector Allowing Physiologically Regulated Membrane-anchored and Secreted Antibody Expression Depending on B-cell Maturation Status. *Molecular Therapy,* 23**,** 1734-1747.

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. *J Virol,* 78**,** 6381-8.

GAO, G. P., ALVIRA, M. R., WANG, L., CALCEDO, R., JOHNSTON, J. & WILSON, J. M. 2002. Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc Natl Acad Sci U S A,* 99**,** 11854-9.

GEORGE, L. A., SULLIVAN, S. K., GIERMASZ, A., RASKO, J. E. J., SAMELSON-JONES, B. J., DUCORE, J., CUKER, A., SULLIVAN, L. M., MAJUMDAR, S., TEITEL, J., MCGUINN, C. E., RAGNI, M. V., LUK, A. Y., HUI, D., WRIGHT, J. F., CHEN, Y., LIU, Y., WACHTEL, K., WINTERS, A., TIEFENBACHER, S., ARRUDA, V. R., VAN DER LOO, J. C. M., ZELENAIA, O., TAKEFMAN, D., CARR, M. E., COUTO, L. B., ANGUELA, X. M. & HIGH,

K. A. 2017. Hemophilia B Gene Therapy with a High-Specific-Activity Factor IX Variant. *N Engl J Med,* 377**,** 2215-2227.

GHOSH, K., GUO, F. & VAN DUYNE, G. D. 2007. Synapsis of sites by cre recombinase. *Journal of Biological Chemistry,* 282**,** 24004-24016.

GIACOMONI, J., ÅKERBLOM, M., HABEKOST, M., FIORENZANO, A., KAJTEZ, J., DAVIDSSON, M., PARMAR, M. & BJöRKLUND, T. 2024. Identification and validation of novel engineered AAV capsid variants targeting human glia. *Frontiers in Neuroscience,* 18.

GIRARD-GAGNEPAIN, A., AMIRACHE, F., COSTA, C., LEVY, C., FRECHA, C., FUSIL, F., NEGRE, D., LAVILLETTE, D., COSSET, F. L. & VERHOEYEN, E. 2014. Baboon envelope pseudotyped LVs outperform VSV-G-LVs for gene transfer into early-cytokine-stimulated and resting HSCs. *Blood,* 124**,** 1221-31.

GIROD, A., WOBUS, C. E., ZADORI, 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. *J Gen Virol,* 83**,** 973-978.

GLIWA, K., HULL, J., KANSOL, A., ZEMBRUSKI, V., LAKSHMANAN, R., MIETZSCH, M., CHIPMAN, P., BENNETT, A. & MCKENNA, R. 2025. Biophysical and structural insights into AAV genome ejection. *Journal of Virology,* 99.

GOMEZ-MANRIQUEZ, J., HERNANDEZ-BELLO, J., MUNOZ-VALLE, J. F., SIFUENTES-FRANCO, S., GRACIANO-MACHUCA, O. & MORALES-NUNEZ, J. J. 2025. B cell development: transcriptional regulation and immunological mechanisms in homeostasis. *Front Immunol,* 16**,** 1593338.

GREINER, V., BOU PUERTO, R., LIU, S., HERBEL, C., CARMONA, E. M. & GOLDBERG, M. S. 2019. CRISPR-Mediated Editing of the B Cell Receptor in Primary Human B Cells. *iScience,* 12**,** 369-378.

GRILLO-LóPEZ, A. J. 2000. Rituximab: the first monoclonal antibody approved for the treatment of lymphoma. *Annals of Oncology,* 11**,** 3-5.

GRIMM, D., LEE, J. S., WANG, L., DESAI, T., AKACHE, B., STORM, T. A. & KAY, M. A. 2008. In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *Journal of Virology,* 82**,** 5887-5911.

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.

GUTIERREZ-GUERRERO, A., ABREY RECALDE, M. J., MANGEOT, P. E., COSTA, C., BERNADIN, O., PERIAN, S., FUSIL, F., FROMENT, G., MARTINEZ-TURTOS, A., KRUG, A., MARTIN, F., BENABDELLAH, K., RICCI, E. P., GIOVANNOZZI, S., GIJSBERS, R., AYUSO, E., COSSET, F. L. & VERHOEYEN, E. 2021. Baboon Envelope Pseudotyped "Nanoblades" Carrying Cas9/gRNA Complexes Allow Efficient Genome Editing in Human T, B, and CD34(+) Cells and Knock-in of AAV6-Encoded Donor DNA in CD34(+) Cells. *Front Genome Ed,* 3**,** 604371.

HACEIN-BEY-ABINA, S., VON KALLE, C., SCHMIDT, M., LE DEIST, F., WULFFRAAT, N., MCINTYRE, E., RADFORD, I., VILLEVAL, J. L., FRASER, C. C., CAVAZZANA-CALVO, M. & FISCHER, A. 2003. A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med,* 348**,** 255-6.

HALBERT, C. L., ALLEN, J. M. & MILLER, A. D. 2001. Adeno-associated virus type 6 (AAV6) vectors mediate efficient transduction of airway epithelial cells in mouse lungs compared to that of AAV2 vectors. *J Virol,* 75**,** 6615-24.

HAN, Y. F., LIN, Y. J. B., AN, W. F., XU, J., YANG, H. C., O'CONNELL, K., DORDAI, D., BOEKE, J. D., SILICIANO, J. D. & SILICIANO, R. F. 2008. Orientation-dependent regulation of integrated HIV-1 expression by host gene transcriptional readthrough. *Cell Host & Microbe,* 4**,** 134-146.

HANDA, A., MURAMATSU, S. I., QIU, J., MIZUKAMI, H. & BROWN, K. E. 2000. Adeno-associated virus (AAV)-3-based vectors transduce haematopoietic cells not susceptible to transduction with AAV-2-based vectors. *J Gen Virol,* 81**,** 2077-2084.

HANLON, K. S., KLEINSTIVER, B. P., GARCIA, S. P., ZABOROWSKI, M. P., VOLAK, A., SPIRIG, S. E., MULLER, A., SOUSA, A. A., TSAI, S. Q., BENGTSSON, N. E., LOOV, C., INGELSSON, M., CHAMBERLAIN, J. S., COREY, D. P., ARYEE, M. J., JOUNG, J. K., BREAKEFIELD, X. O., MAGUIRE, C. A. & GYORGY, B. 2019. High levels of AAV vector integration into CRISPR-induced DNA breaks. *Nat Commun,* 10**,** 4439.

HARTWEGER, H., GAUTAM, R., NISHIMURA, Y., SCHMIDT, F., YAO, K. H., ESCOLANO, A., JANKOVIC, M., MARTIN, M. A. & NUSSENZWEIG, M. C. 2023. Gene Editing of Primary Rhesus Macaque B Cells. *J Vis Exp*.

HARTWEGER, H., MCGUIRE, A. T., HORNING, M., TAYLOR, J. J., DOSENOVIC, P., YOST, D., GAZUMYAN, A., SEAMAN, M. S., STAMATATOS, L., JANKOVIC, M. & NUSSENZWEIG, M. C. 2019. HIV-specific humoral immune responses by CRISPR/Cas9-edited B cells. *J Exp Med,* 216**,** 1301-1310.

HAVLIK, L. P., SIMON, K. E., SMITH, J. K., KLINC, K. A., TSE, L. V., OH, D. K., FANOUS, M. M., MEGANCK, R. M., MIETZSCH, M. & KLEINSCHMIDT, J. 2020. Coevolution of Adeno-associated Virus Capsid Antigenicity and Tropism through a Structure-Guided Approach. *Journal of Virology,* 94.

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. *Proc Natl Acad Sci U S A,* 81**,** 6466-70.

HERRMANN, A. K. & GRIMM, D. 2018. High-Throughput Dissection of AAV-Host Interactions: The Fast and the Curious. *J Mol Biol,* 430**,** 2626-2640.

HERRMANN, A. K., GROSSE, S., BöRNER, K., KRäMER, C., WIEDTKE, E., GUNKEL, M. & GRIMM, D. 2019. Impact of the Assembly-Activating Protein on Molecular Evolution of Synthetic Adeno-Associated Virus Capsids. *Human Gene Therapy,* 30**,** 21-35.

HETTICH, E., JANZ, A., ZEIDLER, R., PICH, D., HELLEBRAND, E., WEISSFLOG, B., MOOSMANN, A. & HAMMERSCHMIDT, W. 2006. Genetic design of an optimized packaging cell line for gene vectors transducing human B cells. *Gene Ther,* 13**,** 844-56.

HIRRLINGER, J., REQUARDT, R. P., WINKLER, U., WILHELM, F., SCHULZE, C. & HIRRLINGER, P. G. 2009. Split-CreERT2: temporal control of DNA recombination mediated by split-Cre protein fragment complementation. *PLoS One,* 4**,** e8354.

HO, A., ORTON, R., TAYLER, R., ASAMAPHAN, P., HERDER, V., DAVIS, C., TONG, L., SMOLLETT, K., MANALI, M., ALLAN, J., RAWLIK, K., MCDONALD, S. E., VINK, E., POLLOCK, L., GANNON, L., EVANS, C., MCMENAMIN, J., ROY, K., MARSH, K., DIVALA, T., HOLDEN, M. T. G., LOCKHART, M., YIRRELL, D., CURRIE, S., O'LEARY, M., HENDERSON, D., SHEPHERD, S. J., JACKSON, C., GUNSON, R., MACLEAN, A.,

MCINNES, N., BRADLEY-STEWART, A., BATTLE, R., HOLLENBACH, J. A., HENDERSON, P., ODAM, M., CHIKOWORE, P., OOSTHUYZEN, W., CHAND, M., HAMILTON, M. S., ESTRADA-RIVADENEYRA, D., LEVIN, M., AVRAMIDIS, N., PAIRO-CASTINEIRA, E., VITART, V., WILKIE, C., CONSORTIUM, D., INVESTIGATORS, I. C., PALMARINI, M., RAY, S., ROBERTSON, D. L., DA SILVA FILIPE, A., WILLETT, B. J., BREUER, J., SEMPLE, M. G., TURNER, D., BAILLIE, J. K. & THOMSON, E. C. 2023. Adeno-associated virus 2 infection in children with non-A-E hepatitis. *Nature,* 617**,** 555-563.

HOFFMANN, M. D., GALLANT, J. P., LEBEAU, A. M. & SCHMIDT, D. 2024. Unlocking precision gene therapy: harnessing AAV tropism with nanobody swapping at capsid hotspots. *NAR Molecular Medicine,* 1**,** ugae008.

HOGGAN, M. D., BLACKLOW, N. R. & ROWE, W. P. 1966. Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. *Proc Natl Acad Sci U S A,* 55**,** 1467-74.

HORDEAUX, J., WANG, Q., KATZ, N., BUZA, E. L., BELL, P. & WILSON, J. M. 2018. The Neurotropic Properties of AAV-PHP.B Are Limited to C57BL/6J Mice. *Molecular Therapy,* 26**,** 664-668.

HORLICK, R. A., MACOMBER, J. L., BOWERS, P. M., NEBEN, T. Y., TOMLINSON, G. L., KRAPF, I. P., DALTON, J. L., VERDINO, P. & KING, D. J. 2013. Simultaneous Surface Display and Secretion of Proteins from Mammalian Cells Facilitate Efficient

Selection and Maturation of Antibodies. *Journal of Biological Chemistry,* 288**,** 19861-19869.

HUANG, D., TRAN, J. T., OLSON, A., VOLLBRECHT, T., TENUTA, M., GURYLEVA, M. V., FULLER, R. P., SCHIFFNER, T., ABADEJOS, J. R., COUVRETTE, L., BLANE, T. R., SAYE, K., LI, W., LANDAIS, E., GONZALEZ-MARTIN, A., SCHIEF, W., MURRELL, B., BURTON, D. R., NEMAZEE, D. & VOSS, J. E. 2020. Vaccine elicitation of HIV broadly neutralizing antibodies from engineered B cells. *Nature Communications,* 11**,** 5850.

HUDIS, C. A. 2007a. Drug therapy: Trastuzumab - Mechanism of action and use in clinical practice. *New England Journal of Medicine,* 357**,** 39-51.

HUDIS, C. A. 2007b. Trastuzumab—mechanism of action and use in clinical practice. *New England Journal of Medicine,* 357**,** 39-51.

HUNG, K. L., MEITLIS, I., HALE, M., CHEN, C.-Y., SINGH, S., JACKSON, S. W., MIAO, C. H., KHAN, I. F., RAWLINGS, D. J. & JAMES, R. G. 2018. Engineering Protein-Secreting Plasma Cells by Homology-Directed Repair in Primary Human B Cells. *Molecular Therapy,* 26**,** 456-467.

HURWITZ, H., FEHRENBACHER, L., NOVOTNY, W., CARTWRIGHT, T., HAINSWORTH, J., HEIM, W., BERLIN, J., BARON, A., GRIFFING, S., HOLMGREN, E., FERRARA, N., FYFE, G., RATLIFF, L., STALTER, T., KERN, K. & KABBINAVAR, F. 2004. Bevacizumab plus irinotecan, fluorouracil, and leucovorin for metastatic colorectal cancer. *New England Journal of Medicine,* 350**,** 2335-2342.

HUSE, K., BAI, B. Y., HILDEN, V. I., BOLLUM, L. K., VåTSVEEN, T. K., MUNTHE, L. A., SMELAND, E. B., IRISH, J. M., WäLCHLI, S. & MYKLEBUST, J. H. 2022. Mechanism of CD79A and CD79B Support for IgM plus B Cell Fitness through B Cell Receptor Surface Expression. *Journal of Immunology,* 209**,** 2042-2053.

HUSER, D., GOGOL-DORING, A., CHEN, W. & HEILBRONN, R. 2014. Adeno-associated virus type 2 wild-type and vector-mediated genomic integration profiles of human

diploid fibroblasts analyzed by third-generation PacBio DNA sequencing. *J Virol,* 88**,** 11253-63.

HUSER, D., GOGOL-DORING, 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 Pathog,* 6**,** e1000985.

HüTTERMANN, L., SCHRöDER, L. C., SHETTY, P. M. V., JONKER, T., HILLE, S. S., REMES, A. K., MATZEN, A., BOENDER, A. R., GRIMM, D., FRANK, D., BOINK, G. J. J., ESCHENHAGEN, T., SCHADE, D. & MüLLER, O. J. 2025. Directed Evolution of AAV9 for Efficient Gene Expression in Cardiomyocytes

and. *Human Gene Therapy,* 36**,** 101-115.

IM, D. S. & MUZYCZKA, N. 1990. The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. *Cell,* 61**,** 447-57.

INDRACCOLO, S., MINUZZO, S., ZAMARCHI, R., CALDERAZZO, F., PIOVAN, E. & AMADORI, A. 2002. Alternatively spliced forms of Igα and Igβ prevent B cell receptor expression on the cell surface. *European Journal of Immunology,* 32**,** 1530-1540.

JANOVITZ, T., OLIVEIRA, T., SADELAIN, M. & FALCK-PEDERSEN, E. 2014. Highly divergent integration profile of adeno-associated virus serotype 5 revealed by high-throughput sequencing. *J Virol,* 88**,** 2481-8.

JOHNSON, J. S. & SAMULSKI, R. J. 2009. Enhancement of adeno-associated virus infection by mobilizing capsids into and out of the nucleolus. *J Virol,* 83**,** 2632-44.

JOHNSON, M. J., LAOHARAWEE, K., LAHR, W. S., WEBBER, B. R. & MORIARITY, B. S. 2018. Engineering of Primary Human B cells with CRISPR/Cas9 Targeted Nuclease. *Scientific Reports,* 8**,** 12144.

JUNG, D., GIALLOURAKIS, C., MOSTOSLAVSKY, R. & ALT, F. W. 2006. Mechanism and control of V(D)J recombination at the immunoglobulin heavy chain locus. *Annu Rev Immunol,* 24**,** 541-70.

KALUDOV, N., BROWN, K. E., WALTERS, R. W., ZABNER, J. & CHIORINI, J. A. 2001. Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J Virol,* 75**,** 6884-93.

KANG, L., JIANG, D. W., ENGLAND, C. G., BARNHART, T. E., YU, B., ROSENKRANS, Z. T., WANG, R. F., ENGLE, J. W., XU, X. J., HUANG, P. & CAI, W. B. 2018. ImmunoPET imaging of CD38 in murine lymphoma models using 89Zr-labeled daratumumab. *European Journal of Nuclear Medicine and Molecular Imaging,* 45**,** 1372-1381.

KASSAMBARA, A., HERVIOU, L., OVEJERO, S., JOURDAN, M., THIBAUT, C., VIKOVA, V., PASERO, P., ELEMENTO, O. & MOREAUX, J. M. 2021. RNA-sequencing data-driven dissection of human plasma cell differentiation reveals new potential transcription regulators. *Leukemia,* 35**,** 1451-1462.

KEIZER, R. J., HUITEMA, A. D., SCHELLENS, J. H. & BEIJNEN, J. H. 2010. Clinical pharmacokinetics of therapeutic monoclonal antibodies. *Clin Pharmacokinet,* 49**,** 493-507.

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. *Mol Ther Methods Clin Dev,* 2**,** 15047.

KERN, A., SCHMIDT, K., LEDER, C., MULLER, O. J., WOBUS, C. E., BETTINGER, K., VON DER LIETH, C. W., KING, J. A. & KLEINSCHMIDT, J. A. 2003. Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. *J Virol,* 77**,** 11072-81.

KIM, H., KIM, M., IM, S. K. & FANG, S. 2018. Mouse Cre-LoxP system: general principles to determine tissue-specific roles of target genes. *Lab Anim Res,* 34**,** 147-159.

KIMURA, T., FERRAN, B., TSUKAHARA, Y., SHANG, Q. F., DESAI, S., FEDOCE, A., PIMENTEI, D. R., LUPTAK, I., ADACHI, T., IDO, Y., MATSUI, R. & BACHSCHMID, M. M. 2019. Production of adeno-associated virus vectors for in vitro and in vivo applications. *Scientific Reports,* 9.

KOERBER, J. T., JANG, J. H. & SCHAFFER, D. V. 2008a. DNA shuffling of adeno-associated virus yields functionally diverse viral progeny. *Molecular Therapy,* 16**,** 1703-1709.

KOERBER, J. T., JANG, J. H. & SCHAFFER, D. V. 2008b. DNA shuffling of adeno-associated virus yields functionally diverse viral progeny. *Mol Ther,* 16**,** 1703-9.

KöHLER, G. & MILSTEIN, C. 1975. Continuous cultures of fused cells secreting antibody of predefined specificity. *Nature,* 256**,** 495-497.

KOIKE, T., FUJII, K., KOMETANI, K., BUTLER, N. S., FUNAKOSHI, K., YARI, S., KIKUTA, J., ISHII, M., KUROSAKI, T. & ISE, W. 2023. Progressive differentiation toward the long-lived plasma cell compartment in the bone marrow. *J Exp Med,* 220.

KöRBELIN, J., HUNGER, A., ALAWI, M., SIEBER, T., BINDER, M. & TREPEL, M. 2017. Optimization of design and production strategies for novel adeno-associated viral display peptide libraries. *Gene Therapy,* 24**,** 470-481.

KOSAKA, M., SAJIKI, A., FUJITA, K., YAMADA, K., KAWANO, K., HIRAZAWA, K., MATSUNO, T., TAKADA, M., SHIMOZAWA, N., INOUE, K. I., YUKI, K. & NISHIGUCHI, K. M. 2025. Evaluation of the loading capacity and patterns of packaged DNA in AAV genomes of different sizes using long-read sequencing. *Molecular Therapy Methods & Clinical Development,* 33.

KOTIN, R. M., SINISCALCO, M., SAMULSKI, R. J., ZHU, X. D., HUNTER, L., LAUGHLIN, C. A., MCLAUGHLIN, S., MUZYCZKA, N., ROCCHI, M. & BERNS, K. I. 1990. Site-specific integration by adeno-associated virus. *Proc Natl Acad Sci U S A,* 87**,** 2211-5.

KRAUS, M., ALIMZHANOV, M. B., RAJEWSKY, N. & RAJEWSKY, K. 2004. Survival of resting mature B lymphocytes depends on BCR signaling via the Igα/β heterodimer. *Cell,* 117**,** 787-800.

KRONENBERG, S., BOTTCHER, B., VON DER LIETH, C. W., BLEKER, S. & KLEINSCHMIDT, J. A. 2005. A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini. *J Virol,* 79**,** 5296-303.

LALANNE, J. B., MICH, J. K., HUYNH, C., HUNKER, A. C., MCDIARMID, T. A., LEVI, B. P., TING, J. T. & SHENDURE, J. 2025. Extensive length and homology dependent chimerism in pool-packaged AAV libraries. *bioRxiv*.

LAM, K. P., KUHN, R. & RAJEWSKY, K. 1997. In vivo ablation of surface immunoglobulin on mature B cells by inducible gene targeting results in rapid cell death. *Cell,* 90**,** 1073-1083.

LAUBER, K., BLUMENTHAL, S. G., WAIBEL, M. & WESSELBORG, S. 2004. Clearance of apoptotic cells: Getting rid of the corpses. *Molecular Cell,* 14**,** 277-287.

LEE, G. K., MAHESHRI, N., KASPAR, B. & SCHAFFER, D. V. 2005. PEG conjugation moderately protects adeno-associated viral vectors against antibody neutralization. *Biotechnology and Bioengineering,* 92**,** 24-34.

LEE, Z., LU, M., IRFANULLAH, E., SOUKUP, M. & HU, W. S. 2022. Construction of an rAAV producer cell line through synthetic biology. *ACS Synthetic Biology,* 11**,** 3285-3295.

LEVY, C., FUSIL, F., AMIRACHE, F., COSTA, C., GIRARD-GAGNEPAIN, A., NEGRE, D., BERNADIN, O., GARAULET, G., RODRIGUEZ, A., NAIR, N., VANDENDRIESSCHE, T., CHUAH, M., COSSET, F. L. & VERHOEYEN, E. 2016. Baboon envelope pseudotyped lentiviral vectors efficiently transduce human B cells and allow active factor IX B cell secretion in vivo in NOD/SCIDgammac(-/-) mice. *J Thromb Haemost,* 14**,** 2478-2492.

LI, W., ASOKAN, A., WU, Z., VAN DYKE, T., DIPRIMIO, N., JOHNSON, J. S., GOVINDASWAMY, L., AGBANDJE-MCKENNA, M., LEICHTLE, S., REDMOND, D. E., MCCOWN, T. J., PETERMANN, K. B., SHARPLESS, N. E. & SAMULSKI, R. J. 2008. Engineering and selection of shuffled AAV genomes: A new strategy for producing targeted biological nanoparticles. *Molecular Therapy,* 16**,** 1252-1260.

LIN, S., STAAHL, B., ALLA, R. K. & DOUDNA, J. A. 2014. Enhanced homology-directed human genome engineering by controlled timing of CRISPR/Cas9 delivery. *Elife,* 3.

LIND, P., LARSSON, K., SPIRA, J., SYDOW-BACKMAN, M., ALMSTEDT, A., GRAY, E. & SANDBERG, H. 1995. Novel forms of B-domain-deleted recombinant factor VIII molecules. Construction and biochemical characterization. *Eur J Biochem,* 232**,** 19-27.

LISOWSKI, L., DANE, A. P., CHU, K., ZHANG, Y., CUNNINGHAM, S. C., WILSON, E. M., NYGAARD, S., GROMPE, M., ALEXANDER, I. E. & KAY, M. A. 2014. Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature,* 506**,** 382-+.

LIU, K., JIN, H. & ZHOU, B. 2020. Genetic lineage tracing with multiple DNA recombinases: A user's guide for conducting more precise cell fate mapping studies. *J Biol Chem,* 295**,** 6413-6424.

LOONSTRA, A., VOOIJS, M., BEVERLOO, H. B., AL ALLAK, B., VAN DRUNEN, E., KANAAR, R., BERNS, A. & JONKERS, J. 2001. Growth inhibition and DNA damage induced by Cre recombinase in mammalian cells. *Proceedings of the National Academy of Sciences of the United States of America,* 98**,** 9209-9214.

LU, L. L., SUSCOVICH, T. J., FORTUNE, S. M. & ALTER, G. 2018. Beyond binding: antibody effector functions in infectious diseases. *Nature Reviews Immunology,* 18**,** 46-61.

LU, M. & HU, W. S. 2024. Enhancing the production of recombinant adeno-associated virus in synthetic cell lines through systematic characterization. *Biotechnology and Bioengineering*.

LU, M., LEE, Z. & HU, W. S. 2024. Multi-omics kinetic analysis of recombinant adeno-associated virus production by plasmid transfection of HEK293 cells. *Biotechnology Progress,* 40.

LUO, B., ZHAN, Y., LUO, M., DONG, H., LIU, J., LIN, Y., ZHANG, J., WANG, G., VERHOEYEN, E., ZHANG, Y. & ZHANG, H. 2020. Engineering of α-PD-1 antibody-expressing long-lived plasma cells by CRISPR/Cas9-mediated targeted gene integration. *Cell Death & Disease,* 11**,** 973.

LUO, X. M., MAARSCHALK, E., O'CONNELL, R. M., WANG, P., YANG, L. & BALTIMORE, D. 2009. Engineering human hematopoietic stem/progenitor cells to produce a broadly neutralizing anti-HIV antibody after in vitro maturation to human B lymphocytes. *Blood,* 113**,** 1422-1431.

LUSBY, E., FIFE, K. H. & BERNS, K. I. 1980. Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. *J Virol,* 34**,** 402-9.

MACAULEY, M. S., KAWASAKI, N., PENG, W. J., WANG, S. H., HE, Y., ARLIAN, B. M., MCBRIDE, R., KANNAGI, R., KHOO, K. H. & PAULSON, J. C. 2015. Unmasking of CD22 Co-receptor on Germinal Center B-cells Occurs by Alternative Mechanisms in Mouse and Man. *Journal of Biological Chemistry,* 290**,** 30066-30077.

MAGUIRE, A. M., SIMONELLI, F., PIERCE, E. A., PUGH, E. N., JR., MINGOZZI, F., BENNICELLI, J., BANFI, S., MARSHALL, K. A., TESTA, F., SURACE, E. M., ROSSI, S., LYUBARSKY, A., ARRUDA, V. R., KONKLE, B., STONE, E., SUN, J., JACOBS, J., DELL'OSSO, L., HERTLE, R., MA, J. X., REDMOND, T. M., ZHU, X., HAUCK, B., ZELENAIA, O., SHINDLER, K. S., MAGUIRE, M. G., WRIGHT, J. F., VOLPE, N. J., MCDONNELL, J. W., AURICCHIO, A., HIGH, K. A. & BENNETT, J. 2008. Safety and efficacy of gene transfer for Leber's congenital amaurosis. *N Engl J Med,* 358**,** 2240-8.

MAHESHRI, N., KOERBER, J. T., KASPAR, B. K. & SCHAFFER, D. V. 2006. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nature Biotechnology,* 24**,** 198-204.

MAHLANGU, J., KACZMAREK, R., VON DRYGALSKI, A., SHAPIRO, S., CHOU, S. C., OZELO, M. C., KENET, G., PEYVANDI, F., WANG, M., MADAN, B., KEY, N. S., LAFFAN, M., DUNN, A. L., MASON, J., QUON, D. V., SYMINGTON, E., LEAVITT, A. D., OLDENBURG, J., CHAMBOST, H., REDING, M. T., JAYARAM, K., YU, H., MAHAJAN, R., CHAVELE, K. M., REDDY, D. B., HENSHAW, J., ROBINSON, T. M., WONG, W. Y., PIPE, S. W. & GROUP, G. E.-T. 2023. Two-Year Outcomes of Valoctocogene Roxaparvovec Therapy for Hemophilia A. *N Engl J Med,* 388**,** 694-705.

MANNO, C. S., PIERCE, G. F., ARRUDA, V. R., GLADER, B., RAGNI, M., RASKO, J. J., OZELO, M. C., HOOTS, K., BLATT, P., KONKLE, B., DAKE, M., KAYE, R., RAZAVI, M., ZAJKO, A., ZEHNDER, J., RUSTAGI, P. K., NAKAI, H., CHEW, A., LEONARD, D., WRIGHT, J. F., LESSARD, R. R., SOMMER, J. M., TIGGES, M., SABATINO, D., LUK, A., JIANG, H., MINGOZZI, F., COUTO, L., ERTL, H. C., HIGH, K. A. & KAY, M. A. 2006. Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response. *Nat Med,* 12**,** 342-7.

MARTENS, J. A., LAPRADE, L. & WINSTON, F. 2004. Intergenic transcription is required to repress the Saccharomyces cerevisiae SER3 gene. *Nature,* 429**,** 571-574.

MARTINCIC, K., ALKAN, S. A., CHEATLE, A., BORGHESI, L. & MILCAREK, C. 2009. Transcription elongation factor ELL2 directs immunoglobulin secretion in plasma cells by stimulating altered RNA processing. *Nature Immunology,* 10**,** 1102-U82.

MASON, J. O., WILLIAMS, G. T. & NEUBERGER, M. S. 1985. Transcription Cell Type Specificity Is Conferred by an Immunoglobulin-Vh Gene Promoter That Includes a Functional Consensus Sequence. *Cell,* 41**,** 479-487.

MATSUSHITA, N., KATO, S., NISHIZAWA, K., SUGAWARA, M., TAKEUCHI, K., MIYASAKA, Y., MASHIMO, T. & KOBAYASHI, K. 2023. Highly selective transgene expression through the flip-excision switch system by using a unilateral spacer sequence. *Cell Rep Methods,* 3**,** 100393.

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 Ther,* 5**,** 938-45.

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 Rep,* 23**,** 1817-1830.

MAURER, A. C. & WEITZMAN, M. D. 2020. Adeno-Associated Virus Genome Interactions Important for Vector Production and Transduction. *Hum Gene Ther,* 31**,** 499-511.

MAYOR, H. D., JAMISON, R. M., JORDAN, L. E. & MELNICK, J. L. 1965. Structure and Composition of a Small Particle Prepared from a Simian Adenovirus. *J Bacteriol,* 90**,** 235-42.

MCCARTY, D. M., CHRISTENSEN, M. & MUZYCZKA, N. 1991. Sequences Required for Coordinate Induction of Adenoassociated Virus-P19 and Virus-P40 Promoters by Rep Protein. *Journal of Virology,* 65**,** 2936-2945.

MCCARTY, D. M., MONAHAN, P. E. & SAMULSKI, R. J. 2001. Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Therapy,* 8**,** 1248-1254.

MCLAUGHLIN, S. K., COLLIS, P., HERMONAT, P. L. & MUZYCZKA, N. 1988. Adeno-associated virus vector which expresses a foreign gene. *Journal of Virology,* 62**,** 1963-1973.

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., BRAUN, L., LIKHITE, S., MIRANDA, C., MEYER, K., FOUST, K. D., BURGHES, A. H. M. & KASPAR, B. K. 2017. Single-Dose Gene-Replacement Therapy for Spinal Muscular Atrophy. *N Engl J Med,* 377**,** 1713-1722.

MENDELL, J. R., SAHENK, Z., LEHMAN, K., NEASE, C., LOWES, L. P., MILLER, N. F., IAMMARINO, M. A., ALFANO, L. N., NICHOLL, A., AL-ZAIDY, S., LEWIS, S., CHURCH, K., SHELL, R., CRIPE, L. H., POTTER, R. A., GRIFFIN, D. A., POZSGAI, E., DUGAR, A., HOGAN, M. & RODINO-KLAPAC, L. R. 2020. Assessment of Systemic Delivery of rAAVrh74.MHCK7.micro-dystrophin in Children With Duchenne Muscular Dystrophy: A Nonrandomized Controlled Trial. *JAMA Neurol,* 77**,** 1122-1131.

MéVEL, M., BOUZELHA, M., LERAY, A., PACOURET, S., GUILBAUD, M., PENAUD-BUDLOO, M., ALVAREZ-DORTA, D., DUBREIL, L., GOUIN, S. G. & COMBAL, J. P. 2020. Chemical modification of the adeno-associated virus capsid to improve gene delivery. *Chemical Science,* 11**,** 1122-1131.

MEYER, N. L., HU, G., DAVULCU, O., XIE, Q., NOBLE, A. J., YOSHIOKA, C., GINGERICH, D. S., TRZYNKA, A., DAVID, L., STAGG, S. M. & CHAPMAN, M. S. 2019. Structure of the gene therapy vector, adeno-associated virus with its cell receptor, AAVR. *Elife,* 8.

MIETZSCH, M., HSI, J., NELSON, A. R., KHANDEKAR, N., HUANG, A. M., SMITH, N. J., ZACHARY, J., POTTS, L., FARRAR, M. A., CHIPMAN, P., GHANEM, M., ALEXANDER, I. E., LOGAN, G. J., HUISKONEN, J. T. & MCKENNA, R. 2025. Structural characterization of antibody-responses following Zolgensma treatment for AAV capsid engineering to expand patient cohorts. *Nat Commun,* 16**,** 3731.

MILELLA, M., TRISCIUOGLIO, D., BRUNO, T., CIUFFREDA, L., MOTOLESE, M., CIANCIULLI, A., COGNETTI, F., ZANGEMEISTER-WITTKE, U., DEL BUFALO, D. & ZUPI, G. 2004. Trastuzumab down-regulates Bcl-2 expression and potentiates apoptosis induction by Bcl-2/Bcl-XL bispecific antisense oligonucleotides in HER-2 gene-amplified breast cancer cells. *Clinical Cancer Research,* 10**,** 7747-7756.

MINGOZZI, F., MAUS, M. V., HUI, D. J., SABATINO, D. E., MURPHY, S. L., RASKO, J. E., RAGNI, M. V., MANNO, C. S., SOMMER, J., JIANG, H., PIERCE, G. F., ERTL, H. C. & HIGH, K. A. 2007. CD8(+) T-cell responses to adeno-associated virus capsid in humans. *Nat Med,* 13**,** 419-22.

MOFFETT, H. A.-O., HARMS, C. A.-O., FITZPATRICK, K. A.-O., TOOLEY, M. A.-O., BOONYARATANAKORNKIT, J. & TAYLOR, J. A.-O. 2019. B cells engineered to express pathogen-specific antibodies protect against infection. LID - 10.1126/sciimmunol.aax0644 [doi] LID - eaax0644.

MORFOPOULOU, S., BUDDLE, S., TORRES MONTAGUTH, O. E., ATKINSON, L., GUERRA-ASSUNCAO, J. A., MORADI MARJANEH, M., ZENNEZINI CHIOZZI, R., STOREY, N., CAMPOS, L., HUTCHINSON, J. C., COUNSELL, J. R., POLLARA, G., ROY, S., VENTURINI, C., ANTINAO DIAZ, J. F., SIAM, A., TAPPOUNI, L. J., ASGARIAN, Z., NG, J., HANLON, K. S., LENNON, A., MCARDLE, A., CZAP, A., ROSENHEIM, J., ANDRADE, C., ANDERSON, G., LEE, J. C. D., WILLIAMS, R., WILLIAMS, C. A., TUTILL, H., BAYZID, N., MARTIN BERNAL, L. M., MACPHERSON, H., MONTGOMERY, K. A., MOORE, C., TEMPLETON, K., NEILL, C., HOLDEN, M., GUNSON, R., SHEPHERD, S. J., SHAH, P., COORAY, S., VOICE, M., STEELE, M., FINK, C., WHITTAKER, T. E., SANTILLI, G., GISSEN, P., KAUFER, B. B., REICH, J., ANDREANI, J., SIMMONDS, P., ALRABIAH, D. K., CASTELLANO, S., CHIKOWORE, P., ODAM, M., RAMPLING, T., HOULIHAN, C., HOSCHLER, K., TALTS, T., CELMA, C., GONZALEZ, S., GALLAGHER, E., SIMMONS, R., WATSON, C., MANDAL, S., ZAMBON, M., CHAND, M., HATCHER, J., DE, S., BAILLIE, K., SEMPLE, M. G., CONSORTIUM, D., CONSORTIUM, P., INVESTIGATORS, I. C., MARTIN, J., USHIRO-LUMB, I., NOURSADEGHI, M., DEHERAGODA, M., HADZIC, N., GRAMMATIKOPOULOS, T., BROWN, R., KELGERI, C., THALASSINOS, K., WADDINGTON, S. N., JACQUES, T. S., THOMSON, E., LEVIN, M., BROWN, J. R. & BREUER, J. 2023. Genomic investigations of unexplained acute hepatitis in children. *Nature,* 617**,** 564-573.

MüLLER, O. J., KAUL, F., WEITZMAN, M. D., PASQUALINI, R., ARAP, W., KLEINSCHMIDT, J. A. & TREPEL, M. 2003. Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors. *Nature Biotechnology,* 21**,** 1040-1046.

MULLICK, A., XU, Y., WARREN, R., KOUTROUMANIS, M., GUILBAULT, C., BROUSSAU, S., MALENFANT, F., BOURGET, L., LAMOUREUX, L., LO, R., CARON, A. W., PILOTTE, A. & MASSIE, B. 2006. The cumate gene-switch: a system for regulated expression in mammalian cells. *Bmc Biotechnology,* 6.

MUNZ, C. 2025. Epstein-Barr virus pathogenesis and emerging control strategies. *Nat Rev Microbiol,* 23**,** 667-679.

NAHMAD, A. D., LAZZAROTTO, C. R., ZELIKSON, N., KUSTIN, T., TENUTA, M., HUANG, D. L., REUVENI, I., NATAF, D., RAVIV, Y., HOROVITZ-FRIED, M., DOTAN, I., CARMI, Y., ROSIN-ARBESFELD, R., NEMAZEE, D., VOSS, J. E., STERN, A., TSAI, S. Q. &

BARZEL, A. 2022. In vivo engineered B cells secrete high titers of broadly neutralizing anti-HIV antibodies in mice. *Nature Biotechnology,* 40**,** 1241-+.

NAHMAD, A. D., RAVIV, Y., HOROVITZ-FRIED, M., SOFER, I., AKRIV, T., NATAF, D., DOTAN, I., CARMI, Y., BURSTEIN, D., WINE, Y., BENHAR, I. & BARZEL, A. 2020. Engineered B cells expressing an anti-HIV antibody enable memory retention, isotype switching and clonal expansion. *Nature Communications,* 11**,** 5851.

NATHWANI, A. C., TUDDENHAM, E. G., RANGARAJAN, S., ROSALES, C., MCINTOSH, J., LINCH, D. C., CHOWDARY, P., RIDDELL, A., PIE, A. J., HARRINGTON, C., O'BEIRNE, J., SMITH, K., PASI, J., GLADER, B., RUSTAGI, P., NG, C. Y., KAY, M. A., ZHOU, J., SPENCE, Y., MORTON, C. L., ALLAY, J., COLEMAN, J., SLEEP, S., CUNNINGHAM, J. M., SRIVASTAVA, D., BASNER-TSCHAKARJAN, E., MINGOZZI, F., HIGH, K. A., GRAY, J. T., REISS, U. M., NIENHUIS, A. W. & DAVIDOFF, A. M. 2011. Adenovirus-associated virus vector-mediated gene transfer in hemophilia B. *N Engl J Med,* 365**,** 2357-65.

NAUMER, M., SONNTAG, F., SCHMIDT, K., NIETO, K., PANKE, C., DAVEY, N. E., POPA-WAGNER, R. & KLEINSCHMIDT, J. A. 2012. Properties of the adeno-associated virus assembly-activating protein. *J Virol,* 86**,** 13038-48.

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. *Mol Ther Methods Clin Dev,* 21**,** 642-655.

NICOLSON, S. C. & SAMULSKI, R. J. 2014. Recombinant adeno-associated virus utilizes host cell nuclear import machinery to enter the nucleus. *J Virol,* 88**,** 4132-44.

NONNENMACHER, M., VAN BAKEL, H., HAJJAR, R. J. & WEBER, T. 2015a. High Capsid-Genome Correlation Facilitates Creation of AAV Libraries for Directed Evolution. *Molecular Therapy,* 23**,** 675-682.

NONNENMACHER, M. & WEBER, T. 2011. Adeno-associated virus 2 infection requires endocytosis through the CLIC/GEEC pathway. *Cell Host Microbe,* 10**,** 563-76.

NONNENMACHER, M. E., CINTRAT, J. C., GILLET, D. & WEBER, T. 2015b. Syntaxin 5-dependent retrograde transport to the trans-Golgi network is required for adeno-associated virus transduction. *J Virol,* 89**,** 1673-87.

NONNENMACHER, M. E., CINTRAT, J. C., GILLET, D. & WEBER, T. 2015c. Syntaxin 5-Dependent Retrograde Transport to the trans-Golgi Network Is Required for Adeno-Associated Virus Transduction. *Journal of Virology,* 89**,** 1673-1687.

NUTT, S. L., HEAVEY, B., ROLINK, A. G. & BUSSLINGER, M. 1999. Commitment to the B-lymphoid lineage depends on the transcription factor Pax5. *Nature,* 401**,** 556-562.

O'DONNELL, J., TAYLOR, K. A. & CHAPMAN, M. S. 2009. Adeno-associated virus-2 and its primary cellular receptor--Cryo-EM structure of a heparin complex. *Virology,* 385**,** 434-43.

OBERDOERFFER, P., OTIPOBY, K. L., MARUYAMA, M. & RAJEWSKY, K. 2003. Unidirectional Cre-mediated genetic inversion in mice using the mutant loxP pair lox66/lox71. *Nucleic Acids Research,* 31.

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**,** 1139-1143.

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. *J Virol,* 74**,** 3494-504.

OJALA, D. S., SUN, S., SANTIAGO-ORTIZ, J. L., SHAPIRO, M. G., ROMERO, P. A. & SCHAFFER, D. V. 2018. In Vivo Selection of a Computationally Designed SCHEMA AAV Library Yields a Novel Variant for Infection of Adult Neural Stem Cells in the SVZ. *Mol Ther,* 26**,** 304-319.

OWENS, R. A., WEITZMAN, M. D., KYOSTIO, S. R. & CARTER, B. J. 1993. Identification of a DNA-binding domain in the amino terminus of adeno-associated virus Rep proteins. *J Virol,* 67**,** 997-1005.

OZELO, M. C., MAHLANGU, J., PASI, K. J., GIERMASZ, A., LEAVITT, A. D., LAFFAN, M., SYMINGTON, E., QUON, D. V., WANG, J. D., PEERLINCK, K., PIPE, S. W., MADAN, B., KEY, N. S., PIERCE, G. F., O'MAHONY, B., KACZMAREK, R., HENSHAW, J., LAWAL, A., JAYARAM, K., HUANG, M., YANG, X., WONG, W. Y., KIM, B. & GROUP, G. E.-T. 2022. Valoctocogene Roxaparvovec Gene Therapy for Hemophilia A. *N Engl J Med,* 386**,** 1013-1025.

PAWLUCZKOWYCZ, A. W., BEURSKENS, F. J., BEUM, P. V., LINDORFER, M. A., VAN DE WINKEL, J. G., PARREN, P. W. & TAYLOR, R. P. 2009. Binding of submaximal C1q promotes complement-dependent cytotoxicity (CDC) of B cells opsonized with anti-CD20 mAbs ofatumumab (OFA) or rituximab (RTX): considerably higher levels of CDC are induced by OFA than by RTX. *J Immunol,* 183**,** 749-58.

PEKRUN, K., DE ALENCASTRO, G., LUO, Q. J., LIU, J., KIM, Y., NYGAARD, S., GALIVO, F., ZHANG, F. J., SONG, R., TIFFANY, M. R., XU, J. P., HEBROK, M., GROMPE, M. & KAY, M. A. 2019. Using a barcoded AAV capsid library to select for clinically relevant gene therapy vectors. *Jci Insight,* 4.

PENZES, J. J., CHIPMAN, P., BHATTACHARYA, N., ZEHER, A., HUANG, R., MCKENNA, R. & AGBANDJE-MCKENNA, M. 2021. Adeno-associated Virus 9 Structural Rearrangements Induced by Endosomal Trafficking pH and Glycan Attachment. *J Virol,* 95**,** e0084321.

PERABO, L., BUNING, H., KOFLER, D. M., RIED, M. U., GIROD, A., WENDTNER, C. M., ENSSLE, J. & HALLEK, M. 2003. In vitro selection of viral vectors with modified tropism: the adeno-associated virus display. *Mol Ther,* 8**,** 151-7.

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. *J Virol,* 71**,** 1079-88.

PETERSON, M. L. 2011. Immunoglobulin heavy chain gene regulation through polyadenylation and splicing competition. *Wiley Interdisciplinary Reviews-Rna,* 2**,** 92-105.

PETRUK, S., SEDKOV, Y., RILEY, K. M., HODGSON, J., SCHWEISGUTH, F., HIROSE, S., JAYNES, J. B., BROCK, H. W. & MAZO, A. 2006. Transcriptional elongation of non-coding bxd RNAs promoted by the Trithorax TAC1 complex represses Ubx by a transcriptional interference mechanism. *Cell,* 127**,** 1209-1221.

PHAN, T. G., AMESBURY, M., GARDAM, S., CROSBIE, J., HASBOLD, J., HODGKIN, P. D., BASTEN, A. & BRINK, R. 2003. B Cell Receptor–independent Stimuli Trigger Immunoglobulin (Ig) Class Switch Recombination and Production of IgG Autoantibodies by Anergic Self-Reactive B Cells. *Journal of Experimental Medicine,* 197**,** 845-860.

PILLAY, S., MEYER, N. L., PUSCHNIK, A. S., DAVULCU, O., DIEP, J., ISHIKAWA, Y., JAE, L. T., WOSEN, J. E., NAGAMINE, C. M., CHAPMAN, M. S. & CARETTE, J. E. 2016. An essential receptor for adeno-associated virus infection. *Nature,* 530**,** 108-12.

PILLAY, S., ZOU, W., CHENG, F., PUSCHNIK, A. S., MEYER, N. L., GANAIE, S. S., DENG, X., WOSEN, J. E., DAVULCU, O., YAN, Z., ENGELHARDT, J. F., BROWN, K. E., CHAPMAN, M. S., QIU, J. & CARETTE, J. E. 2017. Adeno-associated Virus (AAV) Serotypes Have Distinctive Interactions with Domains of the Cellular AAV Receptor. *J Virol,* 91.

PROUDFOOT, N. J. 1986. Transcriptional Interference and Termination between Duplicated Alpha-Globin Gene Constructs Suggests a Novel Mechanism for Gene-Regulation. *Nature,* 322**,** 562-565.

RANGARAJAN, S., WALSH, L., LESTER, W., PERRY, D., MADAN, B., LAFFAN, M., YU, H., VETTERMANN, C., PIERCE, G. F., WONG, W. Y. & PASI, K. J. 2017. AAV5-Factor VIII Gene Transfer in Severe Hemophilia A. *N Engl J Med,* 377**,** 2519-2530.

RAPER, S. E., CHIRMULE, N., LEE, F. S., WIVEL, N. A., BAGG, A., GAO, G. P., WILSON, J. M. & BATSHAW, M. L. 2003. Fatal systemic inflammatory response syndrome in a ornithine transcarbamylase deficient patient following adenoviral gene transfer. *Mol Genet Metab,* 80**,** 148-58.

ROGERS, G. L., HUANG, C., CLARK, R. D. E., SECLéN, E., CHEN, H.-Y. & CANNON, P. M. 2021. Optimization of AAV6 transduction enhances site-specific genome editing of primary human lymphocytes. *Molecular Therapy - Methods & Clinical Development,* 23**,** 198-209.

ROSALES, A., BLONDEL, L. O., HULL, J., GAO, Q., AYKUN, N., PEEK, J. L., VARGAS, A., FERGIONE, S., SONG, M. & WILSON, M. H. 2025. Evolving adeno-associated viruses for gene transfer to the kidney via cross-species cycling of capsid libraries. *Nature Biomedical Engineering,* 9**,** 1086-1100.

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. *Proc Natl Acad Sci U S A,* 64**,** 863-9.

ROSE, J. A. & KOCZOT, F. 1972. Adenovirus-associated virus multiplication. VII. Helper requirement for viral deoxyribonucleic acid and ribonucleic acid synthesis. *J Virol,* 10**,** 1-8.

ROSE, J. A., MAIZEL, J. V., JR., INMAN, J. K. & SHATKIN, A. J. 1971. Structural proteins of adenovirus-associated viruses. *J Virol,* 8**,** 766-70.

ROSSI, A., DUPATY, L., AILLOT, L., ZHANG, L., GALLIEN, C., HALLEK, M., ODENTHAL, M., ADRIOUCH, S., SALVETTI, A. & BUNING, H. 2019. Vector uncoating limits adeno-associated viral vector-mediated transduction of human dendritic cells and vector immunogenicity. *Sci Rep,* 9**,** 3631.

RUSSELL, C. B., VETTERMANN, C., AGARWAL, S., WITT, E., CLARK, W., ARENS, J., FRONZA, R., MOSS, K. M. O., KASPRZYK, T., ROBINSON, T. M., TRAN, H., KENET, G., RAHEJA, P., LESTER, W., EGGAN, K. & ZOOG, S. 2025. Recombinant Adeno-Associated Virus Integration Profiles in Nonhuman Primates and Gene Therapy Participants after Treatment with Valoctocogene Roxaparvovec. *Human Gene Therapy,* 36**,** 945-955.

RUSSELL, S., BENNETT, J., WELLMAN, J. A., CHUNG, D. C., YU, Z. F., TILLMAN, A., WITTES, J., PAPPAS, J., ELCI, O., MCCAGUE, S., CROSS, D., MARSHALL, K. A., WALSHIRE, J., KEHOE, T. L., REICHERT, H., DAVIS, M., RAFFINI, L., GEORGE, L. A., HUDSON,

F. P., DINGFIELD, L., ZHU, X., HALLER, J. A., SOHN, E. H., MAHAJAN, V. B., PFEIFER, W., WECKMANN, M., JOHNSON, C., GEWAILY, D., DRACK, A., STONE, E., WACHTEL, K., SIMONELLI, F., LEROY, B. P., WRIGHT, J. F., HIGH, K. A. & MAGUIRE, A. M. 2017. Efficacy and safety of voretigene neparvovec (AAV2-hRPE65v2) in patients with RPE65-mediated inherited retinal dystrophy: a randomised, controlled, open-label, phase 3 trial. *Lancet,* 390**,** 849-860.

SAEKI, H. & SVEJSTRUP, J. Q. 2009. Stability, Flexibility, and Dynamic Interactions of Colliding RNA Polymerase II Elongation Complexes. *Molecular Cell,* 35**,** 191-205.

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. *J Virol,* 63**,** 3822-8.

SAMULSKI, R. J. & SHENK, T. 1988. Adenovirus-E1b 55-Mr Polypeptide Facilitates Timely Cytoplasmic Accumulation of Adeno-Associated Virus Messenger-Rnas. *Journal of Virology,* 62**,** 206-210.

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**,** 135-43.

SANTIAGO-ORTIZ, J., OJALA, D. S., WESTESSON, O., WEINSTEIN, J. R., WONG, S. Y., STEINSAPIR, A., KUMAR, S., HOLMES, I. & SCHAFFER, D. V. 2015. AAV ancestral reconstruction library enables selection of broadly infectious viral variants. *Gene Therapy,* 22**,** 934-946.

SCHMIT, P. F., PACOURET, S., ZINN, E., TELFORD, E., NICOLAOU, F., BROUCQUE, F., ANDRES-MATEOS, E., XIAO, R., PENAUD-BUDLOO, M., BOUZELHA, M., JAULIN, N., ADJALI, O., AYUSO, E. & VANDENBERGHE, L. H. 2020. Cross-Packaging and Capsid Mosaic Formation in Multiplexed AAV Libraries. *Molecular Therapy-Methods & Clinical Development,* 17**,** 107-121.

SCHRöDER, L. C., HüTTERMANN, L., REMES, A. K., VORAN, J. C., HILLE, S., SOMMER, W., LUTTER, G., WARNECKE, G., FRANK, D., SCHADE, D. & MüLLER, O. J. 2025. AAV library screening identifies novel vector for efficient transduction of human aorta. *Gene Therapy,* 32**,** 154-162.

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. *J Virol,* 81**,** 12936-45.

SCHWEIGGERT, F., HABECK, G., MOST, P., BUSCH, M. & SCHWEIGGERT, J. 2024. ShuffleAnalyzer: A Comprehensive Tool to Visualize DNA Shuffling. *ACS Synth Biol,* 13**,** 2625-2628.

SERAFINI, M., NALDINI, L. & INTRONA, M. 2004. Molecular evidence of inefficient transduction of proliferating human B lymphocytes by VSV-pseudotyped HIV-1-derived lentivectors. *Virology,* 325**,** 413-424.

SERVELLITA, V., SOTOMAYOR GONZALEZ, A., LAMSON, D. M., FORESYTHE, A., HUH, H. J., BAZINET, A. L., BERGMAN, N. H., BULL, R. L., GARCIA, K. Y., GOODRICH, J. S., LOVETT, S. P., PARKER, K., RADUNE, D., HATADA, A., PAN, C. Y., RIZZO, K., BERTUMEN, J. B., MORALES, C., OLUNIYI, P. E., NGUYEN, J., TAN, J., STRYKE, D., JABER, R., LESLIE, M. T., LYONS, Z., HEDMAN, H. D., PARASHAR, U., SULLIVAN, M., WROBLEWSKI, K., OBERSTE, M. S., TATE, J. E., BAKER, J. M., SUGERMAN, D., POTTS, C., LU, X., CHHABRA, P., PEDIATRIC HEPATITIS OF UNKNOWN ETIOLOGY WORKING, G., INGRAM, L. A., SHIAU, H., BRITT, W., GUTIERREZ SANCHEZ, L. H.,

CIRIC, C., ROSTAD, C. A., VINJE, J., KIRKING, H. L., WADFORD, D. A., RABORN, R. T., ST GEORGE, K. & CHIU, C. Y. 2023. Adeno-associated virus type 2 in US children with acute severe hepatitis. *Nature,* 617**,** 574-580.

SHARMA, P., GOSWAMI, S., RAYCHAUDHURI, D., SIDDIQUI, B. A., SINGH, P., NAGARAJAN, A., LIU, J., SUBUDHI, S. K., POON, C., GANT, K. L., HERBRICH, S. M., ANANDHAN, S., ISLAM, S., AMIT, M., ANANDAPPA, G. & ALLISON, J. P. 2023. Immune checkpoint therapy-current perspectives and future directions. *Cell,* 186**,** 1652-1669.

SHEN, S., BRYANT, K. D., BROWN, S. M., RANDELL, S. H. & ASOKAN, A. 2011. Terminal N-linked galactose is the primary receptor for adeno-associated virus 9. *J Biol Chem,* 286**,** 13532-40.

SIMIONI, P., TORMENE, D., TOGNIN, G., GAVASSO, S., BULATO, C., IACOBELLI, N. P., FINN, J. D., SPIEZIA, L., RADU, C. & ARRUDA, V. R. 2009. X-linked thrombophilia with a mutant factor IX (factor IX Padua). *N Engl J Med,* 361**,** 1671-5.

SLAMON, D. J., LEYLAND-JONES, B., SHAK, S., FUCHS, H., PATON, V., BAJAMONDE, A., FLEMING, T., EIERMANN, W., WOLTER, J., PEGRAM, M., BASELGA, J. & NORTON, L. 2001a. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *New England Journal of Medicine,* 344**,** 783-792.

SLAMON, D. J., LEYLAND-JONES, B., SHAK, S., FUCHS, H., PATON, V., BAJAMONDE, A., FLEMING, T., EIERMANN, W., WOLTER, J., PEGRAM, M., BASELGA, J. & NORTON, L. 2001b. Use of chemotherapy plus a monoclonal antibody against HER2 for metastatic breast cancer that overexpresses HER2. *New England Journal of Medicine,* 344**,** 783-792.

SNYDER, R. O., IM, D. S. & MUZYCZKA, N. 1990. Evidence for covalent attachment of the adeno-associated virus (AAV) rep protein to the ends of the AAV genome. *J Virol,* 64**,** 6204-13.

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. *J Virol,* 67**,** 6096-104.

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. *J Virol,* 80**,** 11040-54.

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**,** 12686-12697.

SONNTAG, F., SCHMIDT, K. & KLEINSCHMIDT, J. A. 2010. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc Natl Acad Sci U S A,* 107**,** 10220-5.

SRIVASTAVA, A., LUSBY, E. W. & BERNS, K. I. 1983. Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J Virol,* 45**,** 555-64.

STAHNKE, S., LUX, K., UHRIG, S., KREPPEL, F., HOSEL, M., COUTELLE, O., OGRIS, M., HALLEK, M. & BUNING, H. 2011. Intrinsic phospholipase A2 activity of adeno-associated virus is involved in endosomal escape of incoming particles. *Virology,* 409**,** 77-83.

STAMATAKI, M., LüSCHOW, J., SCHLUMBOHM, C., ALAWI, M., LUNDING, L., FUCHS, E., TREPEL, M., SCHWANINGER, M. & KöRBELIN, J. 2025. Identification of AAV variants with improved transduction of human vascular endothelial cells by screening AAV capsid libraries in non-human primates (aug, 10.1038/s41434-025-00563-4, 2025). *Gene Therapy*.

STEMMER, W. P. C. 1994. DNA Shuffling by Random Fragmentation and Reassembly - in-Vitro Recombination for Molecular Evolution. *Proceedings of the National Academy of Sciences of the United States of America,* 91**,** 10747-10751.

STROBEL, B., DUCHS, M. J., BLAZEVIC, D., RECHTSTEINER, P., BRAUN, C., BAUM-KROKER, K. S., SCHMID, B., CIOSSEK, T., GOTTSCHLING, D., HARTIG, J. S. & KREUZ, S. 2020. A Small-Molecule-Responsive Riboswitch Enables Conditional Induction of Viral Vector-Mediated Gene Expression in Mice. *ACS Synth Biol,* 9**,** 1292-1305.

SU, W., PATRICIO, M. I., DUFFY, M. R., KRAKOWIAK, J. M., SEYMOUR, L. W. & CAWOOD, R. 2022a. Self-attenuating adenovirus enables production of recombinant adeno-associated virus for high manufacturing yield without contamination. *Nat Commun,* 13**,** 1182.

SU, W. H., PATRíCIO, M. I., DUFFY, M. R., KRAKOWIAK, J. M., SEYMOUR, L. W. & CAWOOD, R. 2022b. Self-attenuating adenovirus enables production of recombinant adeno-associated virus for high manufacturing yield without contamination. *Nature Communications,* 13.

SUCHY, F. P., KARIGANE, D., NAKAUCHI, Y., HIGUCHI, M., ZHANG, J., PEKRUN, K., HSU, I., FAN, A. C., NISHIMURA, T., CHARLESWORTH, C. T., BHADURY, J., NISHIMURA, T., WILKINSON, A. C., KAY, M. A., MAJETI, R. & NAKAUCHI, H. 2025. Genome engineering with Cas9 and AAV repair templates generates frequent concatemeric insertions of viral vectors. *Nat Biotechnol,* 43**,** 204-213.

SUMMERFORD, C. & SAMULSKI, R. J. 1998a. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *Journal of Virology,* 72**,** 1438-1445.

SUMMERFORD, C. & SAMULSKI, R. J. 1998b. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J Virol,* 72**,** 1438-45.

SUNG, H., FERLAY, J., SIEGEL, R. L., LAVERSANNE, M., SOERJOMATARAM, I., JEMAL, A. & BRAY, F. 2021. Global cancer statistics 2020: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. *CA: A Cancer Journal for Clinicians,* 71**,** 209-249.

SYEDA, M. Z., HONG, T., HUANG, C., HUANG, W. & MU, Q. 2024. B cell memory: from generation to reactivation: a multipronged defense wall against pathogens. *Cell Death Discovery,* 10**,** 1-15.

TABEBORDBAR, M., LAGERBORG, K. A., STANTON, A., KING, E. M., YE, S., TELLEZ, L., KRUNNFUSZ, A., TAVAKOLI, S., WIDRICK, J. J. & MESSEMER, K. A. 2021. Directed evolution of a family of AAV capsid variants enabling potent muscle-directed gene delivery across species. *Cell,* 184**,** 4919-4938.

TAKAGAKI, Y., SEIPELT, R. L., PETERSON, M. L. & MANLEY, J. L. 1996. The polyadenylation factor CstF-64 regulates alternative processing of IgM heavy chain pre-mRNA during B cell differentiation. *Cell,* 87**,** 941-952.

TAYLOR, N. K., GUGGENBILLER, M. J., MISTRY, P. P., KING, O. D. & HARPER, S. Q. 2024. A self-complementary AAV proviral plasmid that reduces cross-packaging and ITR promoter activity in AAV vector preparations. *Molecular Therapy Methods & Clinical Development,* 32.

THEISS, H. D., KOFLER, D. M., BüNING, H., ALDENHOFF, A. L., KAESS, B., DECKER, T., BAUMERT, J., HALLEK, M. & WENDTNER, C. M. 2003. Enhancement of gene transfer with recombinant adeno-associated virus (rAAV) vectors into primary B-cell chronic lymphocytic leukemia cells by CpG-oligodeoxynucleotides. *Experimental Hematology,* 31**,** 1223-1229.

THOMAS, C. E., STORM, T. A., HUANG, Z. & KAY, M. A. 2004. Rapid uncoating of vector genomes is the key to efficient liver transduction with pseudotyped adeno-associated virus vectors. *J Virol,* 78**,** 3110-22.

TOLEDO-STUARDO, K., RIBEIRO, C. H., GONZALEZ-HERRERA, F., MATTHIES, D. J., LE ROY, M. S., DIETZ-VARGAS, C., LATORRE, Y., CAMPOS, I., GUERRA, Y., TELLO, S., VASQUEZ-SAEZ, V., NOVOA, P., FEHRING, N., GONZALEZ, M., RODRIGUEZ-SIZA, J., VASQUEZ, G., MENDEZ, P., ALTAMIRANO, C. & MOLINA, M. C. 2024. Therapeutic antibodies in oncology: an immunopharmacological overview. *Cancer Immunol Immunother,* 73**,** 242.

TORDO, J., O'LEARY, C., ANTUNES, A., PALOMAR, N., ALDRIN-KIRK, P., BASCHE, M., BENNETT, A., D'SOUZA, Z., GLEITZ, H., GODWIN, A., HOLLEY, R. J., PARKER, H., LIAO, A. Y., ROUSE, P., YOUSHANI, A. S., DRIDI, L., MARTINS, C., LEVADE, T., STACEY, K. B., DAVIS, D. M., DYER, A., CLEMENT, N., BJORKLUND, T., ALI, R. R., AGBANDJE-MCKENNA, M., RAHIM, A. A., PSHEZHETSKY, A., WADDINGTON, S. N., LINDEN, R. M., BIGGER, B. W. & HENCKAERTS, E. 2018. A novel adeno-associated virus capsid with enhanced neurotropism corrects a lysosomal transmembrane enzyme deficiency. *Brain,* 141**,** 2014-2031.

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. *Mol Cell Biol,* 6**,** 2884-94.

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. *Mol Cell Biol,* 4**,** 2072-81.

TREMPE, J. P. & CARTER, B. J. 1988. Alternate mRNA splicing is required for synthesis of adeno-associated virus VP1 capsid protein. *J Virol,* 62**,** 3356-63.

TSAO, L. C., CROSBY, E. J., TROTTER, T. N., AGARWAL, P., HWANG, B. J., ACHARYA, C., SHUPTRINE, C. W., WANG, T., WEI, J. P., YANG, X., LEI, G. J., LIU, C. X., RABIOLA, C. A., CHODOSH, L. A., MULLER, W. J., LYERLY, H. K. & HARTMAN, Z. C. 2019. CD47 blockade augmentation of trastuzumab antitumor efficacy dependent on antibody-dependent cellular phagocytosis. *Jci Insight,* 4.

TSE, L. V., KLINC, K. A., MADIGAN, V. J., CASTELLANOS RIVERA, R. M., WELLS, L. F., HAVLIK, L. P., SMITH, J. K., AGBANDJE-MCKENNA, M. & ASOKAN, A. 2017. Structure-guided evolution of antigenically distinct adeno-associated virus variants for immune evasion. *Proc Natl Acad Sci U S A,* 114**,** E4812-E4821.

UEDA, N., CAHEN, M., LEONARD, J., DELEURME, L., DREANO, S., SIRAC, C., GALY, A., MOREAUX, J., DANGER, Y. & COGNé, M. 2024. Single-hit genome editing

optimized for maturation in B cells redirects their specificity toward tumor antigens. *Scientific Reports,* 14.

URABE, M., DING, C. & KOTIN, R. M. 2002. Insect cells as a factory to produce adeno-associated virus type 2 vectors. *Hum Gene Ther,* 13**,** 1935-43.

VAN DER VOORT, R., KEEHNEN, R. M. J., BEULING, E. A., SPAARGAREN, M. & PALS, S. T. 2000. Regulation of cytokine signaling by B cell antigen receptor and CD40-controlled expression of heparan sulfate proteoglycans. *Journal of Experimental Medicine,* 192**,** 1115-1124.

VANRELL, L., DI SCALA, M., BLANCO, L., OTANO, I., GIL-FARINA, I., BALDIM, V., PANEDA, A., BERRAONDO, P., BEATTIE, S. G., CHTARTO, A., TENENBAUM, L., PRIETO, J. & GONZALEZ-ASEGUINOLAZA, G. 2011. Development of a liver-specific Tet-on inducible system for AAV vectors and its application in the treatment of liver cancer. *Mol Ther,* 19**,** 1245-53.

VENKATAKRISHNAN, B., YARBROUGH, J., DOMSIC, J., BENNETT, A., BOTHNER, B., KOZYREVA, O. G., SAMULSKI, R. J., MUZYCZKA, N., MCKENNA, R. & AGBANDJE-MCKENNA, M. 2013. Structure and dynamics of adeno-associated virus serotype 1 VP1-unique N-terminal domain and its role in capsid trafficking. *J Virol,* 87**,** 4974-84.

VERMA, S., MILES, D., GIANNI, L., KROP, I. E., WENDE, C. M., REHMETT, M., WELSLAU, M., BASECKE, J., O'SHAUGHNESSY, J., DIERAS, V., GUARDINO, E. & GUAN, Z. 2012. Trastuzumab emtansine for HER2-positive advanced breast cancer. *New England Journal of Medicine,* 367**,** 1783-1791.

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**,** 1897-1905.

VINCENT-FABERT, C., FIANCETTE, R., PINAUD, E., TRUFFINET, V., COGNé, N., COGNé, M. & DENIZOT, Y. 2010. Genomic deletion of the whole IgH 3' regulatory region (hs3a, hs1,2, hs3b, and hs4) dramatically affects class switch recombination and Ig secretion to all isotypes. *Blood,* 116**,** 1895-1898.

VINEY, L., BüRCKSTüMMER, T., EDDINGTON, C., MIETZSCH, M., CHOUDHRY, M., HENLEY, T. & AGBANDJE-MCKENNA, M. 2021. Adeno-associated Virus (AAV) Capsid Chimeras with Enhanced Infectivity Reveal a Core Element in the AAV Genome Critical for both Cell Transduction and Capsid Assembly. *Journal of Virology,* 95.

VOIGT, C. A., MARTINEZ, C., WANG, Z. G., MAYO, S. L. & ARNOLD, F. H. 2002. Protein building blocks preserved by recombination. *Nature Structural Biology,* 9**,** 553-558.

VON DRYGALSKI, A., GIERMASZ, A., CASTAMAN, G., KEY, N. S., LATTIMORE, S. U., LEEBEEK, F. W. G., MIESBACH, W. A., RECHT, M., GOMEZ, E., GUT, R. & PIPE, S. W. 2020. Etranacogene Dezaparvovec (AAV5-Padua hFIX variant), an Enhanced Vector for Gene Transfer in Adults with Severe or Moderate-Severe Hemophilia B: Two Year Data from a Phase 2b Trial. *Blood,* 136**,** 13-13.

VOSS, J. E., GONZALEZ-MARTIN, A., ANDRABI, R., FULLER, R. P., MURRELL, B., MCCOY, L. E., PORTER, K., HUANG, D., LI, W., SOK, D., LE, K., BRINEY, B., CHATEAU, M., ROGERS, G., HANGARTNER, L., FEENEY, A. J., NEMAZEE, D., CANNON, P. & BURTON, D. R. 2019. Reprogramming the antigen specificity of B cells using genome-editing technologies. *eLife,* 8**,** e42995.

WAGNER, J. A., MESSNER, A. H., MORAN, M. L., DAIFUKU, R., KOUYAMA, K., DESCH, J. K., MANLEY, S., NORBASH, A. M., CONRAD, C. K., FRIBORG, S., REYNOLDS, T., GUGGINO, W. B., MOSS, R. B., CARTER, B. J., WINE, J. J., FLOTTE, T. R. & GARDNER, P. 1999. Safety and biological efficacy of an adeno-associated virus vector-cystic fibrosis transmembrane regulator (AAV-CFTR) in the cystic fibrosis maxillary sinus. *Laryngoscope,* 109**,** 266-74.

WANG, D., LI, S., GESSLER, D. J., XIE, J., ZHONG, L., LI, J., TRAN, K., VAN VLIET, K., REN, L. & SU, Q. 2018. A Rationally Engineered Capsid Variant of AAV9 for Systemic CNS-Directed and Peripheral Tissue-Detargeted Gene Delivery in Neonates. *Molecular Therapy: Methods & Clinical Development,* 9**,** 234-246.

WANG, J., DECLERCQ, J. J., HAYWARD, S. B., LI, P. W., SHIVAK, D. A., GREGORY, P. D., LEE, G. & HOLMES, M. C. 2016. Highly efficient homology-driven genome editing in human T cells by combining zinc-finger nuclease mRNA and AAV6 donor delivery. *Nucleic Acids Res,* 44**,** e30.

WANG, J., EXLINE, C. M., DECLERCQ, J. J., LLEWELLYN, G. N., HAYWARD, S. B., LI, P. W., SHIVAK, D. A., SUROSKY, R. T., GREGORY, P. D., HOLMES, M. C. & CANNON, P. M. 2015. Homology-driven genome editing in hematopoietic stem and progenitor cells using ZFN mRNA and AAV6 donors. *Nat Biotechnol,* 33**,** 1256-1263.

WASSERMAN, S. R., MURON, S., LEE, H. R., ROUTH, M. L., HEPPERLA, A. J., SCOVILLE, D. M., HUBER, A., UMANA, J. D., PEREIRA, N. E., FOLEY, C. A., JAMES, L. I. & HATHAWAY, N. A. 2025. Dynamic activation of rAAV transgene expression by a small molecule that recruits endogenous transcriptional machinery. *Nucleic Acids Res,* 53.

WEITZMAN, M. D., KYOSTIO, S. R., CARTER, B. J. & OWENS, R. A. 1996. Interaction of wild-type and mutant adeno-associated virus (AAV) Rep proteins on AAV hairpin DNA. *J Virol,* 70**,** 2440-8.

WENNHOLD, K., THELEN, M., LEHMANN, J., SCHRAN, S., PREUGSZAT, E., GARCIA-MARQUEZ, M., LECHNER, A., SHIMABUKURO-VORNHAGEN, A., ERCANOGLU, M. S., KLEIN, F., THANGARAJAH, F., EIDT, S., LöSER, H., BRUNS, C., QUAAS, A., VON BERGWELT-BAILDON, M. & SCHLöSSER, H. A. 2021. CD86+ Antigen-Presenting B Cells Are Increased in Cancer, Localize in Tertiary Lymphoid Structures, and Induce Specific T-cell Responses. *Cancer Immunology Research,* 9**,** 1098-1108.

WILKINSON, I., ANDERSON, S., FRY, J., JULIEN, L. A., NEVILLE, D., QURESHI, O., WATTS, G. & HALE, G. 2021. Fc-engineered antibodies with immune effector functions completely abolished. *Plos One,* 16.

WILSON, J. M. & FLOTTE, T. R. 2020. Moving Forward After Two Deaths in a Gene Therapy Trial of Myotubular Myopathy. *Hum Gene Ther,* 31**,** 695-696.

WU, C., LIU, Y., ZHAO, Q., CHEN, G. M., CHEN, J. H., YAN, X. M., ZHOU, Y. H. & HUANG, Z. H. 2010. Soluble CD40 ligand-activated human peripheral B cells as surrogated antigen presenting cells: A preliminary approach for anti-HBV immunotherapy. *Virology Journal,* 7.

WU, L. N., LIU, Y., ZHU, X. H., ZHANG, L., CHEN, J. F., ZHANG, H., HAO, P., ZHANG, S., HUANG, J., ZHENG, J., ZHANG, Y. M., ZHANG, Y. H. & QIU, X. Y. 2014. The immunoglobulin heavy chain VH6-1 promoter regulates Ig transcription in non-B cells. *Cancer Cell International,* 14.

WU, Z., MILLER, E., AGBANDJE-MCKENNA, M. & SAMULSKI, R. J. 2006. Alpha2,3 and alpha2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *J Virol,* 80**,** 9093-103.

XIAO, X., LI, J. & SAMULSKI, R. J. 1998. Production of high-titer recombinant adeno-associated virus vectors in the absence of helper adenovirus. *J Virol,* 72**,** 2224-32.

XIE, Q., BU, W., BHATIA, S., HARE, J., SOMASUNDARAM, T., AZZI, A. & CHAPMAN, M. S. 2002. The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc Natl Acad Sci U S A,* 99**,** 10405-10.

YAKES, F. M., CHINRATANALAB, W., RITTER, C. A., KING, W., SEELIG, S. & ARTEAGA, C. L. 2002. Herceptin-induced inhibition of phosphatidylinositol-3 kinase and Akt is required for antibody-mediated effects on p27, cyclin D1, and antitumor action. *Cancer Research,* 62**,** 4132-4141.

YAMAGUCHI, Y., SHIMOJO, S., IKEDA, T., FUKUHARA, M., TSUNAKA, Y., SHIBUYA, R., ROCAFORT, M. A. V., NAKATSUKA, R., HIROHATA, K., TORISU, T. & UCHIYAMA, S. 2025. Unfolding of viral protein 1 N-termini facilitates genome ejection from recombinant adeno-associated virus serotype 8. *Molecular Therapy Methods & Clinical Development,* 33.

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. *J Virol,* 73**,** 9468-77.

YAO, Y., WANG, J., LIU, Y., QU, Y., WANG, K., ZHANG, Y., CHANG, Y., YANG, Z., WAN, J. & LIU, J. 2022. Variants of the adeno-associated virus serotype 9 with enhanced penetration of the blood–brain barrier in rodents and primates. *Nature Biomedical Engineering,* 6**,** 1257-1271.

YU, K. K., AGUILAR, K., TSAI, J., GALIMIDI, R., GNANAPRAGASAM, P., YANG, L. L. & BALTIMORE, D. 2012. Use of Mutated Self-Cleaving 2A Peptides as a Molecular Rheostat to Direct Simultaneous Formation of Membrane and Secreted Anti-HIV Immunoglobulins. *Plos One,* 7.

ZADORI, Z., SZELEI, J., LACOSTE, M. C., LI, Y., GARIEPY, S., RAYMOND, P., ALLAIRE, M., NABI, I. R. & TIJSSEN, P. 2001. A viral phospholipase A2 is required for parvovirus infectivity. *Dev Cell,* 1**,** 291-302.

ZENT, C. S., SECRETO, C. R., LAPLANT, B. R., BONE, N. D., CALL, T. G., SHANAFELT, T. D., JELINEK, D. F., TSCHUMPER, R. C. & KAY, N. E. 2008. Direct and complement dependent cytotoxicity in CLL cells from patients with high-risk early-intermediate stage chronic lymphocytic leukemia (CLL) treated with alemtuzumab and rituximab. *Leuk Res,* 32**,** 1849-56.

ZHANG, R., CAO, L., CUI, M., SUN, Z., HU, M., ZHANG, R., STUART, W., ZHAO, X., YANG, Z., LI, X., SUN, Y., LI, S., DING, W., LOU, Z. & RAO, Z. 2019a. Adeno-associated virus 2 bound to its cellular receptor AAVR. *Nat Microbiol,* 4**,** 675-682.

ZHANG, R., XU, G., CAO, L., SUN, Z., HE, Y., CUI, M., SUN, Y., LI, S., LI, H., QIN, L., HU, M., YUAN, Z., RAO, Z., DING, W., RAO, Z. & LOU, Z. 2019b. Divergent engagements between adeno-associated viruses with their cellular receptor AAVR. *Nat Commun,* 10**,** 3760.

ZHAO, J., PENG, H., GAO, J., NONG, A., HUA, H. M., YANG, S. L., CHEN, L. Y., WU, X. S., ZHANG, H. & WANG, J. P. 2021. Current insights into the expression and functions of tumor-derived immunoglobulins. *Cell Death Discovery,* 7.

ZHONG, L., LI, B., MAH, C. S., GOVINDASAMY, L., AGBANDJE-MCKENNA, M., COOPER, M., HERZOG, R. W., ZOLOTUKHIN, I., WARRINGTON, K. H. & WEIGEL-VAN AKEN, K. A. 2008. Next generation of adeno-associated virus 2 vectors: Point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proceedings of the National Academy of Sciences,* 105**,** 7827-7832.

ZHU, D. Q., BROOKES, D. H., BUSIA, A., CARNEIRO, A., FANNJIANG, C., POPOVA, G., SHIN, D., DONOHUE, K. C., LIN, L. F., MILLER, Z. M., WILLIAMS, E. R., CHANG, E. F., NOWAKOWSKI, T. J., LISTGARTEN, J. & SCHAFFER, D. V. 2024. Optimal trade-off control in machine learning-based library design, with application to adeno-associated virus (AAV) for gene therapy. *Science Advances,* 10.

ZHU, X. H., WU, L. N., ZHANG, L., HAO, P., ZHANG, S. A., HUANG, J., ZHENG, J., LIU, Y. N., LI, W. J., ZHANG, Y. M., ZHOU, C. Y., ZHANG, Y. H., YIN, C. C. & QIU, X. Y. 2010. Distinct regulatory mechanism of immunoglobulin gene transcription in epithelial cancer cells. *Cellular & Molecular Immunology,* 7**,** 279-287.

ZINN, E., PACOURET, S., KHAYCHUK, V., TURUNEN, H. T., CARVALHO, L. S., ANDRES-MATEOS, E., SHAH, S., SHELKE, R., MAURER, A. C. & PLOVIE, E. 2015. In silico reconstruction of the viral evolutionary lineage yields a potent gene therapy vector. *Cell Reports,* 12**,** 1056-1068.

ZUCCARINO-CATANIA, G. V., SADANAND, S., WEISEL, F. J., TOMAYKO, M. M., MENG, H. L., KLEINSTEIN, S. H., GOOD-JACOBSON, K. L. & SHLOMCHIK, M. J. 2014. CD80 and PD-L2 define functionally distinct memory B cell subsets that are independent of antibody isotype. *Nature Immunology,* 15**,** 631-637.

13 Appendix

13.1 AAV titration qPCR DNA standard sequence.

GAGGCTGTGTTCCCTGATGTGGACTATGTGAACAGCACTGAGGCTGAAACCATCCTGGACAA
CATCACCCAGAGCACCCAGAGCTTCAATGACTTCACCAGGGTGGTGGGTTTTGTTTTGGCTG
TTAAAGGCAGTTTGGCTCCAATATCTGGAACAGTTCAAAGTGCTCATCTTATTATAAGATTTGAC
GAAAATGGAGTGCTACTATTTTGGTGAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAG
GACGGCAACATCCTGGGGCACAAGCTGGAGTACAACTACAACAGCCACAACGTCTATATCA
TGGCCGACAAGCAGATTTTCTTTGTCGATCCTACCATCCACTCGACACACCCGCCAGCGGC
CGCTGCCAAGCTTCCGAGCTCTCGAATTCTTTTCCCATATATGGAGTTCCGCGTTACATAACT
TACGGTAAATGGCCCGCCTGGCTGACCGCCCAACGACCCCCGCCCATTGACGTCAATAAT
GACGTATGTTCCCATAGTAACGCCAATAGGGTTTTGACACTCATATTTCTTACATGCCCACTATT
AAGGAAGGTAACTCACGAGAACTAATGGGCCAACAATCTATGCCCAACAGGCCTAATTTTGA
ATGAGGTGGGAATTGAAGGACTCACTACTCGCAAGCTGGCACAGAAGCTGGGCGTCGAGC
AGCCAACGCTGTACTGGCATGTGAAGAATAAACGGGCGCTCCTAGACGCGCTTGCCATCG
AAATGCTGGACTTTTCTAGGAACCCCTAGTGATGGAGTTGGCCACTCCCTCTCTGCGCGCTC
GCTCGCTCACTGAGGCCGGGCTTTTCCAGGCCTCATACATCTCCTTCAATGCGGCCTCCAA
CTCGCGGTCCCAAATCAAGGCTGCCTTGGACAATGCGGGAAAGATTATGAGCCTGACTAAT
CTAGAGATTCCGACCCAAGAGACTCAACTTCAAGCTCTTTAACATTCAAGTCAAAGAGGTCAC
GCAGAATGACGGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTCAGGTTTTTACCG
CTTCGTGAGCACAAATAACACTGGCGGAGTCCAGTTCAACAAGAACCTGGCCGGGAGATAC
GCCAACACCTACAAAAACTTTTCTTGATGAATCCTGGACCTGCTATGGCCAGCCACAAAGAA
GGAGAGGACCGTTTCTTTCCTTTGTCTGGATCTTTAATTTTTGGCAAACAAGGAACTGGAAGA
GACAACGTGTTACACATGTAAGCGACGGATGTGGCAAAAGTGACGTTTTTGGTGTGCGCCG
GTGTACACAGGAAGTGACAATTTTCGCGCGGTTTTAGGCGGATGTTGTAGTAAATTTGGGCGT
AACCGAGTAAGATTTCGACCGGAGATGTGCATGTTATGGGAGCCTTACCTGGAATGGTGTGG
CAAGACAGAGACGTATACCTGCAGGGTCCTATTTGGGCCAAAATTCCTCACACGGATGGAC
ATTTGAGCTGAATGAAGCCATACCAAACGACGAGCGTGACACCACGATGCCTGTAGCAATG
GCAACAACCTTGCGTAAACTATTAACTGGCGAACTACTTACTCTAGCTTCTTTCATCATGGCTG
ATGCAATGCGGCGGCTGCATACGCTTGATCCGGCTACCTGCCCATTCGACCACCAAGCGA
AACATCGC

13.2 Binding/internalisation assay qPCR DNA standard sequence.

GAACCGCATCGAGCTGAAGGGCATCGACTTCAAGGAGGACGGCAACATCCTGGGGCACA
AGCTGGAGTACAACTACAACAGCCACAACGTCTATATCATGGCCGACAAGCATTTTCCCTTC
ATACCCTCACGTATTCCCCCAGGTTTACATGTTCCAATATGATTCCACCCATGGCAAATTCCAT
GGCACCGTCAAGGCTGAGAACGGGAAGCTTGTCAT

13.3 Additional primer sequences

Table 25. Additional primer sequences.

Primer	Sequence
Subclone Fw	GCAGAGCTGGTTTAGTG
Subclone Rv	GCTACTTATCTACGTAGCC
IgHJ6 Cutsite Fw	TTTCTGCTACTGCCTGTG
IgHJ6 Cutsite Rv	ATCTGCTTCCAGCTTCGC

13.4 Trastuzumab-treated BT474 viability in ADCP assay

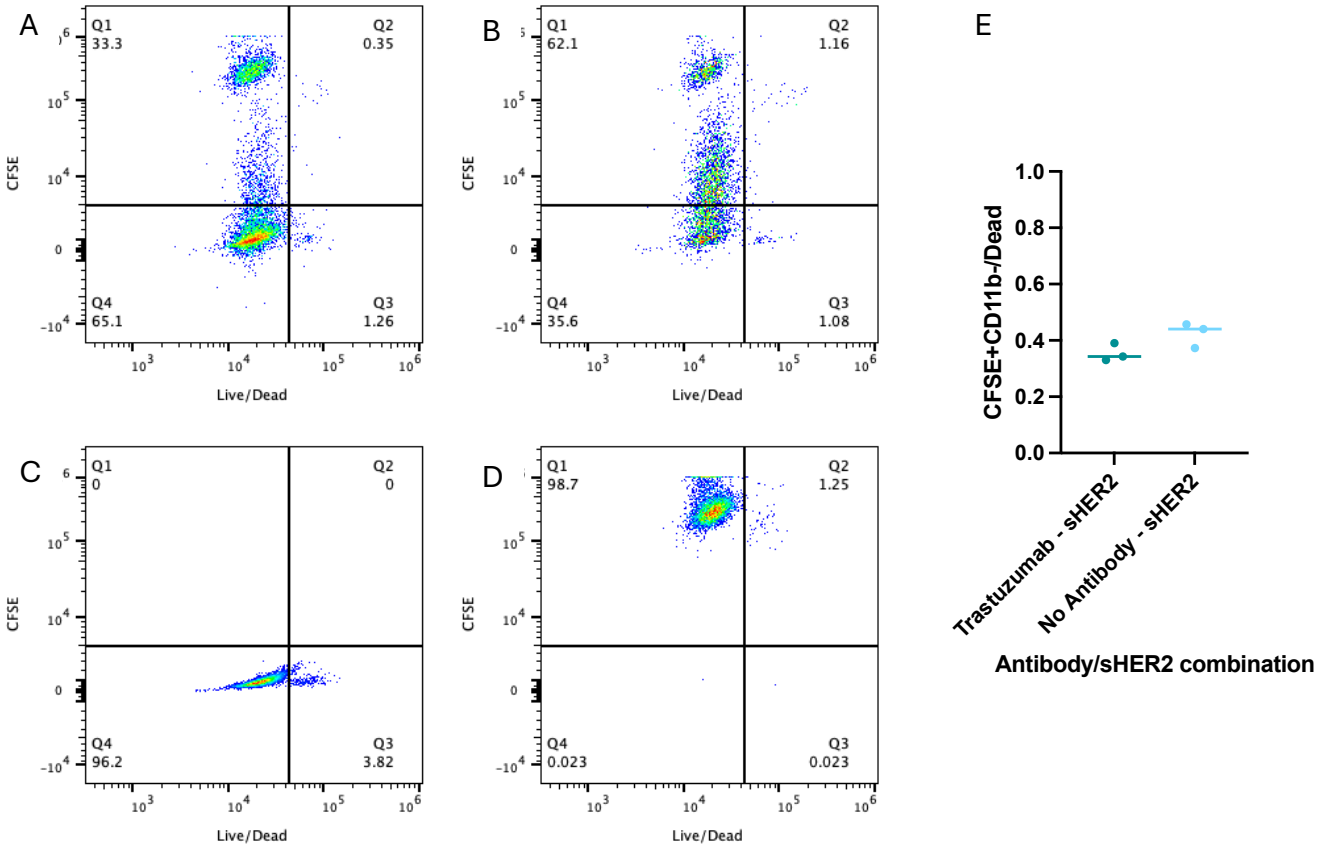


Figure 39 - Survival of BT474 cells treated with Trastuzumab during ADCP assay. *Representative CFSE and LIVE/DEAD quadrant density plots for trastuzumab + sHER2 (a), trastuzumab - sHER2. (b), monocyte only (c) or target cell only (d) conditions during ADCP assay. The frequency dead CFSE+CD11b- (BT474 target) cells is plotted in (e).*