Dissertation
submitted to the
Combined Faculties for the Natural Sciences and for Mathematics
of the Ruperto-Carola University of Heidelberg, Germany
for the degree of
Doctor of Natural Sciences

Presented by

M.Sc. Stefanie Große

born in: Gera, Germany
Oral-examination: 22.04.2016

Small but increasingly mighty: New insights into Adeno-associated virus (AAV) capsid biology and implications for AAV vector optimization

Referees: Prof. Dr. Martin Müller
Dr. Dirk Grimm

Summary

Over the last years, ground-breaking success has been achieved in the field of human gene therapy (i.e., the correction of diseases at the genetic level), culminating in the approval of the first gene therapeutic drug Glybera® in the Western world in 2012. Glybera® is based on Adeno-associated viruses (AAV) which currently form the basis for one of the most promising therapeutic vectors. Their advantages over other vectors are their apathogenicity in humans, the ability to achieve long-term gene expression in non-dividing cells due to episomal persistence, and the limited immune response after vector administration. Moreover, they exhibit a broad natural tropism for a variety of cell types. Still, some tissues cannot be transduced by AAVs and are thus refractory to AAV gene therapies. *Vice versa*, the broad tropism can lead to non-specific off-target expression of the delivered transgene. Consequently, there is a strong need for new AAV capsid variants which can overcome these hurdles.

One way to improve AAV properties is to modify the virus capsid via DNA family shuffling (DFS). This method is based on partial fragmentation of *cap* genes of several AAV serotypes and their homology-based re-assembly into chimeras. Since this process is highly error-prone, a first goal in this work was to optimize DFS on the technical and molecular levels. To this end, we aimed to improve the critical step of DNA fragmentation and hence compared standard DNaseI digestion with DNA shearing by ultrasonication. Of note, albeit the latter was more robust, it was outperformed by DNaseI in terms of shuffling efficiency and crossover formation. Secondly, we noted a dependency of recombination efficiencies on input DNA homologies that was most evident for the diverse *cap* genes of AAV4 and 5. We thus re-synthesized these variants and adapted their sequence to that of AAV2. The resulting 13 % homology increase translated into 2.5 x more crossovers per clone in an AAV2-4-5 *cap* library containing the new sequences as compared to its wild-type counterpart. These benefits were further validated with higher complexity libraries including larger numbers of AAV serotypes.

In addition, we investigated the role of the recently discovered, essential assembly-activating protein AAP on the outcome of the DFS method. As AAP is fully encoded in the AAV *cap* gene and was shown to be necessary for capsid assembly, we hypothesized that AAV shuffling may disrupt AAP integrity. To investigate this theory, we created assembly-deficient AAP knock-out mutants of AAV serotypes 1 – 9 and rh10. Notably, generation of functional particles could be restored in all cases by supplying AAP expression plasmids *in trans* during virus production. We furthermore tested 46 shuffled and randomly chosen AAPs from different AAV *cap* libraries in our new model system. Remarkably, 39 (84.8 %) were able to rescue the AAP knock-out, to extents varying from 5 to 100 %. Of note, most of the deficient AAPs originated from shuffled libraries containing AAV4 or 5 capsid genes. We thus produced virus particles from a shuffled capsid library of AAV2-4-5-8-9 in the presence or absence of a mixture of all corresponding AAP expression plasmids. Strikingly, AAP addition had no effect on virus titers albeit 60 % of the encoded AAPs were predicted to be non-functional due

to inadvertent recombination or disruption during shuffling. As a whole, our data strongly suggest that fortunately, unintended AAP shuffling is no major limiting factor for DFS-based generation of new AAV capsids.
Finally, our AAP knock-out constructs and a newly produced anti-AAP2 polyclonal antibody allowed us to study AAP biology in greater detail. Surprisingly, we detected an influence of AAP2 on viral *cap* protein (VP) stability, indicating the occurrence of post-translational VP modifications. Moreover, we studied AAP2 expression by microscopy and compared it to AAPs of other serotypes. Conspicuously, cellular localization varied between the AAPs, implying unique hotspots of viral capsid assembly.
Collectively, the work presented here has substantially improved the method of AAV DFS and should thus contribute to a much wider use of this powerful technology in the future. Concomitantly, our data provide novel insights into AAP biology and, together with our new constructs and models, open up a multitude of avenues for future research into the fundamental and applied aspects of AAV.

Zusammenfassung

Innerhalb der letzten Jahre konnten im Bereich der humanen Gentherapie (d.h. der Behandlung von Krankheiten auf genetischer Ebene) bahnbrechende Erfolge erzielt werden. Diese gipfelten 2012 in der Zulassung des ersten gentherapeutischen Medikamentes (Glybera®) in der westlichen Welt. Glybera® basiert auf Adeno-assoziierten Viren (AAV), welche die Grundlage für einen der momentan vielversprechendsten therapeutischen Vektoren bilden. Die Vorteile von AAVs gegenüber anderen Vektoren liegen in ihrer fehlenden Pathogenität im Menschen, einer beschränkten Immunreaktion nach Vektor-Verabreichung sowie der Fähigkeit, eine Langzeit-Expression in sich nicht-teilenden Zellen auf Grund von episomaler Persistenz zu erzielen. Zudem zeichnen sie sich durch einen natürlichen, breiten Tropismus in einer Vielzahl von Zellen aus. Dennoch können einige Gewebe nicht mit AAVs transduziert werden und sind deshalb von einer AAV-basierten Gentherapie ausgeschlossen. Umgekehrt kann der breite Tropismus zu einer unspezifischen, fehlgeleiteten Expression des transportierten Transgens führen. Infolgedessen existiert ein großer Bedarf an neuen AAV Kapsid-Varianten, die diese Hürden überwinden können.

Ein Weg, um die Eigenschaften von AAVs zu verbessern, ist die Modifikation des Virus-Kapsides durch „DNA family shuffling“ (DFS). Diese Methode basiert auf der partiellen Fragmentierung von Kapsid-Genen verschiedener AAV-Serotypen sowie deren Homologie-basierter Rekombination zu Chimären. Da dieser Prozess sehr fehleranfällig ist, war das erste Ziel der vorliegenden Arbeit, eine Optimierung von DFS auf technischer sowie molekularer Ebene zu erreichen. Insbesondere sollte der kritische Schritt der DNA-Fragmentierung verbessert werden. Dazu wurde der standardmäßig durchgeführte DNaseI-Verdau mit der DNA-Scherung durch Ultraschall verglichen. Obwohl sich dabei die letztere Methode als deutlich robuster herausstellte, übertraf der DNase-Verdau diese hinsichtlich der Erzeugung von Serotyp-Wechseln („Crossovers“) sowie Shuffling-Effizienz. Darüber hinaus konnte festgestellt werden, dass die Rekombinations-Effizienz unmittelbar von der Homologie der eingebrachten DNA abhängt. Dies offenbarte sich besonders deutlich bei den unterschiedlichen Kapsid-Genen von AAV4 und 5, weshalb beide neu synthetisiert und ihre Sequenzen an die von AAV2 adaptiert wurden. Daraus resultierte eine Homologie-Steigerung von 13 %. Der direkte Vergleich einer Kapsid-Bibliothek aus AAV2-4-5, welche die neuen Sequenzen beinhaltete, mit ihrem ursprünglichen Gegenstück zeigte daraufhin eine 2,5-fach höhere Zahl an Crossovers pro Klon. Auch in komplexeren Bibliotheken mit einer größeren Menge an AAV-Serotypen konnten diese Verbesserungen beobachtet werden.

Des Weiteren wurde in dieser Arbeit die Rolle des kürzlich entdeckten, essentiellen AAV-Proteins AAP („assembly-activating protein“) auf den Erfolg der Methode des DFS untersucht. AAP ist vollständig im Kapsid-Gen kodiert und wird für die Kapsid-Zusammensetzung benötigt. Es wurde deshalb vermutet, dass AAV-Shuffling möglicherweise die Integrität von AAP zerstören könnte. Um diese Theorie zu untersuchen, wurden AAP-Mutationen in die

AAV-Serotypen 1 – 9 sowie rh10 eingefügt. Diese waren daraufhin nicht mehr in der Lage, Kapside auszubilden. Die Erzeugung funktioneller Partikel konnte in allen Fällen durch die Zugabe von AAP-Expressionsplasmiden während der Virusproduktion wiederhergestellt werden. Als nächstes wurden 46 „geshuffelte“ und zufällig ausgewählte AAPs aus verschiedenen AAV-Kapsid-Bibliotheken mit diesem Modell-System getestet. Bemerkenswerterweise waren 39 davon (84,8 %) in der Lage, den AAP-Defekt zu 5 – 100 % zu kompensieren. Der Großteil der nicht funktionierenden AAPs stammte dabei aus Kapsid-Bibliotheken, welche AAV4 und 5 enthielten. Dies wurde zum Anlass genommen, Viruspartikel einer Kapsid-Bibliothek mit AAV2-4-5-8-9 in der An- sowie Abwesenheit einer AAP-Plasmid-Mischung der verwendeten Serotypen zu produzieren. Interessanterweise hatte die Zugabe von AAP keinen Effekt auf die Virus-Titer obwohl 60 % der kodierten AAPs auf Grund von Mutationen oder Beeinträchtigungen durch „Shuffling“ als nicht-funktional klassifiziert wurden. Insgesamt weisen diese Daten darauf hin, dass unbeabsichtigtes AAP-„Shuffling“ glücklicherweise keine wesentliche Beeinträchtigung für die DFS-basierte Generierung von neuen AAV-Kapsiden darstellt.
Zum Abschluss erlaubten die generierten AAV-AAP-Mutanten sowie ein neu produzierter, polyklonaler anti-AAP2 Antikörper tiefere Einblicke in die AAP-Biologie. Überraschenderweise wurde hierbei ein Einfluss von AAP2 auf die Stabilität der AAV2-Kapsid-Proteine festgestellt, was auf eine posttranslationale Modifikation dieser Proteine schließen lässt. Weiterhin wurde die Expression von AAP2 mikroskopisch untersucht und mit AAPs anderer Serotypen verglichen. Auffällig war dabei, dass die zelluläre Lokalisation zwischen den verschiedenen AAPs variierte, was spezifische Orte für die virale Kapsid-Zusammensetzung impliziert.
Zusammenfassend verbessert die hier präsentierte Arbeit die Methode des AAV-Shufflings substantiell und wird daher zu einer breiteren Nutzung dieser Technologie in der Zukunft beitragen. Gleichzeitig liefern die vorgestellten Daten neue Einblicke in die AAP-Biologie und eröffnen zusammen mit den hergestellten Konstrukten und neuen Modellen eine Vielzahl an Wegen für die zukünftige Erforschung grundlegender sowie angewandter Aspekte von AAV.

Contents

Nomenclature

4_{O}	sequence-optimized version of AAV4 *cap*
$4_{O\ repaired}$	sequence-optimized version of AAV4 *cap* with repaired AAP frame
5_{O}	sequence-optimized version of AAV5 *cap*
$5_{O\ repaired}$	sequence-optimized version of AAV5 *cap* with repaired AAP frame and *rep* stop codon
5_{O_rep}	sequence-optimized version of AAV5 *cap* with repaired *rep* stop codon
A	Alanine
A1	A1 acceptor site in AAV transcripts
A2	A2 acceptor site in AAV transcripts
aa	Amino acid
AAP	Assembly-activating protein
AAP2	Assembly-activating protein of AAV serotype 2
AAV	Adeno-associated virus
AF	Alexa Fluor®
APS	Ammonium persulfate
bp	Basepairs
BSA	Bovine serum albumin
CD34	Hematopoietic progenitor cell antigen CD34
cDNA	Complementary DNA
CFU	Colony forming units
ChIP	Chromatin immunoprecipitation
CMV	Cytomegalovirus
D	Aspartic acid
DARPins	Designed ankyrin repeat proteins
ddH_2O	Double-destilled water
DFS	DNA family shuffling
DMEM	Dulbecco's Modified Eagle Medium
DMSO	Dimethyl sulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxynucleoside triphosphate
ds	Double-stranded

E. coli	*Escherichia coli*
EDTA	Ethylenediaminetetraacetic acid
EM	Electron microscopy
ER	Endoplasmic reticulum
FBS	Fetal bovine serum
G	Glycine
g	Gram or Relative centrifugal force (context-dependent)
GFP	Green fluorescent protein
h	Hour
H. sapiens	*Homo sapiens*
HCl	Hydrochloric acid
HPV	Human papillomavirus
HRP	Horseradish peroxidase
HSPG	Heparan sulfate proteoglycan
HSV	Herpes simplex virus
IF	Immunofluorescence
Ig	Immunoglobulin
IP	Immunoprecipitation
ITR	Inverted terminal repeat
IVIG	Pooled human immunoglobulins
K	Lysine
KAc	Potassium acetate
kb	Kilobases
kDa	Kilodalton
L	Leucine
LB	Lysogeny broth
LPLD	Lipoprotein lipase deficiency
M	Methionine or Molar (context-dependent)
mA	Milliampere
MEM-NEAA	Minimum Essential Medium Non-Essential Amino Acids
miRNA	MicroRNA
ml	Milliliter
mM	Millimolar

MOI	Multiplicity of infection
mut	Mutant
N	Asparagine
NAc	Sodium acetate
NaCl	Sodium chloride
NaOH	Sodium hydroxide
NEM	N-Ethylmaleimide
NHP	Non-human primate
NLS	Nuclear localization signal
NoLS	Nucleolar localization signal
NPC	Nuclear pore complex
nt	Nucleotides
OD	Optical density
ORF	Open reading frame
P	Proline
PBS	Phosphate-buffered saline
PCR	Polymerase chain reaction
PEG	Polyethylene glycol
PEI	Polyethylenimine
PLA2	Phospholipase A_2 domain of VP1
po.1	Porcine 1
Q	Glutamine
R	Arginine
RBE	Rep binding element
RBS	Rep binding site
rh10	Rhesus 10
RI	Refraction index
RNA	Ribonucleic acid
rpm	Revolutions per minute
S	Serine
scAAV	Self-complementary AAV
SCID	Severe combined immunodeficiency
SDS	Sodium dodecyl sulfate

shRNA	Short-hairpin RNA
ss	Single-stranded
SUMO	Small ubiquitin-related modifier
T	Threonine
trs	Terminal resolution site
TBS	Tris-buffered saline
TEMED	Tetramethylethylenediamine
TGS	Tris/Glycine/SDS
U	Units
V	Valine or Volt (context-dependent)
v/v	Volume per volume
vg	Viral genomes
VP	Viral (capsid) protein
VR	Variable region
w/v	Weight per volume
WB	Western blot
wt	Wild-type
µg	Microgram
µl	Microliter
µm	Micrometer

List of Figures

List of Tables

1 Introduction

Over the last years, gene therapy approaches for the treatment of diseases like hemophilia, cancer, congenital blindness or central nervous system and immunodeficiency disorders were extensively developed and tested in numerous clinical studies. These types of therapies are characterized by the introduction of nucleic acids into patient cells to knock down, add or modify genes [1]. In order to deliver recombinant DNA, RNA or editing enzymes to the cells or organs of interest physical, chemical or biological methods can be selected [2]. However, physical treatment (e.g. electroporation) and chemical modifications (e.g. cationic lipids) are limited in *in vivo* applications and/or efficiency [3]. Therefore, most of the reported clinical trials chose viral vectors as a delivery tool for their transgene into the target cells. Widely used vector systems are gammaretro- and lentiviral vectors (both belong to the family of retroviruses), adenoviruses and Adeno-associated viruses (AAV) [4]. Retroviruses are mainly used for applications which require a stable integration of the transgene into the host genome to maintain protein expression. In addition to gammaretroviral vectors, lentiviruses can also infect non-dividing cells and therefore broaden the range of cell types for gene therapy [5]. One example for successful gene therapy with retroviruses is the treatment of patients suffering from SCID (severe combined immunodeficiency). In detail, haematopoietic stem cells were isolated from these patients, transduced *ex vivo* for the introduction of the transgene and transplanted back into the patient. The advantage of this treatment is ascribed to less off-targeting compared to an *in vivo* application and is more unlikely to trigger immune responses. Clinical trials conducted so far reported stable transgene expression for more than 10 months [6,7]. However, some of the treated individuals developed acute lymphoblastic leukemia caused by integration of the delivered transgene into a proto-oncogenic locus [8,9] which represents the major drawback of retroviral vectors. In contrast to retroviruses, adenoviral vectors and AAVs (mainly) persist episomally in the cell nucleus [5]. Thus, negative side effects due to integration events in regulatory regions or active genes are less likely to occur. Adenoviruses are valued therapeutic tools in terms of delivery efficiency and transgenic expression. On the other hand they can cause very strong immune reactions after systemic administration which unfortunately resulted in the death of at least one patient [10]. New vector generations already led to an improvement of immunogenicity, but applications are restricted to e.g. cancer gene therapy where immune stimulation is advantageous [2,5]. Therefore, one of the most aspiring gene therapeutic vectors are AAVs since they are not associated with a disease, mediate long-term transgenic expression due to episomal persistence in non-dividing tissue

and do not induce acute inflammatory responses since engineered vectors lack endogenous viral protein expression (see also chapter 1.2) [1, 2, 5]. Their great promise is supported by the first AAV-based gene therapy approved in Europe in 2012 (known as Glybera®) which is designed for patients with lipoprotein lipase deficiency (LPLD) [11]. Next to the successful approval of Glybera®, numerous clinical trials for treatment of e.g. hemophilia, congenital blindness or heart failure are currently ongoing [12] and underline the relevance of the AAV vector system for gene therapy. Thus, the following chapters will provide more details on AAV biology, life cycle and vector engineering which can be used for the generation of new, highly efficient AAV variants.

1.1 Adeno-associated virus (AAV)

1.1.1 AAV classification and genome organization

AAVs are small non-pathogenic viruses which belong to the family of *Parvoviridae* and the genus of Dependoviruses. The classification already implies the need of AAVs for a helper virus in order to undergo a complete replication cycle. Hence it is not surprising that they were discovered as contaminant of an Adenovirus preparation $\sim$ 50 years ago [13]. Next to Adenovirus also herpes simplex virus (HSV) or human papillomavirus (HPV) can mediate helper functions for efficient AAV replication [14–16].
As is typical for parvoviruses, the genome of AAVs consists of single-stranded (ss) DNA with a length of $\sim$ 4700 nucleotides (nt). The coding region of the genome is flanked by two identical palindromic regions (ITRs – inverted terminal repeats) that form a T-shaped hairpin-like structure [17], serve as viral DNA replication initiators [18] and possess transcriptional activity [19,20]. Besides the two ITRs, AAVs contain a *rep* and a *cap* gene which encode non-structural and structural proteins, respectively (see figure 1.1). These genes are driven by three promoters, namely p5, p19 and p40, and are terminated by the same polyA signal 3' of *cap* [21]. In detail, *rep* encodes four proteins [22] involved in viral DNA replication, integration and packaging (see also section 1.1.2). Two of them are transcribed from the p5 promoter (Rep78 and Rep68) whereas the other two start at the p19 promoter (Rep52 and Rep40). All of the transcripts contain an intron which is spliced out via one donor and one of two existing acceptor sites (A1 and A2). Unspliced transcripts lead to Rep78 and Rep52 proteins, spliced variants to Rep68 and Rep40, respectively [21]. The *cap* gene is transcribed from the p40 promoter resulting in a RNA that also contains the abovementioned intron. The RNA variant spliced through A1 (A1-RNA) encodes for structural protein VP1 while A2-spliced variants (A2-RNA) encode VP2 and VP3 [23, 24] (figure 1.1). Notably, A1-RNA generation seldom occurs compared to A2-RNA splicing and thus determines the relative protein ratio of VP1:VP2:VP3 with 1:1:10. In addition, VP2 is initiated by the non-consensus start codon ACG and therefore less frequently translated than VP3 with the conventional start codon AUG. In 2010, an additional protein was discovered to be translated from the p40-transcribed

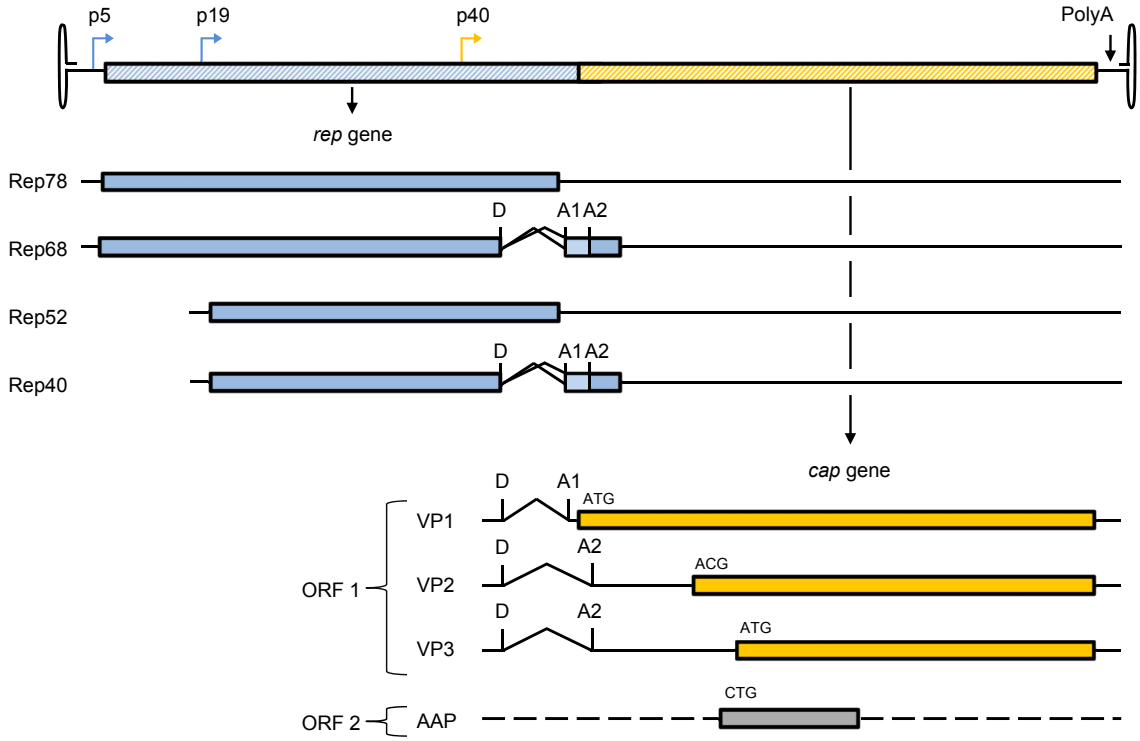

Figure 1.1: Genome organization and mRNA transcripts of AAV. The structure of the complete AAV genome is depicted on top and shows the two genes *rep* (blue) and *cap* (orange) flanked by the ITRs. The colors of the viral promoters p5, p19 and p40 indicate the transcription of the assigned genes. Viral transcripts for *rep* and *cap* are displayed below. All of them contain one splice donor site (D) and two acceptor sites (A1 and A2). In case of *rep* transcription, splicing results in two shorter proteins (Rep68 and Rep40) independent of the acceptor site used. In contrast, *cap* transcript splicing determines the ratio of viral capsid protein translation since A1 spliced RNA (encodes for VP1) appears less frequently than A2 spliced RNA (encodes for VP2 and VP3). Moreover, VP2 is driven by the non-conventional start codon ACG. Next to the VPs, *cap* also encodes AAP in its second ORF (grey) which is initiated by a CTG start codon. AAP = assembly-activating protein, ORF = open-reading frame, D = splice donor, A = splice acceptor.

RNA [25]. In contrast to VP1-VP3, it is encoded in the second open-reading frame (ORF) of the *cap* gene and utilizes the non-conventional initiation codon CUG which is located between the VP2 and VP3 start codons. The protein was named assembly-activating protein (AAP) and is reported to be required for capsid assembly (see section 1.1.2.3 for more details).

1.1.2 AAV life cycle

1.1.2.1 Cell entry, intracellular trafficking and nuclear transport

At the beginning of the AAV life cycle the non-enveloped AAV capsid binds to a cell-surface receptor and initiates endocytosis. This process is based on the different capsid structures and their diverse receptor binding capabilities (see chapter 1.1.3.2 for details). Depending on the bound receptor different routes of capsid internalization are activated. Although many studies have been conducted so far the process of AAV endocytosis is not fully understood.

Usually, AAV serotype 2 was used for cell entry studies and therefore is the best investigated AAV. It was shown that AAV2 can utilize at least two routes of infection which are clathrin-mediated endocytosis [26, 27] and the CLIC/GEEC pathway [28]. Other serotypes such as AAV5 can also internalize through caveolar endocytosis [29]. Generally, it is suggested that AAVs can enter the host cell through different pathways which can then direct the particles to an infectious (nuclear translocation) or non-infectious pathway (lysosomal / proteasomal degradation) [30]. Indeed, addition of proteasome inhibitors results in an enhanced transduction [31–33] which is associated with a prevention of capsid degradation potentially mediated by phosphorylation and ubiquitination [31,34,35]. However, this effect was cell-type specific and so far only shown for AAV serotypes 2, 5, 7 and 8.
Successful AAV infection requires not only efficient endocyctosis but also effective intracellular transport to the nucleus. It was shown that AAV2 trafficks through late and recycling endosomes towards the nucleus [36]. During this transport AAV2 particles undergo a conformational change in the internalized endosomes [27,32] and thereby expose their VP1 and VP2 N-termini to the outer capsid [37]. This process is required for efficient transduction and is based on the phospholipase A_2 (PLA2) domain located in the VP1 protein sequence [38,39]. In detail, the PLA2 domain leads to an endosomal escape of the virions from the golgi apparatus or the endoplasmic reticulum (ER) [40–43]. Moreover, VP1/2 N-termini contain basic regions which act as nuclear localization signals. They are responsible for the translocation of viral particles from the cytoplasm to the nucleus after endosomal escape [43, 44]. The transport of AAVs into the nucleus happens through the nuclear pore complex (NPC) from where the virions pass the nucleoplasm in direction to the nucleoli [27,45,46]. Intact particles are preserved until they escape again from the nucleoli and release their viral DNA.

1.1.2.2 Uncoating, integration and DNA replication

Previous reports have shown that AAV particles which enter the nucleus are intact and translocate to the nucleoli (see previous section). While the process of uncoating and viral genome release has not been completely elucidated so far, it is assumed that DNA release takes place in the nucleoplasm and is not dependent on the complete disassembly of the particle [45,47,48]. Once the viral DNA is unpackaged it has to be transformed from a ss to a ds (double-stranded) form prior to transcription. This can happen through annealing of positive and negative DNA strands which are packaged as single copies in the AAV capsid [49]. On the other hand, complementary strand synthesis occurs and is initiated from the free 3' hydroxyl group of the ITR which serves as primer for the host DNA polymerase [50]. During this process AAV DNA is also replicated for virus particle production. This so-called "rolling hairpin replication" is depicted in more detail in figure 1.2. Importantly, the process of AAV replication and transcription is dependent on the presence of a helper virus. In the absence of e.g. adenovirus, AAV integrates into the host genome which requires the AAV ITRs and one of the viral proteins Rep78 or Rep68. In detail, AAV Rep binds simultaneously to sequences in the ITR and in the AAVS1 locus of human chromosome 19 and thereby tethers the AAV genome to its

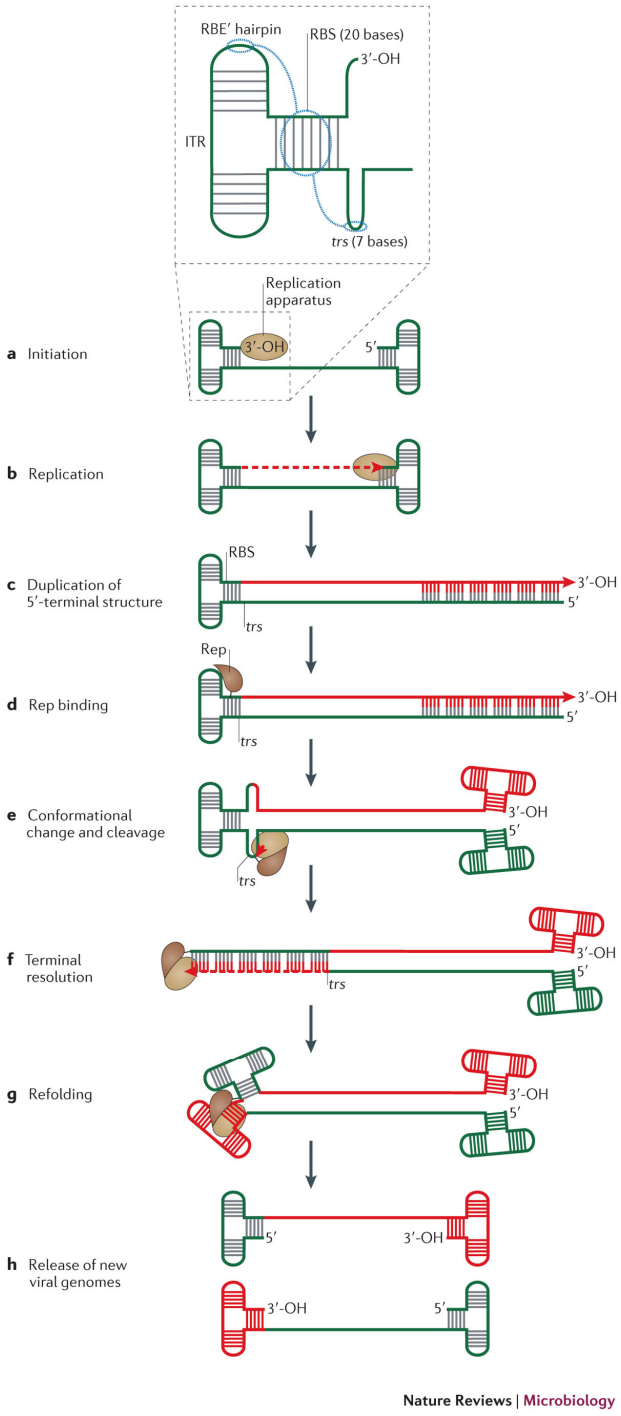

Figure 1.2: Schematic depiction of the AAV genome rolling hairpin replication (Chandler, M. *et al.*, 2013 [51]). Viral replication starts at the free 3' hydroxyl group of the ITR which serves as primer for the host DNA polymerase **(a)**. Elongation continues to the 5' end of the ss DNA **(b, c)** followed by the recruitment of Rep proteins to the 3'ITR **(d)** where they bind to the Rep binding element (RBE) and the Rep binding site (RBS) site (Rep interaction sites are marked with blue circles in the close-up picture on top of the figure). The interaction of Rep and ITR is assumed to induce a conformational change resulting in the nick of the DNA at the terminal resultion site (*trs*) mediated by Rep **(e)** and thereby allows replication of the 3' ITR **(f)**. Subsequently, ITRs refold **(g)** and single-stranded DNA is displaced resulting in two copies of ss viral DNA which provide new free 3' OH groups for a next replication cycle **(h)**. Note that introduction of a mutation in the *trs* of one ITR leads to a failure of Rep nicking and to dimer formation [52, 53]. As a result self-complementary AAV genomes are generated and packaged (not depicted here, see chapter 1.2.1 for more details and figure 1.5 for the resulting genome conformation).
The parental DNA strand is marked in green, the newly synthesized strand in red. ITRs are displayed as T-shaped structures.

integration site [54–56]. Once the host cell is infected with a helper virus, latent AAV copies are activated and replication can start in order to produce new viral particles.

1.1.2.3 Protein expression, virus assembly and the involvement of AAP

Immediately after viral ds DNA is synthesized, protein expression can occur. First, viral Rep proteins are translated since they regulate DNA replication, gene expression and viral

integration. Subsequently, structural proteins VP1-VP3 are expressed and virus particles assemble [57]. Since 2010 it has been known that an additional protein (named AAP) is required for the assembly process [25]. AAP is also encoded in the *cap* gene (see section 1.1.1) and is highly conserved among the different AAV serotypes. Typically, AAV2-AAP (AAP2) localizes in the cell nucleoli as it contains nuclear and nucleolar localization signals [58]. Moreover, it targets the VPs to this organelle where particle assembly takes place. It has been suggested that AAP acts as a scaffold or a chaperone for the VPs in order to allow viral capsid formation. Probably also intracellular proteins like nucleophosmin or nucleolin are involved in this process [59,60]. After the capsid has been formed, large Rep proteins target the ss viral DNA by simultaneous binding of ITRs and VPs to the particle [61]. Rep40 and Rep52 helicase activities are further required for translocation of the DNA through the pores at the five-fold symmetry axes of the capsid [62,63]. During this procedure only one copy of ss DNA (either plus or minus strand) is packaged per particle. Finally, newly formed AAV viruses spread from the nucleus throughout the cell [57] and are released from the cell due to lysis which occurs in the context of helper virus co-infection.

1.1.3 AAV capsid characteristics

1.1.3.1 Capsid structure

The AAV capsid consists of 60 VP units and forms a $\sim$ 18 – 26 nm icosahedral particle with a T=1 symmetry (figure 1.3b) [64,65]. Each particle encapsidates one copy of ss viral DNA (either plus or minus strand). The different cap proteins VP1, VP2 and VP3 are represented in the capsid in a ratio of 1:1:10 which is determined by the RNA splicing frequency and the relative translation efficiency due to alternative start codons (see also section 1.1.1). VP amino acid (aa) sequences overlap completely in the C terminus; hence the VP3 sequence is fully included in VP1 and VP2. VP1/2 share additional 65 aa. Moreover, VP1 comprises a 137 aa unique N-terminal region which mediates the phospholipase A2 activity mentioned in section 1.1.2. Crystal structures, cryo-electron microscopy studies and image reconstructions of different AAV capsids have shown that only the shared aa residues of all VPs constitute the outer capsid surface. N-terminal extensions of VP1 and VP2 as well as the first $\sim$ 15 aa of VP3 could not be resolved and are hidden in the inner particle. So far structures of nine AAV serotypes (AAV1 – AAV9) have been determined in detail, with AAV2 remaining the best characterized [66–76].

Within the particle the single VP subunits are arranged in a two-, three- and five-fold symmetry which results in the characteristic AAV capsid surface composed of cylindrical channels surrounded by canyons at the five-fold axes, protrusions at the three-fold axes and depressions at the two-fold axes (see figure 1.3b). The three-fold protrusions are spike-like structures formed by exposed loops from the VP subunits. Each of these subunits contains one conserved α-helix and one β-barrel with eight anti-parallel β-sheets (βB – βI) connected by large loops (figure 1.3a) [66,67]. Loops are named according to their flanking β-sheets, e.g. HI-loop.

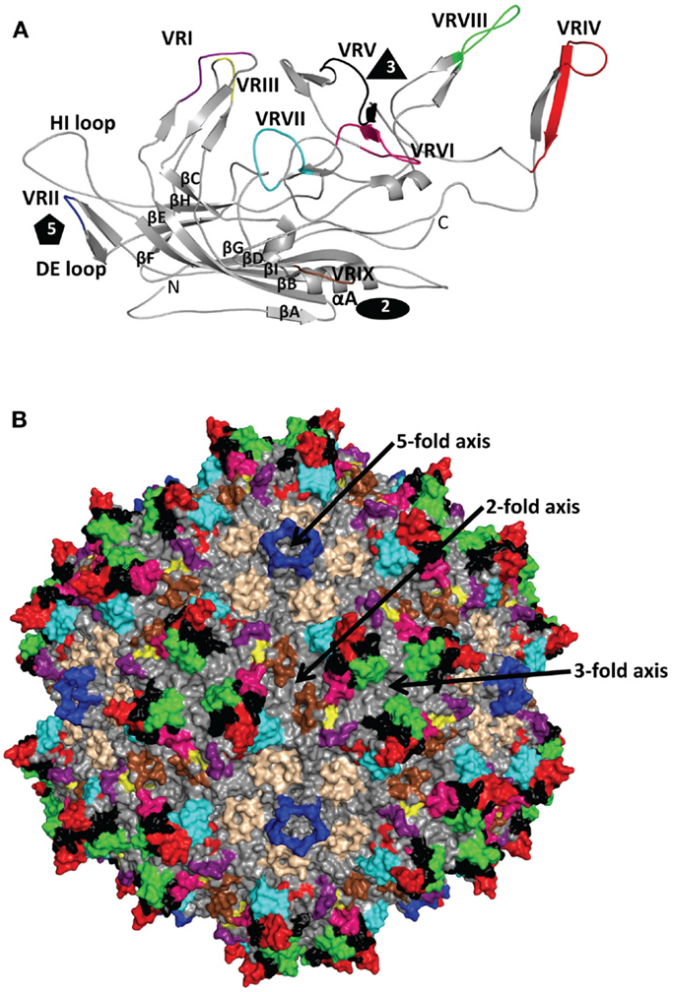

Figure 1.3: AAV2 capsid structure and variable regions (Tseng, M. & Agbandje-McKenna, M., 2014 [79]). **a)** Ribbon structure of the resolved overlapping AAV VP3 C-terminus. The protein contains one conserved α-helix (αA), one single β-sheet (βA) and one β-barrel with eight anti-parallel β-sheets (βB – βI) connected by large loops. Loop structures encompass the nine identified variable regions (VRs) and are named according to their flanking β-sheets (loop DE or HI are labeled as examples). Variable regions are colored as follows: VR-I = purple, VR-II = blue, VR-III = yellow, VR-IV = red, VR-V = black, VR-VI = pink, VR-VII = cyan, VR-VIII = green, VR-IX = brown, HI loop = wheat. N- and C-terminal ends are marked with N and C, respectively. Two-, three- and five-fold symmetry axes are indicated by a black oval, triangle and pentagon, respectively. **b)** Surface rendition of the AAV2 capsid. Icosahedral symmetry axes are labeled. Colors of variable regions match with a). Note the cylindrical channels at the five-fold and the protrusions at the three-fold axes (see text for more details).

The comparison between AAV serotype 2 and 4 capsids revealed a high variability within these loop structures. Thus, they were named variable region (VR) I – IX [69,77]. Resulting differences in surface topology are responsible for varying antibody interactions and receptor bindings. For instance, the GH loop encompasses VR-IV and -VIII which build the aforementioned peak structure at the three-fold axes and mainly mediate AAV2 receptor binding (see also sections 1.1.3.2 and 1.2.2.2).

The cylindrical pores at the five-fold symmetry axes connect the inner part of the capsid with the outside. The N-terminal regions of VP1 and VP2 which have not been accessible for structural determination so far are potentially located at the base of these channels [66, 78]. Due to conformational changes of the capsid during intracellular trafficking, VP1/2 N termini are exported through the five-fold channel and exposed on the outer capsid surface (see also section 1.1.2). Next to the externalization of VP domains, also the encapsidation of viral DNA occurs through the five-fold pore [63].

Of note, not only virus packaging or tissue tropism are influenced by the capsid structure but also the interaction with the immune system. In detail, depending on their surface epitopes AAVs can cause B-cell mediated antibody responses [80] which entails neutralization of virus.

This is especially important for AAV vector applications in gene therapeutic approaches and was thus a matter of intense investigation: It has been shown that the prevalence of pre-existing antibodies in the human population reaches up to 70 % depending on the AAV serotype [79]. Anti-AAV2 antibodies could be detected in $\sim$ 60 – 70 % of examined individuals followed by $\sim$ 35 – 70 % with anti-AAV1 antibodies, $\sim$ 50 % anti-AAV9, $\sim$ 50 % anti-AAV6, $\sim$ 40 % anti-AAV8 and $\sim$ 40 % anti-AAV5 antibodies [81,82]. For capsids with high homology (e.g. AAV2 and 3) also antibody cross-reactivities could be observed [83].

1.1.3.2 AAV serotypes and their tropism

Since the discovery of AAVs $\sim$ 50 years ago at least 12 serotypes and more than 100 isolates of AAVs were described. They are characterized by a different cell tropism and/or immunogenicity and classified according to their similarities in VP sequences. Based on these protein analyses, nine clades of AAVs are defined (clade A – I, figure 1.4) [84,85]. AAVs assigned to clade A – C and F originate from humans whereas AAVs belonging to clade D or G – I were isolated from non-human primate (NHP) or porcine tissue, respectively. An exception is clade

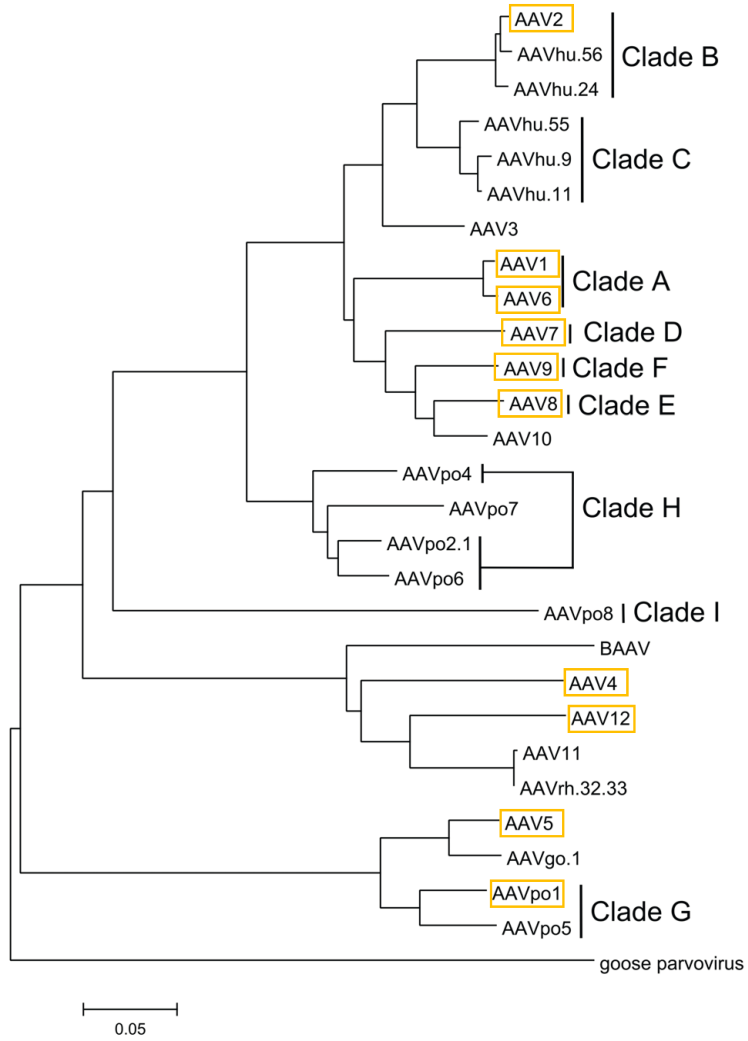

Figure 1.4: AAV clades (modified from Bello, A. *et al.*, 2014 [84]). Phylogenetic tree of important AAV isolates based on the comparison of AAV capsid protein sequences. Designated AAV clades are indicated by vertical lines next to representative members. AAVs marked in orange are variants used in this study. Importantly, AAV3, 4, 5 and 12 are not assigned to any clade. Serotypes AAV3b and AAVrh10, which are not depicted in the figure, are closely related to AAV3 and AAV10 (6 / 12 aa difference), respectively. The phylogenetic distance between the AAV variants is indicated by the scale bar at the lower left. hu. = human, po. = porcine, rh. = rhesus macaque, go. = goat.

E that contains AAV isolates from human as well as NHP tissue. AAV serotypes 3, 4, 5 and 12 are not assigned to any clade. Out of these, especially AAV4, 5 and 12 substantially differ from the sequences of other AAV variants and show a sequence identity of only ~ 60 % (see also figure 3.3). In contrast, AAV1 – 3, 6 – 9 and rh10 exhibit between 78 – 97 % sequence homology. However, already these small differences in the aa sequence can influence AAV cell tropism due to variations in receptor binding. In detail, AAV2 and 3b use heparan sulfate proteoglycan (HSPG) as main receptor for cell entry [86–88] whereas AAV1, 4 and 5 bind N- or O-linked sialic acids [89,90]. AAV6 which is assumed to be a chimera between AAV1 and AAV2 [91] binds both, HSPG and sialic acids [89,92]. For AAV7, 8, 12, rh10 and po.1 the primary cell receptor has not been identified so far. Interestingly, AAV9 is the only variant that uses N-linked glycans with terminal galactosyl residues as binding partner [93]. Thus, AAV9 transduction can be increased by neuraminidase treatment of the cells prior to infection since sialic acids mask N-linked galactose on the cell surface [93,94]. Recently, Pillay *et al.* discovered a new universal receptor for different AAV serotypes [95]. This transmembrane protein was named AAVR and seems to be essential for cellular transduction of AAV1, 2, 3b, 5, 6, 8 and 9. Next to the binding of AAV particles, AAVR was shown to enable virus trafficking from the cell membrane to the *trans*-Golgi network. Hence, AAVR appears to play a critical role in the infectious life cycle of human and NHP AAVs.
Next to primary receptors, also co-receptors were identified to be necessary for efficient viral cell entry. Particularly for AAV2 different co-receptors have been described and studied in more detail, i.e. $\alpha V\beta 5$ and $\alpha 5\beta 1$ integrins [96,97], fibroblast growth factor receptor 1 (FGFR1) [87], hepatocyte growth factor receptor (HGFR) [98] and laminin receptor (LamR) [99]. In addition, AAV8 and 9 were reported to use LamR whereas AAV3 also binds FGFR1 [99,100]. Table 1.1 summarizes the known receptors and co-receptors for the mentioned AAV variants and depicts reported AAV cell specificities ranging from liver, heart and kidney to skeletal muscle and the central nervous system. However, although many cell types can be transduced with AAVs others like T- or B-cells are less or not permissive for AAV transfer and thus not accessible for gene therapeutic applications. Of note, modifications of the viral capsid can overcome these hurdles. In order to provide further insights into the generation of new AAV variants, the next chapter will explain the options for genomic and capsid surface engineering in more detail.

Table 1.1: Mainly used AAV serotypes or variants and their tropism. Informations were gathered from publications [30] and [65] as well as serveral original articles.

AAV serotype	**Origin**	**Receptor**	**Co-receptor**	**Tropism[a]**	**Sequence reference**
AAV1	Laboratory isolate (probably human)	N-linked sialic acid	Unknown	SM, CNS, retina, pancreas	Hoggan *et al.* (1966) [101]
AAV2	Human	HSPG	FGFR1, HGFR, LamR, CD9 tetraspanin	VSMC, SM, CNS, liver, kidney	Hoggan *et al.* (1966) [101]
AAV3b	Human	HSPG	FGFR1, HGFR, LamR	SM, human liver (x.m.m.)	Rutledge *et al.* (1998) [102]
AAV4	NHP (African green monkey)	O-linked sialic acid	Unknown	CNS, retina	Parks *et al.* (1967) [103]
AAV5	Human	N-linked sialic acid	PDGFR	SM, CNS, lung, retina	Bantel-Schaal *et al.* (1984) [104]
AAV6	Laboratory isolate (probably human)	N-linked sialic acid, HSPG	EGFR	SM, heart, lung	Rutledge *et al.* (1998) [102]
AAV7	NHP (rhesus monkey)	Unknown	Unknown	SM, retina, CNS	Gao *et al.* (2002) [105]
AAV8	NHP (rhesus monkey)	Unknown	LamR	Liver, SM, CNS, retina, pancreas, heart	Gao *et al.* (2002) [105]
AAV9	Human	N-linked, terminal galactose	LamR	Liver, lung, pancreas, kidney, heart (i.v), brain (i.v.), SM (i.v.)	Gao *et al.* (2004) [85]
AAVrh10	NHP (rhesus monkey)	Unknown	Unknown	CNS (r)	Gao *et al.* (2004) [85]

AAV serotype	Origin	Receptor	Co-receptor	Tropism[a]	Sequence reference
AAVpo.1	Porcine	Unknown	Unknown	SM, retina	Bello *et al.* (2009) [106]
AAV12	NHP (Vervet monkey)	Unknown	Unknown	Lung (t.c.), colon (t.c.), nasal epithelia	Schmidt *et al.* (2008) [107]

[a]Tissue tropism after local delivery unless indicated otherwise
Abbreviations: AAV = Adeno-associated virus, CNS = central nervous system, EGFR = epidermal growth factor receptor, FGFR1 = fibroblast growth factor receptor 1, HGFR = hepatocyte growth factor receptor; HSPG = heparan sulfate proteoglycan, i.v. = intravenous injection, LamR = laminin receptor, NHP = non-human primate, PDGFR = platelet-derived growth factor receptor, r = rodent model, SM = skeletal muscle, t.c. = tissue culture, VSMC = vascular smooth muscle cell, x.m.m. = xenograft mouse model.

1.2 Recombinant AAV vectors for gene therapy

1.2.1 Genome engineering

Naturally, the AAV genome consists of a *rep* and a *cap* gene flanked by two ITRs (see chapter 1.1.1 for more details). In order to use AAVs as vectors for gene therapeutic approaches a transgene has to be encoded and packaged within the viral capsid. This can be simply achieved by exchanging *rep* and *cap* with a transgene of interest under the control of a selected promoter as well as a polyA signal. *Rep* and *cap* can be provided *in trans* during vector production and hence ITRs are the only *cis*-acting elements required for efficient vector generation [108]. In addition, adenoviral genes (namely E2A, E4, VA-RNA and E1A/E1B) have to be supplied for stimulation of productive AAV replication in the absence of infectious helper-virus [109]. The advantage of the described system is the complete elimination of AAV genes which results in replication-incompetent particles with an improved safety profile. Moreover, it has been shown that the absence of Rep proteins averts the directed integration of AAV vector genomes into the AAVS1 locus [110, 111]. Instead of integrating, recombinant AAV genomes persist preferentially episomally [110, 112, 113] and therefore mediate long-term expression especially in non-dividing cells. However, AAVs were also reported to integrate randomly to a small extent (< 1 %) into the host cell genome [1, 112, 114]. Of note, neither integration hotspots could be detected in AAV treated patient samples [113] nor a correlation between integration sites and oncogenesis could be observed in mice [115]. The latter was recently questioned by a publication which postulated an oncogenic role of wild-type AAV2 in human hepatocellular carcinomas [116]. Yet, this presumption was critically discussed and disproved by leading AAV researchers who again emphasized the safety of AAV as a therapeutic tool [117, 118].
Since AAV expression is not only influenced by efficient gene transfer or intracellular trafficking but also by the effective conversion of ss to ds DNA [119, 120], researchers have improved the AAV vector design towards enhanced gene expression. Therefore, self-complementary AAV

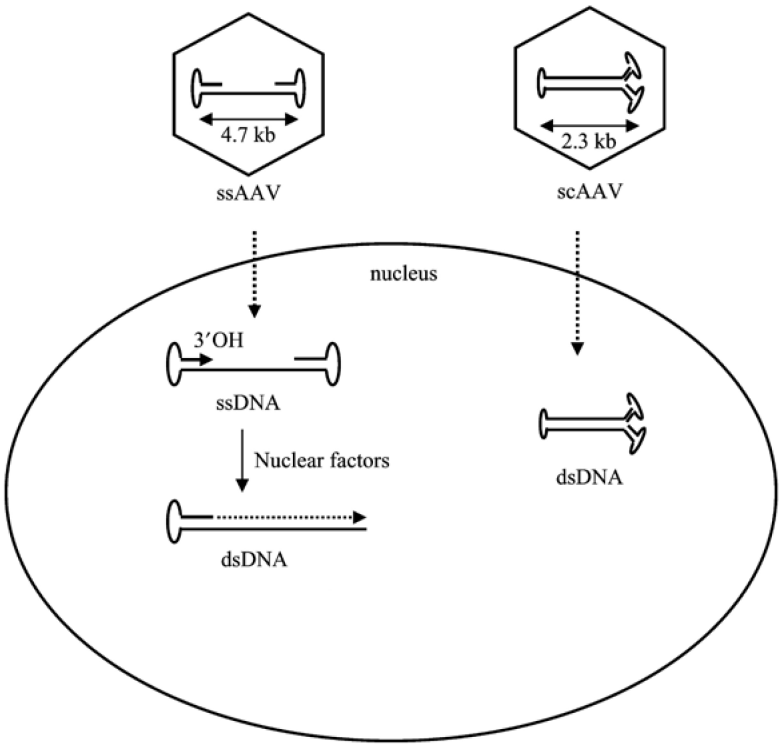

Figure 1.5: AAV genome conformations (modified from Le Bec, C. & Douar, A., 2006 [128]). Schematic depiction of single-strand (ss - left side) and self-complementary (sc - right side) AAV vector genomes. Note that the sc genome is half the size of the ssAAV. After uncoating in the nucleus, ss genomes have to be converted into ds DNA by nuclear factors for later transcription and replication. In contrast, sc genomes re-fold and directly serve as transcription template.

vectors (scAAV) were developed which carry ds instead of ss transgene expression cassettes (figure 1.5) [121]. For this purpose, one ITR of the AAV genome was mutated at the *trs* resulting in a failure of DNA nicking mediated by Rep (see also figure 1.2) [52, 53]. As a consequence, dimeric AAV genomes are created which can re-anneal after AAV uncoating in the nucleus and directly serve as transcription template independent of the hosts replication machinery. It is important to note that due to this modification the packaging capacity of scAAVs is reduced from $\sim$ 4.4 kb to $\sim$ 2.2 kb (without ITRs) in comparison to ssAAVs. However, 2.2 kb are still sufficient to encode small proteins or RNA-based therapeutics such as microRNAs (miRNA) or short-hairpin RNAs (shRNA).

In order to increase AAV packaging capabilities (mainly for ss vectors), dual AAV vectors have been developed. Therefore, oversized transgenic cDNA is split into two parts and packaged into separate vectors. Under these conditions the virus with the 5' part of the transgene also encodes a promoter whereas the vector with the 3' fragment encodes the polyA signal. Subsequent expression of the full-length protein can be achieved by homologous recombination in the cell due to overlapping regions in the cDNA [122–124]. Moreover, splice donor and acceptor sites can be added to the ends of the cDNA to allow for efficient recombination on the RNA level [124, 125]. Both methods were also combined and have yielded increased transgenic expression [126, 127].

1.2.2 Capsid engineering

1.2.2.1 DNA family shuffling

To overcome AAV gene therapy limitations caused by pre-existing immunity or a lack of suitable serotypes for specific tissues, the AAV capsid can be modified to obtain new variants with desired characteristics. One possibility to generate new particles is DNA family shuffling. The general method was first described in 1994 [129] and has the potential to create a large library of different DNA variants originating from a few parental sequences. In detail, DNA family shuffling is based on the fragmentation of input parental DNA and its re-assembly to full-length genes in a primer-less PCR reaction. A prerequisite for recombination during PCR is an initial DNA homology of at least 50 % of the parental sequences in order to allow smaller fragments to act as primers in the amplification reaction. During these priming events fragments of different parents bind each other and thereby generate crossovers which lead to highly diverse shuffled DNA variants.

In 2000 researchers started to apply the method to viruses and generated viral capsids with new tropisms [131, 132]. In 2008 AAV *cap* genes were successfully modified for the first time by Grimm and colleagues [133] followed by several other reports [134–136]. Each of the published studies used between 7 and 10 different AAV capsids as input material for library generation. In this respect it is advantageous that all AAV capsids exhibit a homology of more than 50 % (see section 1.1.3.2) which allows to include any existing capsid in the library design. The genes are then further fragmented by using DNaseI enzyme followed by a first re-assembly PCR and a second PCR with flanking primers to specifically amplify the full-length *cap* gene (figure 1.6). After purification of the specific product, the library is cloned into a wild-type AAV construct containing the *rep* gene and two flanking ITRs. Randomly picked chimeras of the generated library are sequenced in order to validate shuffling efficiency and library diversity. The latter typically ranges between $7x10^5$ and 10^7 different *cap* sequences per approach [133–135] which is normally determined by counting bacterial colony numbers. Finally, virus particles are produced by transfection of cells with the generated *rep-cap* library constructs and a plasmid providing all necessary adenoviral genes for replication, followed by selection of the shuffled AAV capsids in different cell lines or mouse models. The selection process is characterized by several rounds of transduction in the presence or absence of helpervirus infection and ideally enriches one or several functional AAV *cap* variants (see section 1.2.3 for more details). Two of the most promising selected and evaluated shuffled AAVs so far are AAV-DJ [133] and AAV-LK03 [136]. The first one is a chimera of AAV2, 8 and 9 *cap* sequences which outperformed eight naturally existing AAV serotypes in various cell lines as well as primary cells. Additional studies further validated the potential of AAV-DJ as gene therapeutic vector in various tissues [137,138] and in 2012 also the capsid structure could be resolved via cryo-EM [139]. The second AAV chimera, LK03, was selected *in vivo* in mouse livers which were partially repopulated with primary human hepatocytes. It mainly consists of AAV3b but also contains parts of AAV *cap* 1, 2, 4, 6, 8 and 9. LK03 was reported to specifically

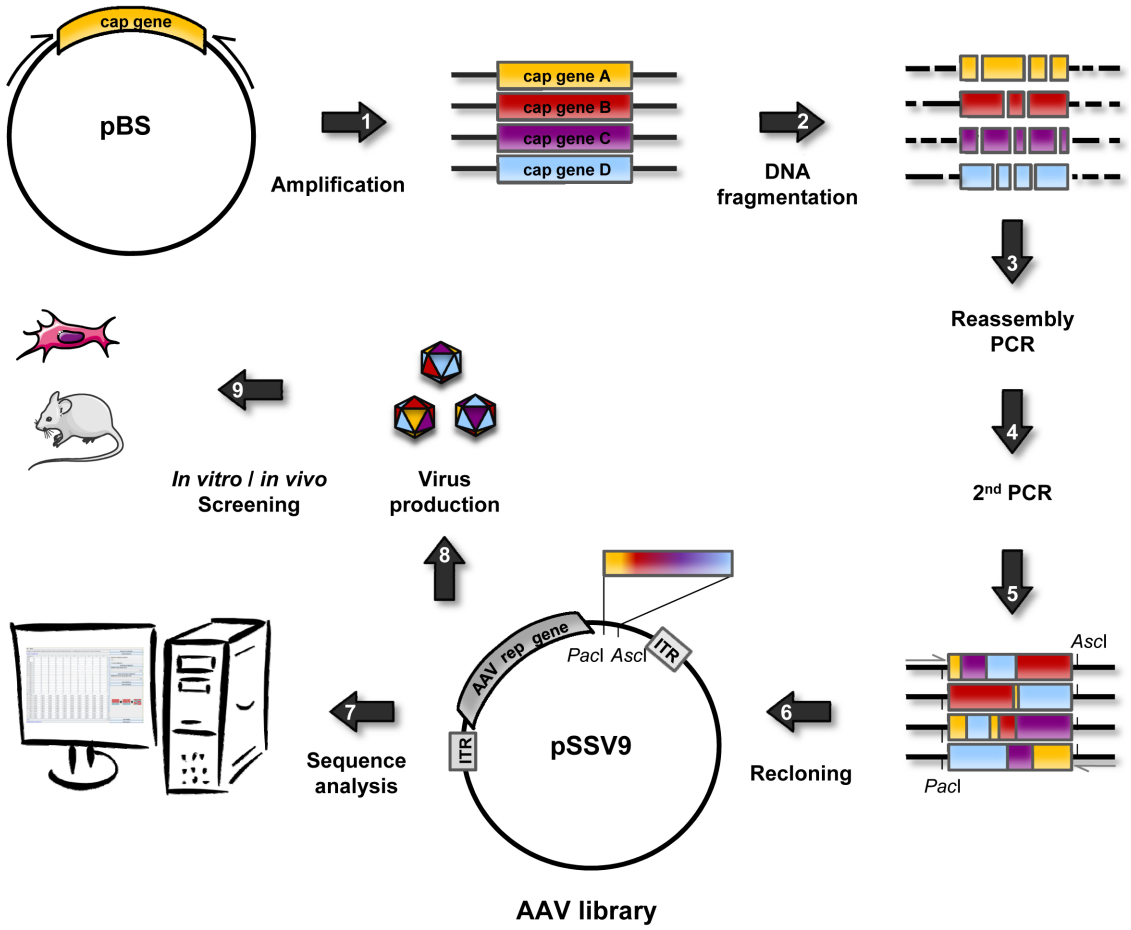

Figure 1.6: Schematic representation of AAV capsid DNA family shuffling. At the beginning **(1)** several AAV *cap* genes are selected and PCR-amplified from a plasmid construct. *Cap* genes are further purified and **(2)** fragmented e.g. with DNase enzyme. DNA fragments serve as input for a **(3)** primer-less re-assembly PCR followed by a **(4)** second PCR with flanking primers. Full-length shuffled *cap* genes are **(5)** purified and **(6)** re-cloned (in this study via PacI/AscI restriction sites) into an AAV wild-type plasmid containing AAV *rep* and ITRs. In order to validate the generated AAV library, **(7)** random clones are sequenced and analyzed. Successful validation is followed by **(8)** virus production and **(9)** several selection rounds *in vitro* or *in vivo*. pBS = Bluescript plasmid with AAV *cap* gene, pSSV9 = ss AAV plasmid carrying *rep* and *cap* flanked by two ITRs of AAV2 [130].

transduce human but not mouse hepatocytes *in vitro* and *in vivo* and thereby represents a leading candidate for human gene therapeutic approaches in the liver. In summary, AAV-DJ and LK03 exemplify the great potential of DNA family shuffling in the context of AAV capsids. Thus, shuffling is a promising tool for the generation of more AAV variants with specific features and expanded tropism.

1.2.2.2 Peptide display

Apart from DNA family shuffling the insertion of small peptides offers another option for the modification of AAV capsids. Typically, these peptides are inserted into exposed regions of the capsid surface and thereby (I) destroy the natural tropism and (II) mediate new recep-

tor binding. In contrast to DNA family shuffling, the capsid backbone comprises only one serotype. First attempts of ligand introduction into AAV were made in 1999 by Girod and colleagues [140]. They inserted a 14 aa peptide into several regions of the AAV2 capsid based on a hypothetical model of the virus. Of note, one of the generated capsid variants could efficiently transduce B16F10 cells which are normally not susceptible to AAV2 infection. Several subsequent studies modified the AAV2 particle in similar ways which resulted in altered vector tropism but also revealed more insights into structure-function relationships [141–144]. The fact that (random) aa insertions into the capsid can also influence viral assembly, DNA packaging or intracellular trafficking underlined the importance of basic knowledge of virus biology. Hence, capsid modifications were facilitated in 2002 when the atomic structure of AAV2 was elucidated [67] which enabled site-directed mutagenesis or rational design of insertion sites into specific positions of the viral surface. Interestingly, already the 1999 study determined a region (aa pos. 587) important for receptor binding which was later identified to be located in close proximity to the heparin binding motif in VR-VIII of AAV2 (aa R585 and R588) [145, 146]. Therefore, further efforts often inserted peptides between aa 587 and 589 in order to destroy the HSPG binding pattern of AAV2 and re-direct receptor binding to naturally non-target cells [147–149]. To moreover allow for a high-throughput analysis of different aa sequences, a library of randomly generated peptides is inserted into capsid positions of interest (named peptide display) and screened for lead candidates [150, 151]. Interesting peptide sequences that were identified by this method and lead to re-direction of AAV2 capsids are NDVRAVS and NSSRDLG in primary human endothelial cells [151], LATQVGQKTA (with an additional V719M mutation) in neural stem cells [152] or PSVSPRP in mouse heart cells [153], amongst others. However, although peptide display is an efficient method to redirect AAV capsids it is limited by the restricted knowledge of receptor binding motifs and 3D structures of other AAV serotypes. Thus, it is difficult to transfer lead peptide candidates into other capsids without the need for a new selection [154–156]. Moreover, the modifications achieved by peptide display are limited to less than 1 % of the capsid surface [157] and are rarely applicable for alterations of complex multi-domain interactions.

1.2.2.3 Other methods

Instead of DNA family shuffling or peptide display also other methods can be exploited for the modification of the AAV capsid surface. For example, with error-prone PCR random mutations are inserted into the *cap* gene during DNA amplification [158, 159]. Of note, the used polymerase misses proof-reading ability which favors point mutations. As a result, generated *cap* libraries carry single aa substitutions at random positions. In contrast to the already described methods, the diversity of achieved mutations per capsid is rather low [158] and thereby also limited in terms of complexity. Improvements of the method can be accomplished by restriction of mutagenesis to specific regions of the capsid like surface receptors or antibody binding sites [160, 161].

Next to error-prone PCR also site-directed mutagenesis can modify the capsid surface efficiently. This method resulted in e.g. the identification of the HSPG receptor of AAV2 before the 3D structure was solved [145,146] or immune escape *cap* variants with reduced binding of neutralizing antibodies [160,162]. Additionally, deliberated deletions of tyrosine residues at the capsid surface led to an improved transduction efficiency of several AAV serotypes due to an increased cellular trafficking of viral particles from the cell surface to the nucleus [163–167]. The latter is related to a decrease in virus capsid phosphorylation and subsequent ubiquitination, which in turn circumvents proteasomal degradation. Despite the efficient modification of several AAV capsids in the described studies, site-directed mutagenesis still requires detailed knowledge of the 3D particle structure. Since only a few AAV isolated have been characterized sufficiently so far, also this method is limited regarding diversity generation.
Another different approach for capsid modification is the incorporation of exogenous molecules into one of the viral proteins. It has been reported that VP2 tolerates a fusion of at least 238 aa (GFP) which is successfully integrated into the capsid and displayed at the surface [168,169]. Others used this ability to expose e.g. designed ankyrin repeat proteins (DARPins) on the AAV surface. DARPins are small, highly specific binders which are pre-selected on molecules of interest [170–172]. By integration of DARPins binding specifically to the CD4 receptor, efficient re-direction of AAV capsids to CD4-positive cells could be observed [173]. Another report introduced the variable region of an anti-CD34 antibody into the viral capsid and achieved an improved transduction of a $CD34^{+}$ human myoleukemia cell line [174]. However, although VP2 fusions can induce successful re-direction of vector tropism, the production of these modified AAVs is highly inefficient and needs additional purification steps.
Recently, an innovative strategy for the generation of new AAV capsids called "ancestral sequence reconstruction" was published [175,176]. Here the authors first analyzed isolated AAV variants and calculated their phylogenic relation. Second, they identified common nodes and reconstructed a library of ancestor sequences based on computational analysis. Impressively, one of the identified ancestral capsids (Anc80L65) yielded higher transduction rates than AAV8 in the liver and retina of injected mice and also outperformed this AAV in the liver of rhesus macaques [175]. Thus, ancestral reconstruction represents a powerful method for AAV capsid modification which can result in promising vector candidates for gene therapeutic applications.

1.2.3 Selection strategies

After AAV *cap* library generation with one of the aforementioned methods (see chapter 1.2.2), lead candidates have to be selected in order to obtain efficiently transducing vector particles. Therefore, two different selection strategies depending on the intended outcome can be applied. (I) AAV libraries can be selected *in vitro* on cells of interest that best match the future application. In case the chosen cell line is susceptible for adenovirus, a helpervirus co-infection can accelerate the selection process due to the induction of active AAV replication [133,150, 177]. If not applicable, PCR rescue has to be performed to recover potent AAV sequences

followed by new virus productions. Ideally, after 3 – 5 screening rounds at least one lead AAV candidate can be characterized. In order to obtain AAV variants that are resistant to antibody neutralization, an additional selection pressure by adding e.g. IVIGs (pooled human immunoglobulins) can be applied [133]. Therefore, IVIGs have to be pre-incubated with the library followed by cell transduction. Despite the simplicity and speed of *in vitro* selection, some drawbacks diminish the potential of the strategy. On the one hand it is not sure whether the evolved capsids target the same cell types under *in vivo* conditions [133,178]. On the other hand, efficient vectors often transduce more cell types apart from their selection tissue and thus cause unintended off-target effects [133]. To overcome these problems, (II) AAV libraries can be selected directly *in vivo* and interesting AAV sequences can be PCR-rescued from organs of interest. This strategy outperforms the *in vitro* system especially regarding the complexity of cellular interactions and the presence of the immune system which cannot be mimicked with a cell line. Normally, mice or rats serve as model system for *in vivo* selections [179–181]. However, since rodents are not necessarily predictive for the future application in human gene therapy [182, 183] also screenings in NHP can be considered [184, 185]. It is important to note that the latter is more cost-intensive and has to cope extensive administration processes, logistics as well as ethical barriers. Moreover, *in vivo* selections generally require AAV libraries with high virus titers and can last several months before a lead capsid candidate is enriched. Thus, one has to weigh the costs versus benefits before decisions about the selection system can be made.

1.3 Aim of thesis

Due to the need for AAV capsids that efficiently transduce specific tissues in gene therapeutic approaches, this thesis focused on the technique of DNA family shuffling for the generation of new AAV *cap* variants. We especially wanted to optimize the currently used protocol for DNA family shuffling and aimed to standardize specific steps (i.e. DNA fragmentation) which are highly susceptible to external influences. Moreover, our goal was to improve the incorporation of AAV *cap* 4 and 5 into shuffled libraries since several reports exemplified the challenges in integrating these distinct capsids. Towards this aim we planned to sequence-optimize these *cap* sequences and to compare their incorporation into different shuffled AAV libraries with varying complexities. In this context, we also wanted to determine the role of AAP in the shuffling procedure. Because of the limited knowledge of AAP biology it seemed plausible that AAP shuffling affects the outcome of the method and the formation of particles. Therefore, we planned to first develop a model system in order to be able to test functionalities of AAPs derived from different AAV serotypes. Second, we wanted to clone a number of shuffled AAPs from various AAV libraries which should be validated in this system. Based on the developed model we also intended to investigate AAP biology in more detail. Thus, we aimed to examine the effect of AAP knock-out on AAV RNA and protein levels. Moreover, we planned to obtain a polyclonal antibody against AAP2 which would allow us to perform detailed microscopy

analyses to gain more insights into AAP localization and expression. Apart from AAP2, also the localization of AAPs from other serotypes should be determined. In summary, this thesis wanted to optimize DNA family shuffling in order to allow improved AAV capsid modifications and wished to contribute to the elucidation of AAP functions.

2

Materials and methods

2.1 Materials

Table 2.1: Oligonucleotides

Internal number	Name	Sequence
#36	Mid – Cap 1-3/6-9, rh10, po.1, 12	GAAATTGGCATTGCGATTCC
#37	Mid – Cap 4/5	GATTGGCATTGCGATTCCAC
#178	DJRev	GTCGCAAAACACTCACGTGACCTC
#459	SAFor	GACTACAAGGACGACGATGACAAG
#460	SARev	CACTGAATTCTCATCAGGCGAAG
#641	Fw_SSV9_rep2_ HindIII	ACGCGGAAGCTTCGATCAAC
#642	Rev_SSV9_Spe/Xba	gatactctagaGTTCGACCGCAGCCTTTCGAATGTC CGGTTTATTactagtGGCGCGCCCTGGACTC
#780	mCherry-For BamHI	ctggcaggatccattgctcATGGTGAGCAAGGGCGAGG AGGATAACATG
#781	mCherry-Rev NotI/XhoI/SalI	ccgacgtcgacggttagcggccgcattctccgctcgagcggattaCTT GTACAGCTCGTC
#792	pEGFP _C2_FP	GATCACATGGTCCTGCTG
#822	LSeqFor	GATCTGGTCAATGTGGATTTG
#823	LSeqRev	GACCGCAGCCTTTCGAATGTC
#827	M13For	GTAAAACGACGGCCAGTGAG
#828	M13Rev	GGAAACAGCTATGACCATG
#833	Rep2For	AGACGCGGAAGCTTCGATCAA
#835	CMV for	CGCAAATGGGCGGTAGGCGTG
#964	SwaI_4GA For	ctgatttaaatCAGGTATGACCGATGGTTATCTTC
#965	SwaI_5GA For	ctgatttaaatCAGGTATGTCTTTCGTTGATCATC
#966	AAV4GA_PacI For	ccttaattaaCAGGTATGACCGATGGTTATCTTC

Internal number	Name	Sequence
#967	AAV4GA_AscI_SpeI Rev	ggactagtggcgcgccTTACAGATGATGAGTCAGGTATC
#968	Start-Leu repair AAP4GA For	GTTTTTGAGGACGAGACTGGCGCAGGCGACGGACC
#969	Start-Leu repair AAP4GA Rev	GGTCCGTCGCCTGCGCCAGTCTCGTCCTCAAAAAC
#970	1st STOP remove AAP4GA For	CTGGCGGAGCAGCGGTCGAGGGTGGCCAGGG
#971	1st STOP remove AAP4GA Rev	CCCTGGCCACCCTCGACCGCTGCTCCGCCAG
#972	2nd STOP remove AAP4GA For	GGGAATGCGACCCAAGGCAATGCGGGTCAAGATCTTTAACATTCAAG
#973	2nd STOP remove AAP4GA Rev	CTTGAATGTTAAAGATCTTGACCCGCATTGCCTTGGGTCGCATTCCC
#974	STOP repair AAP4GA For	GTACGAGCTCCCGTACGTGATGGACGCGGGGCAGG
#975	STOP repair AAP4GA Rev	CCTGCCCCGCGTCCATCACGTACGGGAGCTCGTAC
#976	TGA_repair 5GA_PacI For	ccttaattaaCAGGTATGTCTTTCGTTGATCATCCTCCAGATTGGCTCGAGGAGGTTGGTGAAGGACTAAGAGAGTTC
#977	AAV5GA short_PacI For	ccttaattaaCAGGTATGTCTTTCGTTGATCATC
#978	AAV5GA AscI_SpeI Rev	ggactagtggcgcgccTTACAGAGGACGAGTCAGG
#979	Start-Leu repair AAP5GA For	CTAGCGACGCAGAGGCTGGACCTAGCGGATCCCAG
#980	Start-Leu repair AAP5GA Rev	CTGGGATCCGCTAGGTCCAGCCTCTGCGTCGCTAG
#981	STOP remove AAP5GA For	GAAATTAAGAGCGGATCAGTCGACGGCTCGAACGCCAATG
#982	STOP remove AAP5GA Rev	CATTGGCGTTCGAGCCGTCGACTGATCCGCTCTTAATTTC
#983	AAP2 NotI For	ataagaatgcggccgcATGGAGACGCAGACTCAGTAC
#984	AAP2 / 3b EcoRI Rev	ccggaattcTCAGGGTGAGGTATCCATAC
#985	AAP4 NotI For	ataagaatgcggccgcATGGAGCAGGCGACGGACCC

Internal number	Name	Sequence
#986	AAP4 EcoRI Rev	ccggaattcTCACGTACGGCAGTTCGTAC
#987	AAP5 NotI For	ataagaatgcggccgcATGGACCCAGCGGATCCCAGC
#988	AAP5 EcoRI Rev	ccggaattcTCAGCGTCGCGTAACCGTAC
#989	AAP8 NotI For	ataagaatgcggccgcATGGCGACTCAGAGTCAGTTC
#990	AAP8 / 1 / 6 EcoRI Rev	ccggaattcTCATGAACACGTCCGCCGGG
#991	AAP9 NotI For	ataagaatgcggccgcATGGCGACACAGAGTCAGTCC
#992	AAP9 EcoRI Rev	ccggaattcTCATGAAAACGTCCGCTGG
#993	AAV2 mut Fw Rev	GGCTGTCCGAGAGGCTGGGGGTCAGGTACT GAGTCTGCGTCTCCGGTCTGACCAAAATTC
#994	AAV2 mut Rev Fw	CCAGCCTCTCGGACAGCCACCAGCAGCCCCC TCTGGTCTAGGAACTAATACGATGGCTAC
#1014	AAV8 mut Fw	TCAACCTCTCGGAGAACCTCCAGCAGCGCCC TCTGGTGTAGGACCTAATACAATGGCTGC
#1015	AAV8 mut Rev	GGTTCTCCGAGAGGTTGAGGGTCTGGAACT GACTCTGAGTCGCCGGTCTGACCAAAATTG
#1016	AAV9 mut Fw	TCAACCAATCGGAGAACCTCCCGCAGCCCCC TCAGGTGTAGGATCTCTTACAATGGCTTC
#1017	AAV9 mut Rev	GGTTCTCCGATTGGTTGAGGGTCTGGGACT GACTCTGTGTCGCCGGTCTGACCGAAATTG
#1038	SwaI_Cap289 Fw	ctgatttaaatCAGGTATGGCTGCCGATGGTTATC
#1039	Swa_Cap4 Fw	ctgatttaaatCAGGTATGACTGACGGTTACCTTC
#1040	SwaI_Cap5 Fw	ctgatttaaatCAGGTATGTCTTTTGTTGATCACC
#1041	SwaI_Cap4SO Fw	ctgatttaaatCAGGTATGACCGATGGTTATCTTC
#1042	SwaI_Cap5SO Fw	ctgatttaaatCAGGTATGTCTTTCGTTGATCATC
#1107	qPCR Cap2 rep78/68 For	GAGGCCCTTTTCTTTGTGCA
#1108	qPCR Cap2 rep78/68 Rev	CGATCCCGCGGTAAATTCTC
#1109	qPCR Cap2 rep52/40 For	CCTTCTACGGGTGCGTAAAC
#1110	qPCR Cap2 rep52/40 Rev	CCGAGGACTTGCATTTCTGG
#1111	qPCR Cap2 VP1 For	GCTGCCGATGGTTATCTTCC
#1112	qPCR Cap2 VP1 Rev	ACTTGTACCCAGGAAGCACA
#1113	qPCR Cap2 VP2/3+AAP For	AGGAACCTGTTAAGACGGCT

Internal number	Name	Sequence
#1114	qPCR Cap2 VP2/3+AAP Rev	GGTACTGAGTCTGCGTCTCC
#1117	GAPDH human qPCR Fw	ACCCACTCCTCCACCTTTGAC
#1118	GAPDH human qPCR Rev	TGTTGCTGTAGCCAAATTCGTT
#1161	GFP-AAP2 Rev	CAGCTCCTCGCCCTTGCTCACGGGTGAGGTA TCCATAC
#1162	AAP2-GFP Fw	GTATGGATACCTCACCCGTGAGCAAGGGCGA GGAGCTG
#1163	GFP EcoRI Rev	ccggaattcTTACTTGTACAGCTCGTCCATG
#1242	AAP4SO NotI For	ataagaatgcggccgcATGGCGCAGGCGACGGACCTC
#1243	AAP4SO EcoRI Rev	ccggaattcTCACGTACGGGAGCTCGTAC
#1244	AAP5SO NotI For	ataagaatgcggccgcATGGACCTAGCGGATCCCAG
#1245	AAP5SO EcoRI Rev	ccggaattcTCAGGGTGGCGTATCCATAC
#1282	AAV1 mut Fw	ACAACCTCTCGGAGAACCTCCAGCAACCCCC GCTGCTGTAGGACCTACTACAATGGCTTC
#1283	AAV1 mut Rev	GGTTCTCCGAGAGGTTGTGGATCGGGGACT GACTCTGAGTCGCCGGTCTGACCAAAATTG
#1284	AAV3b mut Fw	TCAACCTCTCGGAGAACCACCAGCAGCCCCC ACAAGTTTAGGATCTAATACAATGGCTTC
#1285	AAV3b mut Rev	TTCTCCGAGAGGTTGAGGGTCTGGGACTGAC TCTGAGTCGCCGGTCTGACCGAAATTTAG
#1286	AAV6 mut Fw	ACAACCTCTCGGAGAACCTCCAGCAACCCCC GCTGCTGTAGGACCTACTACAATGGCTTC
#1287	AAV6 mut Rev	GGTTCTCCGAGAGGTTGTGGGTCGGGGACT GACTCTGAGTCGCCGGTCTGACCAAAATTG
#1288	AAV7 mut Fw	TCAACCTCTCGGAGAACCTCCAGCAGCGCCC TCTAGTGTAGGATCTGGTACAGTGGCTGC
#1289	AAV7 mut Rev	GGTTCTCCGAGAGGTTGAGGGTCGGGGACT GACTCTGAGTCGCCGGTCTGACCGAAATTG
#1290	AAVrh10 mut Fw	CTCAACCAATCGGAGAACCCCC CGCAGGCCC-CTCTGGTCTAGGATCTGGTACAATGGCTG
#1291	AAVrh10 mut Rev	GGTTCTCCGATTGGTTGAGGGTCGGGCACT GACTCTGAGTCGCCGGTCTGCCCAAAGTTG
#1292	AAP1/3/6/7 NotI For	ataagaatgcggccgcATGGCGACTCAGAGTCAGTCC

Internal number	Name	Sequence
#1293	AAPrh10 NotI For	ataagaatgcggccgcATGGCGACTCAGAGTCAGTGC
#1294	AAP7/rh10 EcoRI Rev	ccggaattcTCATGAAGACGTCCGCCGGG
#1318	Rep2 rev, pos. 1827	TCATCCAAATCCACATTGAC
#1319	Rep2 rev, pos. 944	CGTGGCCCATCCCAGAAAG

Oligonucleotides used in this work were synthetized by the companies IDT and Sigma
Overhangs are written in small letters
Enzyme restriction sites are underlined
Introduced mutations are marked in red

Table 2.2: Plasmids

Internal number	Name	Description	Reference / origin
#68	pIRES(blast) FLAG/HA-hAgo2-GA	CMV-hAgo2 (codon-optimized), HA-tag, IRES, blasticidin	Cloned by N. Schürmann
#182	WHc1(SpeI)	helper plasmid with AAV2 *rep*, AAV1 *cap*, *Spe*I site	Cloned by E. Kienle
#183	WHc2(SpeI)	helper plasmid with AAV2 *rep*, AAV2 *cap*, SpeI site	Cloned by E. Kienle
#184	WHc4(SpeI)	helper plasmid with AAV2 *rep*, AAV4 *cap*, SpeI site	Cloned by E. Kienle
#185	WHc5(SpeI)	helper plasmid with AAV2 *rep*, AAV5 *cap*, SpeI site	Cloned by E. Kienle
#186	WHc6(SpeI)	helper plasmid with AAV2 *rep*, AAV6 *cap*, SpeI site	Cloned by E. Kienle
#187	WHc7(SpeI)	helper plasmid with AAV2 *rep*, AAV7 *cap*, SpeI site	Cloned by E. Kienle
#188	WHc8(SpeI)	helper plasmid with AAV2 *rep*, AAV8 *cap*, SpeI site	Cloned by E. Kienle
#189	WHc9(SpeI)	helper plasmid with AAV2 *rep*, AAV9 *cap*, SpeI site	Cloned by E. Kienle
#190	WHcrh10(SpeI)	helper plasmid with AAV2 *rep*, AAVrh10 *cap*, SpeI site	Cloned by E. Kienle
#545	pBS-sds-decoy-empty	AAV expression vector with ITR2, CMV-GFP, ITR4	Cloned by S. Mockenhaupt

Internal number	Name	Description	Reference / origin
#714	pSSV9_pSi	AAV expression vector with ITR2, hRluc, hluc+, ITR2	Cloned by D. Niopek
#778	pSSV9_Pac_Asc	AAV expression vector with ITR2, AAV2 *rep*, PacI site, AscI site, ITR2	Cloned during this study
#1079	pBS-sds-CMV-mCherry	AAV expression vector with ITR2, CMV-mCherry, ITR4	Cloned during this study
#1080	pIRES(blast) FLAG-AAP2	CMV-AAP2 expression vector with FLAG / HA tag	Cloned during this study
#1081	pIRES(blast) FLAG-AAP4	CMV-AAP4 expression vector with FLAG / HA tag	Cloned during this study
#1082	pIRES(blast) FLAG-AAP5	CMV-AAP5 expression vector with FLAG / HA tag	Cloned during this study
#1109	pIRES(blast) FLAG-AAP8	CMV-AAP8 expression vector with FLAG / HA tag	Cloned during this study
#1110	pIRES(blast) FLAG-AAP9	CMV-AAP9 expression vector with FLAG / HA tag	Cloned during this study
#1111	Adeno helper plasmid	helper plasmid with Ad5 genes E2A, E4, VA RNA	Matsushita et al. (1998) [109]
#1134	WHC2mutAAP	helper plasmid with AAV2 *rep*, AAV2 *cap* with mutated AAP2	Cloned during this study
#1189	WHC4GA (4_O)	helper plasmid with AAV2 *rep*, AAV4 *cap* (sequence optimized)	Cloned during this study
#1190	WHC5GA (5_O)	helper plasmid with AAV2 *rep*, AAV5 *cap* (sequence optimized)	Cloned during this study
#1194	WHc3(SpeI)	helper plasmid with AAV2 *rep*, AAV3 *cap*, SpeI site	Cloned by E. Kienle
#1290	WHC4GA_w/o PacI	helper plasmid with AAV2 *rep*, AAV4 *cap* (sequence optimized), no PacI site between *rep* and *cap*	Cloned during this study

Internal number	Name	Description	Reference / origin
#1291	WHC4GA repaired ($4_{\text{O repaired}}$)	helper plasmid with AAV2 *rep*, AAV4 *cap* (sequence optimized), mutations in AAP frame repaired, no PacI site	Cloned during this study
#1292	WHC5GA repaired ($5_{\text{O repaired}}$)	helper plasmid with AAV2 *rep*, AAV5 *cap* (sequence optimized), mutations in AAP frame and *rep* stop repaired, no PacI site	Cloned during this study
#1293	WHC5GA only repSTOP ($5_{\text{O\_rep}}$)	helper plasmid with AAV2 *rep*, AAV5 *cap* (sequence optimized), mutation of *rep* stop codon repaired, no PacI site	Cloned during this study
#1326	WHC8mutAAP	helper plasmid with AAV2 *rep*, AAV8 *cap* with mutated AAP8, no PacI site	Cloned during this study
#1327	WHC9mutAAP	helper plasmid with AAV2 *rep*, AAV9 *cap* with mutated AAP9, no PacI site	Cloned during this study
#1452	pBSc1	AAV1 *cap*	Cloned by E. Kienle
#1453	pBSc2	AAV2 *cap*	Cloned by E. Kienle
#1454	pBSc3	AAV3 *cap*	Cloned by E. Kienle
#1455	pBSc4	AAV4 *cap*	Cloned by E. Kienle
#1456	pBSc5	AAV5 *cap*	Cloned by E. Kienle
#1457	pBSc6	AAV6 *cap*	Cloned by E. Kienle
#1458	pBSc7	AAV7 *cap*	Cloned by E. Kienle
#1459	pBSc8	AAV8 *cap*	Cloned by E. Kienle
#1460	pBSc9	AAV9 *cap*	Cloned by E. Kienle
#1461	pBScrh10	AAVrh10 *cap*	Cloned by E. Kienle
#1462	pBScpo1	AAVpo.1 *cap*	Cloned by E. Kienle
#1463	pBSc12	AAV12 *cap*	Cloned by E. Kienle
#1483	pIRES(blast)_ AAP2_GFP	CMV-AAP2-GFP fusion expression vector with FLAG / HA	Cloned during this study

Internal number	Name	Description	Reference / origin
#1492	pBSc4SO (4_O)	AAV4 *cap* (sequence optimized)	Cloned by E. Kienle
#1493	pBSc5SO (5_O)	AAV5 *cap* (sequence optimized)	Cloned by E. Kienle
#1494	pBSc4SO repaired ($4_{O\ repaired}$)	AAV4 *cap* (sequence optimized), mutations in AAP frame repaired	Cloned during this study
#1495	pBSc5SO repaired ($5_{O\ repaired}$)	AAV5 *cap* (sequence optimized), mutations in AAP frame and *rep* stop repaired	Cloned during this study
#1500	WH_empty_Hind/Spe	helper plasmid with AAV2 *rep* until HindIII site, SpeI site	Cloned during this study
#1544	WHC_ccdb_PacI_AscI	helper plasmid with AAV2 *rep*, PacI site, ccdB, AscI site	Cloned by A-K. Herrmann
#1608	pSSV9_Pac_Asc_ccdB	AAV expression vector with ITR2, AAV2 *rep*, PacI site, ccdB, AscI site, ITR2	Cloned by A-K. Herrmann
#1661	WHC1mutAAP	helper plasmid with AAV2 rep, AAV1 *cap* with mutated AAP1, no PacI site	Cloned during this study
#1662	WHC3mutAAP	helper plasmid with AAV2 rep, AAV3 *cap* with mutated AAP3, no PacI site	Cloned during this study
#1663	WHC6mutAAP	helper plasmid with AAV2 *rep*, AAV6 *cap* with mutated AAP6, no PacI site	Cloned during this study
#1664	WHC7mutAAP	helper plasmid with AAV2 *rep*, AAV7 *cap* with mutated AAP7, no PacI site	Cloned during this study
#1665	WHCrh10mutAAP	helper plasmid with AAV2 *rep*, AAVrh10 *cap* with mutated AAPrh10, no PacI site	Cloned during this study
#1666	pIRES(blast) FLAG-AAP1	CMV-AAP1 expression vector with FLAG / HA tag	Cloned during this study

Internal number	Name	Description	Reference / origin
#1667	pIRES(blast) FLAG-AAP3	CMV-AAP3 expression vector with FLAG / HA tag	Cloned during this study
#1668	pIRES(blast) FLAG-AAP6	CMV-AAP6 expression vector with FLAG / HA tag	Cloned during this study
#1669	pIRES(blast) FLAG-AAP7	CMV-AAP7 expression vector with FLAG / HA tag	Cloned during this study
#1670	pIRES(blast) FLAG-AAPrh10	CMV-AAPrh10 expression vector with FLAG / HA tag	Cloned during this study

Table 2.3: Chemicals and reagents

Product	Company
Acetic acid	VWR (Radnor, USA)
Agarose	Biozym Scientific GmbH (Hessisch Oldendorf, Germany)
Albumin fraction V (BSA)	Roth (Karlsruhe, Germany)
Ampicillin	Roth (Karlsruhe, Germany)
Ammonium persulfate (APS)	Grüssing (Filsum, Germany)
Bacto^{TM} agar	BD (Franklin Lakes, USA)
Bacto^{TM} tryptone	BD (Franklin Lakes, USA)
Bacto^{TM} yeast extract	BD (Franklin Lakes, USA)
Bromophenol blue	Waldeck GmbH (Münster, Germany)
Calcium chloride ($CaCl_2$)	Roth (Karlsruhe, Germany)
Cesium chloride (CsCl)	Roth (Karlsruhe, Germany)
Disodium phosphate (Na_2HPO_4 x $2H_2O$)	AppliChem (Darmstadt, Germany)
Dimethyl sulfoxide (DMSO)	Merck (Darmstadt, Germany)
Dodecylsulfate-Na-salt pellets (SDS)	SERVA Electrophoresis GmbH (Heidelberg, Germany)
Ethylenediaminetetraacetic acid (EDTA) disodium salt	Sigma-Aldrich (St. Louis, USA)
Ethanol	Sigma-Aldrich (St. Louis, USA)
Ethidium bromide	Roth (Karlsruhe, Germany)
Ficoll type 400	Sigma-Aldrich (St. Louis, USA)
Gelatin	Sigma-Aldrich (St. Louis, USA)
Glucose	Merck (Darmstadt, Germany)
Glycerol	VWR (Radnor, USA)
HEPES	Roth (Karlsruhe, Germany)

Product	**Company**
Höchst 33258	Invitrogen / Life Technologies (Paisley, UK)
Hydrochloric acid (HCl)	Sigma-Aldrich (St. Louis, USA)
Immunoselect Antifading Mounting Media	Dianova (Hamburg, Germany)
Iodixanol (Optiprep^{TM})	Axis-Shield (Oslo, Norway)
Isopropanol	Sigma-Aldrich (St. Louis, USA)
β-mercaptoethanol	Roth (Karlsruhe, Germany)
Methanol	Sigma-Aldrich (St. Louis, USA)
Milk powder	Roth (Karlsruhe, Germany)
Magnesium chloride ($MgCl_2$)	AppliChem (Darmstadt, Germany)
Magnesium sulfate ($MgSO_4$)	Merck (Darmstadt, Germany)
N-Ethylmaleimide (NEM)	Sigma-Aldrich (St. Louis, USA)
Nuclease-free water	QIAGEN (Hilden, Germany)
Orange G	Waldeck GmbH (Münster, Germany)
PEG8000	Promega (Madison, USA)
Phenol red	Merck (Darmstadt, Germany)
Ponceau S solution	Sigma-Aldrich (St. Louis, USA)
Potassium acetate (KAc)	AppliChem (Darmstadt, Germany)
Potassium chloride (KCl)	AppliChem (Darmstadt, Germany)
Potassium dihydrogen phosphate (KH_2PO_4)	AppliChem (Darmstadt, Germany)
Protease inhibitor cocktail (cOmplete^{TM}, EDTA-free)	Roche (Penzberg, Germany)
Protein A/G PLUS-Agarose	Santa Cruz Biotechnology (Heidelberg, Germany)
QIAzol lysis reagent	QIAGEN (Hilden, Germany)
Rotiphorese® Gel 40 (19:1)	Roth (Karlsruhe, Germany)
Sodium bicarbonate ($NaHCO_3$)	Merck (Darmstadt, Germany)
Sodium acetate (NAc)	Grüssing (Filsum, Germany)
Sodium carbonate (Na_2CO_3)	Merck (Darmstadt, Germany)
Sodium chloride (NaCl)	Sigma-Aldrich (St. Louis, USA)
Sodium hydroxide (NaOH)	Sigma-Aldrich (St. Louis, USA)
Sulfuric acid (H_2SO_4)	J.T.Baker / Avantor Performance Materials (Center Valley, USA)
TE buffer	Invitrogen / Life Technologies (Paisley, UK)
TEMED (Ultra Pure^{TM})	Invitrogen / Life Technologies (Paisley, UK)

Product	Company
TGS (Tris/Glycine/SDS) buffer (10x)	BioRad (Hercules, USA)
Tris	Roth (Karlsruhe, Germany)
Tris-HCl	Roth (Karlsruhe, Germany)
Tween^{TM}20	Roth (Karslruhe, Germany)

Table 2.4: Buffers and reagents

Buffer / Solution	Composition	
Ampicillin stock solution	5 % (w/v)	Ampicillin (50 mg/ml)
Benzonase buffer (pH 8.5)	50 mM	Tris-HCl (pH 8.5)
	150 mM	NaCl
	2 mM	$MgCl_2$
Blocking buffer (Western blot)	5 % (w/v)	milk powder
	(or 1 % [w/v])	Albumin fraction V)
	solvent	1x TBS-T
$CaCl_2$ solution	1 M	$CaCl_2$
Coating buffer (pH 9.6, ELISA)	13 mM	Na_2CO_3
	35 mM	$NaHCO_3$
DNA loading dye (6x, orange)	0.25 % (w/v)	Orange G
	15 % (w/v)	Ficoll type 400
DNA loading dye (10x, blue)	0.25 % (w/v)	bromophenol blue
	49.5 mM	Tris-HCl (pH 7.6)
	60 % (v/v)	glycerol
Freezing media	10 % (v/v)	DMSO
	90 % (v/v)	FBS
IF blocking buffer	1 % (w/v)	Albumin fraction V
	49.2 µM	$MgCl_2$
	90.1 µM	$CaCl_2$
	solvent	1x PBS
Iodixanol (15 %)	25 % (v/v)	Iodixanol (Optiprep)
	75 % (v/v)	PBS-MK-NaCl
Iodixanol (25 %, red color)	41.56 % (v/v)	Iodixanol (Optiprep)
	58.19 % (v/v)	PBS-MK
	0.25 % (v/v)	phenol red solution
Iodixanol (40 %)	66.67 % (v/v)	Iodixanol (Optiprep)
	33.33 % (v/v)	PBS-MK

Buffer / Solution	Composition	
Iodixanol (60 %, yellow color)	99.75 % (v/v)	Iodixanol (Optiprep)
	0.25 % (v/v)	phenol red solution
Lysogeny broth (LB) media[1]	1 % (w/v)	tryptone
	0.5 % (w/v)	yeast extract
	1 % (w/v)	NaCl
Lysogeny broth (LB) agar[2]	1 % (w/v)	tryptone
	0.5 % (w/v)	yeast extract
	1 % (w/v)	NaCl
	1.5 % (w/v)	agar
NaHEPES resuspension buffer	50 mM	HEPES
	150 mM	NaCl
	25 mM	EDTA
PBS (1x)	137 mM	NaCl
	3 mM	KCl
	10 mM	Na_2HPO_4
	2 mM	KH_2PO_4
PBS-MK	solvent	1 x PBS
	1 mM	$MgCl_2$
	2.5 mM	KCl
PBS-MK-NaCl	solvent	1 x PBS
	1 mM	$MgCl_2$
	2.5 mM	KCl
	1 M	NaCl
P1 resuspension buffer (pH 8)	50 mM	Tris-HCl
	10 mM	EDTA
	100 µg/ml	RNase
P2 lysis buffer	200 mM	NaOH
	1 % (w/v)	SDS
P3 neutralization buffer (pH 5.1)	2.8 M	KAc
40 % PEG / NaCl solution	40 % (w/v)	PEG8000
	1.915 M	NaCl
Phenolred solution	0.5 %	phenol red
Running gel buffer (pH 8.8)	1.5 M	Tris
	0.4 %	SDS

[1] supplemented with 50 µg/ml ampicillin before use
[2] supplemented with 100 µg/ml ampicillin before use

Buffer / Solution	**Composition**	
SDS sample buffer (2x)	2 mM	EDTA
	100 mM	Tris-HCl pH 7.5
	4 %	SDS
	20 % (v/v)	glycerol
	10 %	β-mercaptoethanol
	0.02 %	bromophenol blue
SOB media	2 % (w/v)	tryptone
	0.5 % (w/v)	yeast extract
	0.05 % (w/v)	NaCl
	2.5 mM	KCl
	10 mM	$MgCl_2$[1]
	10 mM	$MgSO_4$[1]
SOC media	20 mM	glucose[2]
	solvent	SOB media
Stacking gel buffer (pH 6.8)	500 mM	Tris
	0.4 % (w/v)	SDS
TAE buffer (50x, pH 8.3)	2 M	Tris
	50 mM	EDTA
	5.71 % (v/v)	acetic acid
TBS (10x, pH 7.5)	250 mM	Tris
	1.25 M	NaCl
TBS-T (1x)	solvent	1x TBS
	0.05 %	Tween
TFBI (pH 5.8)	30.6 mM	KAc
	80 mM	$MgCl_2$
	100 mM	KCl
	16 mM	$CaCl_2$
	13.2 % (v/v)	glycerol
TFBII (pH 8.0)	4.8 mM	MOPS
	76 mM	$CaCl_2$
	10 mM	KCl
	13.2 % (v/v)	glycerol
Topping solution (RI = 1.3710)	3.27 M	CsCl (0.55 g/ml)
	solvent	NaHEPES resuspension buffer

[1] addition after autoclaving
[2] sterile-filtered

Buffer / Solution	Composition	
Transfer buffer	20 % solvent	methanol 1x TGS

Table 2.5: Cell culture media, additives and transfection reagents

Component	Company
Dulbecco's Modified Eagle Medium (DMEM)	Gibco / Life Technologies (Paisley, UK)
Dulbecco's phosphate-buffered saline (DPBS, 1x)	Gibco / Life Technologies (Paisley, UK)
Fetal bovine serum (FBS)	Sigma-Aldrich (St. Louis, USA)
MG-132 proteasome inhibitor	Santa Cruz Biotechnology (Heidelberg, Germany)
Minimum Essential Medium Non-Essential Amino Acids (MEM-NEAA, 100x)	Gibco / Life Technologies (Paisley, UK)
Polyethylenimine (PEI, linear, MW ~25 000)	Polyscience (Warrington, USA)
Penicillin-Streptomycin (10,000 U/mL)	Gibco / Life Technologies (Paisley, UK)

Table 2.6: Cell lines

Cell line	Origin	Tissue type	Reference
HEK293T	*H. sapiens*	Embryonic kidney	DuBridge et al. (1987) [186]
Huh7	*H. sapiens*	Hepatocellular carcinoma	Nakabayashi et al. (1982) [187]
MCF-7	*H. sapiens*	Mammary carcinoma	Soule et al. (1973) [188]
SF-539	*H. sapiens*	Gliosarcoma	Rutka et al. (1986) [189]

Table 2.7: Bacterial strains

Bacterial strain	Description	Source
E. coli ccdB Survival^{TM} T1^{R}	chemically competent	Life Technologies GmbH (Paisley, UK)
E. coli MAX Efficiency DH5α^{TM}	chemically competent	Life Technologies GmbH (Paisley, UK)
E. coli MegaX DH10B^{TM} T1^{R}	electro-competent	Life Technologies GmbH (Paisley, UK)

Table 2.8: Enzymes

Enzyme	Company
Antarctic phosphatase	NEB (Frankfurt am Main, Germany)
Benzonase	Merck (Darmstadt, Germany)
DNaseI	Life Technologies (Paisley, UK)
HotStar HiFidelity Polymerase	Qiagen (Hilden, Germany)
Phusion Hot Start II DNA polymerase	Thermo Fisher Scientific (Waltham, USA)
Restriction enzymes	NEB (Frankfurt am Main, Germany)
RNase (100 mg/ml)	Qiagen (Hilden, Germany)
T4 DNA ligase	NEB (Frankfurt am Main, Germany)

Table 2.9: Primary antibodies

Name	Antigen	Source	Purity	Application	Dilution	Reference / Company
A1	AAV VP1	Mouse, monoclonal	hybridoma cell culture supernatant	WB	1:10	Kind gift from J. Kleinschmidt lab (DKFZ), published 1997 [57]
A20	Assembled AAV2 capsids	Mouse, monoclonal	hybridoma cell culture supernatant	IF	undil.	Kind gift from J. Kleinschmidt lab (DKFZ), published 1995 [190]
AAP[1]	AAP2	Rabbit, polyclonal	purified	IF WB	1:250 1:500	Perbio / Thermo Fisher Scientific (Waltham, USA)
B1	AAV VP1, VP2, VP3	Mouse, monoclonal	hybridoma cell culture supernatant	IF IP WB	undil. 1:11 1:20-30	Kind gift from J. Kleinschmidt lab (DKFZ), published 1995 [190]

[1] This antibody was ordered and produced during this work

Name	Antigen	Source	Purity	Application	Dilution	Reference / Company
β-actin (C4)	mouse, rat, human, avian, bovine, canine, porcine and rabbit β-actin	Mouse, monoclonal	purified	WB	1:1000	Santa Cruz Biotechnology (Heidelberg, Germany)
FK2	mono- and polyubiquitinylated proteins	Mouse, monoclonal	purified	IP WB	1:3125 1:800	Enzo Life Science (Lörrach, Germany)
HA	HA tag	Mouse, monoclonal	purified	IF IP WB	1:300-1:400 1:1250 1:200-300	Santa Cruz Biotechnology (Heidelberg, Germany)
Oct-3/4 (N-19)	Human, mouse and rat Oct-3/4	Goat, polyclonal (IgG)	purified	IP	1:1250	Santa Cruz Biotechnology (Heidelberg, Germany)
SSEA-1 (480)	Human, mouse and rat SSEA-1	Mouse monoclonal (IgM)	purified	IP	1:1250	Santa Cruz Biotechnology (Heidelberg, Germany)
303.9	AAV Rep proteins	Mouse, monoclonal	hybridoma cell culture supernatant	IF WB	undil. 1:10	Kind gift from J. Kleinschmidt lab (DKFZ)), published 1995 [190]

Abbreviations: IF = Immunofluorescence, IP = Immunoprecipitation, WB = Western blot.

Table 2.10: Secondary antibodies

Name / Antigen	Source	Application	Dilution	Company
Alexa Fluor® 488 anti-mouse IgG (H+L)	Goat	IF	1:500	Life Technologies (Paisley, UK)
Alexa Fluor® 568 anti-rabbit IgG (H+L)	Goat	IF	1:1000	Life Technologies (Paisley, UK)
Alexa Fluor® 647 anti-mouse IgG (H+L)	Goat	IF	1:750-1500	Life Technologies (Paisley, UK)
Anti-mouse IgG + IgM, HRP-conjugated	Goat	WB	1:10000	Jackson ImmunoResearch Laboratories (West Grove, USA)
Anti-rabbit IgG, HRP-conjugated	Donkey	WB	1:10000	GE Healthcare (Chalfont St. Giles, UK)

Abbreviations: HRP = Horseradish peroxidase, IF = Immunofluorescence, WB = Western blot.

Table 2.11: Kits

Kit	Company
AAV2 Titration ELISA	PROGEN Biotechnik (Heidelberg, Germany)
Agilent DNA 1000 Kit	Agilent Technologies (Santa Clara, USA)
NucleoBond® Xtra Maxi	Macherey-Nagel (Düren, Germany)
PureYield™ Plasmid Midiprep System	Promega (Madison, USA)
Qubit® dsDNA BR Assay Kit	Invitrogen / Life Technologies (Paisley, UK)
RNeasy Mini Kit	QIAGEN (Hilden, Germany)
SensiMix™ II Probe No-ROX Kit	Bioline (London, UK)
SensiMix™ SYBR No-ROX Kit	Bioline (London, UK)
QIAquick Gel Extraction Kit	QIAGEN (Hilden, Germany)
Tetro cDNA Synthesis Kit	Bioline (London, UK)
Western Lightning® Plus-ECL	PerkinElmer (Waltham, USA)

Table 2.12: Standard markers

Marker	Company
100 bp DNA ladder	Life Technologies GmbH (Paisley, UK)
1 kb DNA ladder	Life Technologies GmbH (Paisley, UK)
PageRuler^{TM} plus prestained protein ladder	Thermo Fisher Scientific (Waltham, USA)

Table 2.13: Materials

Material	Company
Amersham Hyperfilm ECL (X-ray film, 18 × 24 cm)	GE Healthcare (Buckinghamshire, UK)
Amicon® Ultra Centrifugal Filter Units, 15ml, 100000 NMWL	Millipore / Merck (Darmstadt, Germany)
Blotting paper (3 mm, 46 × 57 cm)	neoLab (Heidelberg, Germany)
Cell culture dishes (Ø 6 cm / 15 cm)	Greiner Bio-One (Frickenhausen, Germany), Nunc / Thermo Fisher Scientific (Waltham, USA)
Cell culture flasks (Ø 75 cm^2 / 175 cm^2)	Greiner Bio-One (Frickenhausen, Germany)
Cell culture plates (6-well / 96-well)	Greiner Bio-One (Frickenhausen, Germany)
Cell lifter (18 cm)	Corning (New York, USA)
Centrifuge tubes (15 ml / 50 ml / 500 ml)	BD (Franklin Lakes, USA), Corning (New York, USA), Greiner Bio-One (Frickenhausen, Germany)
Combitips	Eppendorf (Hamburg, Germany)
Covaris reaction tubes (microTUBE AFA Fiber Snap-Cap 6x16mm)	Covaris (Woburn, USA)
Counting chambers	Invitrogen / Life Technologies (Paisley, UK)
Cuvettes (Polystyrene, 10 x 4 x 45 mm)	Sarstedt (Nümbrecht, Germany)
Dialysis tubing visking (Typ 20/32 inch, wall thickness 0.020 mm, MWCO 14 000)	Roth (Karlsruhe, Germany)
Electroporation cuvettes (1 mm gap)	Peqlab (Erlangen, Germany)
Erlenmeyer flasks	Fisher Scientific (Schwerte, Germany)

Material	**Company**
Filter tips	Biozym Scientific GmbH (Hessisch Oldendorf, Germany), Greiner Bio-One (Frickenhausen, Germany), Sarstedt (Nümbrecht, Germany)
Glass bottles	DURAN Group (Wertheim/Main, Germany)
Glass culture tubes (16 x 160 mm)	DURAN Group (Wertheim/Main, Germany)
LabTek® II Chamber Slide™ System	Nunc / Thermo Fisher Scientific (Waltham, USA)
Needles (0.8 x 40 mm / 0.9 x 40 mm)	BD (Franklin Lakes, USA)
Nitrocellulose blotting membrane	neoLab (Heidelberg, Germany)
Microscopy well plates (µ-Plate ibiTreat)	ibidi (Planegg / Martinsried, Germany)
OptiSeal™ Ultracentrifuge tubes – cesium chloride purification (26x77 mm)	Beckman Coulter (Brea, USA)
Pasteur capillary pipettes (230 mm)	neoLab (Heidelberg, Germany)
PCR tubes	STARLAB (Hamburg, Germany)
Petri dishes	Greiner Bio-One (Frickenhausen, Germany)
Pipette tips	Greiner Bio-One (Frickenhausen, Germany), Kisker (Steinfurt, Germany)
qPCR tubes and lids	QIAGEN (Hilden, Germany)
Qubit reaction tubes	Invitrogen / Life Technologies (Paisley, UK)
Reaction tubes (0.5 / 1.5 / 2 / 5 ml)	Eppendorf (Hamburg, Germany), Sarstedt (Nümbrecht, Germany)
Rotor disc™ 100	QIAGEN (Hilden, Germany)
Rotor disc™ heat sealing film	QIAGEN (Hilden, Germany)
Serological pipettes	Greiner Bio-One (Frickenhausen, Germany)
Slide-A-Lyzer dialysis cassette (MWCO 20000)	Thermo Fisher Scientific (Waltham, USA)
Steritop filter (0.22µm)	Millipore / Merck (Darmstadt, Germany)
Syringes (3 ml / 50 ml)	BD (Franklin Lakes, USA)
Syringe filter units (0.2 / 0.45 µm pore size)	Whatman / GE Healthcare (Buckinghamshire, UK)
Ultracentrifuge tubes - iodixanol purification (16x76 mm)	Seton Scientific (Petaluma, USA)

Material	Company
VacConnectors	QIAGEN (Hilden, Germany)

Table 2.14: Equipment

Material	Company
Bacterial incubator Heraeus Function Line	Thermo Fisher Scientific (Waltham, USA)
2100 Bioanalyzer	Agilent Technologies (Santa Clara, USA)
Biophotometer Plus	Eppendorf (Hamburg, Germany)
CaptairBio PCR workstation	Erlab (Köln, Germany)
Centrifuge 5415 R	Eppendorf (Hamburg, Germany)
Centrifuge Avanti J-26 XP	Beckman Coulter (Brea, USA)
Confocal microscope TCS SP5	Leica (Wetzlar, Germany)
Countess™ Automated Cell Counter	Invitrogen / Life Technologies (Paisley, UK)
Covaris focused-ultrasonicator S2	Covaris (Woburn, USA)
HERAcell 150 CO2 Incubator	Heraeus / Thermo Fisher Scientific (Waltham, USA)
HERAsafe® sterile workbench	Thermo Fisher Scientific (Waltham, USA)
Fixed angle type 70 Ti rotor	Beckman Coulter (Brea, USA)
Fixed angle type 70.1 Ti rotor	Beckman Coulter (Brea, USA)
Fixed angle JA-10 rotor	Beckman Coulter (Brea, USA)
FlexCycler (PCR thermocycler)	Analytik Jena (Jena, Germany)
Flow cytometer FC500 MPL	Beckman Coulter (Brea, USA)
Gas cartridge CV470 plus	Camping Gaz (Hattersheim, Germany)
Gene-Disc Heat Sealer	Corbett Life Science (Sydney, Australia) / QIAGEN (Hilden, Germany)
Gel caster (DNA electrophoresis)	BioRad (Hercules, USA)
Gel Doc™ XR	BioRad (Hercules, USA)
Gene Pulser Xcell™ Electroporation System	BioRad (Hercules, USA)
Hamilton syringe, model #705 (50 µl)	Hamilton (Reno, USA)
Inverted microscope CKX-41	Olympus (Hamburg, Germany)
Inverted fluorescence microscope IX-81	Olympus (Hamburg, Germany)
Mastercycler® pro	Eppendorf (Hamburg, Germany)
Microwave	Sharp Electronics GmbH (Hamburg, Germany)
Mini-PROTEAN® comb 10-well (30 µl) / 15-well (15 µl)	BioRad (Hercules, USA)

Material	Company
Mini-PROTEAN® spacer plates (1 mm)	BioRad (Hercules, USA)
Mini-PROTEAN® short plates	BioRad (Hercules, USA)
Mini-PROTEAN® tetra cell casting module	BioRad (Hercules, USA)
Mini-PROTEAN® tetra vertical electrophoresis cell	BioRad (Hercules, USA)
Mini-Sub® cell GT / Sub-Cell® GT combs (1-/2-/8-/10-/15-/20-/30-well)	BioRad (Hercules, USA)
Mini-Sub® cell GT / Sub-Cell® GT DNA electrophoresis system	BioRad (Hercules, USA)
Mini-Sub® cell GT / Sub-Cell® GT UV transparent gel trays	BioRad (Hercules, USA)
Multiskan™ FC Microplate Spectrophotometer	Thermo Fisher Scientific (Waltham, USA)
NanoDrop 2000 UV-Vis Spectrophotometer	Thermo Fisher Scientific (Waltham, USA)
NanoVue Spectrophotometer	GE Healthcare (Buckinghamshire, UK)
pH meter PB-11	Sartorius (Göttingen, Germany)
Pipettes	Eppendorf (Hamburg, Germany), Gilson (Middleton, USA)
PowerPac™ Basic / HC / HV Power Supply	BioRad (Hercules, USA)
Qubit® fluorometer	Invitrogen / Life Technologies (Paisley, UK)
Refractometer	Exacta Optech (San Prospero, Italy)
Rotor-Gene 6000	Corbett Life Science (Sydney, Australia) / QIAGEN (Hilden, Germany)
Shaker DOS-10L	neoLab (Heidelberg, Germany)
Shaker DRS-12	neoLab (Heidelberg, Germany)
Shaking Incubator Multitron	INFORS HT (Basel, Switzerland)
ThermoCell Mixing Block MB-102	Bioer (Hangzhou, China)
Trans-Blot® SD Semi-Dry Transfer Cell	BioRad (Hercules, USA)
Triangular support with rod and burette clamp	Roth (Karlsruhe, Germany)
Tube rotator	VWR (Radnor, USA)
Tube sealer	Beckman Coulter (Brea, USA)
Ultracentrifuge Optima L-90K	Beckman Coulter (Brea, USA)
Ultrasonic bath	BANDELIN (Berlin, Germany)
UV transilluminator UST-30M-8E	biostep GmbH (Burkhardtsdorf, Germany)
Vortex Genie 2	Scientific Industries (Bohemia, USA)

Material	Company
Water bath TW12	Julabo (Seelbach, Germany)
X-OMAT 2000 processor	KODAK (Rochester, NY, USA)

Table 2.15: Software

Software	Application	Company / Reference
2100 Expert Software	Agilent software	Agilent Technologies (Santa Clara, USA)
CellProfiler	Fluorescence picture quantification	Carpenter *et al.* (2006) [191], Kamentsky *et al.* (2011) [192]
Clustal W / Clustal X	DNA / protein alignments	Larkin *et al.* (2007) [193]
Fiji	Fluorescence picture processing	Schindelin *et al.* (2012) [194]
GraphPad Prism	Graphical and statistical analysis	GraphPad Software, Inc. (La Jolla, USA)
Microsoft Excel	Coloring and analysis of Salanto-derived files	Microsoft Corporation (Redmond, USA)
MXP software	Flow cytometry software	Beckman Coulter (Brea, USA)
Quantity One 1-D Analysis Software 4.6.9	Gel documentation software	BioRad (Hercules, USA)
Rotor Gene 6000 Series Software 1.7	qPCR analysis	QIAGEN (Hilden, Germany)
Salanto[1]	Shuffling alignment analysis	Schürmann *et al.* (2013) [195]
Vector NTI	Plasmid maps visualization / editing and DNA / protein alignments	Life Technologies GmbH (Paisley, UK)

[1] This program was further developed in cooperation with C. Bender during this study

2.2 Methods

2.2.1 Cloning procedures

2.2.1.1 Cloning of the pSSV9_PacI_AscI construct

A pSSV9 vector (pSub201 [130]) with PacI, AscI and SpeI sites was generated in order to allow cloning of created shuffled AAV *cap* sequences (see section 2.2.3.7) downstream of the *rep* sequence. For construction of the pSSV9_PacI_AscI vector, a PacI site containing fragment was amplified with the primers #641 and #642 from the pTRUF3 *cap* plasmid which was generated by Grimm et al. [133]. In addition, a SpeI and an AscI restriction site were introduced with primer #642. Subsequently, the purified and digested PCR product was cloned via HindIII / XbaI into a pSSV9 backbone available in the lab. Figure 2.1 depicts the final construct after insertion of a *cap* gene. After transfection into HEK293T cells together with an adeno-viral helper construct (#1111), plasmids will produce replication competent AAVs which package the encoded viral genome. The inserted SpeI site allowed an easy cloning of *cap* from _pSSV9_PacI_AscI into an ITR-less *rep-cap* helper which permits the packaging of a reporter gene and the control of capsid infectivity (see also next section).

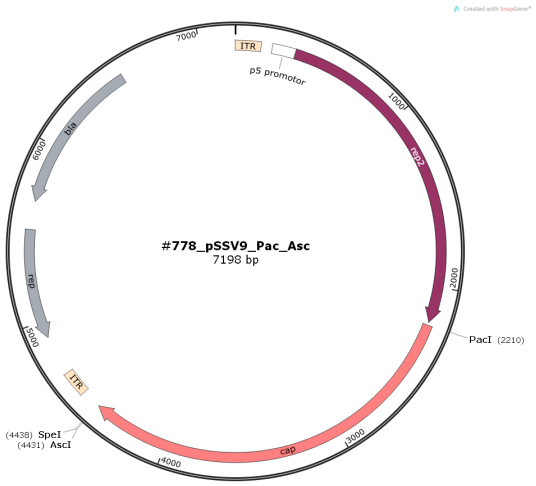

Figure 2.1: Plasmid map of construct no. 778_pSSV9_PacI_AscI. The depicted pSSV9 vector contains the AAV2 *rep* gene (dark red) driven by the AAV2 p5 promoter (white) and followed by a *cap* gene (light red) inserted via PacI and AscI site. Both are flanked by AAV2 ITRs which enable viral genome packaging into the encoded capsid. The plasmid map was generated with the program SnapGene.

2.2.1.2 Subcloning of shuffled AAV *cap* sequences into a helper construct

Shuffled AAV *cap* sequences were cloned from the single-stranded pSSV9 construct (see section 2.2.1.1) into a *rep-cap* helper plasmid without ITRs (generated by E. Kienle [196]) in order to package a gfp reporter into the capsids and assess the functionality of single clones. Therefore, *cap* genes were amplified from picked and analyzed shuffled AAV clones (see 2.2.3.7 for details) with the primers #823 and #1038 – #1042 (see table 2.16 for conditions). The latter were selected based on the AAV serotype at the 5' end of the shuffled sequence and introduced a SwaI site in front of the *cap* gene. Next, PCR fragments were purified and sequentially digested. First, SpeI digestion was performed over night at 37°C with NEB buffer 1.1. Second,

the DNA was incubated with SwaI enzyme at 25°C for at least 3 hours after addition of NEB buffer 3.1. No intermediate purification step was needed. Digested PCR products were purified and ligated into SpeI / SwaI digested plasmid #188. In case SwaI cloning did not work, *cap* sequences were amplified with primers #833 and #823 and cloned into the *rep-cap* helper plasmid via HindIII / SpeI or HindIII / AscI digestion.

Table 2.16: PCR mix and PCR program for shuffled AAV *cap* amplification

PCR reaction		**PCR program**		
Components	Amount	Temperature	Time	
DNA template	0.5 µl			
Forward primer	15 pmol	98°C	30 sec	
Reverse primer	15 pmol	98°C	10 sec	
DMSO	1.5 µl	58°C	15 sec	35 cycles
2.5 mM dNTPs	1 µl	72°C	70 sec	
5x polymerase buffer HF	10 µl	72°C	10 min	
Phusion Hot Start II DNA Polymerase	0.5 µl	4°C	∞	
H_2O	33.5 µl			

2.2.1.3 Sequence optimization of AAV *cap* 4 / *cap* 5 and repair of the AAP frame

Cap sequences of AAV4 and AAV5 were modified by the company GeneArt in order to increase their homology to AAV *cap* 2. Therefore, single nucleotides were mutated without changing the viral protein sequence (see also figure 3.4). The resulting capsids were named 4_O and 5_O, respectively. Since the assembly-activating protein (AAP) was not discovered at the time of sequence generation, several mutations were inadvertently introduced into the AAP frame. In detail, $AAP4_O$ lost its start and stop codon. Moreover, two stop codons were generated within $AAP4_O$. $AAP5_O$ lost its start codon and contained one additional stop codon after modification. Unintentionally, it also lost the stop codon of the viral proteins Rep68 and Rep40 which is located in the VP frame. In order to repair the created mutations, overlap extension PCR was used (see section 2.2.3.1 for details). First, five or three *cap* gene fragments were amplified from $AAV4_O$ or $AAV5_O$, respectively, using the primers #966 – #982 which carried the correct nucleotides to repair the generated mutations. Second, primers #964 and #967, or #965 and #978, were used to amplify the full-length AAV *cap* gene of 4_O and 5_O, respectively. The *cap* fragments generated in the first PCR served as input. The final corrected *cap* sequence was cloned via SwaI / SpeI into a *rep-cap* helper construct (e.g. #189, table 2.1).

2.2.1.4 Construction of AAP expression vectors

For generation of AAP expression vectors, AAP sequences were amplified from AAV *cap* genes with primers #983 – #992, #1242 – #1245 and #1292 – #1294 depending on the serotype.

The used primers introduced a NotI restriction site in front and an EcoRI restriction site behind AAP, and allowed cloning into a pre-existing backbone (#68, elimination of hAgo2 was achieved by NotI / EcoRI digestion). The latter already contained a CMV promoter followed by a FLAG-tag, an HA-tag and the two restriction sites. In addition the plasmid carried an IRES followed by a blasticidin gene.
An AAP-GFP expression vector was generated by overlap extension PCR (see section 2.2.3.1 for details) with the primers #1161 and #1162. Thus, the AAP stop codon as well as the GFP start codon were removed and GFP was fused C-terminally to AAP. The obtained sequence was cloned via NotI / EcoRI into the same backbone mentioned above.

2.2.1.5 Cloning of AAVmut constructs

AAP knock-out AAV *cap* genes were generated through overlap extension PCR (section 2.2.3.1). The used primers (#993 – #994, #1014 – #1017, #1283 – #1291) introduced mutations that led to an exchange of the AAP start codon from CTG to CCG and introduced a stop codon within the AAP frame 66 bp downstream of the start codon (TGG → TAG). Both modifications did not change the AAV-VP amino acid sequences. In total, eight AAV serotypes were mutated (AAV1, 2, 3b, 6, 7, 8, 9, rh10). Primers #178 and #833 served as flanking primers during the second PCR and allowed for cloning of the *cap* gene into a *rep-cap* helper plasmid (e.g. #188, table 2.1) via HindIII / SpeI.

2.2.2 Microbiological methods

2.2.2.1 Production of chemo-competent bacteria

DH5α *E. coli* were streaked on an antibiotic-free agar plate and grown over night at 37°C. The following day, 10 ml antibiotic-free LB media were inoculated with one single colony picked from the plate and again grown overnight at 37°C. Two hundred ml SOC media were inoculated with 1 ml of this culture and incubated until the OD_{600} value reached 0.5. Bacteria were then centrifuged for 15 min at 1800 x g and 4°C, resuspended in 40 ml TFBI buffer and incubated on ice for 10 min. Centrifugation was repeated, and bacteria were resuspended in 10 ml TFBII buffer and again incubated on ice for 10 min. Finally, bacteria were frozen in 50 µl aliquots in liquid nitrogen and stored at -80°C.

2.2.2.2 Heat-shock transformation

Fifty µl of chemo-competent bacteria were mixed with 10 µl ligation mix and incubated for 30 min on ice. After 45 sec at 42°C, bacteria were again incubated for 2 min on ice and agitated for 1 hour at 37°C with 1 ml antibiotic-free LB media. Afterwards, the reaction was centrifuged for 5 min at 400 x g, and the bacterial pellet was resuspended in 100 µl LB media and plated on antibiotic-containing LB agar plates. For ampicillin containing plasmids, the transformation reaction was directly plated after 2 min incubation on ice. In case of plasmid re-transformations, 2 µl of DNA were incubated for 1 – 5 min with 50 µl bacteria on ice. Plates were incubated over night at 37°C and single colonies picked for DNA preparations.

2.2.2.3 Production of electro-competent bacteria

For production of electro-competent bacteria, 30 ml of antibiotic-free LB media were inoculated under sterile conditions from a glycerol stock of commercial Mega X DH10B bacteria and agitated at 180 rpm over night at 37°C. The next day, four flasks with 400 ml antibiotic-free LB media were inoculated with 5 ml of the pre-culture and incubated for 2 – 3 hours until the OD_{600} value reached 0.5 – 0.55. Bacterial growth was stopped by 15 – 20 min incubation on ice, and cultures were subsequently centrifuged for 15 min at 4°C and 4424 x g. The bacterial pellets were resuspended in a total volume of 60 ml cold ddH_2O, transferred to four pre-cooked dialysis tubings (Typ 20/32 inch, wall thickness 0.020mm) and dialyzed against 4 – 8 liters of sterile ddH_2O over night at 4°C. The next day, bacteria were centrifuged for 15 min at 4°C and 4424 x g, pellets were resuspended in 10 % glycerin and OD_{600} set up to 1 (which corresponds to $2.5x10^{10}$ cells) by addition of 10 % glycerin. Aliquots of 40, 100 and 300 µl bacteria were frozen for electroporation experiments. Efficiency of bacteria was assessed by electroporation of 15 pg pUC19 control plasmid into 30 µl bacteria (see section 2.2.2.4 for details). The next day, colonies on the plates with the 1:10 and 1:100 dilution were counted and efficiency (Colony forming units [CFU] / µg) calculated with the following formula:

$$\frac{CFU}{\mu g} = \frac{\text{number of colonies}}{\text{15 pg pUC DNA}} x \frac{1x10^6 \text{pg}}{\mu g} x \frac{\text{volume of transformants}}{\text{plated volume}} x \text{ dilution factor}$$

Efficiency had to reach at least $5x10^8$ CFU/µg to be sufficient for transformation experiments.

2.2.2.4 Electroporation

Electro-competent bacteria were thawed on ice. Then, 30 µl of bacteria were mixed with 0.5 – 2 µl of ligation mix and transferred to a pre-cooled electroporation cuvette (1 mm gap). It was made sure that all bubbles in the cuvette were removed by gentle tapping and that metal contacts were dry. Subsequently, the cuvette was placed into a Gene Pulser Xcell™ Electroporation System machine and pulsed with 1800 V, 25 µF and 200 Ω. One ml pre-warmed SOC media was used to resuspend the bacteria and to transfer them to a 1.5 ml reaction tube. Electroporated bacteria were agitated for 1 hour at 37°C and 450 rpm in a mixing block. Afterwards, 100 µl bacteria suspension were plated on antibiotic-containing LB agar plates and the rest of the reaction was centrifuged for 5 min at 400 xg. The pellet was resuspended in 100 µl LB media and also plated. Plates were incubated over night at 37°C and single colonies picked for DNA preparations. In this thesis, electroporation was mainly used for generation of shuffled AAV *cap* libraries (see also chapter 2.2.3.7 for more details) during this work.

2.2.3 Molecular biology methods

2.2.3.1 Polymerase chain reaction (PCR)

Typically, 50 µl PCR reactions were prepared to amplify specific DNA fragments from plasmids or genomic DNA. Therefore (unless mentioned otherwise), 10 µl 5 x HF reaction buffer, 1 µl 2.5 mM dNTPs, 1.5 µl DMSO, 15 – 25 pmol forward primer, 15 – 25 pmol reverse primer, 0.5 µl Phusion Hot Start II DNA polymerase and 0.5 – 2 µl template DNA were mixed and filled up with water to the final volume. In the following reaction, DNA was initially denaturated for 30 sec at 98°C followed by 30 – 35 cycles with the conditions 10 sec at 98°C; 15 sec at 50°C – 68°C (depending on the primer annealing temperatures); 30 sec / kb at 72°C and a final annealing step for 5 – 10 min at 72°C. Subsequently, DNA was run on an agarose gel and purified with the QIAquick Gel Extraction Kit according to the manufacturer's protocol. Overlap extension PCR was used to introduce mutations into DNA fragments. For this purpose, one primer carrying the intended mutation was designed and another primer with either the same sequence but reverse complement or at least 15 overlapping nucleotides was created. These primers acted as forward and reverse primer in two separate first PCR reactions together with another forward and reverse primer flanking the gene of interest. PCR components others than the primers are listed above. DNA fragments were subsequently purified and used as templates (0.5 – 2 µl per reaction) for a second (overlap extension) PCR in which the same flanking primer pair was used for amplification. In order to allow the single fragments to anneal properly, a two-step PCR program was programmed in which primers were added only after 15 cycles (see table 2.17). Finally, full-length PCR products were purified, cloned into a target vector and sequenced to validate correct alteration of the sequence.

Table 2.17: PCR program for overlap extension PCR

Temperature	Time	
98°C	30 sec	
98°C	10 sec	
72°C	30 sec / kb	15 cycles
Addition of primers		
98°C	10 sec	
50°C – 68°C[1]	25 sec	20 cycles
72°C	30 sec / kb	
72°C	5 – 10 min	
4°C	∞	

[1] depending on the primer annealing temperatures

2.2.3.2 Plasmid DNA preparation

Plasmid DNA was isolated from bacterial cultures for characterization of newly generated constructs and/or transfections. Therefore, different volumes of antibiotic-containing media were inoculated with single colonies of transformed bacteria (sections 2.2.2.2 and 2.2.2.4), depending on the amount of DNA that was required, and agitated for 14 – 16 hours at 37°C. Plasmid DNA was then purified with the commercial kits NucleoBond® Xtra Maxi (200 – 400 ml culture), PureYield™ Plasmid Midiprep System (80 – 100 ml culture) or with self-made buffers according to the alkaline lysis method [197] (3 – 5 ml culture). The latter was performed as follows: 2 ml of bacterial suspension were pelleted and resuspended in 300 µl P1 Resuspension buffer. Three hundred µl P2 Lysis buffer were added and incubated for 5 min at room temperature. After addition of 300 µl P3 Neutralization buffer and 5 min incubation at room temperature, samples were centrifuged for 10 min at 16100 x g. Supernatant was transferred into a new tube containing 600 µl isopropanol, carefully mixed and centrifuged for 20 min at 16100 x g. Supernatant was discarded and the DNA pellet washed with 500 µl ethanol. After 5 min of centrifugation at 16100 x g, supernatant was discarded, and the DNA pellet was air-dried and resuspended in 50 µl nuclease-free water. DNA concentrations were determined with a NanoVue or NanoDrop 2000 UV-Vis spectrophotometer.

2.2.3.3 Ethanol precipitation

In some cases it was necessary to further purify isolated DNA due to impurities (i.e., ethanol contamination) after medium scale DNA preparations. Therefore, 400 µl DNA were diluted with 200 µl 3 M sodium acetate and 1400 µl 100 % ethanol. After 25 min centrifugation at full speed and 4°C, DNA pellets were washed with 500 µl 70 % ethanol and again centrifuged for 5 min. DNA pellets were air-dried and resuspended in 300 – 500 µl nuclease-free water.

2.2.3.4 Enzymatic digests

DNA was control digested to verify plasmid identities, successful construct generation or ITR integrity of AAV constructs. The latter were digested with XmaI to test ITR2 integrity, or PstI for ITR4, respectively. Moreover, DNA was digested in quantitative amounts (purified PCR products or 6 – 15 µg plasmid DNA) for further cloning procedures. In all cases, NEB enzymes (2 – 20 U per reaction) and their corresponding buffers were used. Digests were performed between 3 – 5 hours or over night at the recommended temperature. For combinatorial digests either compatible buffers were used or sequential digests were performed.

2.2.3.5 Agarose gel electrophoresis

For separation of DNA fragments according to their size, agarose gel electrophoresis was performed. Either PCR or restriction digest products were separated through 1 – 2 % agarose gels in 1 x TAE buffer. Gels were prepared ahead of the run with the Sub-Cell® DNA

electrophoresis system and supplemented with 1 µg / ml ethidiumbromide. Samples were mixed with orange or blue loading dye, depending on the expected size of the DNA fragments, and loaded on the gel together with a standard marker for size determination. Electrophoresis was performed at 80 – 150 V for 30 – 60 min, and DNA was visualized on a Gel Doc™ XR system or a UV transilluminator. If necessary, DNA was cut out and purified with the QIAquick Gel Extraction Kit according to the manufacturer's protocol.

2.2.3.6 Ligation

For ligations of digested DNA fragments, normally 5 – 7.5 µl purified insert and 1 – 4.5 µl purified plasmid backbone were mixed together with 1 µl 10 x ligase buffer and 0.5 µl T4 DNA ligase in a 10 µl reaction. Ligation reactions were incubated 3 – 5 hours at room temperature or over night at 16°C. Subsequently, 0.5 – 2 µl (electroporation) or the complete 10 µl reaction (heat-shock transformation) were used for transformation into bacteria (see also sections 2.2.2.2 and 2.2.2.4).

2.2.3.7 DNA family shuffling of AAV capsids

AAV *cap* amplification and DNA fragmentation via DNaseI digest

For the generation of a shuffled AAV capsid library, the *cap* composition of the library was selected based on the further application. Subsequently, the selected capsids were PCR amplified with the primers #827 and #828 from plasmids #1452 – 1463 or #1492 – 1495 which contained the appropriate *cap* genes. The latter were flanked with standardized sequences containing binding sites for the primers SAFor (#459) and SARev (#460) and the restriction enzymes PacI and AscI which allowed later cloning steps. Table 2.18 depicts the PCR reaction and PCR program that were used in order to yield high amounts of DNA for the following fragmentation. PCR products were run on a 1 % agarose gel and purified with the QIAquick Gel Extraction Kit. For DNA fragmentation, 4 µg of *cap* DNA were mixed with each serotype being represented in equal amounts. In addition, the reaction contained 6 µl 10 x DNase reaction buffer, 0.3 µl DNaseI enzyme and water up to a total volume of 60 µl. Importantly,

Table 2.18: PCR mix and reaction for AAV *cap* amplification from pBS vectors

PCR reaction		**PCR program**		
Components	Amount	Temperature	Time	
Cap plasmid	200 ng	95°C	5 min	
Primer #827 (M13For)	2 µM	94°C	15 sec	
Primer #828 (M13Rev)	2 µM	57°C	30 sec	40 cycles
5x polymerase buffer	10 µl	68°C	3 min	
HotStar HiFidelity Polymerase	2 µl	72°C	10 min	
H_2O	Filled up to 50 µl	4°C	∞	

DNaseI has to be added as last component to the mixture followed by immediate flicking (3 x) of the tube, a short centrifugation step and an incubation of the reaction between 1.5 and 2.5 min at 25°C. The exact incubation time had to be determined in pre-experiments in order to create optimal DNA fragment sizes between 100 and 1200 bp. The reaction was stopped via addition of 6 µl 25 mM EDTA and 10 min incubation at 75°C. Next, 9 µl H_2O and 15 µl 6 x loading dye were added to the mixture and the DNA was separated on a 1 % agarose gel. DNA fragments between 100 and 1200 bp were cut out, purified with the QIAquick Gel Extraction Kit and eluted in 30 µl nuclease-free water.

DNA fragmentation via Covaris

Since DNase digestion is highly error-prone and it is thus difficult to generate DNA fragments of equal quality and size, DNA fragmentation with a Covaris focused-ultrasonicator S2 machine was assessed as an alternative. Therefore, 4 µg *cap* DNA were mixed in nuclease-free water (total volume: 50 µl), and transferred to a Covaris reaction tubes (microTUBE AFA Fiber Snap-Cap 6 x 16 mm). Depending on the aimed fragment size, different Covaris programs were used to shear the DNA (see table 2.19). Next, the DNA was separated on a 1 % agarose gel, fragments were cut out and purified with the QIAquick Gel Extraction Kit and eluted in 30 µl H_2O. To compare Covaris versus DNase-treated samples, DNA was analyzed on an Agilent 2100 Bioanalyzer Chip. The used DNA 1000 kit depicted a size range between 25 – 1000 bp. Sample preparation and chip run were performed according the manufacturer's protocols.

Table 2.19: Covaris settings for DNA fragmentations of different sizes

Fragment size [bp]	**150**	**300**	**800**
Duty cycle	10 %	10 %	5 %
Intensity	5	4	3
Cycles per burst	200	200	200
Time [s]	500	120	80

Re-assembly PCR (1st PCR) and 2nd PCR

After DNA fragmentation, *cap* sequences were re-assembled in a primerless PCR reaction (1st PCR) which is based on partial homology between the parental AAV sequences. Therefore, 500 ng purified DNA fragments, 10 µl 5 x HF buffer, 1.5 µl DMSO, 1 µl 10 mM dNTPs and 0.5 µl Phusion Hot Start II DNA polymerase were mixed in a 50 µl PCR reaction together with H_2O. PCR conditions were: an initial denaturation at 98°C for 30 sec, 40 cycles of 10 sec denaturation at 98°C, 30 sec annealing at low temperature (42°C) to foster self-priming and generation of chimeric sequences, and 45 sec elongation at 72°C. The PCR was finished with a final extension step at 72°C for 10 min. PCR products were directly used without purification

as DNA template for the subsequent 2nd PCR. Two µl of DNA were added to a reaction containing 2 µM SAFor (#459) and 2 µM SARev (#460) primer, 10 µl 5 x polymerase buffer, 0.5 µl 25 mM $MgSO_4$, 2 µl HotStar HiFidelity Polymerase (Qiagen) and 33.5 µl H_2O. In order to obtain enough DNA for further cloning steps and a highly diverse library, 16 x 50 µl PCR reactions were set up for one AAV *cap* library. For small-scale production and analysis of shuffling efficiencies 4 x 50 µl reactions were prepared. PCR conditions were 5 min at 95°C, 40 cycles of 15 sec at 94°C and 3 min at 68°C, and a final extension step at 72°C for 10 min. PCR products were run on a 1 % agarose gel. Re-assembled *cap* fragments with a size of $\sim$ 2.2 kb were cut out and purified with the QIAquick Gel Extraction Kit.

AAV *cap* library cloning and electroporation

After purification of the re-assembled *cap* fragments, DNA was digested with the restriction enzymes PacI and AscI (25 U each) in the appropriate buffer over night at 37°C. The next day, the DNA was run on a 1 % agarose gel and fragments were purified with the QIAquick Gel Extraction Kit. AAV *cap* library fragments were then cloned into an AAV single-strand plasmid backbone containing the AAV2 *rep* gene flanked by two ITR2 sequences (plasmid #778 or #1608). For the ligation, 3 x more *cap* fragment DNA than plasmid backbone DNA (molar ratio) was used. In detail, 861 ng vector backbone DNA, 1139 ng *cap* DNA, 4 µl 10 x ligase buffer and 4 µl T4 DNA ligase were mixed in a 40 µl ligation reaction (resulting in a concentration of 50 ng DNA per µl) and incubated over night at 16°C. After ligation, DNA was transformed into electro-competent bacteria. Therefore, 600 µl bacteria were pooled and carefully mixed with the complete ligation reaction. Thirty-two µl of the mixture were transferred to a chilled electroporation cuvette, pulsed (see 2.2.2.4), resuspended in 1 ml pre-warmed SOC media and transferred to a 1.5 ml reaction tube. The procedure was repeated 20 times. After 1 hour shaking at 37°C, reactions were pooled again and 100 µl bacteria suspension were plated on antibiotic-containing LB agar plates. Additionally, 1:10 and 1:100 dilutions were prepared and plated in order to determine library diversity on the next day and to characterize chimeric clones (see paragraph below) through picking of single colonies. The rest of the bacterial suspension was used to inoculate 400 ml of antibiotic-containing LB media, and was processed for DNA preparation with the NucleoBond® Xtra Maxi Kit after incubation at 37°C over night. In order to create a back-up of the generated library, 800 µl of bacterial suspension were mixed with 200 µl glycerol and stored at -80°C. For a small-scale library production only a 10 µl ligation reaction was prepared with one fourth of the above mentioned components. Moreover, only 1 x 30 µl bacteria were electroporated and no suspension culture was inoculated.

Analysis of shuffled sequences with the Salanto program

For AAV library characterization normally 12 – 20 single colonies were picked after electroporation. DNA was isolated as described above (see 2.2.3.2) and sent for sequencing with the primers #36, or #37, #822 and #823. Therefore, DNA was prepared according the

company's instructions (GATC Biotech). From the resulting data sets, full-length AAV *cap* sequences were assembled since one sequencing reaction covered only a third of the shuffled capsid. Next, a DNA alignment of parental and chimeric sequences was generated with the program Vector NTI and its AlignX function. Alternatively, the freeware ClustalX was used. Subsequently, a FASTA file was created from the alignment and manually corrected for misaligned nucleotides in case it was needed. The resulting file could be opened by Salanto and analyzed with the integrated tools. For example, proportions of parental capsids in the AAV library were determined, numbers of cross-over events within the clones or in the complete library were calculated, and position-wise assignments of chimeras to parental sequences were displayed. The latter was further processed with the Microsoft Excel program, i.e. chimeras were colored according to the parental sequences (see supplementary table 1 for the used macro). More details about Salanto, its function and tools as well as the newest version can be found here: https://bitbucket.org/benderc/salanto/wiki/Home. See also supplementary figure 2 for used features and generated diagrams described in this work.

2.2.3.8 RNA extraction

For RNA extraction, HEK293T cells were seeded in a 6-well plate format with 0.5 x 10^6 cells in 2 ml media per well, transfected one day later (see 2.2.5.2) and harvested 48 hours after transfection. Therefore, cells were washed with PBS, resuspended in 600 µl RLT buffer (RNeasy Mini Kit) and transferred to 1.5 ml reaction tubes. Samples were stored at -80°C until further processing. Subsequently, RNA was purified according to the manufacturer's protocol (section "Purification of Total RNA from Animal Cells using Spin Technology") with inclusion of the on-column DNase digestion step. Finally, RNA concentrations were determined using the NanoVue spectrophotometer.

2.2.3.9 cDNA synthesis and quantitative PCR (qPCR)

To measure mRNA copy numbers in the extracted RNA samples, 5 µg RNA were transcribed with the Tetro cDNA Synthesis Kit using the random hexamer primers according to the manufacturer's protocol. cDNA was then diluted 1:500 and 1 µl was used in a 10 µl qPCR reaction (see table 2.20). qPCR was performed with the SensiMix^{TM} SYBR No-ROX Kit and reactions were run in a Rotor-Gene 6000 machine using the following program: 95°C 10 min, 40 cycles of 95°C 15 sec, 58°C 15 sec, 72°C 15 sec, and a final melting step from 50°C to 99°C in 5-second steps of 1°C each. Samples were measured in duplicates and normalized to a GAPDH housekeeper (primers #1117/#1118). For final analysis, sample values were calculated according to the $2^{-\Delta\Delta CT}$ method [198].

Table 2.20: Composition of one 10 µl reaction for qPCR analysis (using the SensiMix^{TM} SYBR No-ROX Kit)

Component	Volume [µl]
2x SensiMix^{TM} SYBR No-ROX	5
10 µM forward primer	1.25
10 µM reverse primer	1.25
cDNA template (1:500 diluted)	1
H_2O	1.5

2.2.4 Protein biochemistry methods

2.2.4.1 Immunoprecipitation (IP)

HEK293T cells were seeded in 6 cm dishes (1.4 x 10^6 cells per dish, 4.5 ml media) and transfected 24 hours later with PEI (see 2.2.5.2). Twenty-four hours after transfection old media was removed and fresh media containing the proteasome inhibitor MG-132 (final concentration 2 µM) was added. After additional 24 hours cells were harvested in media and centrifuged for 10 min at 1500 x g and 4°C. The supernatant was discarded and 1000 µl PBS containing protease inhibitors (PBS_{PI}; 1 tablet protease inhibitor cocktail was dissolved in 50 ml PBS) was added. Centrifugation was repeated, the supernatant was discarded and the cells were resuspended in 1200 µl PBS_{PI}. Cells were then lysed by 5 x freeze-thaw cycles (2 min liquid nitrogen / 5 min 37°C), sonicated for 1 min and cell debris was removed by centrifugation for 5 min at 10000 x g, 4°C. The supernatant was transferred into a new tube and the volume adjusted to 1300 µl with PBS_{PI}. Subsequently, 50 µl of the samples were aliquoted for later Western blot analysis and served as IP input. Next, specific antibodies were added to the lysate: 115 µl B1 antibody hybridoma supernatant were used to pull down AAV VPs, 0.4 µl anti-Ubiquitin antibody FK2 were used for IP of mono- and polyubiquitinated proteins, and either 1 µl anti-Oct3/4 or 1 µl anti-SSEA1 antibody served as control antibodies. The lysate-antibody mixtures were incubated for $\sim$ 5 hours under rotary agitation at 4°C. Afterwards, 50 µl of Protein A/G PLUS Agarose beads were added per sample and incubated over night at 4°C under rotary agitation. The next day, samples were centrifuged for 3 min at 2500 x g and 50 µl of supernatant were aliquoted for Western blot analysis to serve as non-bound fraction. At this step, proteins of interest should be specifically bound to the antibodies coupled to the beads. The supernatant was then carefully removed from the beads and discarded. Beads were washed four times with ice-cold PBS to remove non-specifically bound proteins. Therefore, samples were gently mixed with PBS and centrifuged for 3 min at 2500 x g, 4°C. After the last washing step, as much buffer as possible was removed and the bead pellet was resuspended in 50 µl 2 x SDS buffer. For dissociation of the immunocomplexes from the beads, samples were incubated for 10 min at 95°C. Beads were then collected by 2 min

centrifugation at 2500 x g, and the supernatant was transferred to a new tube and stored at -20°C until Western blot analysis. Every step during the IP was performed on ice to avoid protein degradation and de-ubiquitination.

2.2.4.2 Immunostaining

HEK293T cells were seeded three days prior to immunostaining in either 96-well glass plates suitable for confocal microscopy with a density of 2.8 x 10^4 cells per well or in LabTek formats. The latter were first coated with 1 % gelatine for 30 min at 37°C and then filled with the same cell number per well. One day later cells were transfected using polyethylenimine (see 2.2.5.2). Therefore, 172.5 ng of an adeno-viral helper construct, a *rep-cap* helper which encodes for the AAV capsid and plasmid #714 (unless mentioned otherwise) carrying the transgene flanked by ITRs were transfected in equal amounts. In addition either 44.2 ng of AAP plasmid or 44.2 ng stuffer DNA were added to the mixture. Forty-eight hours after transfection, cells were washed with 200 µl (LabTek) or 100 µl (96-well) ice-cold PBS, air-dried for $\sim$ 2 min and fixed with 100 % methanol for 10 min at -20°C. After 2 x 5 min washing with IF blocking buffer, cells were incubated for 1 – 2 hours at 37°C with 150 µl (LabTek) or 50 µl (96-well) primary antibodies diluted in blocking buffer (see table 2.9 for dilutions). Subsequently, cells were washed 3 x 5 min with IF blocking buffer and incubated for 1 hour at 37°C with secondary antibodies diluted in blocking buffer (see table 2.10 for dilutions). For quantification of B1 and A20 stainings an anti-mouse AF647-labeled secondary antibody was used. The anti-HA antibody was detected with an anti-mouse AF488-labeled secondary antibody (unless mentioned otherwise). In case of the concurrent use of the anti-AAP2 antibody with either the B1 or the A20 antibody, anti-rabbit AF568 and anti-mouse AF488 antibodies were added to the samples. Next, cells were washed 2 x 5 min with IF blocking buffer, incubated for 5 min with Hoechst stain 33258 diluted 1:3000 in PBS and washed again 2 x 5 min with PBS. For long-term storage, LabTeks were embedded in three drops of mounting media and covered with a cover slip. 96-well plates were stored at 4°C in the dark containing 100 µl PBS per well. In case of time-course experiments, samples were fixed with methanol, washed once with IF blocking buffer and stored in blocking buffer at 4°C until staining procedures were continued as described above.

2.2.4.3 Microscopy and picture quantification

In order to quantify the immunostaining of the abovementioned samples, pictures were taken with an Olympus inverted fluorescence microscope IX-81. Table 2.21 depicts the microscope settings for each color channel. Microscopy pictures were quantified with the open-source software CellProfiler. The implemented protocols are described in supplementary tables 2 – 4. To obtain more detailed pictures a Leica confocal microscope TCS SP5 was used. The program Fiji was used to further process the captured microscopy pictures (e.g. z-stack projections, scale bar insertions).

Table 2.21: Microscope settings for the Olympus inverted fluorescence microscope IX-81.

Channel	**Filter cube**	**Exposure time**	**Intensity**	**Excitation filter**
Blue (Hoechst)	DAPI	10 ms	100 %	empty
Green (GFP/AF488)	GFP	200 – 500 ms	50 – 100 %	empty
Red (AF568)	DAPI/FITC/TxRed	1000 ms	50 %	TxRed 572/23
Dark red (AF647)	Cy5	100 ms	100 %	empty

2.2.4.4 Sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE)

For analysis of viral proteins, samples were first mixed with 2 x SDS sample buffer and denaturated for 5 – 10 min at 95°C. Next, samples were loaded with a Hamilton syringe on either 8 % (VP1, VP2, VP3, actin) or 12 % (AAP) resolving sodium dodecyl sulfate (SDS) polyacrylamide gels for separation by their molecular weight. Gels were prepared as described in table 2.22 with the Mini-PROTEAN® tetra cell casting module. In detail, one-millimeter spacer plates were assembled with short glass plates. After ensuring that the setup was tight and leakproof, the resolving gel was poured between the two glass plates and overlaid with isopropanol. Once the resolving gel was polymerized, the isopropanol was removed, the stacking gel was poured and a comb with the required number of wells (10 or 15) was inserted. As soon as the stacking gel was polymerized, the gel was transferred to the electrophoresis chamber and samples were loaded together with a standard protein ladder (PageRuler™ plus prestained protein ladder). Electrophoresis was performed in 1 x TGS buffer for 2 – 2.5 hours at 90 V. After the run had finished the gels were processed for Western blotting (see 2.2.4.5).

Table 2.22: SDS polyacrylamide gel composition for one gel

Stacking gel Component	**5 %** Volume [ml]	**Resolving gel** Component	**8 %** Volume [ml]	**12 %** Volume [ml]
H_2O	1.46	H_2O	2.65	2.15
Stacking gel buffer (pH 6.8)	0.25	Running gel buffer (pH 8.8)	1.25	1.25
40 % Acrylamide-Bisacrylamide	0.25	40 % Acrylamide-Bisacrylamide	1.0	1.5
10 % SDS	0.02	10 % SDS	0.05	0.05
TEMED	0.02	TEMED	0.05	0.05
10 % APS	0.002	10 % APS	0.003	0.002

2.2.4.5 Western blotting

After proteins were separated via SDS-PAGE (see 2.2.4.4), they were transferred to a nitrocellulose membrane through semi-dry blotting. Therefore, gels were removed from the glass plates, the stacking gel was cut off and the resolving gel as well as the nitrocellulose membrane and an appropriate amount of blotting papers were equilibrated in transfer buffer. Next, four blotting papers, the membrane, the SDS gel and another four blotting papers were assembled on a Trans-Blot® SD Semi-Dry Transfer Cell followed by 1 – 1.5 hours of blotting at 150 – 300 mA at 4°C. The efficiency of the transfer was determined through Ponceau S staining of the membrane. Afterwards, membranes were washed with 1 x TBS-T buffer and blocked for at least one hour at room temperature with either 5 % milk or 1 % albumin fraction V in TBS-T. Primary antibodies were diluted according to table 2.9 in the corresponding blocking buffer (except for the anti-actin antibody all antibodies were diluted in 5 % milk blocking solution) and incubated with the membrane over night at 4°C. The membranes were cut around the 55 kDa ladder band ahead of blocking, in case viral proteins (VPs) and actin were co-stained on one gel. The next day, membranes were washed 3 x 10 min with TBS-T buffer and incubated for at least 1 hour at room temperature with a horseradish peroxidase-conjugated secondary antibody diluted in blocking buffer (see table 2.10 for dilutions). After 3 x 10 min subsequent washing steps in TBS-T buffer, membranes were incubated for 1 min with Western Lightning Plus-ECL reagent to visualize bound antibodies via chemiluminescence and exposed to X-ray films. The latter were developed in an X-OMAT 2000 processor.

2.2.5 Cell culture methods

2.2.5.1 Cell culture

All cell lines were cultured under standard sterile growth conditions and incubated at 37°C with 5 % CO_2. HEK293T, MCF-7 and SF-539 cells were cultured in Dulbecco's Modified Eagle Medium (DMEM) media with GlutaMAX™ supplemented with 10 % FCS and 100 U/mL Penicillin-Streptomycin. Additionally, Huh7 media contained 1 % non-essential amino acids. Cells were split every two to four days at 80 – 100 % confluency (see table 2.23).

Table 2.23: Split conditions for several cell lines

Cell line	Split factor	Split frequency
HEK293T	1:8 – 1:12	2-3 days
Huh7	1:3 – 1:6	3-4 days
MCF-7	1:8 – 1:12	3-4 days
SF-539	1:6 – 1:10	3-4 days

2.2.5.2 Transfection with polyethylenimine (PEI)

For transfection, HEK293T cells were seeded 24 hours before PEI transfection in a required well / plate format and transfected as soon as they reached an optimal confluency of 60 – 75 %. For transfection, a PEI mix and a DNA mix were prepared separately (see table 2.24 for composition of mixtures). Both were mixed together and incubated for 10 min at room temperature to allow the formation of DNA-PEI complexes. Finally, the complete mix was added dropwise to the cells. Forty-eight to 72 hours post-transfection cells were processed for further experiments.

Table 2.24: Composition of PEI and DNA mixes for transfections with PEI in different cell culture (well-) plate formats. Amounts are calculated for one well or one dish, respectively.

Plate format	PEI mix			DNA mix		
	PEI	H_2O	300 mM NaCl	DNA	H_2O	300 mM NaCl
96-well (normal)	1.7 µl	2.1 µl	3.8 µl	200 ng	3.8 µl	3.8 µl
96-well (glass)	2.55 µl	3.15 µl	5.7 µl	300 ng	5.7 µl	5.7 µl
LabTek	1.83 µl	2.25 µl	4.08 µl	216.7 ng	4.08 µl	4.08 µl
6-well	22 µl	27 µl	49 µl	2.6 µg	49 µl	49 µl
6 cm dish	50.8 µl	62.3 µl	113.1 µl	6.0 µg	113.1 µl	113.1 µl
15 cm dish	352 µl	438 µl	790 µl	44 µg	790 µl	790 µl

2.2.6 Virological methods

2.2.6.1 Cell seeding, transfection and harvest for large-scale AAV production

Four million HEK293T cells were seeded in 22 ml of media per 15 cm dish two days before transfection. Normally, 10 dishes per vector for iodixanol gradient purification, or 40 – 60 dishes per vector for cesium chloride purification, were transfected according to table 2.24. For virus production, equal amounts of an adeno-viral helper construct, a *rep-cap* helper which encodes for an AAV capsid and a plasmid carrying the transgene flanked by ITRs were transfected (14.67 µg each per dish). In case of AAV productions supplemented with AAP, 11.7 µg of the abovementioned plasmids were transfected in addition with either 9 µg of an AAP construct or 9 µg of stuffer (sheared salmon sperm) DNA. For shuffled AAV libraries, 22 µg of adeno-viral helper plasmid and 22 µg library plasmid were transfected per dish. Three days after transfection cells were scraped off the plates, transferred to tubes, centrifuged for 10 min at 400 x g and washed with 1 x PBS. Finally, cell pellets were resuspended in benzonase buffer (5 ml per 10 dishes) and AAVs were purified through iodixanol or cesium chloride gradients (see chapter 2.2.6.2 and 2.2.6.3).

2.2.6.2 AAV purification through iodixanol density gradient centrifugation

Cell pellets produced as depicted in chapter 2.2.6.1 were lysed by 5 x freeze-thaw cycles and 1 min sonification followed by a benzonase digest (150 U / ml) for 1 hour at 37°C. Samples were carefully mixed every 10 – 20 min to ensure complete degradation of free DNA and RNA. Then, samples were adjusted to 7 ml and centrifuged for 15 min at 4000 x g and 4°C. The supernatant was transferred to a new tube and centrifuged again. During centrifugation, iodixanol solutions (15 – 60 %) were prepared according to table 2.4. To pour the solutions into an ultracentrifugation tube, a Pasteur capillary pipette was put into the tube. Next, virus-containing supernatant was transferred into the ultracentrifugation tube, followed by 1.5 ml of each iodixanol solution in the order 15 % / 25 % / 40 % / 60 %. Tubes were filled completely with benzonase buffer and sealed after removal of remaining bubbles. Gradients were run in a fixed-angle 70.1Ti Beckman rotor for 2 hours at 50000 rpm and 4°C in an Optima L-90K ultracentrifuge. Afterwards, the 40 % iodixanol phase was collected with a syringe from the side of the gradient while taking care not to transfer the 25 % phase which contained empty particles. Therefore, tubes were placed in a burette clamp fixed at a triangular support with rod. Purified virus was aliquoted and stored at -80°C until further use.

2.2.6.3 AAV purification through cesium chloride density gradient centrifugation

Samples were produced as described in section 2.2.6.1 and treated as detailed in the beginning of chapter 2.2.6.2 to disrupt the cells and degrade free DNA and RNA. After benzonase digest, samples were centrifuged for 15 min at 4000 x g, and the supernatant was transferred to a new tube. Subsequently, 1/39 volumes of 1 M calcium chloride solution were added (final concentration 25 mM) to separate high molecular weight impurities [199]. Samples were then incubated for 1 h on ice and centrifuged for 15 min at 10000 x g and 4°C. Next, 1/4 volumes of 40 % PEG / NaCl solution were added to the sample, which was carefully mixed and incubated 3 – 5 hours or over night on ice to precipitate AAV particles. Afterwards, samples were centrifuged at 2500 x g / 4°C for 30 min and carefully resuspended in 10 ml Na-HEPES resuspension buffer with a pipette. In case precipitation was performed for 3 – 5 hours, resuspension of the pellet took place over night at 4°C on a tube rotator. In an additional centrifugation step (2500 x g, 4°C, 30 min) remaining precipitates were removed, the supernatant was transferred to a new tube and Na-HEPES buffer was added to a volume of 24 ml. After addition of 13.2 g cesium chloride, the solution temperature dropped and samples were incubated until they reached room temperature again. In the meanwhile the refraction index (RI) of the cesium chloride topping solution was assessed and – if necessary – adjusted to 1.3710. Next, the RI of the samples was also adjusted with either Na-HEPES buffer (RI > 1.3710) or cesium chloride (RI < 1.3710) to 1.3710, and samples were transferred with a serological pipette to large OptiSeal™ Ultracentrifuge tubes (26 x 77 mm). Bubbles were removed, and the tubes were filled up with topping solution and closed with appropriate lids. Gradients were run in a fixed-angle 70Ti Beckman rotor for 23 hours at 45000 rpm and

21°C in an Optima L-90K ultracentrifuge. After the centrifugation, tubes were placed in a burette clamp fixed at a triangular support with rod. Needles were used to create holes at the top of the tube for air exchange and at the bottom for sample collection. Different volumes of fractions (3 / 3 / 0.5 / 0.5 / 0.5 / 5 / 0.5 / 0.5 / 0.5 / 3 ml) were collected in 15 ml tubes and the RI of each fraction was determined. Fractions with a RI between 1.3711 and 1.3766 were pooled, and the volume adjusted to 9 ml with PBS. Next, samples were transferred to a Slide-A-Lyzer dialysis cassette and dialyzed against 1 x PBS at 4°C with regularly buffer exchanges (after 30 min, 1 hour, 2 hours, over night, 2 hours and 2 hours). Subsequently, samples were concentrated in equilibrated Amicon® Ultra Centrifugal Filter Units to a volume of 500 – 1200 µl using several centrifugation steps at 400 x g. Finally, purified virus was aliquoted and stored at -80°C until further use.

2.2.6.4 AAV titration

To determine AAV titers, 10 µl cesium chloride or 10 µl iodixanol gradient-purified AAV sample was diluted in 10 µl TE buffer and mixed with 20 µl 2 M NaOH followed by 30 min incubation at 56°C for virus lysis. Samples were neutralized with 38 µl 1 M HCl and further diluted with 922 µl H_2O. In case of iodixanol-purified virus, samples were additionally diluted 1:10 in H_2O since high iodixanol concentrations could inhibit the qPCR reaction. In each titration one negative control with nuclease-free water and one AAV positive control with an already known titer were included and treated as described above. In addition, a standard curve containing an appropriate plasmid was prepared. The latter was 10-fold serially diluted in order to obtain several dilutions with 5 x 10^3 to 5 x 10^9 molecules per reaction. Final qPCR mixtures contained 1 x SensiMix™ II Probe No-ROX mix, 0.4 µM forward primer, 0.4 µM reverse primer, 0.1 µM qPCR probe, 1.43 µl treated sample and were filled up to 10 µl reaction volume with nuclease-free water (see table 2.25 for primer and probe sequences). Samples were measured in triplicates. PCRs were run in a Rotor-Gene 6000 machine with the following program: denaturation for 10 min at 95°C, 40 cycles of 10 sec denaturation at 95°C and 20 sec combined annealing and extension at 60°C. After the run, C_T values were

Table 2.25: Sequences of probe and primers used for AAV titrations. Probes were labeled 5' with the fluorophore FAM and 3' with the quencher BHQ1.

Name	Probe (5‘ → 3‘)	5‘ primer (5‘ → 3‘)	3‘ primer (5‘ → 3‘)
CMV	AGTCATCGCTATTA CCATGG	TGCCCAGTACATG ACCTTATGG	GAAATCCCCGTGAG TCAAACC
GFP	ACGACGGCAACTACA	GAGCGCACCATCT TCTTCAAG	TGTCGCCCTCGAAC TTCAC
Rep	TGATCGTCACCTCC AACA	AAGTCCTCGGCCC AGATAGAC	CAATCACGGCGCAC ATGT

translated into molecule numbers based on the standard curve (expected r^2 value > 0.985) and final sample concentrations were calculated by multiplying the obtained values by 7 (since 1.43 µl sample were added to 10 µl reaction), 100 (dilution through alkaline lysis), 100 (for up-scaling from 10 µl to 1 ml), 2 (only for single-stranded genome viruses) and 10 (only for iodixanol gradient-purified samples which were additionally diluted 1:10).

2.2.6.5 Determination of empty and full AAV capsids

First, cesium chloride gradient purification was performed until the centrifugation step as described above (see 2.2.6.3). Subsequently, the first 5 ml of the gradient were discarded and 0.5 ml fractions were collected. In total, 40 fractions were taken and their RI was determined. Fractions with RIs between 1.3638 and 1.3780 were used for the following experiments. For determination of infectious particles, 5 µl of each fraction were used to transduce HEK293T cells seeded one day before with 2.5 x 10^4 cells per well in 96-well plates. Viral dilutions of 1:10 and 1:100 reduced sample toxicity which was caused by the high cesium chloride concentrations. Cells were analyzed 48 hours later by flow cytometry (see 2.2.6.7). Assembled viral particles were determined by AAV2 Titration ELISA (PROGEN Biotechnik). Therefore, every second fraction was diluted between 1:100 and 1:10000 and measured against a provided standard sample. The ELISA was performed according to the manufacturer's protocol. Finally, measured absorbance values were normalized to a blank and AAV particle concentrations were calculated with the help of a standard curve which was included in the kit.

2.2.6.6 AAV crude cell lysate production

HEK293T cells were seeded in 6-well plates with 0.45 – 0.50 x 10^6 cells in 2 ml media per well one day before transfection. For transfection, PEI was used as transfection reagent (see 2.2.5.2). In case of normal virus production, equal amounts of an adeno-viral helper construct, a *rep-cap* helper which encodes for an AAV capsid and a plasmid carrying the transgene (normally a CMV promoter-driven GFP, unless mentioned otherwise) flanked by ITRs were transfected (866.67 ng each per well). When AAP was added during production, 690 ng per construct were transfected together with either 530 ng of AAP plasmid or 530 ng stuffer DNA. Forty-eight hours after transfection, cells were harvested in media with a pipette and centrifuged for 10 min at 1500 x g. The supernatant was discarded and the cell pellet resuspended in 1 ml PBS. After a second centrifugation, the pellet was resuspended in 500 µl PBS. Sixty µl of this suspension were transferred to a new tube and diluted with 40 µl PBS for later Western blot analysis (see 2.2.4.5). The remaining cell suspension (440 µl) was lysed by 5 x freeze-thaw cycles (37°C / liquid N_2) and 1 min sonification. Cell debris was removed through 10 min centrifugation at 16100 x g and AAV particle containing supernatant was transferred to a new tube. AAV crude cell lysates were stored at -20°C for later transduction of different cell lines.

2.2.6.7 Crude cell lysate transduction and flow cytometry analysis

Cells were seeded in 96-well plates 24 hours ahead of AAV crude lysate transduction (see table 2.26 for cell line specific cell densities). Normally, HEK293T cells were used to test AAV2 constructs, MCF7 cells for AAV4 vectors and SF539 cells for AAV5 particles. Since crude cell lysates cannot be titered due to containing intracellular components which could interfere with the qPCR reaction, equal volumes of vector were used for transduction. Accordingly, 5 µl or 10 µl AAV crude lysate were added per 96-well for adherent or suspension cells, respectively. Additionally, cells were transduced with vector dilutions of 1:10 and 1:100. Forty-eight hours post-transduction, cells were processed for flow cytometry analysis. In detail, cell culture media was discarded and cells were washed with 100 µl PBS. After 5 min incubation at 37°C with 30 µl trypsin, cells were resuspended in 170 µl 1 % BSA in PBS (final volume = 200 µl). The amount of transduced cells was determined through the measurement of GFP-positive events in flow cytometry. Therefore, size and granularity of the cells were monitored by forward and sideward scatter whereby a cutoff gate was set in order to exclude cell debris. Within the displayed events a 'Living cells' fraction was defined and had to be adapted for each cell line. GFP expression of the 'Living cells' fraction was detected in the green channel and plotted against the red channel which was expected to be negative. With the help of a non-transduced sample, the amount of background fluorescence was determined and a 'GFP positive' gate set. Transduction rates were calculated as percentage of positive cells out of all events in the 'Living cells' gate. In total 15000 events were measured or detection was stopped after 1 min acquisition, respectively, when the number of events fell below 15000. All data were acquired with a FC500 MPL flow cytometer and the accompanying MXP software in a 96-well plate format. Recorded data were further processed with an Excel macro (see supplementary table 5) to extract all relevant values. For final analysis, the crude lysate dilution was selected in which the AAV wild-type control showed less than 95 % positive cells in order to avoid the analysis of oversaturated data. Subsequently, all values were normalized to this control which was set to 1.0 and compared with other measurements.

Table 2.26: Cell densities of different cell lines for AAV crude lysate transduction (96-well plate format)

Cell line	Cell numbers per well
HEK293T	1.5 – 2.0 x 10^4 cells
Huh7	0.7 x 10^4 cells
MCF7	0.5 x 10^4 cells
SF539	0.4 x 10^4 cells

3 Results

3.1 Optimization of DNA family shuffling

3.1.1 DNase digest outperforms Covaris fragmentation

Previous studies showed successful generation of chimeric AAV capsids using DNA family shuffling as a method for molecular evolution [133, 136]. However, DNA family shuffling includes critical steps which, when performed improperly or without sufficient training, can influence the quality of the generated library. Amongst others, one crucial procedure during shuffling is the fragmentation of input DNA. Normally, enzymatic DNaseI digests are performed in order to generate DNA fragments < 1000 bp [133–136]. Yet, DNase digest is highly dependent on the enzyme amount used and on the incubation time of the mixture (figure 3.1a). Slightly decreased incubation times already result in undesirable large fragments, while higher enzyme quantities increase fragmentation to an unfavorable extent, respectively. Moreover, enzyme stocks can vary in their activity and therefore have to be tested before each fragmentation approach.

Thus, we aimed to improve DNA fragmentation and compared enzymatic DNase digest with physical treatment regarding their fragment generation and shuffling efficiency. In detail, we applied DNA shearing with a Covaris focused-ultrasonicator machine which is widely used for e.g. next generation sequencing [200, 201]. The system uses acoustic energy to create cavitation zones through which DNA double-strand breaks are generated. By modifying different parameters (e.g. intensity or time) it is possible to define variable fragment lengths. We chose settings to create fragments of 150, 300 and 800 bp and compared them with DNase-treated DNA on an Agilent Bioanalyzer DNA 1000 chip. As input we used AAV capsids of the serotypes 1, 7, 8, 9 and rh10 (from hereon referred to as 1789rh10 AAV *cap* library). Figure 3.1b-c clearly depicts the difference between both methods: Controlled fragmentation with focused ultrasonication yielded defined DNA pieces of specific length depending on the machine settings (figure 3.1c). In contrast, DNase digestion generated a wide range of fragment sizes without a defined peak (figure 3.1b). We next purified the generated fragments in order to check their ability to reconstitute intact AAV *cap* genes and used them as input for the first primer-less re-assembly PCR reaction, followed by a second PCR with flanking primers (see methods section for details). Figure 3.2a shows the result of the second PCR after its separation on an agarose gel. Interestingly, only the DNase-treated sample and the 800 bp fragmentation yielded sufficient amounts of re-assembled capsid genes of ~ 2.2 kb. With 150 or 300 bp

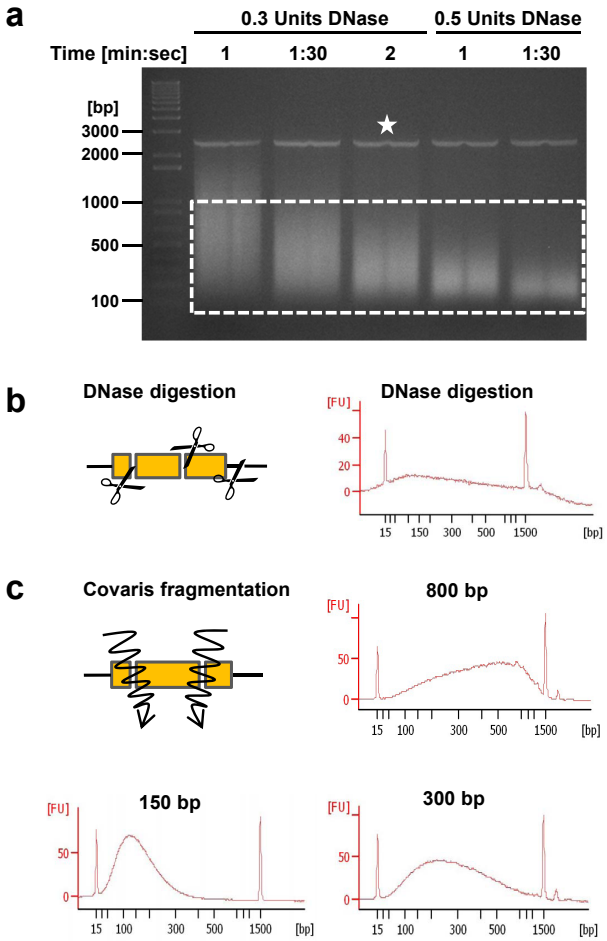

Figure 3.1: Optimization of DNase digest conditions and comparison with controlled fragmentation. a) Examples for DNase digests at 25°C with different incubation times and varying amounts of DNase enzyme, respectively. The optimal condition for DNase digestion is marked with a star and created a wide range of fragments between 100 and 1000 bp (white square). **b) – c)** Schematic depiction of random DNA fragmentation via DNase enzyme (b - left) and controlled focused acoustic DNA fragmentation (c - left). Chromatograms show fragmented DNA analyzed with an Agilent Bioanalyzer DNA 1000 chip. Characteristically, DNase treatment generated a broad range of different DNA sizes without a distinct peak whereas physically treated samples show defined DNA peaks with the intended lengths (150, 300 and 800 bp). Signals at 15 and 1500 bp are marker bands included in the chip. The chromatogram of the DNase digest was generated by N. Schürmann.

DNA fragments no or only a slight capsid DNA band could be detected, respectively. Therefore, we continued library production only with the successfully assembled samples, cloned the capsid fragments into our generated pSSV9_Pac_Asc construct (see section 2.2.1.1) and sequenced 20 library clones after transformation into electro-competent bacteria. Analysis of the shuffled sequences was performed with the Salanto program (see supplementary figure 1 and 2). On top of the previously published tools [195], additional Salanto functions were implemented in order to allow deeper analysis of shuffled AAV *cap* libraries. In detail, the tools "Crossover analysis", "Fragment length analysis", "Show homology" and "Show Crossovercount" were developed and validated in close collaboration with Christian Bender. The latest version of Salanto can be found at https://bitbucket.org/benderc/salanto/wiki/Home. The created 1789rh10 *cap* libraries were further analyzed with the "Fragment length analysis" tool which allows for the determination of number and length of shuffled *cap* fragments within the library. As depicted in figure 3.2b-c, number and length of the generated fragments differed between physically or enzymatically treated samples. In general, DNase digest resulted in a library with smaller fragments (average length = 148.9 bp) which entailed a higher fragment number (average count = 43.5 within the 20 analyzed clones). In contrast, controlled fragmentation with the Covaris resulted in 1.45-fold longer fragments (Ø 216.5 bp) and hence in 1.48-fold

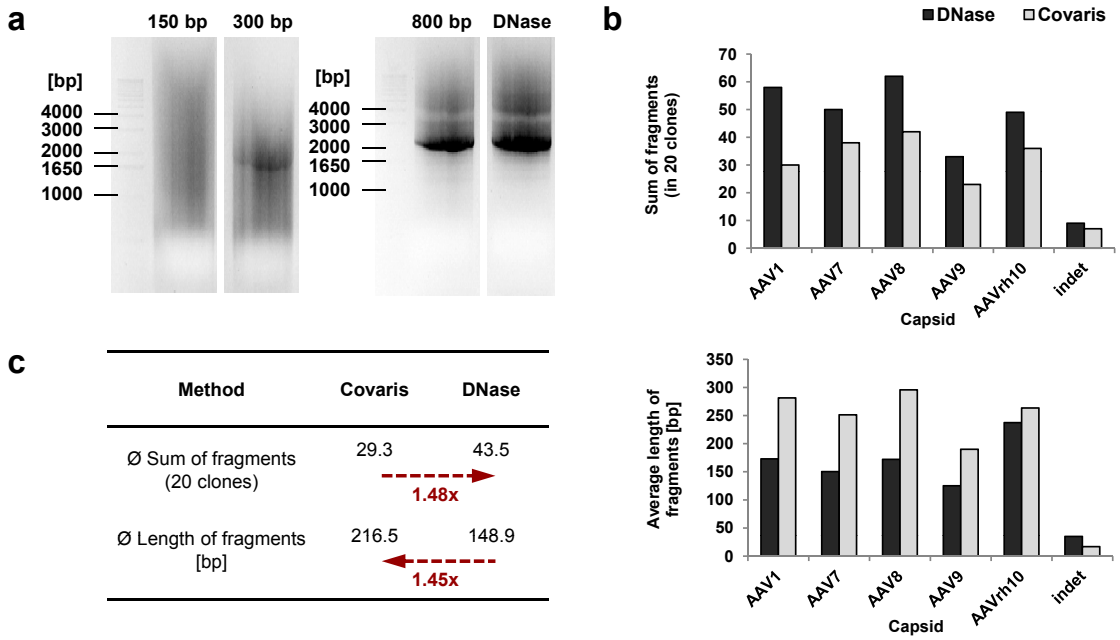

Figure 3.2: Capsid re-assembly after enzymatic or physical DNA treatment and comparison of fragment length and number within generated chimeras. a) Second shuffling PCR reaction with different DNA inputs. From left to the right: 150, 300 or 800 bp fragments created by Covaris treatment or DNase-digested DNA were used as template DNA. PCR reactions were separated on a 1 % agarose gel. Only the DNase treated sample and the 800 bp fragmentation yielded enough DNA for further cloning steps. **b) – c)** Comparison between fragment lengths and numbers of 20 shuffled and analyzed chimeras of a 1789rh10 AAV *cap* library after DNase treatment (black bars in b) or 800 bp Covaris fragmentation (grey bars in b). Red arrows indicate the resulting fold-change of the analyzed values between the two methods (c). Analysis of chimeric clones was done with the Salanto program and the "Fragment length analysis" tool. "Indet" denotes regions in chimeras which could not be assigned unequivocally.

decreased fragment numbers (Ø 29.3). Of note, these data reflect the DNA input characteristics at the beginning of the shuffling reaction. Eight hundred bp fragmentation lacked DNA of smaller size and mainly contained longer fragments whereas the DNase-treated sample showed a broad range of DNA fragments between 100 and 1000 bp (see again figure 3.1b-c). Our data suggest that the fewer small fragments are present in the re-assembly reaction, the lower are the crossover rate and thus the fragment length in the final clones. Since we generally aim to create highly diverse AAV *cap* libraries with many crossovers, we conducted that DNase digest outperformed 800 bp Covaris treatment in our set up. Although Covaris treatment was more robust upon individual experiements, we decided to continue with DNase digestion as fragmentation method for all upcoming AAV shuffling reactions in order to guarantee the highest possible shuffling efficiency.

3.1.2 Rate of sequence homologies influence shuffling efficiency

During our aforementioned experiments we not only observed a correlation between input fragment size and crossover rate but also between homology of parental capsids and crossover

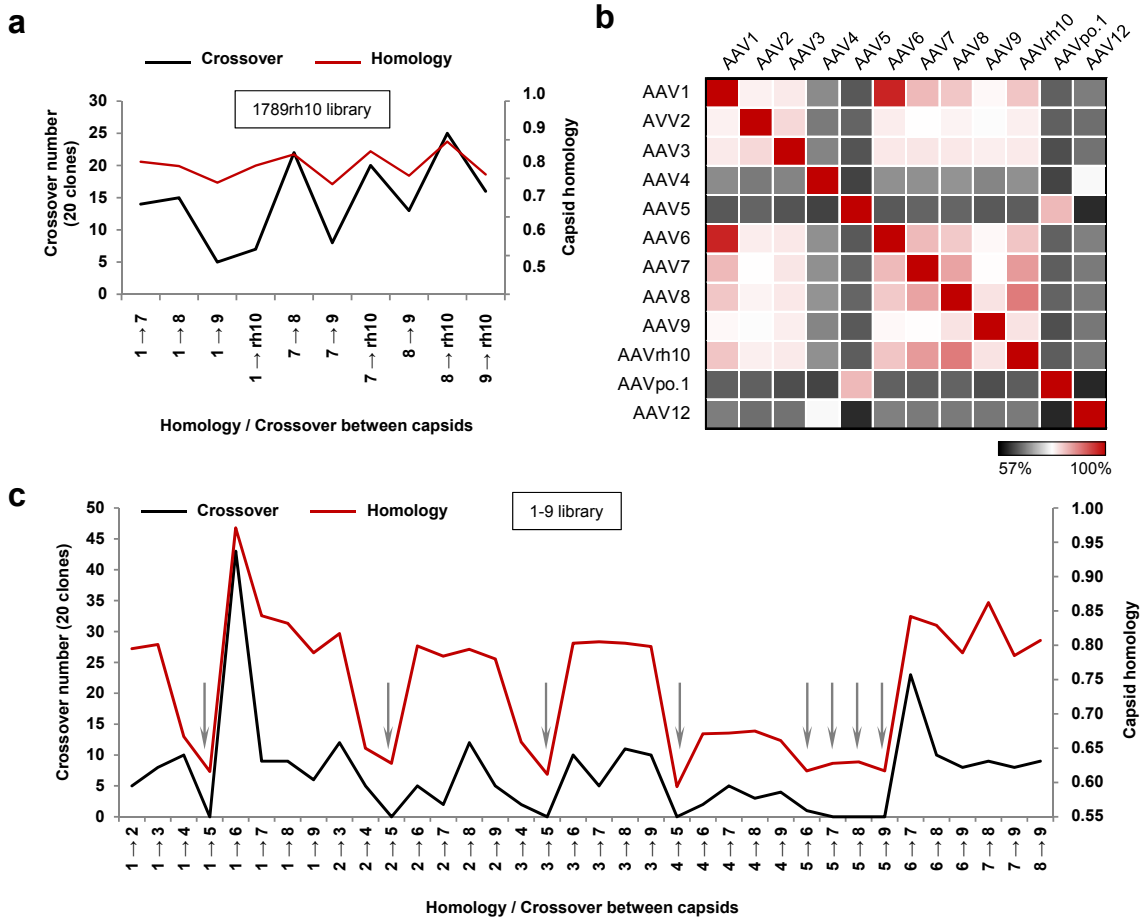

Figure 3.3: Correlation of sequence homology and crossover frequency. a/c) Correlation between crossover frequency (black) and homology (red) of the input capsids after generation of the AAV libraries 1789rh10 (a) and 1-9 (c). Twenty clones per library were analyzed with the Salanto program tools "Crossover analysis" and "Similarity to reference sequences". Grey arrows highlight the low homology and reduced crossover rates of AAV *cap* 4 and 5 to other capsids. **b)** Sequence identities between AAV capsids 1 – 9, rh10, po.1 and 12.

frequency in the created 1789rh10 library (figure 3.3a). Therefore, we analyzed the *cap* sequences of AAV1 – 9, rh10, po.1 and 12 regarding their similarity with the Salanto program. Calculated homologies ranged between 56.6 and 97.1 % (figure 3.3b). The lowest similarity was obtained for AAV *cap* po.1 and cap 12, the highest for AAV *cap* 1 and *cap* 6. Overall, AAV *cap* 1, 2, 3, 6, 7, 8, 9 and rh10 are very similar to each other whereas AAV *cap* 4, 5, po.1 and 12 differ the most. To further investigate the correlation between homology and crossover rate, we generated in addition to the 1789rh10 library an AAV library consisting of capsids 1 – 9. Again we observed the same correlation as described above. The higher the homology between capsids, the more likely they recombine during the shuffling procedure (compare the red and black lines in figure 3.3c). Of note, AAV *cap* 4 and 5 showed the lowest recombination rate amongst all capsids (see grey arrows in figure 3.3c).

3.1.3 Sequence optimizations of AAV *cap* 4 and 5 increase shuffling efficiency

The low recombination rate of AAV *cap* 4 or 5 prompted us to investigate whether increasing capsid homology will enable more crossover events. Therefore, we modified both capsids and adapted them to the *cap* sequence of AAV2. In detail, single nucleotides of AAV4 or 5, respectively, were adjusted to AAV2 without changing the amino acid sequence of all three VPs. Figure 3.4a schematically depicts the process for *cap* 4. Full sequences of the optimized *cap* variants can be found in supplementary figure 3 and 4. Through optimization, homologies between *cap* 2 and 4 or 5 (named from here on *cap* 4_O and 5_O), respectively, increased by 13 % (figure 3.4b-c). Additionally, the similarity between *cap* 4_O and 5_O improved by 11 %. Moreover, capsids 4_O and 5_O showed increased homologies to all other capsids ranging between 0.4 and 4.7 % except for *cap* 12 (*cap* 4_O) or po.1 (*cap* 5_O), respectively.
Next, we investigated the shuffling potential of the optimized *cap* sequences and generated a library containing wild-type AAV capsids 2, 4 and 5 (short: 245) as well as a library containing *cap* 2, 4_O and 5_O (short: 245O). Per library 20 randomly picked clones were sequenced and analyzed with the Salanto tools "Show precise type assignment", "Number of crossovers per clone" and "Crossover analysis". Figure 3.5a shows the colored precise type assignment of the chimeras generated by the program and processed with Microsoft Excel (see section 2.2.3.7 for details). Both libraries clearly differed from each other. The 245 library showed low amounts of crossovers (Ø 3.3 per clone) which clustered preferentially at the 5' part of the

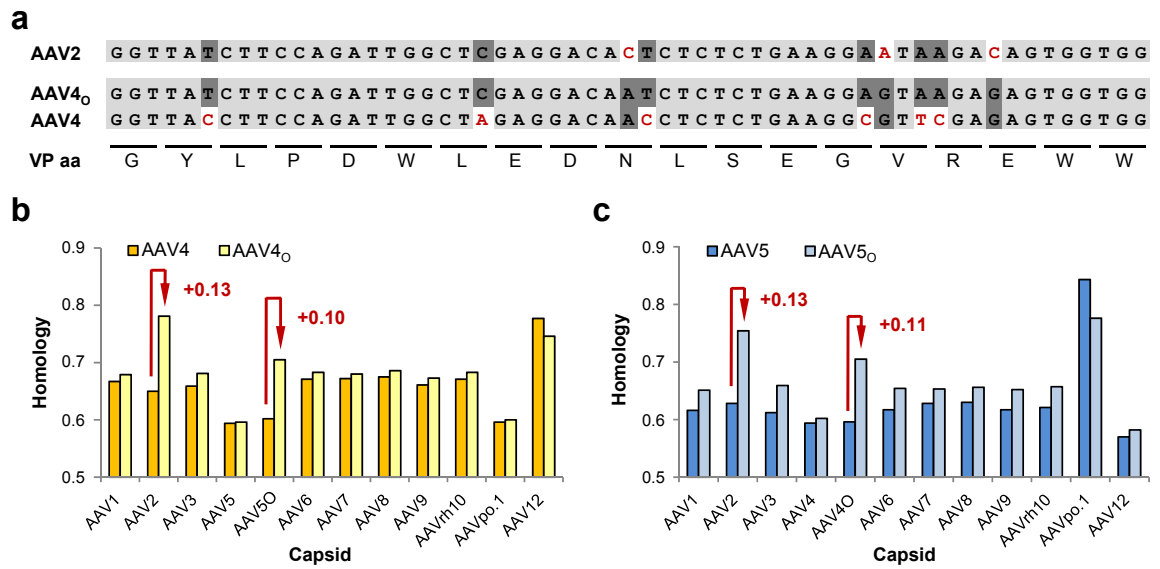

Figure 3.4: Sequence optimization of AAV *cap* 4 and 5. a) Schematic depiction of the optimization process for AAV *cap* 4. From top to bottom: AAV *cap* 2 sequence, optimized $AAV4_O$ sequence with adapted nucleotides, wild-type AAV *cap* 4 sequence, amino acid sequence of AAV *cap* 4 and 4_O VPs. Note that amino acids do not change between wild-type and optimized gene version. **b) – c)** Increase of sequence homologies between AAV *cap* $4/4_O$ (b) and $5/5_O$ (c), respectively, to capsids 1 – 9, rh10, po.1 and 12. Red arrows mark the intended improvement between *cap* $2/4_O$, $2/5_O$ as well as $4_O/5_O$.

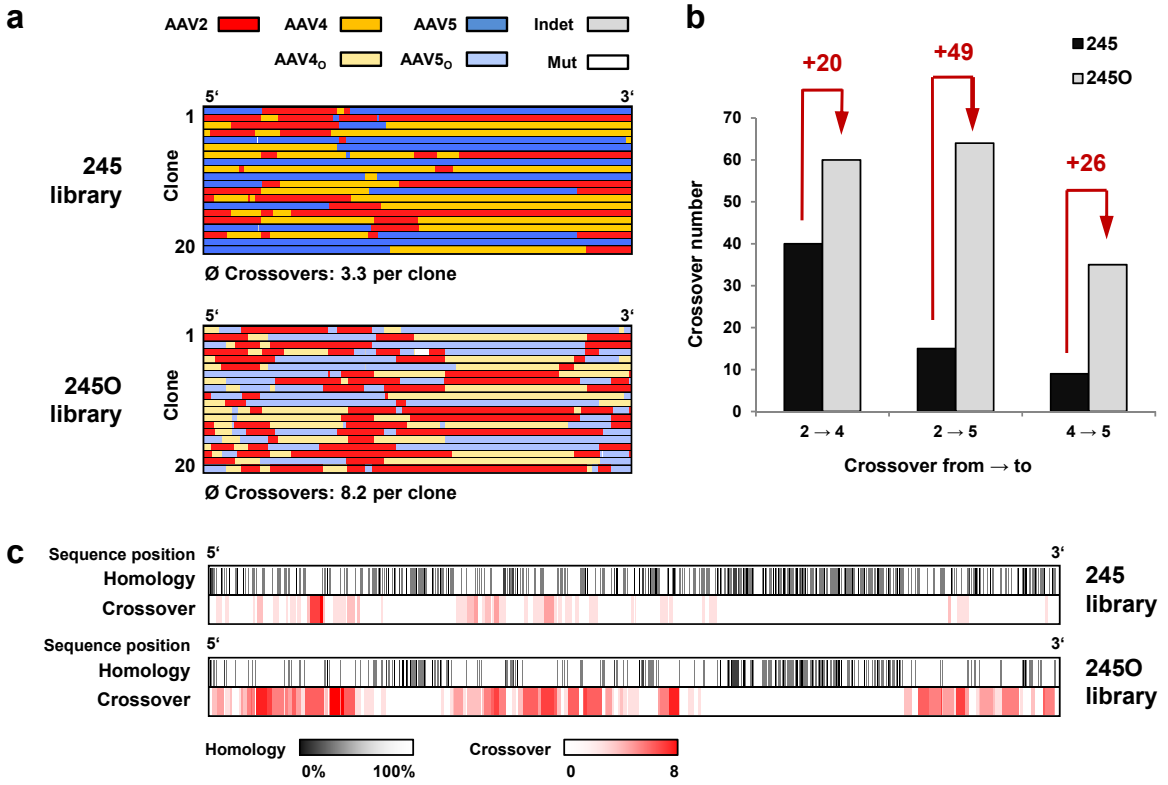

Figure 3.5: Comparison of AAV *cap* libraries 245 and 245O. a) Precise type assignment of AAV *cap* libraries 245 and 245O. Twenty clones per library were analyzed with the Salanto program (see text for details). Use of *cap* 4_O and 5_O yielded higher crossover numbers in comparison to the wild-type library, especially in the 3' part of the *cap* gene. **b)** Detailed crossover analysis between the single *cap* genes in libraries 245 and 245O. Crossover numbers increased in the range of 20 – 49 events within the 245O library. **c)** Correlation of sequence homology and crossover events. Higher crossover numbers (compare white and red) are related to regions with increased sequence homology (compare black and white). Homology calculations between the parental capsids 2, 4 and 5 or 2, 4_O and 5_O were performed with the "Show homology" tool of Salanto and coloured with Excel. Position-related crossover analysis was done with the "Show crossovercount" tool.

cap gene. In contrast, the 245O library contained on average 8.2 crossovers per clone which were distributed almost equally throughout the complete sequences. A detailed analysis of the crossover numbers revealed 20 – 49 more events between the single serotypes after sequence optimization compared to the wild-type capsids (figure 3.5b). Importantly, the location of crossovers correlated with the homology of the capsids. The higher the similarity between the capsid DNA sequences, the more likely crossovers appeared in these regions (figure 3.5c). Altogether, sequence optimization of *cap* 4 and 5 yielded higher shuffling efficiency within a 245O library in direct comparison to the wild-type counterpart.

In order to determine the influence of increased DNA homologies in a more complex library, we first decided to combine the commonly used AAV capsids 8 and 9 with *cap* 2, 4 and 5 (short: 24589) or 4_O and 5_O (short: 24589O), respectively. Again, we observed a higher shuffling efficiency especially in the 3' part of the gene when the sequence-optimized versions

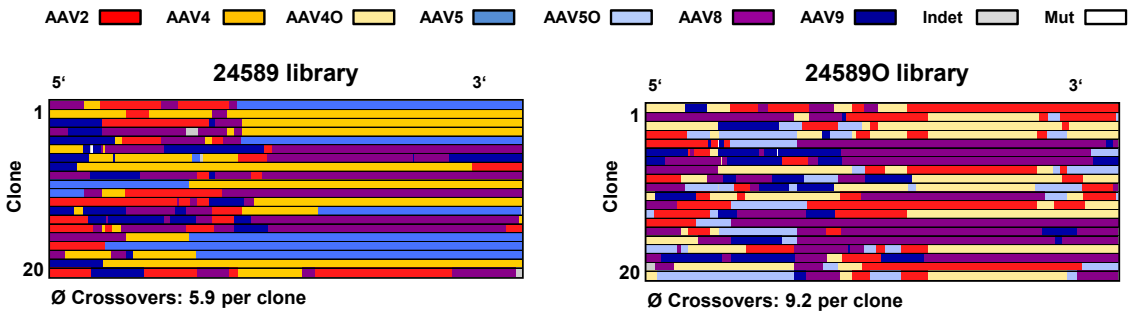

Figure 3.6: Comparison of AAV *cap* libraries 24589 and 24589O. Precise type assignment of AAV *cap* libraries 24589 and 24589O. Twenty clones per library were analyzed with the Salanto program. Use of *cap* 4_O and 5_O yielded higher crossover numbers in comparison to the wild-type library, especially in the 3' part of the *cap* gene.

of *cap* 4 and 5 were used (5.9 versus 9.2 crossovers per clone, see figure 3.6). Detailed analysis revealed higher crossover numbers between *cap* 2, 4_O and 5_O among each other (see table 3.1 for details). Moreover, recombination rates between AAV *cap* 5_O and 8 or 9 were increased. Secondly, we combined *cap* $4/4_O$ and $5/5_O$ with capsids other than *cap* 2 and generated a 145689 or 145689O library. In addition, we used our modified capsid genes in a complex library that contained *cap* genes of the AAV serotypes 1 – 9 (short: 1-9O) and compared it to the wild-type counterpart (1-9). Table 3.1 summarizes the results of the abovementioned libraries (precise type assignments can be found in supplementary figure 8). In general, we determined

Table 3.1: Comparison of crossover numbers between capsids in *cap* 4 / 4_O and 5 / 5_O containing libraries. Twenty clones per library were analyzed with the tools "Numbers of crossovers per clone" and "Crossover analysis" of the Salanto program.

Shuffled library		**24589**	**24589O**	**145689**	**145689O**	**1-9**	**1-9O**
Average number of crossovers per clone		5.9	9.2	9.5	10.1	14.9	16.1
Crossover from → to	2 → 4	12	38	n/a	n/a	6	8
	2 → 5	1	25	n/a	n/a	0	3
	4 → 5	5	16	0	14	0	4
	4 → x[a]	31	22	27	43	26	25
	5 → x[a]	3	20	7	10	3	13

[a] Any capsid in library other than 2, 4 or 5
n/a = not applicable

increased crossover rates between *cap* 4_O and 5_O in all generated libraries. Besides, *cap* 5_O recombined more frequently with all other capsids than AAV5 wild-type. In line with library 24589O, crossover numbers per clone slightly improved also in libraries 145689O and 1-9O. Taken together, our results show that sequence optimization of capsids yields higher crossover

numbers and better shuffling in the 3' part of the *cap* gene. Moreover, this effect persisted even in highly complex libraries.

3.1.4 AAP expression can be reconstituted in sequence-optimized capsids and knock-out mutants

During sequence optimization of AAV *cap* 4 and 5 the first ORF remained unaffected and VP protein sequence was not changed. However, at the same time the second ORF was modified inadvertently and thereby the assembly-activating protein (AAP) mutated (see also section 2.2.1.3 for a more detailed description). As a result, $AAV4_O$ and 5_O were not able to assemble functional viral particles any more (see also figure 3.8, second bar of b/c). In order to study whether the functionality of AAP can be restored upon *trans*-complementation, we generated an AAV2-AAP knock-out mutant (named AAV2mut from hereon) which lacks the AAP2 start codon and contains a stop codon 66 bp downstream of the modified start (figure 3.7a). Of note, the AAV2mut construct was unable to produce functional viral particles (figure 3.7b, second bar from the left).

In parallel to the AAV2 mutant, we cloned an AAP2 expression vector. In detail, the AAP2 start codon was exchanged from CTG to ATG and the DNA sequence was placed behind a CMV promoter. Moreover, AAP2 was N-terminally linked to a FLAG- and an HA-tag in order to allow detection through Western blotting and microscopy. To investigate the functionality of the construct, we tested which amounts of AAP2 plasmid were necessary to restore assembly of the AAV2 mutant. Therefore, we produced AAV crude cell lysates by transfection of HEK293T cells in 6-well format with different amounts of AAP2 plasmid ranging from 0.0625 to 0.65 µg per well. DNA amounts were supplemented with stuffer (sheared salmon sperm) DNA in order to achieve comparable transfection conditions. As a reporter a GFP-encoding transgene was packaged. To test production of assembled and functional viral particles, crude lysates were used to transduce new HEK293T cells. Forty-eight hours later GFP expression was analyzed via flow cytometry. Already the lowest amount of supplemented AAP2 plasmid (0.0625 µg per well) was able to recover AAV2mut assembly by up to 50 % compared to wild-type. The most potent rescue could be observed with 0.5 µg AAP2 DNA which yielded as many GFP-positive cells as AAV2wt (figure 3.7b). Therefore, we decided to use this ratio of AAP2 and helper plasmids for all upcoming experiments (see also section 2.2.6.6).

Next to the crude lysate transduction, we also determined the amount of viral proteins and AAP2 in the cell lysate after transfection via Western blotting. Interestingly, we observed a reduction of VP amounts for AAV2mut which could be restored to wild-type levels upon AAP2 supplementation (see also section 3.2.1). At the same time Rep protein levels stayed constant (figure 3.7c). AAP2 could be detected only for the highest amounts of the transfected plasmid with an anti-HA tag antibody.

The successful recovery of AAV particle assembly after AAP knock-out prompted us to investigate the effect of AAP4wt or 5wt addition during transfection of $AAV4_O$ or 5_O constructs, respectively. In order to test the produced crude lysates, MCF7 cells (AAV4 particles) or

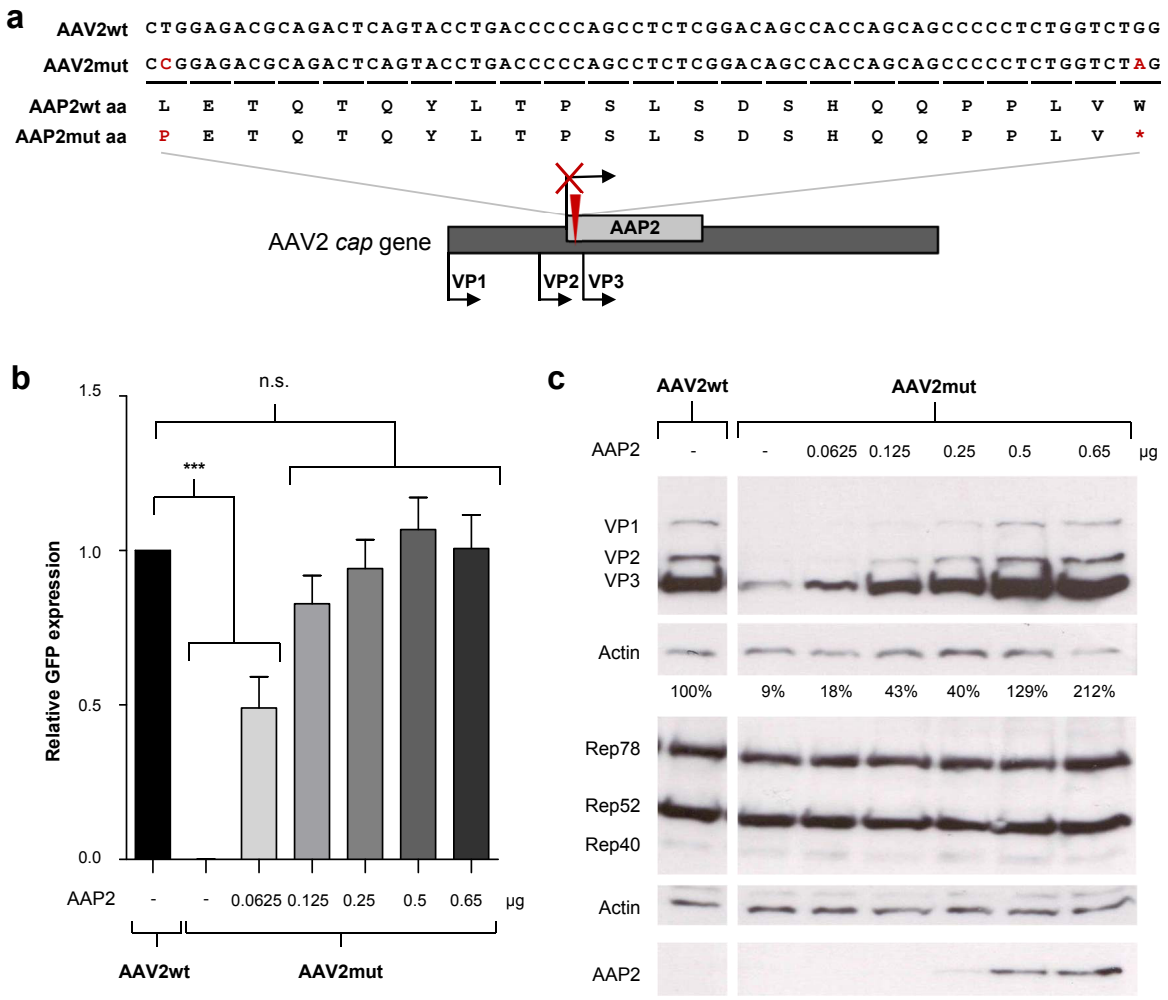

Figure 3.7: Generation of AAV2-AAP knock-out mutant and recovery with AAP2 supplemented *in trans*. **a)** Schematic depiction of introduced AAV2 *cap* modifications in order to knock-out AAP2. The AAP2 start codon CTG was exchanged to CCG and a stop codon was inserted 66 bp downstream. VP1-3 protein sequences were not influenced by the mutations. **b)** *Trans*-complementation assay of the AAV2mut construct with different amounts of AAP2 compared to AAV2wt. HEK293T cells were transfected (6-well format) for the production of GFP-encoding AAV crude cell lysates and supplemented with different amounts of AAP2 plasmid (0.0625 - 0.65 µg per well). Crude cell lysates were subsequently used to transduce new HEK293T cells and GFP expression was measured 48 hours after transduction via flow cytometry. Values were normalized to the AAV2wt control. As statistical analysis one-way ANOVA with Bonferroni's multiple comparisons test was performed. **c)** Western blot analysis of cell lysates 48 hours after transfection described in b). Next to the viral proteins VP1-3, AAV Rep proteins as well as AAP2 were detected. The latter was stained with an anti-HA tag antibody. Relative VP levels (depicted below the actin staining) were quantified with the Fiji program and double-normalized to actin and the AAV2wt control. n.s. = not significant, *** $p < 0.001$, n = 3 (*Trans*-complementation assay).

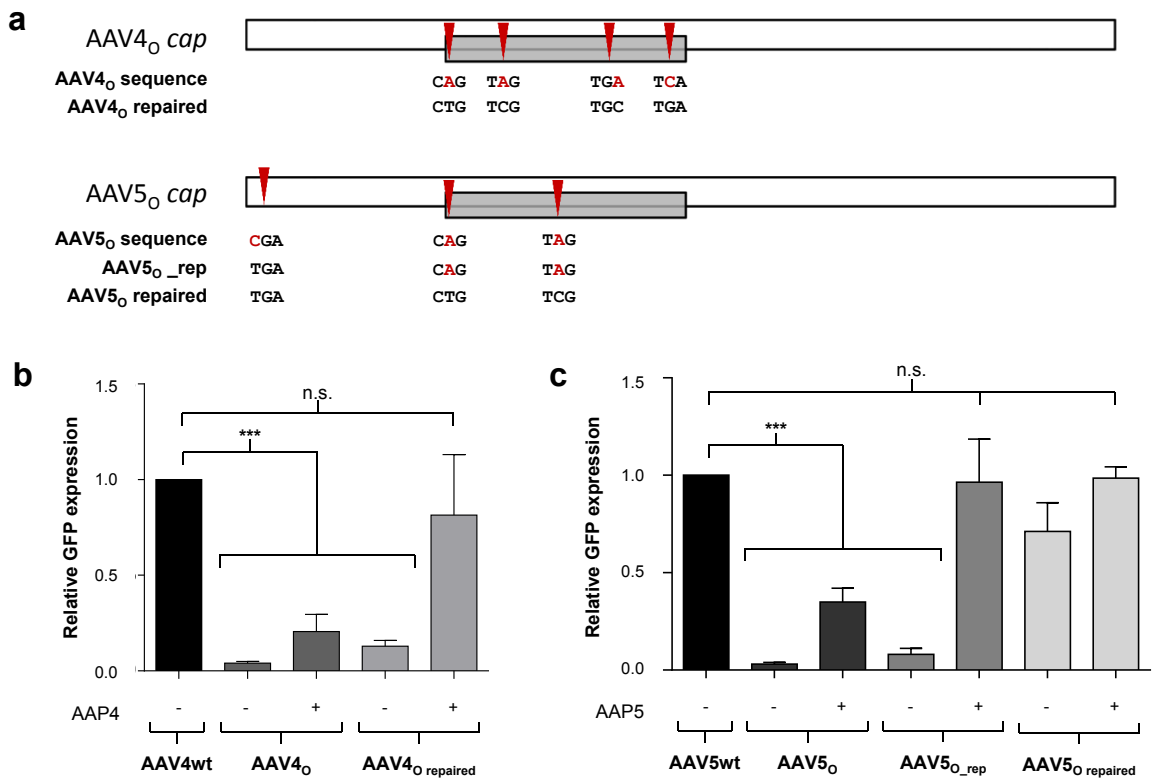

Figure 3.8: Repair of AAV4$_O$ and 5$_O$ capsid genes and *trans*-complementation assays with AAP4wt and 5wt. a) Schematic depiction of repaired AAV4$_O$ and 5$_O$ *cap* sequences. Nucleotides in red mark the mutations that led to an AAP knock-out or the deletion of the *rep* stop codon in AAV5$_O$, respectively. The corrected nucleotide triplets are shown below. **b/c)** *Trans*-complementation assays of AAV4 (b) and AAV5 (c) constructs with or without AAP4wt or 5wt, respectively. Crude cell lysates were tested in MCF7 cells (AAV4) or SF539 cells (AAV5). GFP-positive cells were quantified via flow cytometry and normalized to the wild-type controls. As statistical analysis one-way ANOVA with Bonferroni's multiple comparisons test was performed. n.s. = not significant, *** $p < 0.001$, n = 3.

SF539 cells (AAV5 particles) were transduced. Couriously, we detected only a partial recovery of AAV assembly up to 35 % compared to wild-type after AAP supplementation (figure 3.8 b/c, compare bar one to three). Therefore, we repaired the introduced mutations in AAV4$_O$ and 5$_O$ resulting in *cap* sequences AAV4$_{O\ repaired}$ and 5$_{O_rep}$ / 5$_{O\ repaired}$ (see figure 3.8a). In addition, we corrected the deleted stop codon of AAV5$_O$ which led to extended Rep40 and Rep68 protein sequences (see supplementary figures 5 – 7 for DNA sequences) and tested the new constructs again with or without *trans*-complementation of AAP (figure 3.8b/c). For AAV4$_O$ the repair of all modified nucleotides only led to an 8 % increase of GFP positive cells after crude lysate transduction. However, after addition of AAP4wt during transfection, AAV4$_O$ functionality achieved 81 % of wild-type level. Expression of AAV5$_{O\ repaired}$ alone achieved already 71 % GFP-positive cells compared to wild-type which could be increased to 99 % by addition of AAP5wt during AAV particle production. Of note, co-expression of AAV5$_{O_rep}$ together with AAP5wt recovered virus functionality to 96 % of wild-type.

Altogether, we could show that AAP functionality could be restored by either repair of modified nucleotides ($AAV5_{O\ repaired}$) or addition of AAP wild-type ($AAV4_{O\ repaired}$). Importantly, in all cases achieved levels of GFP expressing cells were comparable to wild-type AAV particle transductions.

3.1.5 AAP does not influence AAV shuffling

Since the AAP frame can be disrupted not only by sequence optimization but also during the AAV shuffling process, we wanted to examine the functionality of individual shuffled AAPs and the role of AAP in a complete library. Towards this aim, we produced the aforementioned AAV *cap* libraries 24589 and 24589O as viral stocks and purified them through an iodixanol density gradient. During production we supplemented both libraries with a mixture of all wild-type AAPs present in the libraries, i.e., AAP2, 4, 5, 8 and 9. Subsequent qPCR titration of the virus stocks revealed that AAP addition did not influence virus titers (table 3.2). Analysis of the shuffled AAPs within 20 randomly selected and sequenced clones showed that in library 24589 all AAPs are theoretically functional (intact start and stop codon, normal AAP length). In contrast, 12 / 20 are not functional in library 24589O due to missing start or stop codons, or newly created early stops. However, *trans*-complementation with AAP during virus production did not increase genome-containing particle numbers for library 24589O. In parallel, also AAP supplementation during crude lysate production of 15 individual, shuffled 24589O library clones, which contained mainly non-functional AAPs, did not lead to the formation of functional AAV particles (see supplementary figure 9).
Based on this observation, we assumed that AAP shuffling is not critical and that other factors influence virus production after AAV shuffling to a larger degree. To further investigate this hypothesis, we cloned 46 randomly selected, shuffled and theoretically functional AAPs from five different AAV *cap* libraries with different homologies and complexities into our CMV expression vector backbone (named AAP-X_1 to AAP-X_{47}, see supplementary figure 10 for protein type assignments). Since these libraries also contained AAV capsids 1, 3, 6, 7, 8, 9 and rh10, we initially determined the recovery potential of their AAPs with the corresponding AAP knock-out mutants in SF539 cells. Figure 3.9a displays the successful rescue of all mutants after *trans*-complementation with their serotype-specific AAPs. Moreover, we

Table 3.2: Genomic titers of AAV *cap* libraries 24589 and 24589O with or without supplementation of an AAP mixture of serotypes 2, 4, 5, 8 and 9 during virus production.

AAV *cap* library	AAP addition	Titer [vg/ml]
24589	–	8.20×10^{11}
24589	+	7.72×10^{11}
24589O	–	1.85×10^{11}
24589O	+	1.78×10^{11}

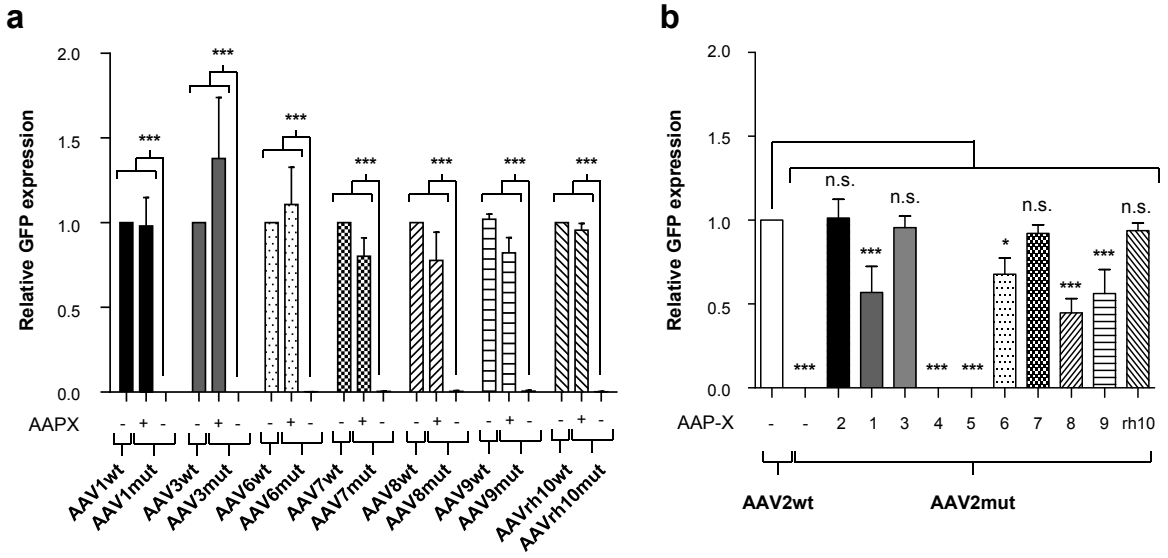

Figure 3.9: *Trans*-complementation assays with various AAV serotypes and AAP cross-reactivity with AAV2mut. a) *Trans*-complementation assays of AAP-deficient AAV1, 3, 6, 7, 8, 9 and rh10 constructs with their corresponding AAPs. Produced crude cell lysates were tested in SF539 cells. Values were normalized to the AAVwt controls. **b)** Cross-reactivity test of all cloned wild-type AAPs together with the AAV2mut construct. Samples were tested in HEK293T cells and values were normalized to AAV2wt. Positive cells were determined via flow cytometry. As statistical analysis one-way ANOVA with Bonferroni's multiple comparisons test was performed in a and b. n.s. = not significant, *** $p < 0.001$, $n = 3$.

investigated the potential cross-reactivity of the cloned wild-type AAPs with the AAV2mut construct. Interestingly, all AAPs except for AAP4 and 5 were able to recover AAV2mut functionality between 45 and 96 % compared to AAP2 (figure 3.9b).

We therefore tested the selected shuffled AAPs first with the AAV2mut construct and, if no recovery was observed, either the $AAV4_{O\ repaired}$ or the $AAV5_{O\_rep}$ capsid were used in further *trans*-complementation assays. We found 37 of 46 AAPs to be functional with AAV2mut (figure 3.10) and classified them as high performers (50 % – 100 % AAV2mut rescue ability compared to AAP2) or intermediate performers (5 % – 50 % AAV2mut rescue ability compared to AAP2). AAPs with less than 5 % recovery potential were estimated as non-functional. Apart from the 37 AAV2mut-compatible variants, two more AAPs showed functionality with $AAV5_{O\_rep}$ (AAP-X_{19} and AAP-X_{46}, data not shown). AAP-X_{44} also reached higher recovery rates with $AAV5_{O\_rep}$ than with AAV2mut. In total, 84.8 % of the tested AAPs were functional. Of note, the highest degree of defect AAPs was observed in libraries which contained less homologous capsids (i.e., *cap* 4 and 5). Library 24589O revealed the lowest number of active AAPs (67 %). Another defect AAP appeared in the complex 1-9 library. In summary, we conclude from these results that AAP shuffling does not influence AAV capsid shuffling to a major extent.

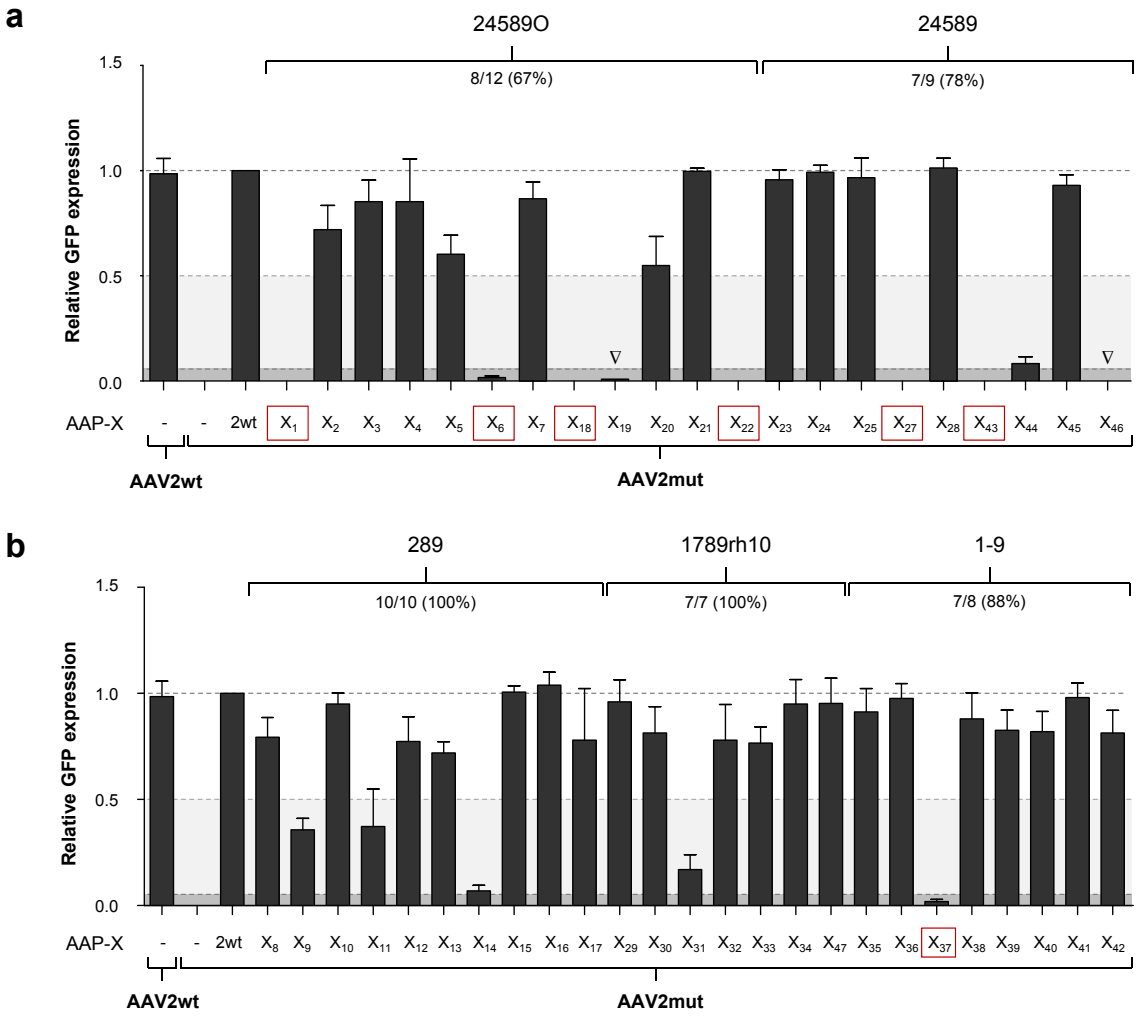

Figure 3.10: Test of randomly picked shuffled AAPs. *Trans*-complementation assays using the AAV2mut construct together with 46 randomly selected shuffled AAPs (AAP-X_1 to AAP-X_{47}). AAPs originate from AAV *cap* libraries 24589 and 24589O (a) or 289, 1789rh10 and 1-9 (b) as indicated above the diagrams. *Cap* library 289 was generated by A.K. Herrmann. Out of 46 tested AAPs 37 were functional with AAV2mut. Two additional AAPs only showed functionality with the $AAV5_{O\_rep}$ construct (marked with triangles). Non-functional AAPs are highlighted with a red square. The cut-off value for AAP functionality was set at 5 % of GFP positive cells (dark grey background) compared to AAV2mut supplemented with AAP2. AAP-X_9, -X_{11}, -X_{14}, -X_{31} and -X_{44} are AAPs with an intermediate rescue ability which was defined between 5 and 50 % GFP positive cells compared to the control (light grey area). Crude cell lysates were tested in HEK293T cells, positive cells were determined by flow cytometry and values were normalized to AAV2mut supplemented with AAP2. Numbers in brackets indicated the percentage of working (high and intermediate phenotype) AAPs per library. Error bars represent experimental differences after three independent transductions of the same crude lysate.

3.1.6 Successful AAV virus production with repaired optimized *cap* genes

In order to prove that our generated shuffled libraries with *cap* 4 and 5 can yield functional virus particles, we produced iodixanol-purified virus stocks and titered them via qPCR. Moreover, we compared libraries containing the wild-type version of AAV4 and 5 with their

sequence-optimized containing counterparts. For the latter we repeated the shuffling procedure and included the repaired versions of *cap* 4_O and 5_O to overcome inadvertent AAP limitations caused by the sequence adaption (libraries are accordingly labeled with O_r). DNA type assignments of the new libraries can be found in supplementary figure 11 and resulting titers are depicted in table 3.3. Notably, virus titers were comparable for wild-type and O_r libraries with the exception of the $245/245O_r$ libraries. In detail, $245O_r$ virus titers were more than 100-fold lower compared to its wild-type counterpart. AAV particle amounts of the other libraries ranged between 10^{11} and 10^{12} particles per milliliter with $24589{\pm}O_r$ being 10-fold lower than $145689{\pm}O_r$ titers. Altogether, shuffled libraries with sequence-optimized AAV *cap* variants generated normal amounts of viral genome-containing particles except for $245O_r$ which was not able to yield high titers.

Table 3.3: Genomic titers of different AAV *cap* libraries ± 4 / $5O_r$ capsids.

AAV *cap* library[a]	**Titer [vg/ml]**
245	1.2 x 10^{12}
$245O_r$	7.1 x 10^{9}
24589	7.1 x 10^{11}
$24589O_r$	7.7 x 10^{11}
145689	2.7 x 10^{12}
$145689O_r$	3.2 x 10^{12}

[a] r indicates that the repaired versions of sequence-optimized AAV *cap* 4 and 5 were used

3.2 AAP biology

3.2.1 AAP knock-out leads to VP degradation

The fact that AAP biology is still unexplored to a large extent prompted us to investigate this protein in more detail. In particular, we were interested in following up our unexpected observation that AAP knock-out led to a reduction of VP levels in the cell lysate detected by Western blotting (see figure 3.7c). The VP signal could be recovered completely by addition of AAP during crude cell lysate production. Since the protein sequence of the VPs was not influenced by our introduced mutations we first hypothesized that VP reductions were caused by a dysregulation on mRNA levels. We therefore isolated mRNA of HEK293T cells after transfection (similar to crude lysate productions) and performed quantitative PCR after reverse transcription. In order to measure the different transcripts of the AAV genome, we designed four primer pairs (#1107 – #1114, see table 2.1) binding at different positions in the viral mRNAs (arrows in figure 3.11a – right side). Data were analyzed according to the $2^{-\Delta\Delta CT}$ method [198] and normalized first to the house-keeper GAPDH and second to the

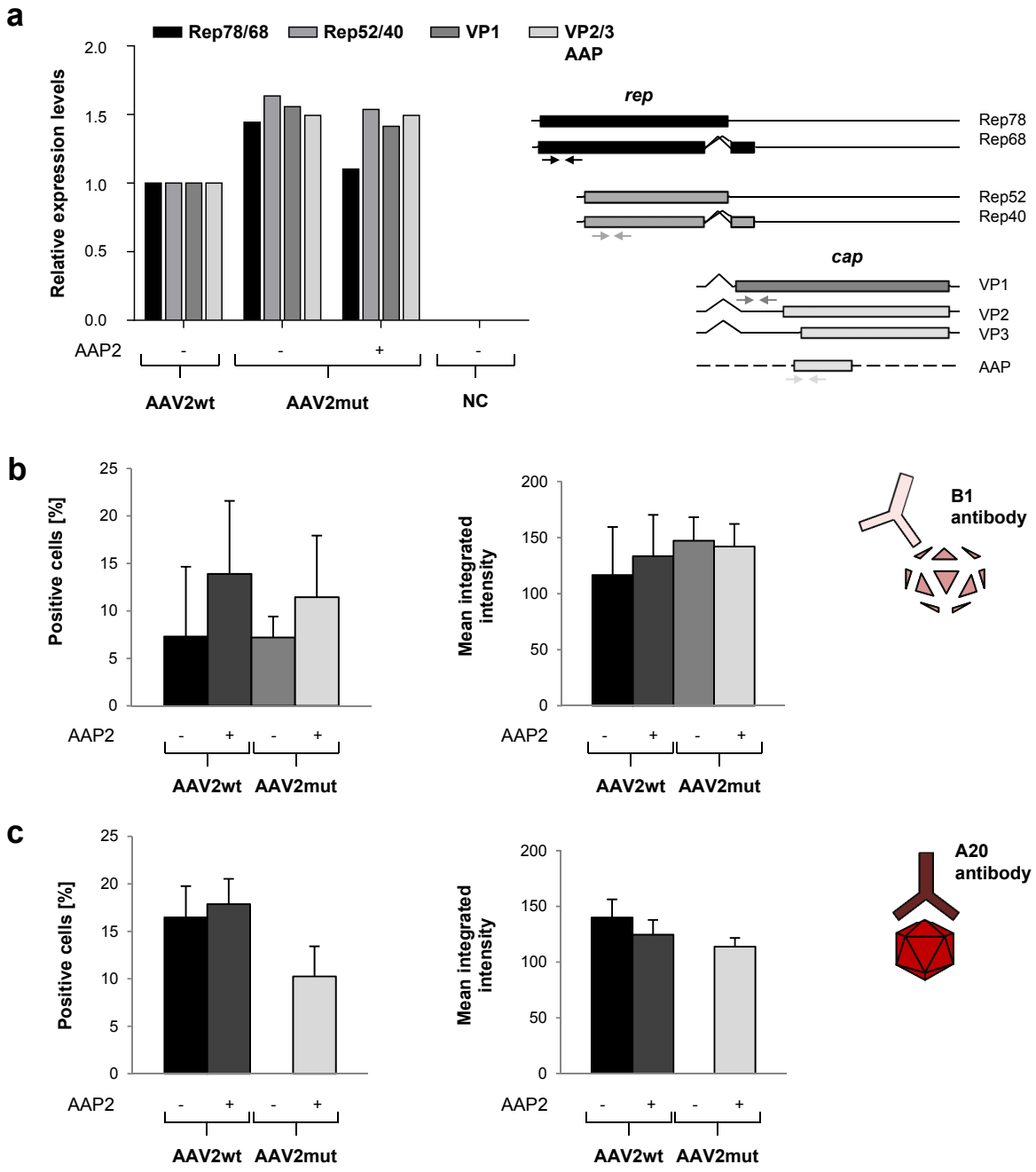

Figure 3.11: VP quantification on mRNA and protein levels. a) Quantitative PCR of AAV genome transcripts. Relative expression levels of the different viral mRNAs are displayed in the diagram. Arrows in the scheme at the right side depict the binding sites of the used primer pairs within the different transcripts. Values were analyzed according to the $2^{-\Delta\Delta CT}$ method and normalized to GAPDH as house-keeper as well as the AAV2wt control. **b/c)** Quantification of AAV-VPs (b) or assembled AAV2 particles (c) in microscopy pictures of cells stained with B1 or A20 antibody, respectively. Schemes on the right side display the specificity of the antibodies. Six pictures per condition were analyzed with the CellProfiler program in order to determine the percentage of positive cells and the mean integrated intensity of the stained cells 48 hours after transfection.

AAV2wt control. As depicted in figure 3.11a, we could not observe a significant dysregulation of mRNA levels after knock-out or supplementation of AAP during AAV production. Hence, we conducted that VP reduction is not caused on mRNA level.
Next, we quantified the amount of VPs and assembled viral particles 48 hours after transfection in the cells via immunostaining. For detection we either used the A20 antibody [190] which only binds assembled AAV2 particles (see scheme in figure 3.11c) or the B1 antibody [190] which is specific for linear, non-assembled AAV VP1, VP2 and VP3 (scheme in 3.11b). For each condition, six microscopy pictures were analyzed with the open-source software CellProfiler (see also supplementary table 2). Next to the percentage of positively stained cells also the mean integrated intensity of the signals was determined. Consistent with the mRNA data, we again detected no major differences in numbers of positive cells and intensities between AAV2wt and AAV2mut ± AAP2. As expected, A20 staining gave no signal for the AAV2mut construct without AAP supplementation. Based on these observations, we hypothesized that the levels of free VP proteins in the cells stay constant independent of AAP2. Hence, increasing VP levels after AAP2 addition by Western blotting potentially revealed from assembled AAV2 particles which got denaturated and detectable with the B1 antibody upon sample treatment. To test whether VPs are degraded in the absence of AAP2, we added the proteasome inhibitor MG-132 (final concentration in the media: 2 µM) five or 24 hours after transfection to the cells (see scheme in figure 3.12a), performed Western blot analysis 48 hours after transfection (figure 3.12b) and quantified the visible bands with the Fiji software. Indeed, we determined increasing VP levels with the AAV2mut construct after MG-132 treatment. When the inhibitor was already supplemented to the media five hours after transfection, protein amounts of AAV2mut were comparable to the AAV2 wild-type control. Twenty-four hours after transfection 44 % VP proteins compared to wild-type could be detected whereas without treatment only 12 % of VPs were present in the cell lysate. Importantly, AAV crude lysates generated from the MG-132 treated cells showed comparable transduction efficiencies relative to the AAV2wt control (figure 3.12c). These results (I) prove that although VPs accumulate in the cell, they are not able to assemble functional virus particles in the absence of AAP2 and (II) indicate that VPs are degraded through the proteasome pathway when not assembled into AAV particles. Interestingly, we also observed a slight, additional band at $\sim$ 200 kDa only for the AAV2mut construct after MG-132 treatment (figure 3.12b, red square) which led us to speculate that these bands might be ubiquitinated viral proteins that were not degraded by the blocked proteasome. In our next experiments we tried to detect this possible VP ubiquitination via a VP-IP or a Ubiquitin-IP. However, thus far, we could not prove our hypothesis due to insufficient pull-down efficiencies or unspecific binding in the IP (data not shown). Nevertheless, the additional band after MG-132 treatment could be robustly reproduced in all conducted experiments.

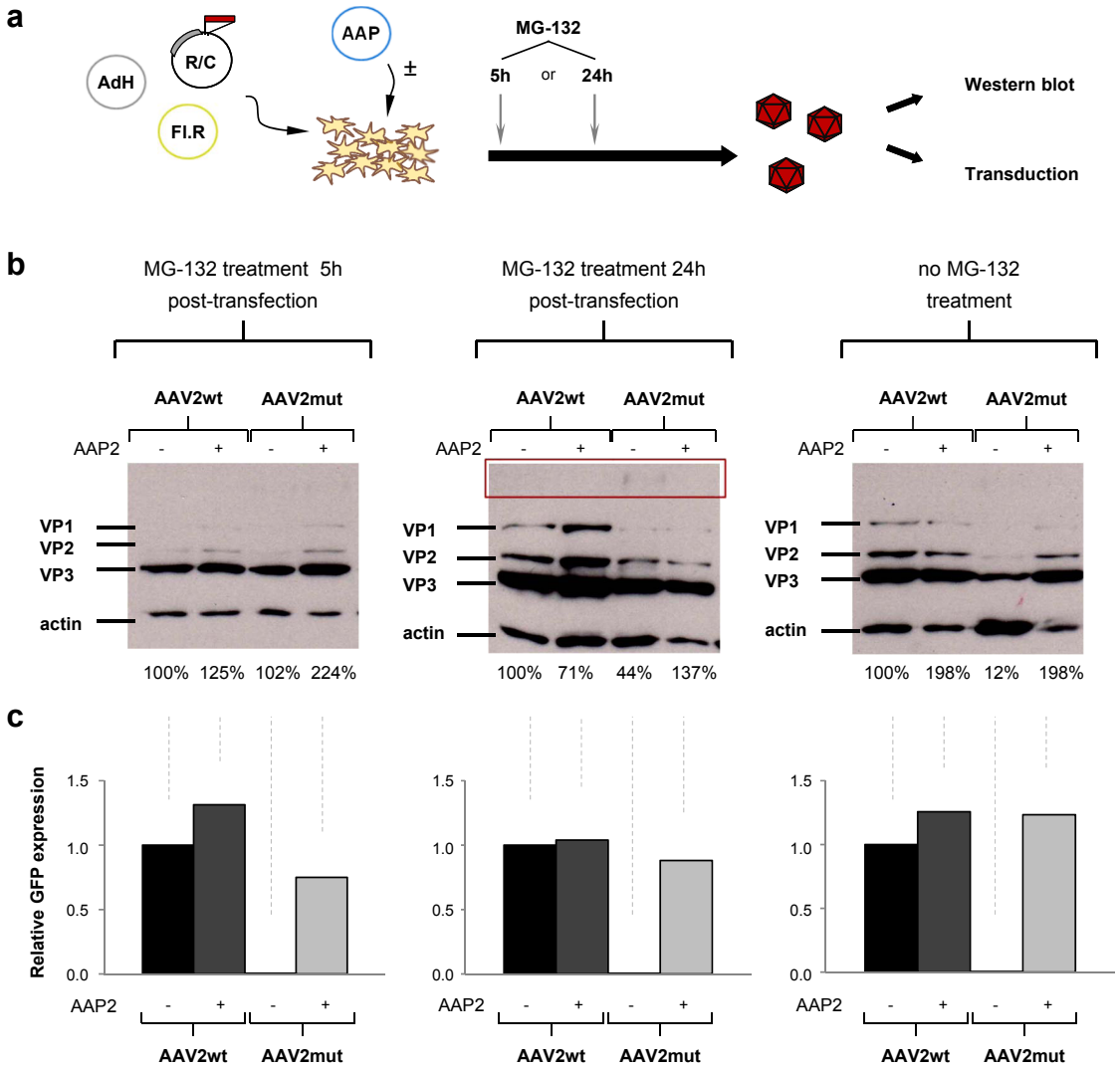

Figure 3.12: MG-132 treatment and quantification of VP signals by Western blotting. **a)** Schematic depiction of the AAV crude cell lysate production with MG-132 treatment five or 24 hours after transfection. Samples were harvested 48 hours after transfection and processed for Western blot analysis or cell transductions. **b)** Western blot analysis of cell lysates after AAV crude lysate production. Cells were treated with MG-132 five (left panel) or 24 hours (middle panel) after transfection and compared to untreated samples (right panel). Viral proteins were detected with the B1 antibody, quantified with the Fiji program and normalized first to actin and second to the AAV2wt control without AAP2. The red square in the middle panel marks an additional, slight band of $\sim$ 200 kDa which was detected only for AAV2mut. **c)** Relative GFP expression of HEK293T cells 48 hours after transduction with crude lysates from samples shown in b). Positive cells were measured via flow cytometry and values normalized to the AAV2wt control without AAP2. AdH = adeno helper plasmid, Fl.R. = fluorescence reporter plasmid encoding for GFP, R/C = AAV *rep-cap* helper plasmid.

3.2.2 Detection of endogenous AAP2 during AAV2 production

Next to the effect of AAP2 on VP expression (see previous chapter), we were also interested in the endogenous expression of AAP and the protein localization in the cell. To further investigate this question, we ordered a polyclonal anti-AAP2 antibody from the company Perbio. In detail, the company immunized rabbits with the previously published short peptide GKDSSTTTGDSDPRDSTS (AAP2: aa 121-138) [25] and affinity-purified the resulting antibody. We used this antibody in immunostainings and detected AAP2 together with VPs

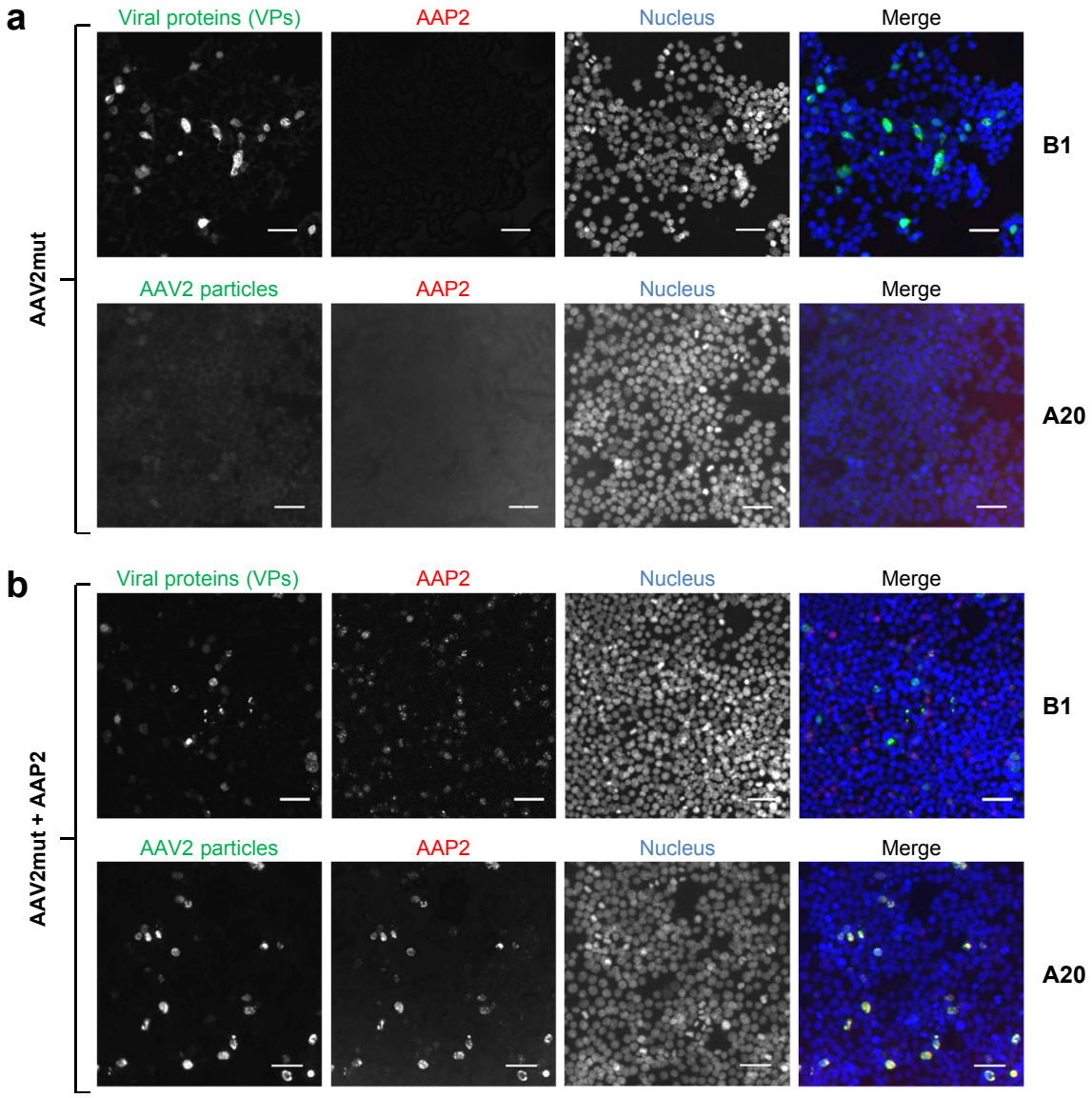

Figure 3.13: Validation of the anti-AAP2 antibody. AAV2mut-transfected HEK293T cells supplemented with stuffer DNA **a)** or an AAP2 expression construct **b)** were immunostained after 48 hours with the new anti-AAP2 antibody together with the B1 (first row) or A20 (second row) antibody, respectively. Transfections and immunostainings were performed in LabTek format. Nuclei were stained with Hoechst. Scale bar = 50 µm.

or assembled particles after cell transfection with an AAV2wt *rep-cap* helper, a transgene (luciferase) and the adenohelper construct. To control for the specificity of the antibody we stained AAV2mut-transfected cells with A20 or B1 antibodies, respectively, and monitored unspecific binding or cross-reactivity. As expected, we did not detect any signal in the AAV2mut-transfected cells which do not express AAP2 (figure 3.13a). However, when AAP2 was co-transfected together with AAV2mut a clear signal was seen with the anti-AAP2 antibody (figure 3.13b) and assembly could be detected with the A20 antibody (second row in figure 3.13b). In order to detect endogenous AAP2 we transfected cells with the AAV2wt construct and quantified the amount of positive cells at different time points in the captured microscopy pictures (four pictures per condition) with the CellProfiler software (see supplementary tables 3 – 4 for protocols). Diagrams in figure 3.14 show the expression of AAP2 and VPs (a) or assembled viral particles (b) over time, respectively. In general, we observed that not all cells which are positively stained for VPs are also positive for AAP2. In contrast, the number of AAP-positive cells which also contained assembled AAV2 particles was almost equal except after 48 hours where more AAV particles were detected. For a more detailed analysis of the AAP localization we also captured images with a Leica confocal microscope. Again we found cells that stained positive for VPs but not for AAP (figure 3.15). However, as soon as VPs accumulated in the nucleus they co-localized with AAP2. In contrast, AAV2 particles and AAP2 always co-localized (figure 3.16). Twenty-four hours after transfection both localizations were limited to the nucleus and expanded throughout the complete cell from 30 hours on. Interestingly, AAP showed a more dispersed phenotype once more particles were detected in the cytoplasm. In summary, AAP2 could be successfully detected with the generated anti-AAP2 antibody and showed different localizations in the cell during virus production.

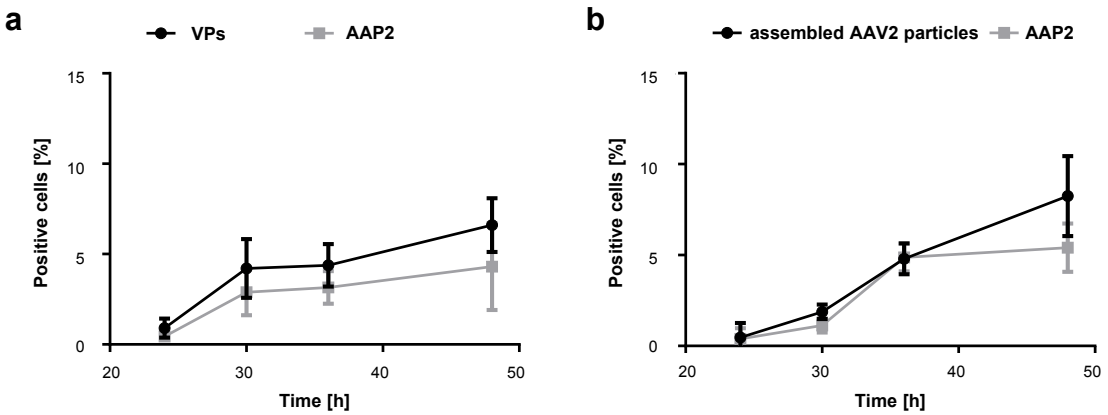

Figure 3.14: Quantification of AAP2-, VP- or virus particle-immunostainings over time. **a)** Quantification of VP- and AAP2-expressing HEK293T cells. **b)** Quantification of AAP2-expressing and AAV2 particle-containing HEK293T cells. Numbers of positive cells were determined 24, 30, 36 or 48 hours after transfection of an AAV2wt *rep-cap* helper, a transgene (luciferase) and the adenohelper construct in LabTek format. Antibodies B1 (a), A20 (b) and anti-AAP2 (a/b) were used for immunostaining. Four pictures per time point were analyzed with the CellProfiler program.

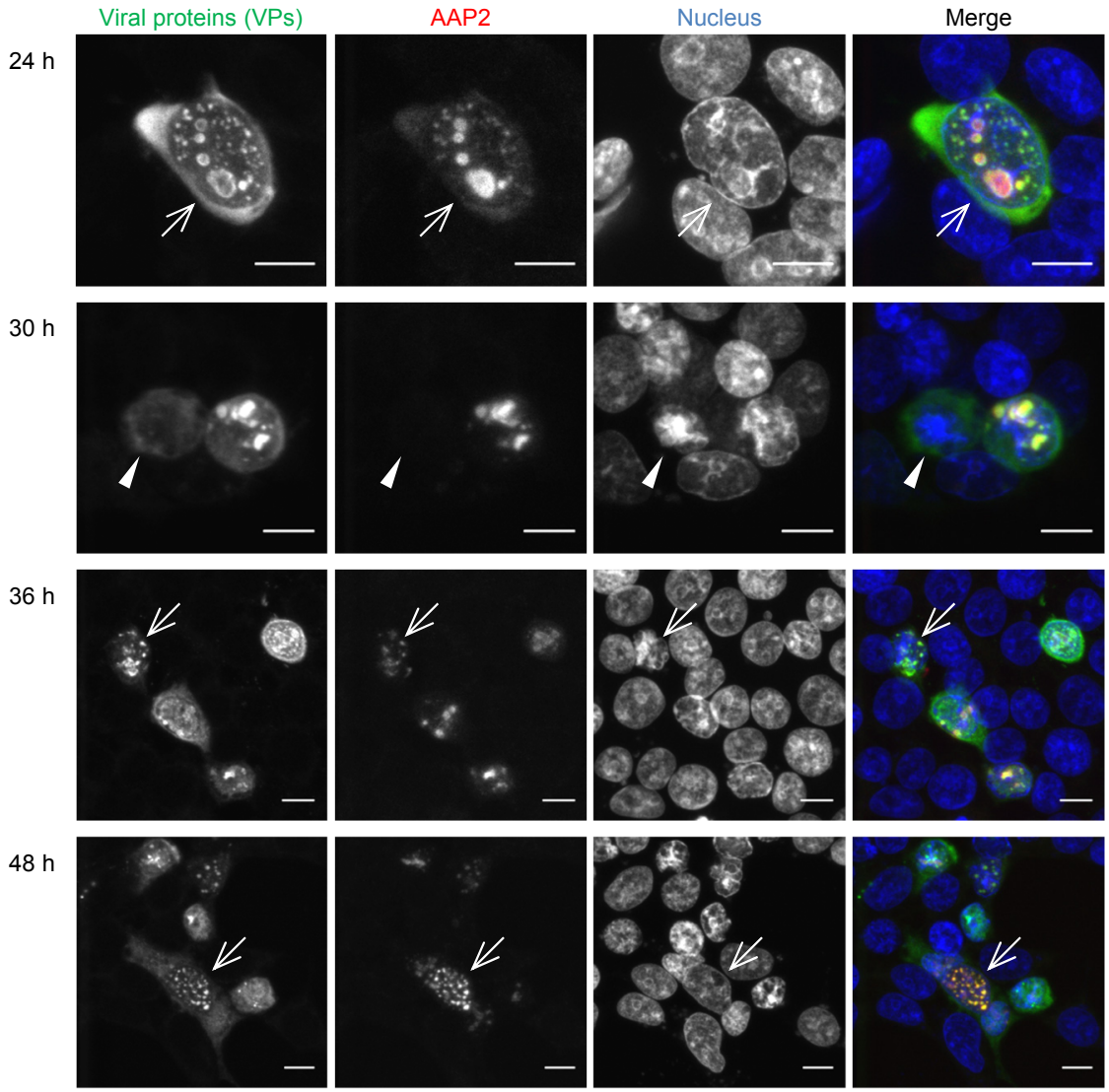

Figure 3.15: Immunostaining of viral proteins and AAP2 at different time points during AAV2 virus production. HEK293T cells were transfected with an AAV2wt *rep-cap* helper, a transgene (luciferase) as well as the adenohelper construct and fixed for immunostaining at the indicated time points. Arrows exemplify the nuclear localization of VPs only when AAP2 is expressed. The triangle in the second row depicts one cell without AAP2 and a distinct nuclear exclusion of VPs. Stainings were performed with the new anti-AAP2 antibody and the VP-detecting antibody B1. Transfections and immunostainings were performed in LabTek format. Pictures were captured with a TCS SP5 confocal microscope. Scale bar = 10 µm.

3.2.3 Localization of AAP4, 5, 8 and 9 differ from AAP2

Next to the intracellular localization of AAP2, we also examined the expression pattern of AAP4, 5, 8 and 9 in the cell. Due to the N-terminal fusion of all cloned AAPs to an HA-tag it was possible to detect the proteins with an anti-HA tag antibody via immunostaining. Analysis of the captured images revealed that localizations between the various serotype AAPs differ greatly from each other. AAP2 displayed a very distinct nucleolar localization

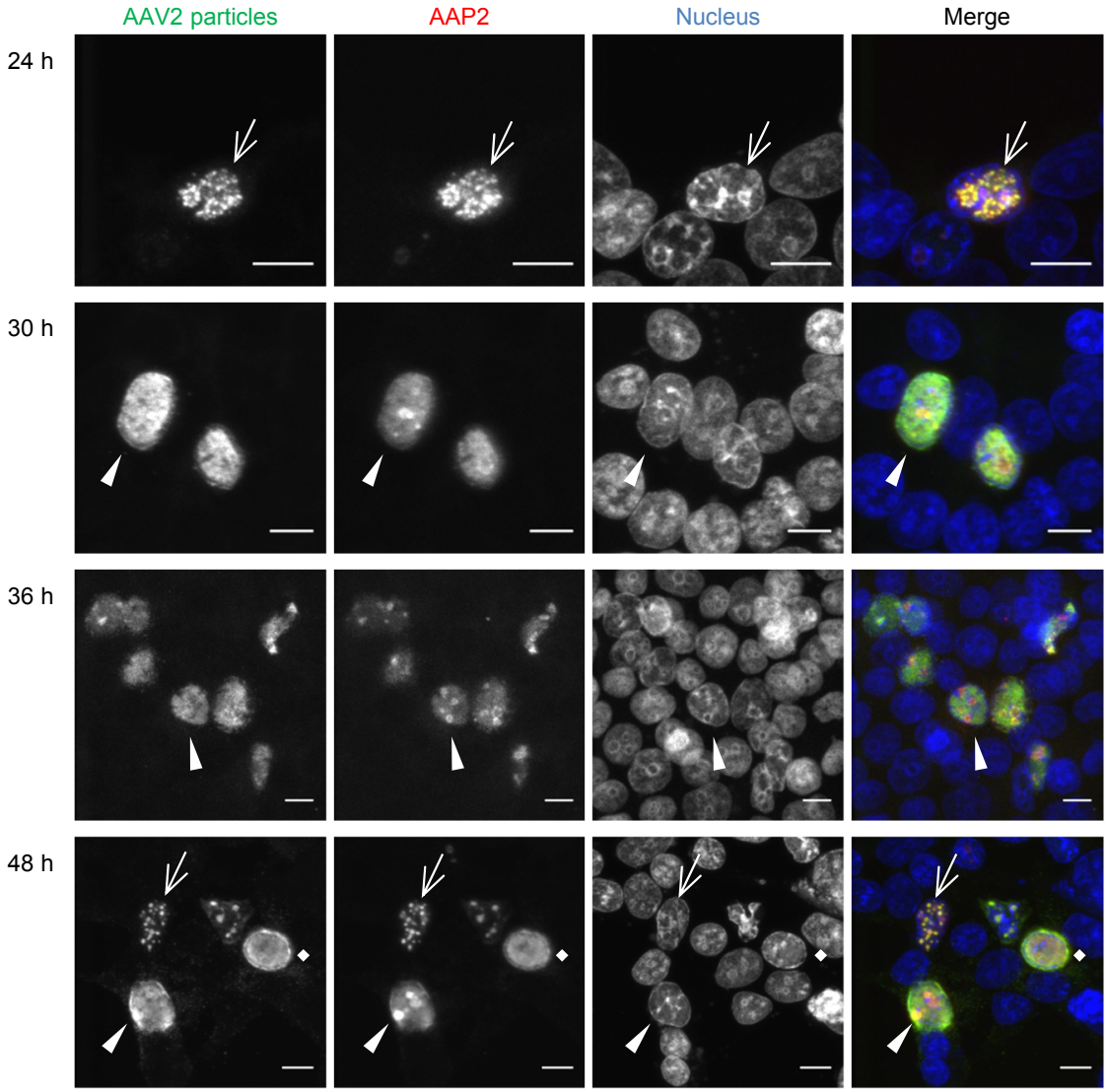

Figure 3.16: Immunostaining of assembled AAV2 particles and AAP2 at different time points during AAV2 virus production. HEK293T cells were transfected with an AAV2wt *rep-cap* helper, a transgene (luciferase) as well as the adenohelper construct and fixed for immunostaining at the indicated time points. Arrows depict the distinct nucleolar co-localization of AAP2 and assembled particles at early time points during virus production. Assembled viruses spread throughout the cell with ongoing time; in parallel AAP2 localized in the nucleoli as well as the nucleoplasm (triangle) or became fully excluded from the nucleoli (rhombus). Stainings were performed with the new anti-AAP2 antibody and the assembled AAV2 particle-detecting antibody A20. Nuclei were stained with Hoechst. Transfections and immunostainings were performed in LabTek format. Pictures were captured with a TCS SP5 confocal microscope. Scale bar = 10 µm.

(figure 3.17) which goes in line with our above described results and published observations by Sonntag *et al.* [25]. However, all other investigated AAPs showed different localizations. AAP8 and 9 were localized in the nucleus but did not accumulate in the nucleoli like AAP2. AAP4 resembles a nucleolar accumulation but was also detected in the cytoplasm. AAP5 appeared predominately only in the cytoplasm and differed the most from AAP2. In this regard,

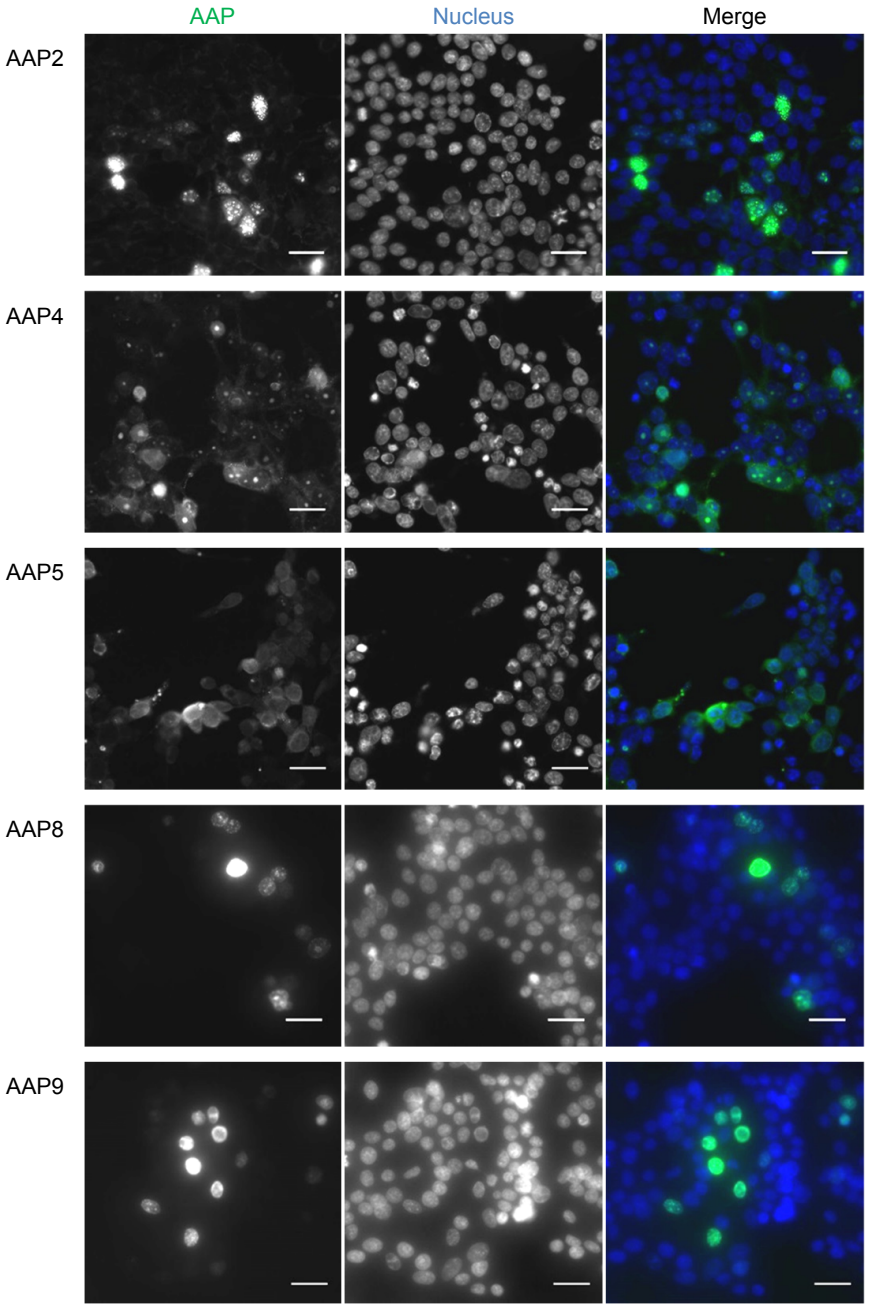

Figure 3.17: Immunostaining of AAPs of different AAV serotypes. AAPs of AAV serotypes 2, 4, 5, 8 and 9 were transfected together with a corresponding *rep-cap* helper, a transgene (luciferase) and the adenohelper construct into HEK293T cells and stained with an anti-HA tag antibody 48 hours after transfection. Transfections and immunostainings were performed in either LabTek (AAP2) or 96-well (AAP4, 5, 8 and 9) format. Scale bar = 30 µm.

more detailed investigations and stainings are planned to validate nucleolar localizations of e.g. AAP4 and gain further insights into the process of capsid assembly for the single AAV serotypes.

4 Discussion

4.1 Optimization of AAV *cap* DNA family shuffling on the technical and molecular level

4.1.1 Comparison of physical and enzymatic fragmentation reveals new prerequisites for successful DNA shuffling

Gene therapeutic approaches which treat genetic disorders like hemophilia, SCID or cystic fibrosis have recently shifted more and more into the focus of researchers and public attention. Due to the rapid gains in knowledge and techniques it is becoming increasingly feasible to treat inherited disorders but also to apply gene therapy to patients suffering from cancer or acquired neurological disorders. In this context, one of the most promising therapeutic delivery tools are AAVs. Their non-pathogenicity, simple vector design and low risk of integration make them very attractive for use in humans [1,2,5]. This was evidenced in 2012 when the first AAV-based gene therapeutic drug (Glybera®) was approved in the Western world [11]. The successful admission of Glybera® represents a milestone in AAV research and will certainly result in the positive evaluation of more clinical trials. Thus, an increasing need for new AAV serotypes with special characteristics for further therapeutic approaches is expected. The requirements for novel capsids are (I) less cross-reactivity to pre-existing or newly evolved antibodies, (II) a defined tissue tropism for new cell types and (III) concurrently a low transduction efficiency in non-target cells. In order to meet these demands, several methods for capsid modification have been published during the last years [133,150,158,173,175]. Amongst these, peptide display and DNA family shuffling (DFS) are the most efficient means to generate AAV particles with new properties [133–136,140,150,151].

This thesis focused on DFS as a technique for capsid modification which is based on recombination of partially homologous genes into chimeric sequences after controlled fragmentation. Within this work we aimed to provide an optimized protocol for DFS which can be easily adapted by other laboratories. Especially, we wanted to improve the critical step of DNA fragmentation which can influence the quality of the generated shuffled library. Normally, fragmentation is achieved by using the enzyme DNaseI. However, we know from our experience that DNase is very sensitive to external influences such as shearing forces and that different stocks can have varying activities. Moreover, we could show that small changes in the protocol (e.g. an increase of incubation time or enzyme amount) can already lead to

different fragment sizes (figure 3.1). This hampers the reproducibility of the approach and requires optimization experiments ahead of each shuffling reaction. To overcome these limitations, we considered implementing the method of ultrasonication-based DNA shearing in our shuffling protocol and aimed to compare it with DNase treatment side-by-side. Physical DNA fragmentation is a widely used method for e.g. next generation and ChIP-sequencing and produces DNA fragments with pre-defined sizes. Although the fragmentation process is not completely unbiased but rather seems to be sequence-dependent [202] it is a robust way to cleave DNA in a controlled manner. Depending on the machine settings, various lengths of DNA fragments can be generated, e.g. 150, 300 and 800 bp which were used here for shuffling. In contrast, DNase digest results in a broad range of DNA fragments (100 – 1000 bp) without a specific peak size. Interestingly, we observed that physically generated fragments of 150 or 300 bp led to no or inefficient recovery of full-length genes. Only 800 bp and DNase-treated samples were able to re-anneal efficiently. These observations go in line with a publication from Maheshri and Schaffer who predicted the re-assembly efficiency of AAV1 and 2 *cap* genes *in silico* [203]. They described a reduction in re-assembly depending on the average fragment size. One hundred bp fragments showed a more than 6-fold reduced re-assembly in comparison to 400 bp fragments. On the other hand, smaller fragments generated 4-fold more crossover events than bigger fragments. Similarly, our 800 bp library contained fewer crossover events and 1.45-fold longer fragments than the DNase-treated library (7.8 versus 12.1 crossovers). This indicates that smaller fragments generated by DNase digest act as primers during the re-assembly PCR and thereby mediate homologous recombination. However, at the same time bigger fragments seem to be needed for the re-assembly process since the exclusive use of small fragments (150 / 300 bp) failed to result in full-length AAV *cap* sequences. They potentially function as a template for the annealing reaction but can also recombine among each other in regions of overlapping homology. As a conclusion, it appears beneficial for the shuffling reaction to include a mixture of smaller and bigger DNA fragments. Indeed, also other AAV shuffling approaches used a mixture of fragments between 50 and 1200 bp for re-assembly PCRs [133–135]. Thus, one possibility to overcome the limited recombination capacity of long fragments generated by ultrasonication could be to mix 150 and 800 bp fragments prior to re-assembly in order to mimic enzymatic conditions. Future experiments should verify the feasibility and efficiency of this attempt in order to further optimize the DFS protocol. Until then, we conclude that DNase digest outperformed ultrasonication-based DNA shearing in terms of crossover efficiency in this work even if the physical method is less sensitive to external and experimental influences and hence more reproducible.

4.1.2 Extension of the Salanto toolbox facilitates analysis of shuffled libraries

For all analyses conducted during this study the previously published Salanto (Shuffling alignment analysis tool) program [195] was used and further extended by several additional tools in close collaboration with Christian Bender. Specifically, the newly implemented functions facilitate the crossover analysis within a generated library. It is important to note that

"Crossover analysis", "Show crossovercount" and "Fragment length analysis" are based on the "Type assignment". Importantly, the latter assigns a crossover at the first differing base between parent x or y whereas e.g. Coco *et al.* mark crossovers in the middle of a homologous region [204]. Next to an extended crossover analysis, the new feature "Show homology" allows for the position-wise comparison of chimeric and parental sequences regarding their homology. The tool "Show crossovercount" combines the position-wise homology analysis with the type assignment in order to enable a correlation of crossover appearance with identical nucleotides. Overall, to our best knowledge Salanto is the only published toolbox so far which is able to automatically analyze complex alignment files and to allow detailed conclusions about chimera composition, homology and recombination frequencies.

4.1.3 Sequence-optimization broadens shuffled AAV *cap* library diversity

As already mentioned before, crossover frequency correlates with the quality of a generated shuffled library and determines its diversity. Importantly, crossovers preferentially occur within regions with sequence identities of at least 5 – 15 nt [205]. Therefore, it is favorable for the shuffling process if parental sequences exhibit homologies greater than 65 %. With homologies below 50 % almost no recombination events can be obtained [206]. AAVs show homologies above 55 % which make them very suitable for shuffling approaches. However, sequence identities between capsids range from 56.6 – 97.1 % and result in different crossover rates. Previously, Joern *et al.* observed that similar parents tend to pair among each other [207] which reduces recombination between parental sequences of lower identity. Indeed, we obtained a correlation between crossover events and capsid homology ranging from 43 crossovers between AAV1 and AAV6 (97 % identical) to zero crossovers between AAV4 and AAV5 (59 % identical) in a 1-9 *cap* library. Thus, the incorporation of serotype 4 and 5 capsids into a shuffled library is limited which simultaneously lowers the theoretical library diversity.

To overcome this problem, we sequence-optimized AAV *cap* 4 and 5 by adapting their sequences to the one of AAV2 and achieved a homology improvement of 13 %. Comparison of a wild-type 245 *cap* library and its optimized counterpart revealed nearly 2.5-fold more crossovers per clone in the latter. Of note, detailed sequence analyses showed that crossovers in the 245O library appeared especially in regions with increased nucleotide identity. We additionally determined the incorporation of *cap* 4 and 5 into more complex libraries ranging from five to nine shuffled serotypes. All of them showed enhanced recombination rates between *cap* 2 and 4 / 5 as well as other serotypes with overall crossover numbers being less influenced in libraries with increasing complexity. Our data suggest that homology-dependent shuffling limitations can be alleviated by increasing sequence identities of individual capsids.

Interestingly, in parallel to a higher shuffling efficiency, we detected a $>$ 100-fold lower virus titer in library $245O_r$ in comparison to the wild-type counterpart. These observations indicate that improved recombination can be at the cost of virus fitness. Although Joern *et al.*

did not observe a correlation between loss-of-function and crossover numbers in their generated protein chimeras, they found clones with higher percentages of one parent to be more functional than chimeras of equal parental proportions [207]. Mechanistically, it is therefore likely that the interaction between dissimilar domains can hinder proper AAV capsid assembly, stability and / or genome packaging. Indeed, previous publications observed a partial impairment of AAV assembly and genome packaging upon domain swapping between different AAV serotypes. Especially, the surface exposed GH loop which encompasses VR-IV to VIII was substantially sensitive for specific modifications, i.e., AAV1-to-AAV2 or AAV2-to-AAV8 domain exchanges [208,209]. In addition, DiPrimio *et al.* reported an essential role for the HI loop in particle assembly and DNA packaging [210]. Although loopexchanges from AAV2 to AAV1 or AAV8, respectively, were tolerated, swapping to AAV4 and AAV5 resulted in a lower packaging efficiency and no capsid assembly for the latter [210]. The complex interaction of capsid surface regions is further illustrated by another publication from 2014. In this report rational aa substitutions in the VRs of AAV2 yielded only a low-titer library [211]. Detailed investigations could show that combinations of modified VR-VIII with VR-III or -VII do not result in AAV particles. Moreover, substitutions in VR-III, -VI, and -IX are less tolerated [211]. Notably, VR-III and -IX are located at the capsid order's two-fold symmetry axis which is assumed to be involved in capsid assembly [212]. Taken together, it seems plausible that combinations of low-homology capsids like 2, 4 and 5 can reduce the fitness of a chimeric AAV. Likewise, our data indicate that detrimental structural effects are less pronounced the more similar capsids are present in the library. Virus titers did not drop once serotypes other than 2, 4 or 5 were shuffled (i.e. *cap* library 24589, 145689). Of note, this can also indicate a "dilution" of capsids containing non-compatible *cap* fragments while more vital chimeras outgrow the others. Former shuffling approaches by various groups point in this direction as well [133–136,179,180]. Although all of them included at least *cap* 4 in their shuffled libraries, only two finally reported a functional chimera containing AAV4 regions after selection. Conspicuously, *cap* 4 parts preferentially appeared in the 5' region of the selected clones [134,136] which is located at the inner part of the assembled capsid and not involved in cell surface receptor binding but contributes in one case (AAV-LK03) to the PLA2 domain of VP1 [136]. Generally, the low frequency of AAV4 or 5 occurrence could originate from an overall low incorporation of *cap* 4 and 5 into the libraries but could also be a result of reduced fitness of AAV4/5-containing chimeras. Despite these potential hurdles, sequence-optimization of less homologous capsids and their improved integration into shuffled libraries represents a useful tool for future applications. Especially, it will enable to investigate AAV structure-function relationships in more detail and potentially yield new capsids with unique properties.

Next to efficient transduction, promising AAV vector candidates ideally also display a low immunogenicity. This is beneficial since pre-existing neutralizing antibodies can reduce the efficacy of applied AAVs. It is reported that 40 – 70 % of the human population already have neutralizing antibodies against specific AAV serotypes [79]. However, so far no publication has reported a prevalence of anti-AAV4 antibodies in the human population. A closely related

AAV *cap* variant named rh32.33 which was also isolated from NHP likewise lacked specific anti-capsid antibodies in humans [81]. Hence, AAV4 capsid chimeras are highly interesting for gene therapy in terms of reduced immunogenicity. Our approach opens new avenues for the generation of *cap* 4-containing vectors with new properties.

4.1.4 AAV DNA family shuffling is not restricted by inadvertent recombination of AAP

In 2010, an essential factor for AAV particle assembly called assembly-activating protein (AAP) was described [25]. Since then, it has been assumed that AAP acts as a scaffold or chaperone for VPs in order to mediate correct particle formation. In this context, the question arose whether shuffled AAP chimeras are impaired in functionality which, if true, would lead to limitations for AAV capsid shuffling. Importantly, we obtained evidence in this work that previous concerns may be unfounded. Out of our randomly tested and representative 46 shuffled AAPs, 39 (84.8 %) were functional. In this context, cross-complementation efficiencies of functional AAPs varied between 7 and 100 % indicating either different protein activities or limited interactions with the tested AAV2 knock-out mutant. Interestingly, libraries containing the capsids AAV4 and 5 comprised the highest amount of deficient AAPs. It is assumed that AAP interacts with as of yet unknown VP domains within the AAV capsid [213]. Comparison of AAP sequences additionally revealed a low homology between AAP4 / 5 and AAP1 – 3b or 6 – 10, respectively [213]. Therefore, it seems plausible that AAV4 and 5 VPs, which display the lowest homology between the common AAV serotypes, cannot interact with AAPs of other origin and *vice versa*. Indeed, two of the shuffled AAPs were only functional with our $AAV5_{O\_rep}$ construct and not compatible with AAV2mut. Additionally, only AAP4 and 5 failed to support AAV2 assembly in a *trans*-complementation assay with our generated AAV2mut construct whereas all other AAPs were interchangeable. These results are in accordance with a publication of Sonntag *et al.* who already reported successful cross-complementation between AAV-AAP1, 2, 8 and 9 but not AAP5 [213]. Here, we further extended this data set with AAV-AAP3b, 4, 6, 7 and rh10. Except for AAP4 all tested AAPs were able to stimulate AAV2 assembly to various degrees.

The above-described data prompted us to investigate whether AAPwt *trans*-complementation during virus production can overcome the serotype specific limitations in case AAV4 and / or 5 are present in a shuffled library. Interestingly, we could not detect an increase in virus titers of our 24589 libraries in the presence of an AAPwt mixture. One explanation could be that cross-complementation of AAPs encoded within the library already counteracts restrictive effects during production. Thus, further AAP addition cannot enhance assembly to a higher extent. On the other hand, non-functional shuffled 24589O capsid chimeras with inactive AAP could not be recovered after AAP addition. The latter data point more into the direction that the function of AAP itself is not the restrictive factor in shuffling. The high functionality of our individually tested AAPs further supports this assumption. We rather hypothesize that the AAP binding site within the AAV capsids is influenced by the shuffling process. To

date, published data only reported that the VP C terminus interacts with AAP and is critical for the assembly process [212]. Especially the point mutations I682S and K688A/K692A abolished VP-AAP interaction and blocked capsid formation [212]. Furthermore, other mutations affecting capsid assembly indicate an interplay of AAP with VP domains at the capsid's two-fold symmetry axis [64, 141, 212]. It is therefore likely that AAP binding could be impaired due to unfavorable VP domain combinations of incompatible AAV serotypes during capsid shuffling which results in assembly-deficient particles. On the other hand, also sterical hindrances within the capsid can potentially lead to non-functional particles independent of AAP. However, neither of these possibilities can be circumvented during AAV capsid shuffling. Conversely, stringent selection of shuffled *cap* libraries yielded functional AAV particles with intact AAP sequence. For instance, AAV-DJ encodes a shuffled AAP which encompasses AAV serotypes 2-9-2-9-8 whereas AAV-LK03 contains only AAP3b [133, 136]. Hence, natural sifting during virus production and further selection enriches functional AAV particles with intact AAPs.

Apart from the generation of novel AAV capsids, DFS can also be exploited to gain more insights into structure-function relationships of proteins. In 2013 work from our group demonstrated the power of DFS by discovering a function-related motif in the human Argonaute protein 2 [195]. There, the shuffling of Argonaute 2 and 3 resulted in the identification of slicing-competent domains. Intriguingly, the described motif contains two small important regions separated by 85 aa and was thus unlikely to be identified through normal site-directed mutagenesis. Knowing that DFS is able to elucidate structure-function relationships, AAP shuffling has the potential to shed light on (I) important AAP domains which are necessary for capsid assembly and (II) the interplay between AAP and VPs in order to resolve their exact binding sites.

4.2 New knock-out model system enables novel insights into AAV biology

4.2.1 AAP stabilizes VPs and prevents proteasomal degradation

After the discovery of the assembly-activating protein AAP, only a few studies have further investigated its role as scaffold or chaperone for AAV particle formation [25, 212, 213]. Thus, little is known about general AAP biology and its influence on other steps of the AAV life cycle. Hence, we aimed to shed more light on the features of AAP and its importance for AAV.

Up to now all previous AAP studies obtained insights into AAP biology through a VP3-based model system [25, 212, 213]. There, instead of a full-length *cap* gene, a VP3-only expression vector driven by a CMV promoter was used. This construct was based on the observation that VP3 subunits are able to form non-infectious virus particles which are morphologically indistinguishable from the wild-type capsids [168, 214]. Notably, these results were obtained

in the presence of AAP whose sequence starts 80 nt upstream of the VP3 start codon and was not affected by introduced mutations. However, once these 80 nt are deleted virus-like particles can only be formed in the presence of AAP supplemented *in trans* [25]. Despite the simplicity and usefulness of the VP3 model system, it neither reflects (I) the endogenous VP expression due to the used CMV promoter nor (II) the interplay between AAP and VPs. Sonntag *et al.* performed only one experiment with the full-length, AAP-deleted AAV genome at the beginning of their study [25]. With our newly developed full-length AAV model system we can overcome the abovementioned limitations and were able to study AAP biology in more detail. Of note, we not only modified the AAV2 genome but also generated AAP-lacking AAV1, 3b, 6, 7, 8, 9 and rh10 constructs. Therefore, we deleted the AAP start codon CTG and inserted a stop codon 66 nt upstream of the AAP start. Both mutations did not change the VP1-3 sequence but were able to completely abolish formation of infectious virus particles. In line with previous publications, AAP supplementation could recover virus formation for all tested serotypes which proved the general applicability of the system to other AAV capsids. Interestingly, the cross-complementation and recovery of infectious particles was dose-dependent for the tested AAV2mut construct suggesting that AAP is a limiting factor during virus assembly. Furthermore, we unexpectedly detected a drastic reduction of AAV2 VP levels in cell lysates lacking AAP2 after Western blotting. Intriguingly, Sonntag *et al.* also observed an alteration of the VP expression pattern with their full-length mutant in the absence of AAP but did not follow up on the phenomenon [25]. We could show in this work that the loss of VP signal is not caused by an alteration of mRNA levels. Moreover, we saw no significant differences in VP levels of the AAV2mut construct in comparison to AAV2wt in microscopy. Thus, we concluded that the decline of VP signal in Western blots is caused by a reduction of assembled particles. At the same time, VPs do not seem to accumulate in the cell but rather are degraded in the absence of virus assembly. Indeed, after addition of the proteasome inhibitor MG-132 AAV2mut protein amounts became comparable to AAV2wt indicating a post-translational modification of VPs and subsequent degradation. Strong additional support for this conclusion was the detection of an additional, VP-specific band of higher molecular weight in AAV2mut-treated cells only after MG-132 treatment. These results led us to postulate a potential ubiquitination of *cap* proteins which is congruent with previous publications. Generally, ubiquitination of proteins is an important mechanism within the cell to e.g. target proteins for degradation or change their intracellular localization [215,216]. This process requires several steps which include (a) an activation of the ubiquitin molecule by the enzyme E1, (b) the transfer to the E2 enzyme and (c) its cross-linking to the final target protein through the ubiquitin ligase E3. Finally, ubiquitinated proteins can be degraded by the 26S proteasome complex [215,216]. AAVs were shown to interact with ubiquitin for the first time in 2000, when Duan *et al.* reported an increase of ubiquitinated AAV VP proteins after transduction in human airway epithelia cells treated with proteasome inhibitor LLnL [31]. Two years later the same group confirmed these results and further described *in vitro* ubiquitination studies with AAV2 and AAV5 particles [34]. Interestingly, heat-denaturated capsids were

better substrates for the conjugation of ubiquitin than intact viruses. Additionally, several publications reported an increase of virus transduction after elimination of tyrosine residues at the capsid surface [217–219] which were shown to be phosphorylated by the epidermal growth factor receptor protein tyrosine kinase (EGFR-PTK) and thereby mediate ubiquitination of the viral capsid [220]. Moreover, Nayak *et al.* described a ubiquitin-dependent proteasomal degradation of transfected AAV5 VPs mediated by the adenovirus 5 E4Orf6 protein [221]. Altogether, these data support our hypothesis that non-assembled VPs undergo post-translational modifications leading to degradation. However, we always detected only one or two very distinct bands of $\sim$ 200 kDa after Western blot analysis of MG-132-treated cells. This is contradictory to patterns of protein ubiquitination normally observed which are characterized by a protein smear. Intriguingly, the pattern of our described bands resembles the denaturated (10 min, 95°C) AAV particles reported by Yan *et al.* [34]. One can assume that due to the absence of AAP in the cells and inhibited proteasomal degradation, VPs aggregate similar to the heat treatment leading to these high-molecular weight clusters that are resistant to denaturing conditions in SDS-PAGE. Since only the heat-denaturated particles were shown to be substrates for ubiquitination [34] also aggregated cellular VPs could be potential targets for ubiquitin. Furthermore, preliminary Western blot results of a collaboration partner showed a similar VP staining which argues against an artifact. In contrast to our system, M. Penaud-Budloo in the group of E. Ayuso tested the aforementioned AAV2mut modifications in a baculovirus-based AAV production system. In line with our results, they observed a reduction of VP protein levels in Western blot analysis and no assembly of infectious particles in the absence of AAP2. Additionally, AAV2mut-expressing insect cells showed not only the high-molecular band we detected but also a smear of VP proteins with increased size (unpublished data, M. Penaud-Budloo & E. Ayuso, IRS, Université de Nantes). The latter additionally supports our hypothesis of post-translational VP modifications. Yet, we were not able to pull down ubiquitinated VPs in various IP attempts so far. Although this could be a result of insufficient protein amounts and problems with the IP *per se*, it could also hint towards a ubiquitin-independent proteasomal degradation or other post-translational modifications apart from ubiquitin. Recently, a siRNA screen associated the small ubiquitin-related modifier (SUMO) with AAVs [222]. SUMO belongs to the group of ubiquitin-like modifiers and alters proteins similar to ubiquitin [223]. Instead of ubiquitination, SUMOylated proteins are normally not degraded but play important roles in transcription, DNA repair or signal transduction [215]. Nevertheless, SUMOylation can also target proteins for degradation. Holscher *et al.* observed that a knock-down of several members of the SUMO pathway led to an increased AAV transduction efficiency [222]. However, these results were obtained for incoming virus particles after transduction and could not dissect at which step of intracellular vector processing SUMO is relevant. Hence, it remains unclear whether VPs interact with SUMO before virus assembly since no publication has thus far determined SUMOylation or ubiquitination of non-assembled VPs during virus production.

4.2.2 Subcellular localization of AAP2 changes during virus production

AAP2 has been shown to contain several nuclear (NLS) and nucleolar localization signals (NoLS) which efficiently translocate AAP2 to the nucleoli [58]. Mutations in these regions lead to abnormal localization patterns and influence AAV capsid production. Thus, nucleolar AAP2 plays an important role for virus assembly. However, all publications only investigated the localization of transfected AAP2 driven by a CMV promoter. It remained unknown whether localization of endogenous AAP shows the same pattern and expression levels. Thus, we aimed to develop an anti-AAP2 antibody which enables the detection of endogenous AAP2 in the context of virus production. In parallel to AAP2, we also determined cellular expression of either free non-assembled VPs or assembled AAV2 particles. Microscopic analysis showed that endogenous AAP2 localized, similar to CMV promoter-driven AAP2 [25, 58], in the nucleoli of the cells. Intriguingly, we observed an increasing de-localization from the nucleoli as soon as viral particles were produced. Furthermore, quantification of the microscopy pictures revealed a decrease of AAP2-positive cells at later time-points (48 h) after transfection. These data are in accordance with reduced AAP2 protein levels in Western blot analysis at later time points of AAV virus production in insect cells (unpublished data, M. Penaud-Budloo & E. Ayuso, IRS, Université de Nantes). Our observations indicate (I) that AAP is used up during the assembly process and cannot be recycled and / or (II) that the nuclear localization signals are surpassed by other molecular interactions. In this context, it could be possible that AAP2 remains associated with assembled virus particles since both signals always co-localized. Yet, this has to be further investigated before drawing a conclusion.

In contrast to AAV2 particles, free VPs did not always co-localize with AAP. In detail, we obtained cells which only expressed VPs but not AAP2. Quantitative analysis confirmed this observation and revealed more VP-positive cells than AAP-positive cells for all analyzed time points after transfection. Notably, in the absence of AAP, VP signals mainly localized to the cytoplasm. VPs only appeared in the nucleoli when cells were also positive for AAP2. Sonntag *et al.* likewise reported a nucleolar exclusion of AAV2 VP3 without AAP which turned into nucleolar accumulation and virus assembly after expression of the protein [25]. In this context, they also showed that nucleolar translocation of VPs fused to a NoLS is not sufficient for assembly but requires AAP.

Collectively, our AAV2 protein stainings are in line with previously published time courses of VP expression and particle assembly [57]. Furthermore, our data expand the knowledge about endogenous AAP expression and can be easily combined with existing data. Thus, our proposed model for the time course of AAV protein expression (see also figure 4.1) encompasses a) the translation of AAV2 VPs 24 early after transfection followed by b) the onswitch of AAP2 expression which leads to a translocation of VPs to the nucleoli due to the nucleolar localization of AAP. Next, c) virus particles assemble in the nucleoli when AAP2 is present and d) spread throughout the cell with time. The latter includes a decrease of AAP2 protein levels and a translocation of the protein from the nucleoli.

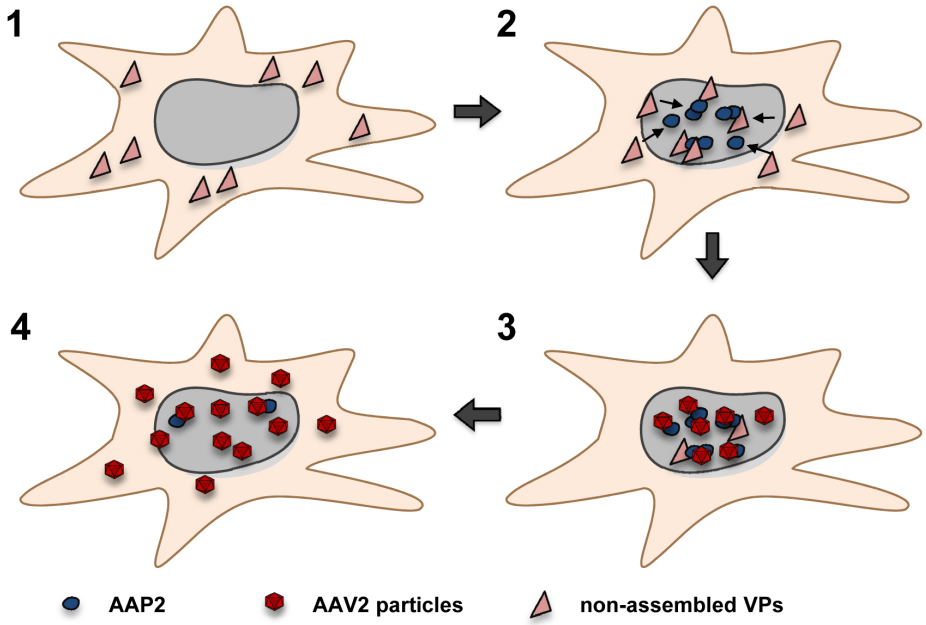

Figure 4.1: AAV2 protein expression model. Schematic depiction of the different steps of AAV2 protein expression during virus production. First (1), VPs are expressed and (2) subsequently translocate to the nucleoli once nucleolar AAP2 is made by the cell. (3) Virus particles assemble in the nucleoli and co-localize with AAP2. Finally (4), newly assembled particles spread throughout the cell. In parallel, AAP2 protein levels are reduced and de-localize from the nucleoli.

4.2.3 Subcellular localization of AAP2 differs from other serotypes

Next to the localization of AAP2, we also assessed intracellular AAP expression of other AAV serotypes. In detail, we compared AAP2, 4, 5, 8 and 9 patterns after transfection of the corresponding CMV-AAP expression plasmids together with other constructs needed for virus production. Since our generated anti-AAP2 antibody showed only minimal cross-reactivity with AAP8 and 9, and none with AAP4 or 5 (data not shown), we used an anti-HA antibody for all AAP stainings. Interestingly, only AAP2 showed the nucleolar localization above described whereas all other investigated serotypes behaved differently. AAP8 and 9 were not exclusively detected in the nucleoli but also in the nucleoplasm. AAP4 exhibited mainly cytoplasmic localization but also seemed to accumulate in the nucleoli. In contrast to all other AAPs, AAP5 showed no nucleolar localization but appeared in the nucleoplasm and cytoplasm. Since all AAPs performed well in our conducted *trans*-complementation assays, we could exclude an impairment of functionality by the varying localizations. Thus, we hypothesize that nucleolar localization of AAP is not a pre-requisite for successful virus assembly and is rather dependent on the AAV serotype. Our observations go in line with an abstract presented at the Annual meeting of the American Society of Gene and Cell Therapy in 2015 where the authors also reported different localizations of different AAPs [224]. Notably, Earley *et al.* speculated in their publication about the dependency of nucleolar localization on the electric charge of basic regions within AAP [58]. They determined the electric charge of

AAP2 to be above 12 which seems to be required for nucleolar localization. In contrast, with a value of 8, AAP5 showed the lowest charge of all analyzed AAPs and led the authors to predict a nucleolar exclusion of AAP5 [58]. Indeed, we could confirm their hypothesis and detected AAP5 only in the cytoplasm and nucleoplasm. Mechanistically, it is therefore likely that AAV assembly does not exclusively appear in the nucleoli but at the site of AAP localization depending on the AAV serotype. In turn, this would mean (and was already indicated by Earley *et al.*) that no nucleolar factors are necessarily needed for assembly and nucleolar localization is rather a side-effect of the AAP sequence than actual function. However, these are speculations and need further validation by future experiments.

4.3 Future perspectives

Within this work we provided several options for the optimization of AAV *cap* DNA family shuffling and discussed the influence of AAP on the method. Concurrently, we raised a series of important new questions which should be addressed in the future.
First, although Covaris fragmentation is less sensitive to external and experimental influences and hence more reproducible, DNaseI digest yielded better shuffling efficiencies and crossover numbers within a generated library. Thus, we recommend to further optimize ultrasonication-based DNA shearing and test a mixture of small and big DNA fragments in order to mimic enzymatic fragmentation. Successful implementation will increase the adaptability of the method for other labs and, moreover, allow for a deliberate fine-tuning of a shuffled library by mixing different DNA input sizes.
Second, our approach of sequence optimization of less homologous *cap* DNA will contribute to an increasing diversification of AAV libraries. In this context, optimization of other AAV serotypes, which are genetically distinct to the commonly used serotypes (i.e. AAV2, 8 or 9), offers promising options. In particular, the incorporation of AAVs of bovine, porcine, avian and caprine origin could potentially yield new AAV variants with improved immunogenicity since they should not exhibit any prevalence for pre-existing antibodies in the human population. Another possibility to increase library diversity would be the partial sequence optimization of *cap* genes in regions of interest followed by the addition of only these parts in the shuffling reaction. This would minimize risks for potential negative effects of recombination of non-compatible VP domains.
Third, optimized DFS cannot only yield new AAV capsids with ameliorated properties but could also contribute to the elucidation of AAP functions. Considering the knowledge about AAP remains very limited, DFS would allow for a more detailed investigation of AAP-VP binding sites and / or the study of single AAP domains regarding their importance for functionality. Since our analyzed shuffled AAPs showed varying efficiencies in the conducted cross-complementation assays, we believe that they are especially suitable for initial structure-function relation approaches. Additionally, our new AAP knock-out model system with the generated mutants enables-serotype specific interaction studies between AAP and AAV *cap*

proteins. Furthermore, microscopic analysis of these mutants will provide more details on the localization of different AAPs and can shed light on the question whether nucleolar localization of AAPs is a pre-requisite for AAV capsid assembly which is, surprisingly, not supported by our results. Finally, our reported data raised new questions about the causes of AAP protein level reductions after AAV capsid assembly and the type of post-translational modification of VPs in the absence of AAP which have to be investigated in more detail in the future.

Publications and presentations

1 Publications

Grosse, S., Herrmann, A.K., Bender, C., Kienle, E., Wiedtke, E., Kreuz, S., Schürmann, N. & Grimm, D. A robust and all-inclusive pipeline for directed molecular evolution of Adeno-associated viral (AAV) gene transfer vectors. (Manuscript in preparation).

Senís, E., Mockenhaupt, S., Rupp, D., Bauer, T., Paramasivam, N., Knapp, B., Gronych, J., **Grosse, S.**, Windisch, M.P., Theis, F.J., Eils, R., Lichter, P., Schlesner, M., Bartenschlager, R. & Grimm, D. TALEN/CRISPR-mediated engineering of a promoterless anti-viral RNAi hairpin into an endogenous miRNA locus. (Manuscript in preparation).

Michler, T., **Grosse, S.**, Mockenhaupt, S., Röder, N., Stückler, F., Knapp, B., Heikenwälder, M. & Protzer, U. Sense strand neutralisation improves potency, safety and specificity of anti-hepatitis B virus short hairpin RNA. *EMBO Mol Med* (In revision).

Liesche, C., Venkatraman, L., Aschenbrenner, S., **Grosse, S.**, Grimm, D., Eils, R. & Beaudouin, J. Death receptor-based enrichment of Cas9-expressing cells. *BMC Biotechnol* **16** (2016). [225]

Mockenhaupt, S., **Grosse, S.**, Rupp, D., Bartenschlager, R. & Grimm, D. Alleviation of off-target effects from vector-encoded shRNAs via codelivered RNA decoys. *Proc Natl Acad Sci* **112**, E4007–16 (2015). [226]

Senís, E.[1], Fatouros, C.[1], **Grosse, S.**[1], Wiedtke, E., Niopek, D., Mueller, A.K., Börner, K. & Grimm D. CRISPR/Cas9-mediated genome engineering: An adeno-associated viral (AAV) vector toolbox. *Biotechnol J* **9**, 1402–1412 (2014). [227]

Kienle, E.[1], Senís, E.[1], Börner, K., Niopek, D., **Grosse, S.**, Wiedtke, E. & Grimm D. Engineering and evolution of synthetic adeno-associated virus (AAV) gene therapy vectors via DNA family shuffling. *J Vis Exp* (2012). [177]

[1] equal contribution

2 Presentations

Grosse, S. & Grimm, D. New serotype-specific functions of AAV-AAP for capsid protein expression and vector particle production. Poster. *Annual meeting of the American Society of Gene and Cell Therapy (ASGCT)*, Washington DC, USA (2014).

Grosse, S., Bender, C., Kienle, E. & Grimm, D. An advanced toolbox for molecular AAV evolution through DNA family shuffling. Poster. *Annual meeting of the European Society of Gene and Cell Therapy (ESGCT)*, Madrid, Spain (2013).

References

[1] Kay, M. A. State-of-the-art gene-based therapies: the road ahead. *Nat Rev Genet* **12**, 316–328 (2011).

[2] Giacca, M. & Zacchigna, S. Virus-mediated gene delivery for human gene therapy. *J Control Release* **161**, 377–388 (2012).

[3] Nayerossadat, N., Maedeh, T. & Ali, P. A. Viral and nonviral delivery systems for gene delivery. *Adv Biomed Res* **1** (2012).

[4] The Journal of Gene Medicine. Gene Therapy Clinical Trials Worldwide (2015). URL `http://www.wiley.com/legacy/wileychi/genmed/clinical/`.

[5] Thomas, C. E., Ehrhardt, A. & Kay, M. a. Progress and problems with the use of viral vectors for gene therapy. *Nat Rev Genet* **4**, 346–358 (2003).

[6] Cavazzana-Calvo, M. *et al.* Gene therapy of human severe combined immunodeficiency (SCID)-X1 disease. *Science* **288**, 669–672 (2000).

[7] Aiuti, A. *et al.* Correction of ADA-SCID by stem cell gene therapy combined with nonmyeloablative conditioning. *Science* **296**, 2410–2413 (2002).

[8] Hacein-Bey-Abina, S. *et al.* A serious adverse event after successful gene therapy for X-linked severe combined immunodeficiency. *N Engl J Med* **348**, 255–256 (2003).

[9] Hacein-Bey-Abina, S. *et al.* LMO2-associated clonal T cell proliferation in two patients after gene therapy for SCID-X1. *Science* **302**, 415–419 (2003).

[10] Lehrman, S. Virus treatment questioned after gene therapy death. *Nature* **401**, 517–518 (1999).

[11] Yla-Herttuala, S. Endgame: Glybera finally recommended for approval as the first gene therapy drug in the European Union. *Mol Ther* **20**, 1831–1832 (2012).

[12] Kaufmann, K. B., Büning, H., Galy, A., Schambach, A. & Grez, M. Gene therapy on the move. *EMBO Mol Med* **5**, 1642–1661 (2013).

[13] Atchison, R. W., Casto, B. C. & Hammon, W. M. Adenovirus-Associated Virus Defective Virus Particles. *Science* **149**, 754–755 (1965).

[14] Mishra, L. & Rose, J. A. Adeno-associated virus DNA replication is induced by genes that are essential for HSV-1 DNA synthesis. *Virology* **179**, 632–639 (1990).

[15] Weindler, F. W. & Heilbronn, R. A subset of herpes simplex virus replication genes provides helper functions for productive adeno-associated virus replication. *J Virol* **65**, 2476–2483 (1991).

[16] You, H. *et al.* Multiple human papillomavirus genes affect the adeno-associated virus life cycle. *Virology* **344**, 532–540 (2006).

[17] Koczot, F. J., Carter, B. J., Garon, C. F. & Rose, J. A. Self-complementarity of terminal sequences within plus or minus strands of adenovirus-associated virus DNA. *Proc Natl Acad Sci USA* **70**, 215–219 (1973).

[18] Berns, K. I. Parvovirus replication. *Microbiol Rev* **54**, 316–329 (1990).

[19] Flotte, T. R. *et al.* Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter. *J Biol Chem* **268**, 3781–3790 (1993).

[20] Haberman, R. P., McCown, T. J. & Samulski, R. J. Novel Transcriptional Regulatory Signals in the Adeno-Associated Virus Terminal Repeat A/D Junction Element. *J Virol* **74**, 8732–8739 (2000).

[21] Qiu, J. & Pintel, D. Processing of adeno-associated virus RNA. *Front Biosci* **13**, 3101–3115 (2008).

[22] Mendelson, E., Trempe, J. P. & Carter, B. J. Identification of the trans-acting Rep proteins of adeno-associated virus by antibodies to a synthetic oligopeptide. *J Virol* **60**, 823–832 (1986).

[23] Becerra, S. P., Koczot, F., Fabisch, P. & Rose, J. a. Synthesis of adeno-associated virus structural proteins requires both alternative mRNA splicing and alternative initiations from a single transcript. *J Virol* **62**, 2745–2754 (1988).

[24] Trempe, J. P. & Carter, B. J. Alternate mRNA splicing is required for synthesis of adeno-associated virus VP1 capsid protein. *J Virol* **62**, 3356–3363 (1988).

[25] Sonntag, F., Schmidt, K. & Kleinschmidt, J. A. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc Natl Acad Sci USA* **107**, 10220–10225 (2010).

[26] Duan, D. *et al.* Dynamin is required for recombinant adeno-associated virus type 2 infection. *J Virol* **73**, 10371–10376 (1999).

[27] Bartlett, J. S., Wilcher, R. & Samulski, R. J. Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J Virol* **74**, 2777–2785 (2000).

[28] Nonnenmacher, M. & Weber, T. Adeno-associated virus 2 infection requires endocytosis through the CLIC/GEEC pathway. *Cell Host Microbe* **10**, 563–576 (2011).

[29] Bantel-Schaal, U., Braspenning-Wesch, I. & Kartenbeck, J. Adeno-associated virus type 5 exploits two different entry pathways in human embryo fibroblasts. *J Gen Virol* **90**, 317–322 (2009).

[30] Nonnenmacher, M. & Weber, T. Intracellular transport of recombinant adeno-associated virus vectors. *Gene Ther* **19**, 649–658 (2012).

[31] Duan, D., Yue, Y., Yan, Z., Yang, J. & Engelhardt, J. F. Endosomal processing limits gene transfer to polarized airway epithelia by adeno-associated virus. *J Clin Invest* **105**, 1573–1587 (2000).

[32] Douar, A. M., Poulard, K., Stockholm, D. & Danos, O. Intracellular trafficking of adeno-associated virus vectors: routing to the late endosomal compartment and proteasome degradation. *J Virol* **75**, 1824–1833 (2001).

[33] Denby, L., Nicklin, S. A. & Baker, A. H. Adeno-associated virus (AAV)-7 and -8 poorly transduce vascular endothelial cells and are sensitive to proteasomal degradation. *Gene Ther* **12**, 1534–1538 (2005).

[34] Yan, Z. *et al.* Ubiquitination of both Adeno-Associated Virus Type 2 and 5 Capsid Proteins Affects the Transduction Efficiency of Recombinant Vectors. *J Virol* **76**, 2043–2053 (2002).

[35] Zhong, L. *et al.* A Dual Role of EGFR Protein Tyrosine Kinase Signaling in Ubiquitination of AAV2 Capsids and Viral Second-strand DNA Synthesis. *Mol Ther* **15**, 1323–1330 (2007).

[36] Ding, W., Zhang, L. N., Yeaman, C. & Engelhardt, J. F. rAAV2 traffics through both the late and the recycling endosomes in a dose-dependent fashion. *Mol Ther* **13**, 671–682 (2006).

[37] Sonntag, F., Bleker, S., Leuchs, B., Fischer, R. & Kleinschmidt, J. A. 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**, 11040–11054 (2006).

[38] Girod, A. *et al.* 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 (2002).

[39] Grieger, J. C., Johnson, J. S., Gurda-Whitaker, B., Agbandje-McKenna, M. & Samulski, R. J. Surface-exposed adeno-associated virus Vp1-NLS capsid fusion protein rescues infectivity of noninfectious wild-type Vp2/Vp3 and Vp3-only capsids but not that of fivefold pore mutant virions. *J Virol* **81**, 7833–7843 (2007).

[40] Stahnke, S. *et al.* Intrinsic phospholipase A2 activity of adeno-associated virus is involved in endosomal escape of incoming particles. *Virology* **409**, 77–83 (2011).

[41] Bantel-Schaal, U., Hub, B. & Kartenbeck, J. Endocytosis of adeno-associated virus type 5 leads to accumulation of virus particles in the Golgi compartment. *J Virol* **76**, 2340–2349 (2002).

[42] Pajusola, K. *et al.* Cell-type-specific characteristics modulate the transduction efficiency of adeno-associated virus type 2 and restrain infection of endothelial cells. *J Virol* **76**, 11530–11540 (2002).

[43] Johnson, J. S. *et al.* Mutagenesis of adeno-associated virus type 2 capsid protein VP1 uncovers new roles for basic amino acids in trafficking and cell-specific transduction. *J Virol* **84**, 8888–8902 (2010).

[44] Grieger, J. C., Snowdy, S. & Samulski, R. J. Separate basic region motifs within the adeno-associated virus capsid proteins are essential for infectivity and assembly. *J Virol* **80**, 5199–5210 (2006).

[45] Johnson, J. S. & Samulski, R. J. Enhancement of Adeno-Associated Virus Infection by Mobilizing Capsids into and Out of the Nucleolus. *J Virol* **83**, 2632–2644 (2009).

[46] Kelich, J. M. *et al.* Super-resolution imaging of nuclear import of adeno-associated virus in live cells. *Mol Ther Methods Clin Dev* **2**, 15047 (2015).

[47] Cotmore, S. F., D'abramo, A. M. J., Ticknor, C. M. & Tattersall, P. Controlled conformational transitions in the MVM virion expose the VP1 N-terminus and viral genome without particle disassembly. *Virology* **254**, 169–181 (1999).

[48] Ros, C., Baltzer, C., Mani, B. & Kempf, C. Parvovirus uncoating in vitro reveals a mechanism of DNA release without capsid disassembly and striking differences in encapsidated DNA stability. *Virology* **345**, 137–147 (2006).

[49] Nakai, H., Storm, T. A. & Kay, M. A. Recruitment of single-stranded recombinant adeno-associated virus vector genomes and intermolecular recombination are responsible for stable transduction of liver in vivo. *J Virol* **74**, 9451–9463 (2000).

[50] Gonçalves, M. a. F. V. Adeno-associated virus: from defective virus to effective vector. *Virol J* **2**, 43 (2005).

[51] Chandler, M. *et al.* Breaking and joining single-stranded DNA: the HUH endonuclease superfamily. *Nat Rev Micro* **11**, 525–538 (2013).

[52] McCarty, D. M. *et al.* Adeno-associated virus terminal repeat (TR) mutant generates self-complementary vectors to overcome the rate-limiting step to transduction in vivo. *Gene Ther* **10**, 2112–2118 (2003).

[53] Wang, Z. *et al.* Rapid and highly efficient transduction by double-stranded adeno-associated virus vectors in vitro and in vivo. *Gene Ther* **10**, 2105–2111 (2003).

[54] Kotin, R. M. *et al.* Site-specific integration by adeno-associated virus. *Proc Natl Acad Sci USA* **87**, 2211–2215 (1990).

[55] Samulski, R. J. *et al.* Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *EMBO J* **10**, 3941–3950 (1991).

[56] McCarty, D. M., Young, S. M. & Samulski, R. J. Integration of Adeno-Associated Virus (AAV) and Recombinant AAV Vectors. *Annu Rev Genet* **38**, 819–845 (2004).

[57] Wistuba, A., Kern, A., Weger, S., Grimm, D. & Kleinschmidt, J. A. Subcellular compartmentalization of adeno-associated virus type 2 assembly. *J Virol* **71**, 1341–1352 (1997).

[58] Earley, L. F. *et al.* Identification and characterization of nuclear and nucleolar localization signals in the adeno-associated virus serotype 2 assembly-activating protein. *J Virol* **89**, 3038–3048 (2015).

[59] Qiu, J. & Brown, K. E. A 110-kDa nuclear shuttle protein, nucleolin, specifically binds to adeno-associated virus type 2 (AAV-2) capsid. *Virology* **257**, 373–382 (1999).

[60] Bevington, J. M. *et al.* Adeno-associated virus interactions with B23/Nucleophosmin: identification of sub-nucleolar virion regions. *Virology* **357**, 102–113 (2007).

[61] Dubielzig, R., King, J. a., Weger, S., Kern, A. & Kleinschmidt, J. A. Adeno-associated virus type 2 protein interactions: formation of pre-encapsidation complexes. *J Virol* **73**, 8989–8998 (1999).

[62] King, J. a., Dubielzig, R., Grimm, D. & Kleinschmidt, J. a. DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J* **20**, 3282–3291 (2001).

[63] Bleker, S., Pawlita, M. & Kleinschmidt, J. A. Impact of Capsid Conformation and Rep-Capsid Interactions on Adeno-Associated Virus Type 2 Genome Packaging Impact of Capsid Conformation and Rep-Capsid Interactions on Adeno-Associated Virus Type 2 Genome Packaging. *J Virol* **80**, 810–820 (2006).

[64] Bleker, S., Sonntag, F. & Kleinschmidt, J. A. Mutational analysis of narrow pores at the fivefold symmetry axes of adeno-associated virus type 2 capsids reveals a dual role in genome packaging and activation of phospholipase A2 activity. *J Virol* **79**, 2528–2540 (2005).

[65] Agbandje-McKenna, M. & Kleinschmidt, J. AAV capsid structure and cell interactions. *Methods Mol Biol* **807**, 47–92 (2011).

[66] Kronenberg, S., Kleinschmidt, J. a. & Böttcher, B. Electron cryo-microscopy and image reconstruction of adeno-associated virus type 2 empty capsids. *EMBO Rep* **2**, 997–1002 (2001).

[67] Xie, Q. *et al.* The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc Natl Acad Sci USA* **99**, 10405–10410 (2002).

[68] Walters, R. W. *et al.* Structure of adeno-associated virus serotype 5. *J Virol* **78**, 3361–3371 (2004).

[69] Padron, E. *et al.* Structure of Adeno-Associated Virus Type 4. *J Virol* **79**, 5048–5058 (2005).

[70] Miller, E. B. *et al.* Production, purification and preliminary X-ray crystallographic studies of adeno-associated virus serotype 1. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **62**, 1271–1274 (2006).

[71] Quesada, O. *et al.* Production, purification and preliminary X-ray crystallographic studies of adeno-associated virus serotype 7. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **63**, 1073–1076 (2007).

[72] Nam, H.-J. *et al.* Structure of adeno-associated virus serotype 8, a gene therapy vector. *J Virol* **81**, 12260–12271 (2007).

[73] Xie, Q., Ongley, H. M., Hare, J. & Chapman, M. S. Crystallization and preliminary X-ray structural studies of adeno-associated virus serotype 6. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **64**, 1074–1078 (2008).

[74] Mitchell, M. *et al.* Production, purification and preliminary X-ray crystallographic studies of adeno-associated virus serotype 9. *Acta Crystallogr Sect F Struct Biol Cryst Commun* **65**, 715–718 (2009).

[75] Lerch, T. F., Xie, Q. & Chapman, M. S. The structure of adeno-associated virus serotype 3B (AAV-3B): insights into receptor binding and immune evasion. *Virology* **403**, 26–36 (2010).

[76] DiMattia, M. A. *et al.* Structural insight into the unique properties of adeno-associated virus serotype 9. *J Virol* **86**, 6947–6958 (2012).

[77] Govindasamy, L. *et al.* Structurally mapping the diverse phenotype of adeno-associated virus serotype 4. *J Virol* **80**, 11556–11570 (2006).

[78] Kronenberg, S. *et al.* A Conformational Change in the Adeno-Associated Virus Type 2 Capsid Leads to the Exposure of Hidden VP1 N Termini A Conformational Change in the Adeno-Associated Virus Type 2 Capsid Leads to the Exposure of Hidden VP1 N Termini. *J Virol* **79**, 5296–5303 (2005).

[79] Tseng, Y.-S. & Agbandje-McKenna, M. Mapping the AAV Capsid Host Antibody Response toward the Development of Second Generation Gene Delivery Vectors. *Front Immunol* **5**, 9 (2014).

[80] Jooss, K. & Chirmule, N. Immunity to adenovirus and adeno-associated viral vectors: implications for gene therapy. *Gene Ther* **10**, 955–963 (2003).

[81] Calcedo, R., Vandenberghe, L. H., Gao, G., Lin, J. & Wilson, J. M. Worldwide epidemiology of neutralizing antibodies to adeno-associated viruses. *J Infect Dis* **199**, 381–390 (2009).

[82] Boutin, S. *et al.* Prevalence of Serum IgG and Neutralizing Factors Against Adeno-Associated Virus (AAV) Types 1, 2, 5, 6, 8 and 9 in the Healthy Population: Implications for Gene Therapy Using AAV Vectors. *Hum Gene Ther* **712**, 704–712 (2010).

[83] Wobus, C. E. *et al.* Monoclonal antibodies against the adeno-associated virus type 2 (AAV-2) capsid: epitope mapping and identification of capsid domains involved in AAV-2-cell interaction and neutralization of AAV-2 infection. *J Virol* **74**, 9281–9293 (2000).

[84] Bello, A. *et al.* Novel Adeno-associated Viruses Derived From Pig Tissues Transduce Most Major Organs in Mice. *Sci Rep* **4** (2014).

[85] Gao, G. *et al.* Clades of Adeno-Associated Viruses Are Widely Disseminated in Human Tissues. *J Virol* **78**, 6381–6388 (2004).

[86] Summerford, C. & Samulski, R. J. Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J Virol* **72**, 1438–1445 (1998).

[87] Qing, K. *et al.* Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. *Nat Med* **5**, 71–77 (1999).

[88] Rabinowitz, J. E. *et al.* Cross-Packaging of a Single Adeno-Associated Virus (AAV) Type 2 Vector Genome into Multiple AAV Serotypes Enables Transduction with Broad Specificity. *J Virol* **76**, 791–801 (2002).

[89] Wu, Z., Miller, E., Agbandje-McKenna, M. & Samulski, R. J. 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**, 9093–9103 (2006).

[90] Kaludov, N., Brown, K. E., Walters, R. W., Zabner, J. & Chiorini, J. A. 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**, 6884–6893 (2001).

[91] Xiao, W. *et al.* Gene therapy vectors based on adeno-associated virus type 1. *J Virol* **73**, 3994–4003 (1999).

[92] Ng, R. *et al.* Structural characterization of the dual glycan binding adeno-associated virus serotype 6. *J Virol* **84**, 12945–12957 (2010).

[93] Shen, S., Bryant, K. D., Brown, S. M., Randell, S. H. & Asokan, A. Terminal N-linked galactose is the primary receptor for adeno-associated virus 9. *J Biol Chem* **286**, 13532–13540 (2011).

[94] Bell, C. L. *et al.* The AAV9 receptor and its modification to improve in vivo lung gene transfer in mice. *J Clin Invest* **121**, 2427–2435 (2011). URL `http://dx.doi.org/10.1172/JCI57367`.

[95] Pillay, S. *et al.* An essential receptor for adeno-associated virus infection. *Nature* [Epub ahead of print] (2016).

[96] Summerford, C., Bartlett, J. S. & Samulski, R. J. AlphaVbeta5 integrin: a co-receptor for adeno-associated virus type 2 infection. *Nat Med* **5**, 78–82 (1999).

[97] Asokan, A., Hamra, J. B., Govindasamy, L., Agbandje-McKenna, M. & Samulski, R. J. Adeno-associated virus type 2 contains an integrin alpha5beta1 binding domain essential for viral cell entry. *J Virol* **80**, 8961–8969 (2006).

[98] Kashiwakura, Y. *et al.* Hepatocyte growth factor receptor is a coreceptor for adeno-associated virus type 2 infection. *J Virol* **79**, 609–614 (2005).

[99] Akache, B. *et al.* The 37/67-kilodalton laminin receptor is a receptor for adeno-associated virus serotypes 8, 2, 3, and 9. *J Virol* **80**, 9831–9836 (2006).

[100] Blackburn, S. D., Steadman, R. A. & Johnson, F. B. Attachment of adeno-associated virus type 3H to fibroblast growth factor receptor 1. *Arch Virol* **151**, 617–623 (2006).

[101] Hoggan, M. D., Blacklow, N. R. & Rowe, W. P. Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. *Proc Natl Acad Sci USA* **55**, 1467–1474 (1966).

[102] Rutledge, E. A., Halbert, C. L. & Russell, D. W. Infectious Clones and Vectors Derived from Adeno-Associated Virus (AAV) Serotypes Other Than AAV Type 2. *J Virol* **72**, 309–319 (1998).

[103] Parks, W. P., Melnick, J. L., Rongey, R. & Mayor, H. D. Physical Assay and Growth Cycle Studies of a Defective Adeno-Satellite Virus. *J Virol* **1**, 171–180 (1967).

[104] Bantel-Schaal, U. & zur Hausen, H. Characterization of the DNA of a defective human parvovirus isolated from a genital site. *Virology* **134**, 52–63 (1984).

[105] Gao, G.-P. *et al.* Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc Natl Acad Sci USA* **99**, 11854–11859 (2002).

[106] Bello, A. *et al.* Isolation and evaluation of novel adeno-associated virus sequences from porcine tissues. *Gene Ther* **16**, 1320–1328 (2009).

[107] Schmidt, M. *et al.* Adeno-Associated Virus Type 12 (AAV12): a Novel AAV Serotype with Sialic Acid- and Heparan Sulfate Proteoglycan-Independent Transduction Activity. *J Virol* **82**, 1399–1406 (2008).

[108] Samulski, R. J., Chang, L. S. & Shenk, T. Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *J Virol* **63**, 3822–3828 (1989).

[109] Matsushita, T. *et al.* Adeno-associated virus vectors can be efficiently produced without helper virus. *Gene Ther* **5**, 938–945 (1998).

[110] Kearns, W. G. *et al.* Recombinant adeno-associated virus (AAV-CFTR) vectors do not integrate in a site-specific fashion in an immortalized epithelial cell line. *Gene Ther* **3**, 748–755 (1996).

[111] Smith, R. H. Adeno-associated virus integration: virus versus vector. *Gene Ther* **15**, 817–822 (2008).

[112] Nakai, H. *et al.* Extrachromosomal recombinant adeno-associated virus vector genomes are primarily responsible for stable liver transduction in vivo. *J Virol* **75**, 6969–6976 (2001).

[113] Kaeppel, C. *et al.* A largely random AAV integration profile after LPLD gene therapy. *Nat Med* **19**, 889–891 (2013).

[114] Inagaki, K., Piao, C., Kotchey, N. M., Wu, X. & Nakai, H. Frequency and spectrum of genomic integration of recombinant adeno-associated virus serotype 8 vector in neonatal mouse liver. *J Virol* **82**, 9513–9524 (2008).

[115] Li, H. *et al.* Assessing the potential for AAV vector genotoxicity in a murine model. *Blood* **117**, 3311–3319 (2010).

[116] Nault, J.-C. *et al.* Recurrent AAV2-related insertional mutagenesis in human hepatocellular carcinomas. *Nat Genet* **47**, 1187–1193 (2015).

[117] Büning, H. & Schmidt, M. Adeno-associated Vector Toxicity—To Be or Not to Be? *Mol Ther* **23**, 1673–1675 (2015).

[118] Berns, K. I. *et al.* Adeno-Associated Virus Type 2 and Hepatocellular Carcinoma? *Hum Gene Ther* **26**, 779–781 (2015).

[119] Ferrari, F. K., Samulski, T., Shenk, T. & Samulski, R. J. Second-strand synthesis is a rate-limiting step for efficient transduction by recombinant adeno-associated virus vectors. *J Virol* **70**, 3227–3234 (1996).

[120] Fisher, K. J. *et al.* Transduction with recombinant adeno-associated virus for gene therapy is limited by leading-strand synthesis. *J Virol* **70**, 520–532 (1996).

[121] McCarty, D. M., Monahan, P. E. & Samulski, R. J. Self-complementary recombinant adeno-associated virus (scAAV) vectors promote efficient transduction independently of DNA synthesis. *Gene Ther* **8**, 1248–1254 (2001).

[122] Duan, D., Yue, Y., Yan, Z. & Engelhardt, J. F. A new dual-vector approach to enhance recombinant adeno-associated virus-mediated gene expression through intermolecular cis activation. *Nat Med* **6**, 595–598 (2000).

[123] Sun, L., Li, J. & Xiao, X. Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization. *Nat Med* **6**, 599–602 (2000).

[124] Duan, D., Yue, Y. & Engelhardt, J. F. Expanding AAV packaging capacity with trans-splicing or overlapping vectors: a quantitative comparison. *Mol Ther* **4**, 383–391 (2001).

[125] Yan, Z., Zhang, Y., Duan, D. & Engelhardt, J. F. Trans-splicing vectors expand the utility of adeno-associated virus for gene therapy. *Proc Natl Acad Sci USA* **97**, 6716–6721 (2000).

[126] Ghosh, A., Yue, Y., Lai, Y. & Duan, D. A hybrid vector system expands adeno-associated viral vector packaging capacity in a transgene-independent manner. *Mol Ther* **16**, 124–130 (2008).

[127] Ghosh, A., Yue, Y. & Duan, D. Efficient transgene reconstitution with hybrid dual AAV vectors carrying the minimized bridging sequences. *Hum Gene Ther* **22**, 77–83 (2011).

[128] Le Bec, C. & Douar, A. M. Gene therapy progress and prospects–vectorology: design and production of expression cassettes in AAV vectors. *Gene Ther* **13**, 805–813 (2006).

[129] Stemmer, W. P. DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. *Proc Natl Acad Sci USA* **91**, 10747–10751 (1994).

[130] Samulski, R. J., Chang, L. S. & Shenk, T. A recombinant plasmid from which an infectious adeno-associated virus genome can be excised in vitro and its use to study viral replication. *J Virol* **61**, 3096–4101 (1987).

[131] Soong, N. W. *et al.* Molecular breeding of viruses. *Nat Genet* **25**, 436–439 (2000).

[132] Powell, S. K. *et al.* Breeding of retroviruses by DNA shuffling for improved stability and processing yields. *Nat Biotechnol* **18**, 1279–1282 (2000).

[133] Grimm, D. *et al.* In vitro and in vivo gene therapy vector evolution via multispecies interbreeding and retargeting of adeno-associated viruses. *J Virol* **82**, 5887–5911 (2008).

[134] Koerber, J. T., Jang, J.-H. & Schaffer, D. V. DNA shuffling of adeno-associated virus yields functionally diverse viral progeny. *Mol Ther* **16**, 1703–1709 (2008).

[135] Li, W. *et al.* Engineering and selection of shuffled AAV genomes: a new strategy for producing targeted biological nanoparticles. *Mol Ther* **16**, 1252–1260 (2008).

[136] Lisowski, L. *et al.* Selection and evaluation of clinically relevant AAV variants in a xenograft liver model. *Nature* **506**, 382–386 (2014).

[137] Melo, S. P. *et al.* Somatic correction of junctional epidermolysis bullosa by a highly recombinogenic AAV variant. *Mol Ther* **22**, 725–733 (2014).

[138] Lakhan, R. *et al.* Local administration of AAV-DJ pseudoserotype expressing COX2 provided early onset of transgene expression and promoted bone fracture healing in mice. *Gene Ther* **22**, 721–728 (2015).

[139] Lerch, T. F. *et al.* Structure of AAV-DJ, a retargeted gene therapy vector: cryo-electron microscopy at 4.5 Å resolution. *Structure* **20**, 1310–1320 (2012).

[140] Girod, A. *et al.* Genetic capsid modifications allow efficient re-targeting of adeno-associated virus type 2. *Nat Med* **5**, 1052–1056 (1999).

[141] Wu, P. *et al.* Mutational Analysis of the Capsid Gene and Construction of AAV2 Vectors with Altered Tropism. *J Virol* **74**, 8635–8647 (2000).

[142] Grifman, M. *et al.* Incorporation of tumor-targeting peptides into recombinant adeno-associated virus capsids. *Mol Ther* **3**, 964–975 (2001).

[143] Nicklin, S. a. *et al.* Efficient and selective AAV2-mediated gene transfer directed to human vascular endothelial cells. *Mol Ther* **4**, 174–181 (2001).

[144] Shi, W., Arnold, G. S. & Bartlett, J. S. Insertional mutagenesis of the adeno-associated virus type 2 (AAV2) capsid gene and generation of AAV2 vectors targeted to alternative cell-surface receptors. *Hum Gene Ther* **12**, 1697–711 (2001).

[145] Opie, S. R., Warrington, K. H., Zolotukhin, S., Agbandje-mckenna, M. & Muzyczka, N. Identification of Amino Acid Residues in the Capsid Proteins of Adeno-Associated Virus Type 2 That Contribute to Heparan Sulfate Proteoglycan Binding. *J Virol* **77**, 6995–7006 (2003).

[146] Kern, A. *et al.* Identification of a Heparin-Binding Motif on Adeno-Associated Virus Type 2 Capsids. *J Virol* **77**, 11072–11081 (2003).

[147] Work, L. M. *et al.* Development of efficient viral vectors selective for vascular smooth muscle cells. *Mol Ther* **9**, 198–208 (2004).

[148] White, S. J. *et al.* Targeted gene delivery to vascular tissue in vivo by tropism-modified adeno-associated virus vectors. *Circulation* **109**, 513–519 (2004).

[149] Waterkamp, D. A., Muller, O. J., Ying, Y., Trepel, M. & Kleinschmidt, J. A. Isolation of targeted AAV2 vectors from novel virus display libraries. *J Gene Med* **8**, 1307–1319 (2006).

[150] Perabo, L. *et al.* In vitro selection of viral vectors with modified tropism: the adeno-associated virus display. *Mol Ther* **8**, 151–157 (2003).

[151] Müller, O. J. *et al.* Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors. *Nat Biotechnol* **21**, 1040–1046 (2003).

[152] Jang, J.-H. *et al.* An Evolved Adeno-associated Viral Variant Enhances Gene Delivery and Gene Targeting in Neural Stem Cells. *Mol Ther* **19**, 667–675 (2011).

[153] Ying, Y. *et al.* Heart-targeted adeno-associated viral vectors selected by in vivo biopanning of a random viral display peptide library. *Gene Ther* **17**, 980–90 (2010).

[154] Michelfelder, S. *et al.* Peptide ligands incorporated into the threefold spike capsid domain to re-direct gene transduction of AAV8 and AAV9 in vivo. *PLoS One* **6**, e23101 (2011).

[155] Varadi, K. *et al.* Novel random peptide libraries displayed on AAV serotype 9 for selection of endothelial cell-directed gene transfer vectors. *Gene Ther* **19**, 800–809 (2012).

[156] Raupp, C. *et al.* The threefold protrusions of adeno-associated virus type 8 are involved in cell surface targeting as well as postattachment processing. *J Virol* **86**, 9396–9408 (2012).

[157] Grimm, D. & Zolotukhin, S. E Pluribus Unum: Fifty years of research, millions of viruses, and one goal - tailored acceleration of AAV evolution. *Mol Ther* 1–13 (2015).

[158] Perabo, L. *et al.* Combinatorial engineering of a gene therapy vector: directed evolution of adeno-associated virus. *J Gene Med* **8**, 155–162 (2006).

[159] Maheshri, N., Koerber, J. T., Kaspar, B. K. & Schaffer, D. V. Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat Biotechnol* **24**, 198–204 (2006).

[160] Maersch, S., Huber, A., Büning, H., Hallek, M. & Perabo, L. Optimization of stealth adeno-associated virus vectors by randomization of immunogenic epitopes. *Virology* **397**, 167–175 (2010).

[161] Pulicherla, N. *et al.* Engineering liver-detargeted AAV9 vectors for cardiac and musculoskeletal gene transfer. *Mol Ther* **19**, 1070–1078 (2011).

[162] Lochrie, M. A. *et al.* Mutations on the External Surfaces of Adeno-Associated Virus Type 2 Capsids That Affect Transduction and Neutralization. *J Virol* **80**, 821–834 (2006).

[163] Li, Z. *et al.* Tyrosine phosphorylation of AAV2 vectors and its consequences on viral intracellular trafficking and transgene expression. *Virology* **381**, 194–202 (2008).

[164] Zhong, L. *et al.* Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc Natl Acad Sci USA* **105**, 7827–7832 (2008).

[165] Li, M. *et al.* High-efficiency transduction of fibroblasts and mesenchymal stem cells by tyrosine-mutant AAV2 vectors for their potential use in cellular therapy. *Hum Gene Ther* **21**, 1527–1543 (2010).

[166] Markusic, D. M. *et al.* High-efficiency transduction and correction of murine hemophilia B using AAV2 vectors devoid of multiple surface-exposed tyrosines. *Mol Ther* **18**, 2048–2056 (2010).

[167] Ku, C. A. *et al.* Gene therapy using self-complementary Y733F capsid mutant AAV2/8 restores vision in a model of early onset Leber congenital amaurosis. *Hum Mol Genet* **20**, 4569–4581 (2011).

[168] Warrington, K. H. *et al.* Adeno-Associated Virus Type 2 VP2 Capsid Protein Is Nonessential and Can Tolerate Large Peptide Insertions at Its N Terminus †. *J Virol* **78**, 6595–6609 (2004).

[169] Lux, K. *et al.* Green Fluorescent Protein-Tagged Adeno-Associated Virus Particles Allow the Study of Cytosolic and Nuclear Trafficking. *J Virol* **79**, 11776–11787 (2005).

[170] Binz, H. K. *et al.* High-affinity binders selected from designed ankyrin repeat protein libraries. *Nat Biotechnol* **22**, 575–582 (2004).

[171] Steiner, D., Forrer, P. & Pluckthun, A. Efficient selection of DARPins with sub-nanomolar affinities using SRP phage display. *J Mol Biol* **382**, 1211–1227 (2008).

[172] Stumpp, M. T., Binz, H. K. & Amstutz, P. DARPins: a new generation of protein therapeutics. *Drug Discov Today* **13**, 695–701 (2008).

[173] Münch, R. C. *et al.* Displaying High-affinity Ligands on Adeno-associated Viral Vectors Enables Tumor Cell-specific and Safe Gene Transfer. *Mol Ther* **21**, 109–118 (2013).

[174] Yang, Q. *et al.* Development of novel cell surface CD34-targeted recombinant adenoassociated virus vectors for gene therapy. *Hum Gene Ther* **9**, 1929–1937 (1998).

[175] Zinn, E. *et al.* In Silico Reconstruction of the Viral Evolutionary Lineage Yields a Potent Gene Therapy Vector. *Cell Rep* **12**, 1056–1068 (2015).

[176] Santiago-Ortiz, J. *et al.* AAV ancestral reconstruction library enables selection of broadly infectious viral variants. *Gene Ther* **22**, 934–946 (2015).

[177] Kienle, E. *et al.* Engineering and evolution of synthetic adeno-associated virus (AAV) gene therapy vectors via DNA family shuffling. *J Vis Exp* (2012).

[178] Michelfelder, S. *et al.* Successful expansion but not complete restriction of tropism of adeno-associated virus by in vivo biopanning of random virus display peptide libraries. *PLoS One* **4**, e5122 (2009).

[179] Yang, L. *et al.* A myocardium tropic adeno-associated virus (AAV) evolved by DNA shuffling and in vivo selection. *Proc Natl Acad Sci USA* **106**, 3946–3951 (2009).

[180] Gray, S. J. *et al.* Directed evolution of a novel adeno-associated virus (AAV) vector that crosses the seizure-compromised blood-brain barrier (BBB). *Mol Ther* **18**, 570–578 (2010).

[181] Dalkara, D. *et al.* In Vivo-Directed Evolution of a New Adeno-Associated Virus for Therapeutic Outer Retinal Gene Delivery from the Vitreous. *Sci Transl Med* **5**, 189ra76 (2013).

[182] Nietupski, J. B. *et al.* Systemic administration of AAV8-alpha-galactosidase A induces humoral tolerance in nonhuman primates despite low hepatic expression. *Mol Ther* **19**, 1999–2011 (2011).

[183] Bell, P. *et al.* Inverse zonation of hepatocyte transduction with AAV vectors between mice and non-human primates. *Mol Genet Metab* **104**, 395–403 (2011).

[184] Gao, G. *et al.* Transendocardial delivery of AAV6 results in highly efficient and global cardiac gene transfer in rhesus macaques. *Hum Gene Ther* **22**, 979–984 (2011).

[185] Gray, S. J., Nagabhushan Kalburgi, S., McCown, T. J. & Jude Samulski, R. Global CNS gene delivery and evasion of anti-AAV-neutralizing antibodies by intrathecal AAV administration in non-human primates. *Gene Ther* **20**, 450–459 (2013).

[186] DuBridge, R. B. *et al.* Analysis of mutation in human cells by using an Epstein-Barr virus shuttle system. *Mol Cell Biol* **7**, 379–387 (1987).

[187] Nakabayashi, H., Taketa, K., Miyano, K., Yamane, T. & Sato, J. Growth of human hepatoma cells lines with differentiated functions in chemically defined medium. *Cancer Res* **42**, 3858–3863 (1982).

[188] Soule, H. D., Vazguez, J., Long, A., Albert, S. & Brennan, M. A human cell line from a pleural effusion derived from a breast carcinoma. *J Natl Cancer Inst* **51**, 1409–1416 (1973).

[189] Rutka, J. T. *et al.* Establishment and characterization of a cell line from a human gliosarcoma. *Cancer Res* **46**, 5893–5902 (1986).

[190] Wistuba, A., Weger, S., Kern, A. & Kleinschmidt, J. A. Intermediates of adeno-associated virus type 2 assembly: identification of soluble complexes containing Rep and Cap proteins. *J Virol* **69**, 5311–5319 (1995).

[191] Carpenter, A. E. *et al.* CellProfiler: image analysis software for identifying and quantifying cell phenotypes. *Genome Biol* **7**, R100 (2006).

[192] Kamentsky, L. *et al.* Improved structure, function and compatibility for CellProfiler: modular high-throughput image analysis software. *Bioinformatics* **27**, 1179–1180 (2011).

[193] Larkin, M. A. *et al.* Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948 (2007).

[194] Schindelin, J. *et al.* Fiji: an open-source platform for biological-image analysis. *Nat Methods* **9**, 676–682 (2012).

[195] Schürmann, N., Trabuco, L. G., Bender, C., Russell, R. B. & Grimm, D. Molecular dissection of human Argonaute proteins by DNA shuffling. *Nat Struct Mol Biol* **20**, 818–826 (2013).

[196] Kienle, E. *Secrets to finding the ideal mate: New insights into parameters that govern successful Adeno-associated virus (AAV) vector evolution.* Ph.D. thesis, Heidelberg University (2014).

[197] Birnboim, H. C. & Doly, J. A rapid alkaline extraction procedure for screening recombinant plasmid DNA. *Nucleic Acids Res* **7**, 1513–1523 (1979).

[198] Livak, K. J. & Schmittgen, T. D. Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* **25**, 402–408 (2001).

[199] Sommer, R. *et al.* Combined polyethylene glycol and CaCl2 precipitation for the capture and purification of recombinant antibodies. *Process Biochem* **49**, 2001–2009 (2014).

[200] Quail, M. A. *et al.* A large genome center's improvements to the Illumina sequencing system. *Nat Meth* **5**, 1005–1010 (2008).

[201] Fisher, S. *et al.* A scalable, fully automated process for construction of sequence-ready human exome targeted capture libraries. *Genome Biol* **12**, R1 (2011).

[202] Poptsova, M. S. *et al.* Non-random DNA fragmentation in next-generation sequencing. *Sci Rep* **4**, 4532 (2014).

[203] Maheshri, N. & Schaffer, D. V. Computational and experimental analysis of DNA shuffling. *Proc Natl Acad Sci U S A* **100**, 3071–3076 (2003).

[204] Coco, W. M. *et al.* DNA shuffling method for generating highly recombined genes and evolved enzymes. *Nat Biotech* **19**, 354–359 (2001).

[205] Moore, G. L. & Maranas, C. D. Modeling DNA Mutation and Recombination for Directed Evolution Experiments. *J Theor Biol* **205**, 483–503 (2000).

[206] Moore, G. L., Maranas, C. D., Lutz, S. & Benkovic, S. J. Predicting crossover generation in DNA shuffling. *Proc Natl Acad Sci USA* **98**, 3226–3231 (2001).

[207] Joern, J. M., Meinhold, P. & Arnold, F. H. Analysis of shuffled gene libraries. *J Mol Biol* **316**, 643–656 (2002).

[208] Hauck, B. & Xiao, W. Characterization of tissue tropism determinants of adeno-associated virus type 1. *J Virol* **77**, 2768–2774 (2003).

[209] Shen, X., Storm, T. & Kay, M. A. Characterization of the relationship of AAV capsid domain swapping to liver transduction efficiency. *Mol Ther* **15**, 1955–1962 (2007).

[210] DiPrimio, N., Asokan, A., Govindasamy, L., Agbandje-McKenna, M. & Samulski, R. J. Surface loop dynamics in adeno-associated virus capsid assembly. *J Virol* **82**, 5178–89 (2008).

[211] Marsic, D. *et al.* Vector Design Tour de Force: Integrating Combinatorial and Rational Approaches to Derive Novel Adeno-associated Virus Variants. *Mol Ther* **22**, 1900–1909 (2014).

[212] Naumer, M. *et al.* Properties of the adeno-associated virus assembly-activating protein. *J Virol* **86**, 13038–13048 (2012).

[213] Sonntag, F. *et al.* The assembly-activating protein promotes capsid assembly of different adeno-associated virus serotypes. *J Virol* **85**, 12686–12697 (2011).

[214] Rabinowitz, J. E., Xiao, W. & Samulski, R. J. Insertional mutagenesis of AAV2 capsid and the production of recombinant virus. *Virology* **265**, 274–285 (1999).

[215] Herrmann, J., Lerman, L. O. & Lerman, A. Ubiquitin and ubiquitin-like proteins in protein regulation. *Circ Res* **100**, 1276–1291 (2007).

[216] Hochstrasser, M. Origin and function of ubiquitin-like proteins. *Nature* **458**, 422–429 (2009).

[217] Zhong, L. *et al.* Next generation of adeno-associated virus 2 vectors: Point mutations in tyrosines lead to high-efficiency transduction at lower doses. *PNAS* **105**, 7827–7832 (2008).

[218] Petrs-Silva, H. *et al.* High-efficiency Transduction of the Mouse Retina by Tyrosine-mutant AAV Serotype Vectors. *Mol Ther* **17**, 463–471 (2008).

[219] Qiao, C. *et al.* Adeno-associated virus serotype 6 capsid tyrosine-to-phenylalanine mutations improve gene transfer to skeletal muscle. *Hum Gene Ther* **21**, 1343–1348 (2010).

[220] Zhong, L. *et al.* Tyrosine-phosphorylation of AAV2 vectors and its consequences on viral intracellular trafficking and transgene expression. *Virology* **381**, 194–202 (2008).

[221] Nayak, R., Farris, K. D. & Pintel, D. J. E4Orf6-E1B-55k-dependent degradation of de novo-generated adeno-associated virus type 5 Rep52 and capsid proteins employs a cullin 5-containing E3 ligase complex. *J Virol* **82**, 3803–3808 (2008).

[222] Hölscher, C. *et al.* The SUMOylation Pathway Restricts Gene Transduction by Adeno-Associated Viruses. *PLoS Pathog* **11**, e1005281 (2015).

[223] Wilson, V. G. & Heaton, P. R. Ubiquitin proteolytic system: focus on SUMO. *Expert Rev Proteomics* **5**, 121–135 (2008).

[224] Powers, J., Earley, L. F., Adachi, K. & Nakai, H. Distinct Intracellular Localization of AAPs Derived from AAV Serotypes 1-11 and Their Ability To Cross-Complement in Heterologous Capsid Assembly. In *Oral Abstr Sess* (Annual meeting of the American Society of Gene and Cell Therapy (ASGCT), New Orleans, 2015).

[225] Liesche, C. *et al.* Death receptor-based enrichment of Cas9-expressing cells. *BMC Biotechnol* **16** (2016).

[226] Mockenhaupt, S., Grosse, S., Rupp, D., Bartenschlager, R. & Grimm, D. Alleviation of off-target effects from vector-encoded shRNAs via codelivered RNA decoys. *Proc Natl Acad Sci USA* **112**, E4007–16 (2015).

[227] Senis, E. *et al.* CRISPR/Cas9-mediated genome engineering: An adeno-associated viral (AAV) vector toolbox. *Biotechnol J* **9**, 1402–1412 (2014).

Supplementary information

1 Supplementary figures

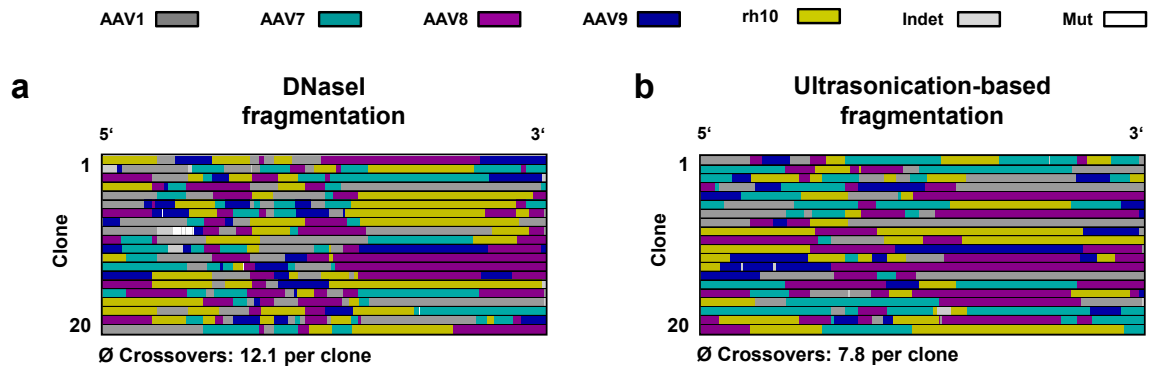

Figure 1: DNA type assignments of AAV *cap* library 1789rh10 produced with DNase or ultrasonication-based fragmentation. a) Type assignment of 20 shuffled AAV *cap* chimeras after DNaseI fragmentation. **b)** Type assignment of 20 shuffled AAV *cap* chimeras after ultrasonication-based (Covaris) fragmentation. All sequences were analyzed with the Salanto program.

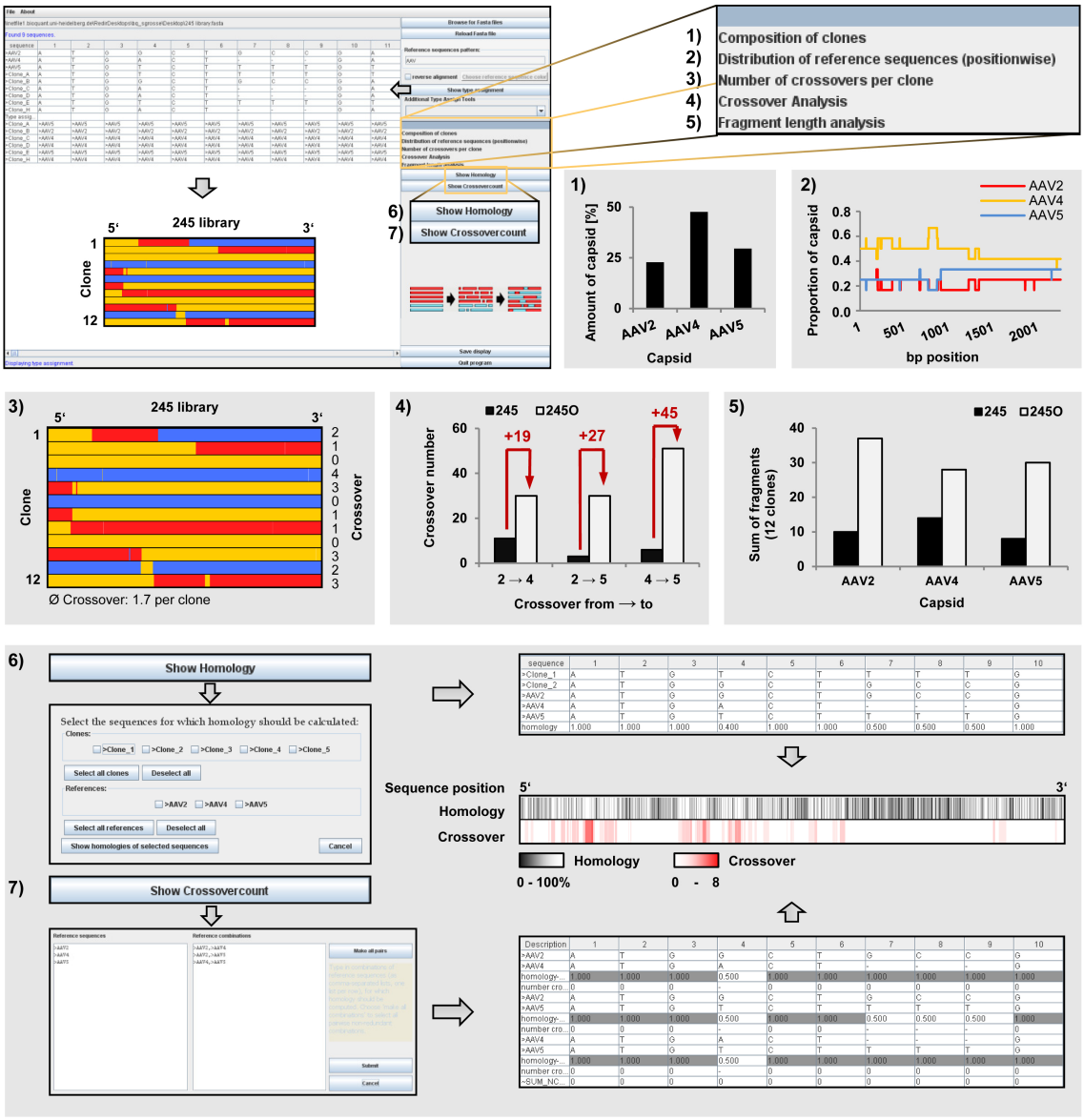

Figure 2: Applied functions of the Salanto program. Screenshot of Salanto with opened and analyzed data set (upper left). The shown type assignment analysis was colored with the Microsoft program Excel to further indicate chimeric clone composition (colored panel in upper left). Additional used functions of Salanto are labeld with 1) - 7) and exemplified with diagrams in the grey boxes. **1)** "Composition of clones" calculates percentages of parental sequences in the final library. **2)** "Distribution of reference sequence (position wise)" depicts the proportion of parents for each nucleotide position in the analyzed library. **3)** "Number of crossovers per clone" calculates crossover events per chimera. **4)** "Crossover analysis" determines crossover events between the single parents. **5)** "Fragment length analysis" shows the number and length of shuffled gene fragments for each parent along the complete library. **6)** "Show homology" calculates position wise homologies of selected parental and / or chimeric sequences. **7)** "Show crossovercount" connects position wise homology with the appearance of crossovers along the sequences. Calculations in 1) – 5) and 7) are based on the type assignment analysis.

ATGACCGATGGTTATCTTCCAGATTGGCTCGAGGACAATCTCTCTGAAGGAGTAAGAGAGTGGTGGG
CGCTCCAACCTGGCGCACCAAAACCAAAGGCCAACCAGCAGCATCAGGACAACGCCAGGGGTCTTGT
GCTTCCTGGGTACAAGTACCTCGGACCCGGCAACGGACTCGACAAGGGAGAGCCGGTCAACGCGGCA
GACGCCGCGGCCCTCGAGCACGACAAAGCCTACGACCAGCAGCTCAAGGCCGGAGACAACCCGTACC
TCAAGTACAACCACGCCGACGCGGAGTTTCAGCAGCGCCTTCAAGGAGATACGTCTTTTGGGGGCAA
CCTCGGACGAGCAGTCTTTCAGGCGAAAAAGAGGGTTCTTGAACCTCTGGGCCTGGTTGAGCAAGCT
GGTGAGACGGCTCCGGGAAAAAAGAGGCCGCTAATCGAGTCTCCTCAGCAGCCAGACTCCTCCACGG
GAATCGGAAAGAAGGGCAAGCAGCCTGCAAAAAAAAAATTGGTTTTTGAGGACGAGACAGGCGCAGG
CGACGGACCTCCTGAGGGATCGACATCAGGAGCCATGTCTGACGATAGTGAGATGCGTGCAGCCGCT
GGCGGAGCAGCGGTAGAGGGTGGCCAGGGCGCCGACGGAGTGGGTAATGCCTCGGGAGATTGGCATT
GCGATTCCACATGGAGCGAGGGCCACGTCACCACCACCAGCACCCGAACCTGGGTCCTGCCCACCTA
CAACAACCACCTCTACAAACGACTTGGCGAATCCTTGCAGAGCAATACCTACAATGGCTTCAGCACC
CCTTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGCAAAGACTCA
TCAACAACAACTGGGGAATGCGACCCAAGGCAATGAGAGTCAAGATCTTTAACATTCAAGTCAAAGA
GGTCACGACGAGTAACGGTGAGACGACGGTTGCCAATAACCTTACCAGCACGGTTCAGATCTTTGCT
GACTCGTCGTACGAGCTCCCGTACGTCATGGACGCGGGGCAGGAAGGAAGCCTCCCGCCGTTCCCAA
ACGACGTCTTTATGGTGCCACAGTATGGATACTGCGGCCTGGTCACCGGGAACACAAGCCAGCAGCA
AACAGACCGCAATGCCTTTTACTGCCTGGAGTACTTTCCTTCTCAGATGCTGCGTACCGGAAACAAC
TTTGAGATCACCTACTCTTTTGAGAAGGTTCCTTTCCACAGCATGTACGCTCACAGCCAGAGTCTGG
ACCGTCTCATGAATCCTCTCATCGACCAGTACCTGTGGGGCTTGCAGAGCACAACCACTGGAACTAC
ACTCAACGCGGGAACGGCAACGACTAACTTTACTAAGCTCAGACCGACTAACTTTAGCAACTTCAAG
AAGAACTGGCTTCCTGGACCCTCTATCAAGCAGCAGGGATTCTCAAAGACAGCTAACCAGAACTACA
AGATTCCAGCCACGGGGTCTGACAGCCTAATAAAATACGAGACCCACTCCACCCTCGATGGCAGATG
GTCTGCGCTGACTCCGGGCCCTCCCATGGCAACCGCCGGGCCCGCTGACTCAAAGTTTTCTAACAGC
CAGCTTATCTTCGCTGGGCCGAAACAGAACGGGAACACAGCTACGGTCCCTGGAACGCTCATCTTTA
CAAGCGAAGAGGAACTCGCGGCAACCAATGCCACGGATACGGACATGTGGGGTAATCTACCTGGCGG
CGACCAGAGCAACAGCAACCTACCAACTGTCGACCGTCTCACCGCACTAGGCGCTGTTCCAGGCATG
GTCTGGCAGAACAGAGATATCTACTATCAGGGGCCCATCTGGGCAAAGATTCCACACACGGACGGAC
ATTTTCACCCCTCTCCCCTCATCGGTGGATTCGGACTTAAACACCCTCCTCCACAGATTTTCATCAA
GAACACCCCGGTACCTGCGAATCCTGCGACCACCTTCAGTTCGACACCGGTTAATTCCTTCATCACA
CAGTACTCCACGGGACAGGTCAGCGTGCAGATCGACTGGGAGATCCAGAAGGAAAGAAGCAAACGCT
GGAATCCCGAAGTTCAGTTCACTTCCAACTACGGCCAGCAGAATAGTCTGCTCTGGGCTCCGGACGC
TGCTGGCAAGTATACAGAGCCTCGCGCCATTGGCACCAGATACCTGACTCATCATCTGTAA

Figure 3: AAV *cap* 4_O sequence. DNA sequence of the sequence optimized AAV4 capsid ($AAV4_O$). The AAP ORF is underlined. Modified nucleotides which lead to an AAP knockout are marked in red.

ATGTCTTTCGTTGATCATCCTCCAGATTGGCTCGAGGAGGTTGGCGAAGGACTAAGAGAGTTCTTGG
GGCTCGAAGCTGGCCCACCAAAACCAAAGCCCAACCAGCAGCATCAGGACCAGGCCAGGGGTCTTGT
GCTTCCTGGGTACAACTACCTCGGACCCGGCAACGGACTCGACAGGGGAGAGCCGGTCAACCGGGCA
GACGAGGTGGCCCGCGAGCACGACATATCCTACAACGAGCAGCTCGAGGCCGGAGACAACCCGTACC
TCAAGTACAACCACGCCGACGCGGAGTTTCAGGAGAAGCTCGCAGACGATACGTCTTTTGGGGGCAA
CCTCGGAAAAGCAGTCTTCCAGGCGAAAAAGAGGGTTCTTGAACCTTTCGGCCTGGTTGAGGAAGGT
GCTAAGACGGCTCCGACAGGAAAGAGGATTGACGACCACTTTCCTAAGCGGAAAAAGGCCCGCACGG
AGGAGGACTCGAAGCCTTCAACGTCTAGCGACGCAGAGGCAGGACCTAGCGGATCCCAGCAGCTCCA
AATCCCAGCACAACCCGCCTCTAGTCTGGGAGCTGATACGATGTCTGCAGGCGGTGGCGGACCACTG
GGAGACAATAACCAGGGCGCCGACGGAGTGGGTAATGCCTCGGGAGATTGGCATTGCGATTCCACAT
GGATGGGCGACAGAGTCGTCACCAAGAGCACCCGAACCTGGGTCCTGCCCAGCTACAACAACCACCA
GTACAGAGAAATTAAGAGCGGATCAGTAGACGGCTCGAACGCCAATGCCTACTTTGGCTACAGCACC
CCTTGGGGGTATTTTGACTTCAACAGATTCCACAGCCACTGGTCACCACGTGACTGGCAAAGACTCA
TCAACAACTACTGGGGATTCCGACCCAGGAGCCTCAGAGTCAAGATCTTTAACATTCAAGTCAAAGA
GGTCACGGTGCAGGACAGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTCAGGTGTTTACT
GACGACGACTACCAGCTCCCGTACGTCGTCGGCAACGGGACTGAAGGATGCCTCCCGGCGTTCCCAC
CACAGGTCTTCACGCTGCCACAGTATGGATACGCCACCCTGAACAGAGACAACACGGAGAACCCAAC
AGAACGCTCTTCATTTTTCTGCCTGGAGTACTTTCCTTCTAAGATGCTGCGTACCGGAAACAACTTT
GAGTTCACCTACAATTTTGAGGAGGTTCCTTTCCACAGCAGCTTCGCTCCCAGCCAGAATCTGTTCA
AGCTCGCGAATCCTCTCGTCGACCAGTACCTGTATCGCTTCGTCAGCACAAACAATACAGGTGGAGT
TCAGTTTAATAAGAACCTAGCGGGTCGCTATGCGAACACGTATAAGAACTGGTTTCCTGGACCCATG
GGCCGCACGCAGGGATGGAACCTGGGATCTGGGGTTAACAGAGCCAGTGTATCCGCGTTCGCTACAA
CTAACAGGATGGAGCTCGAGGGCGCAAGCTATCAGGTGCCTCCGCAGCCGAACGGGATGACCAACAA
CCTCCAGGGATCAAACACTTATGCTCTGGAGAACACTATGATCTTTAACAGCCAACCCGCAAACCCA
GGAACAACAGCTACGTACCTTGAAGGGAACATGCTGATTACAAGCGAATCGGAAACCCAGCCAGTCA
ATCGCGTGGCTTACAACGTGGGTGGTCAGATGGCTACCAACAACCAGAGCAGCACCACAGCACCAGC
TACCGGAACTTACAACCTACAAGAGATTGTTCCAGGCAGCGTCTGGATGGAGAGAGATGTGTACCTT
CAGGGGCCCATCTGGGCAAAGATTCCAGAGACGGGCGCACATTTTCACCCCTCTCCCGCCATGGGTG
GATTCGGACTTAAACACCCTCCTCCAATGATGCTCATCAAGAACACCCCGGTACCTGGGAATATCAC
CAGCTTCAGTGACGTACCGGTTTCTTCCTTCATCACACAGTACTCCACGGGACAGGTCACCGTGGAG
ATGGAGTGGGAGCTGAAGAAGGAAAACAGCAAACGCTGGAATCCCGAAATTCAGTACACTAACAACT
ACAACGACCCTCAGTTTGTGGACTTTGCTCCGGACTCTACTGGCGAGTATAGAACGACTCGCCCCAT
TGGCACCAGATACCTGACTCGTCCTCTGTAA

Figure 4: AAV *cap* 5_O sequence. DNA sequence of the sequence optimized AAV5 capsid ($AAV5_O$). The AAP ORF is underlined. Modified nucleotides which lead to the deletion of the *rep* stop codon and an AAP knockout are marked in red.

ATGACCGATGGTTATCTTCCAGATTGGCTCGAGGACAATCTCTCTGAAGGAGTAAGAGAGTGGTGGG
CGCTCCAACCTGGCGCACCAAAACCAAAGGCCAACCAGCAGCATCAGGACAACGCCAGGGGTCTTGT
GCTTCCTGGGTACAAGTACCTCGGACCCGGCAACGGACTCGACAAGGGAGAGCCGGTCAACGCGGCA
GACGCCGCGGCCCTCGAGCACGACAAAGCCTACGACCAGCAGCTCAAGGCCGGAGACAACCCGTACC
TCAAGTACAACCACGCCGACGCGGAGTTTCAGCAGCGCCTTCAAGGAGATACGTCTTTTGGGGGCAA
CCTCGGACGAGCAGTCTTTCAGGCGAAAAAGAGGGTTCTTGAACCTCTGGGCCTGGTTGAGCAAGCT
GGTGAGACGGCTCCGGGAAAAAAGAGGCCGCTAATCGAGTCTCCTCAGCAGCCAGACTCCTCCACGG
GAATCGGAAAGAAGGGCAAGCAGCCTGCAAAAAAAAAATTGGTTTTTGAGGACGAGACTGGCGCAGG
CGACGGACCTCCTGAGGGATCGACATCAGGAGCCATGTCTGACGATAGTGAGATGCGTGCAGCCGCT
GGCGGAGCAGCGGTCGAGGGTGGCCAGGGCGCCGACGGAGTGGGTAATGCCTCGGGAGATTGGCATT
GCGATTCCACATGGAGCGAGGGCCACGTCACCACCACCAGCACCCGAACCTGGGTCCTGCCCACCTA
CAACAACCACCTCTACAAACGACTTGGCGAATCCTTGCAGAGCAATACCTACAATGGCTTCAGCACC
CCTTGGGGGTATTTTGACTTCAACAGATTCCACTGCCACTTTTCACCACGTGACTGGCAAAGACTCA
TCAACAACAACTGGGGAATGCGACCCAAGGCAATGCGGGTCAAGATCTTTAACATTCAAGTCAAAGA
GGTCACGACGAGTAACGGTGAGACGACGGTTGCCAATAACCTTACCAGCACGGTTCAGATCTTTGCT
GACTCGTCGTACGAGCTCCCGTACGTGATGGACGCGGGGCAGGAAGGAAGCCTCCCGCCGTTCCCAA
ACGACGTCTTTATGGTGCCACAGTATGGATACTGCGGCCTGGTCACCGGGAACACAAGCCAGCAGCA
AACAGACCGCAATGCCTTTTACTGCCTGGAGTACTTTCCTTCTCAGATGCTGCGTACCGGAAACAAC
TTTGAGATCACCTACTCTTTTGAGAAGGTTCCTTTCCACAGCATGTACGCTCACAGCCAGAGTCTGG
ACCGTCTCATGAATCCTCTCATCGACCAGTACCTGTGGGGCTTGCAGAGCACAACCACTGGAACTAC
ACTCAACGCGGGAACGGCAACGACTAACTTTACTAAGCTCAGACCGACTAACTTTAGCAACTTCAAG
AAGAACTGGCTTCCTGGACCCTCTATCAAGCAGCAGGGATTCTCAAAGACAGCTAACCAGAACTACA
AGATTCCAGCCACGGGGTCTGACAGCCTAATAAAATACGAGACCCACTCCACCCTCGATGGCAGATG
GTCTGCGCTGACTCCGGGCCCTCCCATGGCAACCGCCGGGCCCGCTGACTCAAAGTTTTCTAACAGC
CAGCTTATCTTCGCTGGGCCGAAACAGAACGGGAACACAGCTACGGTCCCTGGAACGCTCATCTTTA
CAAGCGAAGAGGAACTCGCGGCAACCAATGCCACGGATACGGACATGTGGGGTAATCTACCTGGCGG
CGACCAGAGCAACAGCAACCTACCAACTGTCGACCGTCTCACCGCACTAGGCGCTGTTCCAGGCATG
GTCTGGCAGAACAGAGATATCTACTATCAGGGGCCCATCTGGGCAAAGATTCCACACACGGACGGAC
ATTTTCACCCCTCTCCCCTCATCGGTGGATTCGGACTTAAACACCCTCCTCCACAGATTTTCATCAA
GAACACCCCGGTACCTGCGAATCCTGCGACCACCTTCAGTTCGACACCGGTTAATTCCTTCATCACA
CAGTACTCCACGGGACAGGTCAGCGTGCAGATCGACTGGGAGATCCAGAAGGAAAGAAGCAAACGCT
GGAATCCCGAAGTTCAGTTCACTTCCAACTACGGCCAGCAGAATAGTCTGCTCTGGGCTCCGGACGC
TGCTGGCAAGTATACAGAGCCTCGCGCCATTGGCACCAGATACCTGACTCATCATCTGTAA

Figure 5: AAV *cap* $4_{\text{O repaired}}$ sequence. DNA sequence of the repaired $AAV4_{O}$ capsid ($AAV4_{O\ repaired}$). The AAP ORF is underlined. Nucleotides which were repaired in order to recover AAP expression are marked in blue.

ATGTCTTTCGTTGATCATCCTCCAGATTGGCTCGAGGAGGTTGGTGAAGGACTAAGAGAGTTCTTGG
GGCTCGAAGCTGGCCCACCAAAACCAAAGCCCAACCAGCAGCATCAGGACCAGGCCAGGGGTCTTGT
GCTTCCTGGGTACAACTACCTCGGACCCGGCAACGGACTCGACAGGGGAGAGCCGGTCAACCGGGCA
GACGAGGTGGCCCGCGAGCACGACATATCCTACAACGAGCAGCTCGAGGCCGGAGACAACCCGTACC
TCAAGTACAACCACGCCGACGCGGAGTTTCAGGAGAAGCTCGCAGACGATACGTCTTTTGGGGGCAA
CCTCGGAAAAGCAGTCTTCCAGGCGAAAAAGAGGGTTCTTGAACCTTTCGGCCTGGTTGAGGAAGGT
GCTAAGACGGCTCCGACAGGAAAGAGGATTGACGACCACTTTCCTAAGCGGAAAAAGGCCCGCACGG
AGGAGGACTCGAAGCCTTCAACGTCTAGCGACGCAGAGGCAGGACCTAGCGGATCCCAGCAGCTCCA
AATCCCAGCACAACCCGCCTCTAGTCTGGGAGCTGATACGATGTCTGCAGGCGGTGGCGGACCACTG
GGAGACAATAACCAGGGCGCCGACGGAGTGGGTAATGCCTCGGGAGATTGGCATTGCGATTCCACAT
GGATGGGCGACAGAGTCGTCACCAAGAGCACCCGAACCTGGGTCCTGCCCAGCTACAACAACCACCA
GTACAGAGAAATTAAGAGCGGATCAGTAGACGGCTCGAACGCCAATGCCTACTTTGGCTACAGCACC
CCTTGGGGGTATTTTGACTTCAACAGATTCCACAGCCACTGGTCACCACGTGACTGGCAAAGACTCA
TCAACAACTACTGGGGATTCCGACCCAGGAGCCTCAGAGTCAAGATCTTTAACATTCAAGTCAAAGA
GGTCACGGTGCAGGACAGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTCAGGTGTTTACT
GACGACGACTACCAGCTCCCGTACGTCGTCGGCAACGGGACTGAAGGATGCCTCCCGGCGTTCCCAC
CACAGGTCTTCACGCTGCCACAGTATGGATACGCCACCCTGAACAGAGACAACACGGAGAACCCAAC
AGAACGCTCTTCATTTTTCTGCCTGGAGTACTTTCCTTCTAAGATGCTGCGTACCGGAAACAACTTT
GAGTTCACCTACAATTTTGAGGAGGTTCCTTTCCACAGCAGCTTCGCTCCCAGCCAGAATCTGTTCA
AGCTCGCGAATCCTCTCGTCGACCAGTACCTGTATCGCTTCGTCAGCACAAACAATACAGGTGGAGT
TCAGTTTAATAAGAACCTAGCGGGTCGCTATGCGAACACGTATAAGAACTGGTTTCCTGGACCCATG
GGCCGCACGCAGGGATGGAACCTGGGATCTGGGGTTAACAGAGCCAGTGTATCCGCGTTCGCTACAA
CTAACAGGATGGAGCTCGAGGGCGCAAGCTATCAGGTGCCTCCGCAGCCGAACGGGATGACCAACAA
CCTCCAGGGATCAAACACTTATGCTCTGGAGAACACTATGATCTTTAACAGCCAACCCGCAAACCCA
GGAACAACAGCTACGTACCTTGAAGGGAACATGCTGATTACAAGCGAATCGGAAACCCAGCCAGTCA
ATCGCGTGGCTTACAACGTGGGTGGTCAGATGGCTACCAACAACCAGAGCAGCACCACAGCACCAGC
TACCGGAACTTACAACCTACAAGAGATTGTTCCAGGCAGCGTCTGGATGGAGAGAGATGTGTACCTT
CAGGGGCCCATCTGGGCAAAGATTCCAGAGACGGGCGCACATTTTCACCCCTCTCCCGCCATGGGTG
GATTCGGACTTAAACACCCTCCTCCAATGATGCTCATCAAGAACACCCCGGTACCTGGGAATATCAC
CAGCTTCAGTGACGTACCGGTTTCTTCCTTCATCACACAGTACTCCACGGGACAGGTCACCGTGGAG
ATGGAGTGGGAGCTGAAGAAGGAAAACAGCAAACGCTGGAATCCCGAAATTCAGTACACTAACAACT
ACAACGACCCTCAGTTTGTGGACTTTGCTCCGGACTCTACTGGCGAGTATAGAACGACTCGCCCCAT
TGGCACCAGATACCTGACTCGTCCTCTGTAA

Figure 6: AAV *cap* 5_{O_rep} sequence. DNA sequence of the partially repaired AAV5_O capsid (AAV5_{O_rep}). The AAP $\overline{\text{O}}$RF is underlined. The nucleotide which was repaired in order to recover the *rep* stop codon is marked in blue.

ATGTCTTTCGTTGATCATCCTCCAGATTGGCTCGAGGAGGTTGGTGAAGGACTAAGAGAGTTCTTGG
GGCTCGAAGCTGGCCCACCAAAACCAAAGCCCAACCAGCAGCATCAGGACCAGGCCAGGGGTCTTGT
GCTTCCTGGGTACAACTACCTCGGACCCGGCAACGGACTCGACAGGGGAGAGCCGGTCAACCGGGCA
GACGAGGTGGCCCGCGAGCACGACATATCCTACAACGAGCAGCTCGAGGCCGGAGACAACCCGTACC
TCAAGTACAACCACGCCGACGCGGAGTTTCAGGAGAAGCTCGCAGACGATACGTCTTTTGGGGGCAA
CCTCGGAAAAGCAGTCTTCCAGGCGAAAAAGAGGGTTCTTGAACCTTTCGGCCTGGTTGAGGAAGGT
GCTAAGACGGCTCCGACAGGAAAGAGGATTGACGACCACTTTCCTAAGCGGAAAAAGGCCCGCACGG
AGGAGGACTCGAAGCCTTCAACGTCTAGCGACGCAGAGGCTGGACCTAGCGGATCCCAGCAGCTCCA
AATCCCAGCACAACCCGCCTCTAGTCTGGGAGCTGATACGATGTCTGCAGGCGGTGGCGGACCACTG
GGAGACAATAACCAGGGCGCCGACGGAGTGGGTAATGCCTCGGGAGATTGGCATTGCGATTCCACAT
GGATGGGCGACAGAGTCGTCACCAAGAGCACCCGAACCTGGGTCCTGCCCAGCTACAACAACCACCA
GTACAGAGAAATTAAGAGCGGATCAGTCGACGGCTCGAACGCCAATGCCTACTTTGGCTACAGCACC
CCTTGGGGGTATTTTGACTTCAACAGATTCCACAGCCACTGGTCACCACGTGACTGGCAAAGACTCA
TCAACAACTACTGGGGATTCCGACCCAGGAGCCTCAGAGTCAAGATCTTTAACATTCAAGTCAAAGA
GGTCACGGTGCAGGACAGTACGACGACGATTGCCAATAACCTTACCAGCACGGTTCAGGTGTTTACT
GACGACGACTACCAGCTCCCGTACGTCGTCGGCAACGGGACTGAAGGATGCCTCCCGGCGTTCCCAC
CACAGGTCTTCACGCTGCCACAGTATGGATACGCCACCCTGAAACAGAGACAACACGGAGAACCCAAC
AGAACGCTCTTCATTTTTCTGCCTGGAGTACTTTCCTTCTAAGATGCTGCGTACCGGAAACAACTTT
GAGTTCACCTACAATTTTGAGGAGGTTCCTTTCCACAGCAGCTTCGCTCCCAGCCAGAATCTGTTCA
AGCTCGCGAATCCTCTCGTCGACCAGTACCTGTATCGCTTCGTCAGCACAAACAATACAGGTGGAGT
TCAGTTTAATAAGAACCTAGCGGGTCGCTATGCGAACACGTATAAGAACTGGTTTCCTGGACCCATG
GGCCGCACGCAGGGATGGAACCTGGGATCTGGGGTTAACAGAGCCAGTGTATCCGCGTTCGCTACAA
CTAACAGGATGGAGCTCGAGGGCGCAAGCTATCAGGTGCCTCCGCAGCCGAACGGGATGACCAACAA
CCTCCAGGGATCAAACACTTATGCTCTGGAGAACACTATGATCTTTAACAGCCAACCCGCAAACCCA
GGAACAACAGCTACGTACCTTGAAGGGAACATGCTGATTACAAGCGAATCGGAAACCCAGCCAGTCA
ATCGCGTGGCTTACAACGTGGGTGGTCAGATGGCTACCAACAACCAGAGCAGCACCACAGCACCAGC
TACCGGAACTTACAACCTACAAGAGATTGTTCCAGGCAGCGTCTGGATGGAGAGAGATGTGTACCTT
CAGGGGCCCATCTGGGCAAAGATTCCAGAGACGGGCGCACATTTTCACCCCTCTCCCGCCATGGGTG
GATTCGGACTTAAACACCCTCCTCCAATGATGCTCATCAAGAACACCCCGGTACCTGGGAATATCAC
CAGCTTCAGTGACGTACCGGTTTCTTCCTTCATCACACAGTACTCCACGGGACAGGTCACCGTGGAG
ATGGAGTGGGAGCTGAAGAAGGAAAACAGCAAACGCTGGAATCCCGAAATTCAGTACACTAACAACT
ACAACGACCCTCAGTTTGTGGACTTTGCTCCGGACTCTACTGGCGAGTATAGAACGACTCGCCCCAT
TGGCACCAGATACCTGACTCGTCCTCTGTAA

Figure 7: AAV *cap* $5_{\text{O repaired}}$ sequence. DNA sequence of the repaired AAV5_{O} capsid (AAV$5_{\text{O repaired}}$). The AAP ORF is underlined. Nucleotides which were repaired in order to recover AAP expression as well as the *rep* stop codon are marked in blue.

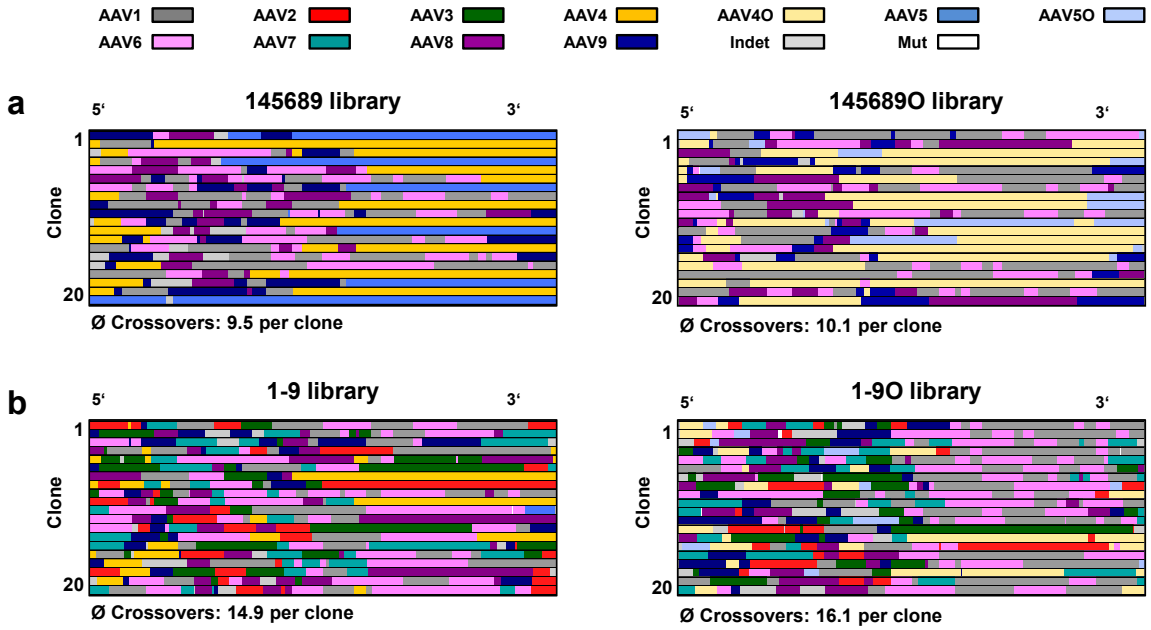

Figure 8: DNA type assignments of libraries with AAV *cap* 4 / 4_O and AAV *cap* 5 / 5_O. **a)** DNA type assignments of libraries 145689 and 145689O. **b)** DNA type assignments of libraries 1-9 and 1-9O. Per library 20 clones were analyzed with the Salanto program. Average crossover numbers per clone within the library are depicted below.

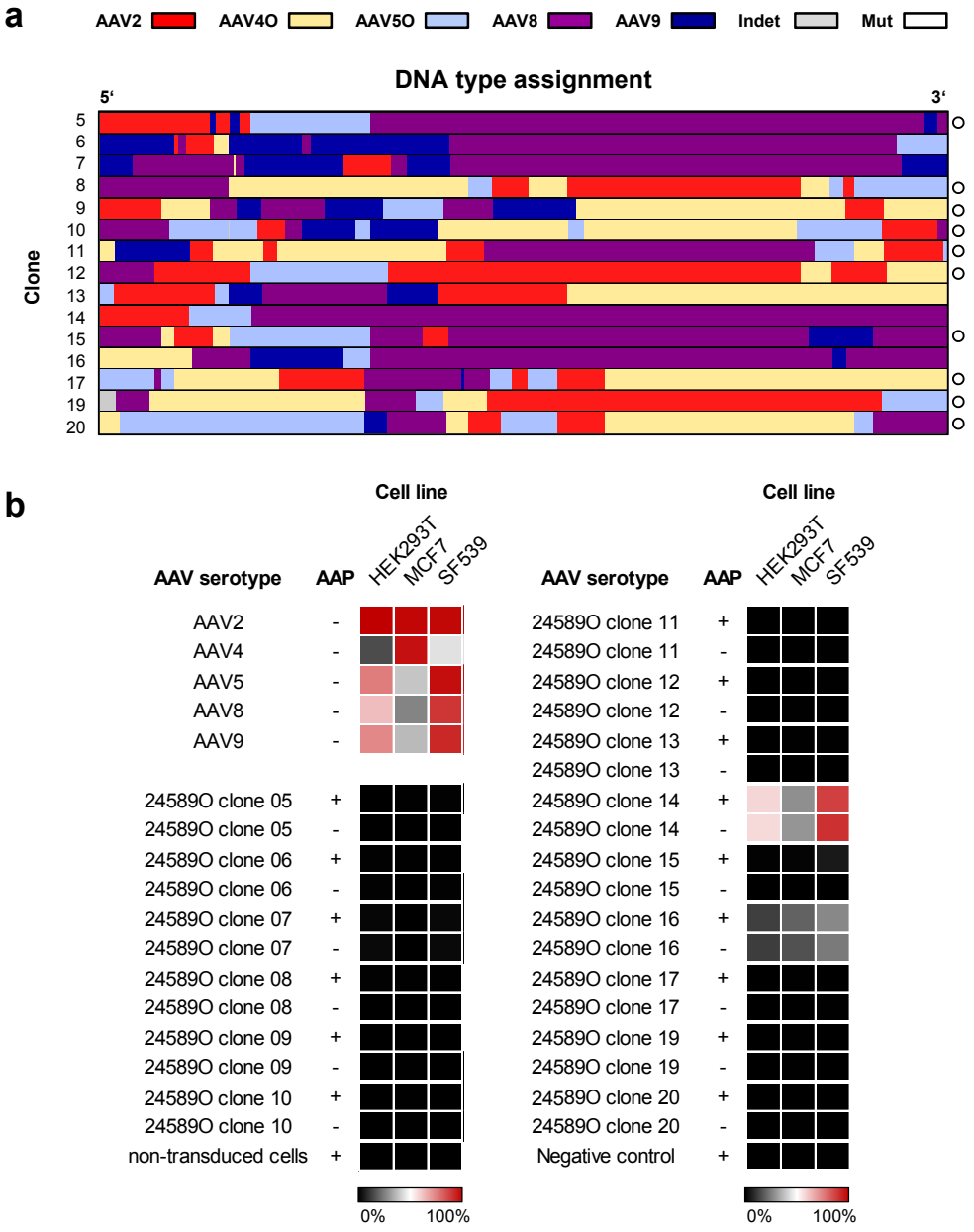

Figure 9: Test of 15 individual, shuffled 24589O clones with or without AAP supplementation. a) DNA type assignment of all tested 24589O library clones. Clone numbers are depicted on the left side, circles on the right side mark clones which carry non-functional AAPs. Clones were analyzed with the Salanto program and are already depicted in figure 3.6. **b)** Heat-map of AAV transduction with the tested 24589O library clones in three different cell lines. Crude cell lysates were produced by transfection of HEK293T cells with the adeno-helper construct, a fluorescence reporter and a *rep-cap* helper encoding for either the wild-type AAV capsids 2, 4, 5, 8 or 9 or one of the shuffled library clones. In additon, either an AAP mixture containing AAP2, 4, 5, 8, and 9 or stuffer (salmon sperm) DNA was added during transfection (indicated with '+' or '-', respectively). Next, HEK293T, MCF7 or SF539 cells were transduced with 5 µl crude lysate and analyzed via flow cytometry 48 hours later. Percentages of positive cells are color-coded from 0 % (black) to 100 % (dark red). Out of 15 tested shuffled clones only two (clone no. 14 and 16) were able to transduce cells. No clone was restored upon AAP supplementation.

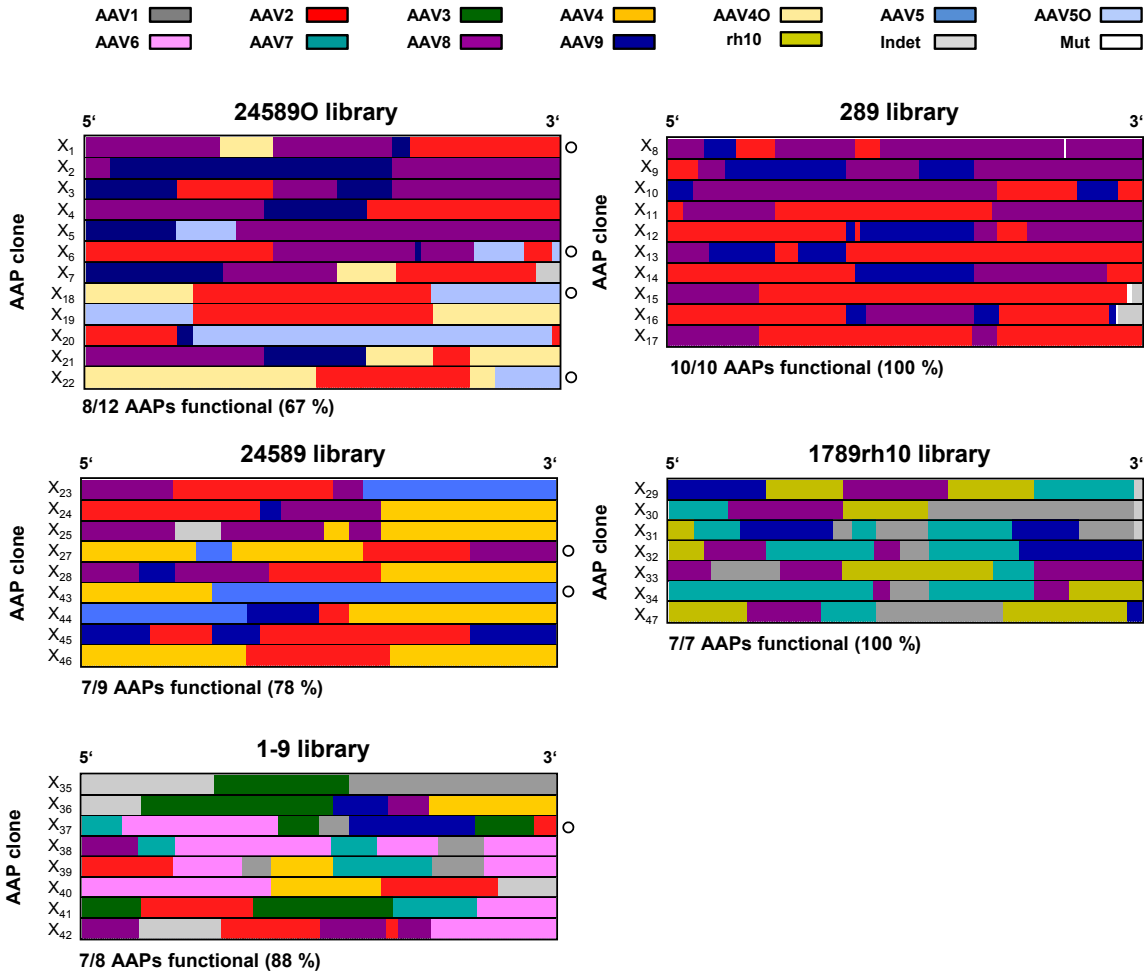

Figure 10: Protein type assignments of all cloned shuffled AAPs. Colored protein type assignment of AAP-X_1 – AAP-X_{47} which were randomly selected from five different shuffled AAV *cap* libraries (depicted on top of each assignment). In total 39/46 (84.8 %) of the tested AAPs were functional in *trans*-complementation assays together with AAV2mut or $AAV5_{O\_rep}$. Non-functional AAPs are marked with a circle. Sequences were analyzed with the Salanto program.

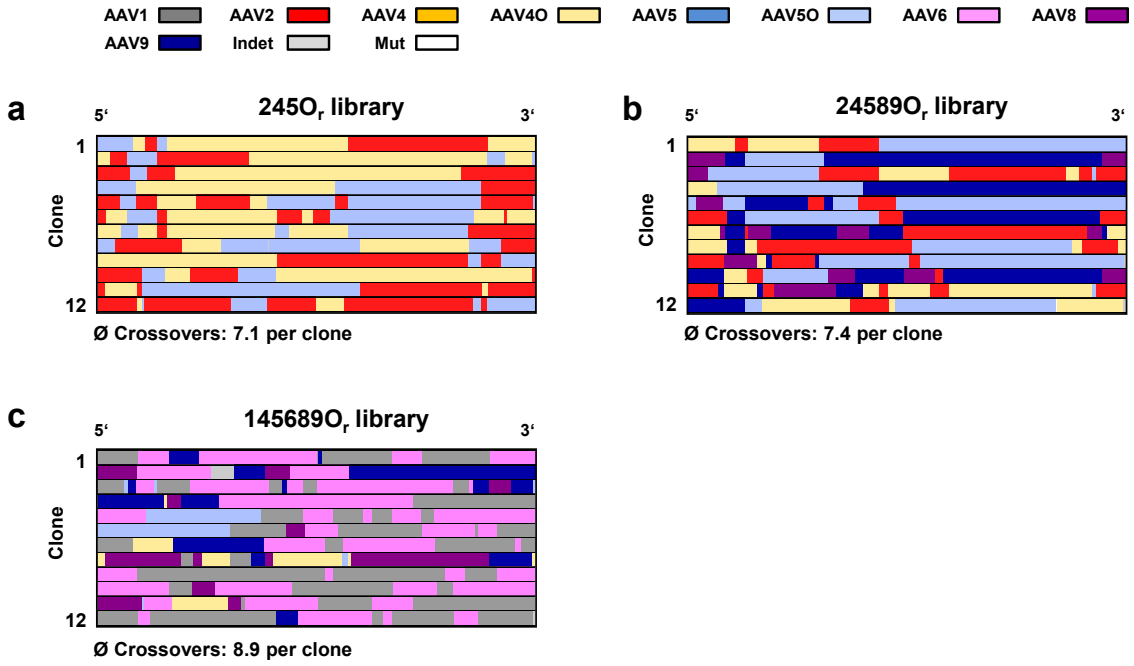

Figure 11: DNA type assignments of *cap* libraries containing the repaired (r) versions of *cap* 4_O and 5_O. Colored DNA type assignment of $245O_r$ **(a)**, $24589O_r$ **(b)** and $145689O_r$ **(c)** AAV *cap* libraries. Twelve randomly selected clones per library were analyzed with the Salanto program. Average crossover numbers per clone within the library are depicted below.

2 Supplementary tables

Table 1: Microsoft excel macro for precise type assignment coloring

Protocol step
Sub color_precise_type_assignment_Salanto()
Dim Zelle As Range
For Each Zelle In ActiveSheet.UsedRange
If Zelle.Value = ">AAV1" Then
Zelle.Interior.ColorIndex = 16
End If
If Zelle.Value = ">AAV1" Then
Zelle.Interior.TintAndShade = 0.2
End If
If Zelle.Value = ">AAV2" Then
Zelle.Interior.ColorIndex = 3
End If
If Zelle.Value = ">AAV2" Then
Zelle.Interior.TintAndShade = 0.1
End If
If Zelle.Value = ">AAV3" Then
Zelle.Interior.ColorIndex = 51
End If
If Zelle.Value = ">AAV3" Then
Zelle.Interior.TintAndShade = 0.1
End If
If Zelle.Value = ">AAV4" Then
Zelle.Interior.ColorIndex = 44
End If
If Zelle.Value = ">AAV5" Then
Zelle.Interior.ColorIndex = 41
End If
If Zelle.Value = ">AAV5" Then
Zelle.Interior.TintAndShade = 0.1
End If
If Zelle.Value = ">AAV6" Then
Zelle.Interior.ColorIndex = 21
End If
If Zelle.Value = ">AAV6" Then
Zelle.Interior.TintAndShade = 0.7
End If
If Zelle.Value = ">AAV7" Then
Zelle.Interior.ColorIndex = 31
End If
If Zelle.Value = ">AAV7" Then
Zelle.Interior.TintAndShade = 0.1
End If

Protocol step
If Zelle.Value = ">AAV8" Then
Zelle.Interior.ColorIndex = 29
End If
If Zelle.Value = ">AAV8" Then
Zelle.Interior.TintAndShade = 0.02
End If
If Zelle.Value = ">AAV9" Then
Zelle.Interior.ColorIndex = 11
End If
If Zelle.Value = ">AAV9" Then
Zelle.Interior.TintAndShade = 0.1
End If
If Zelle.Value = ">AAVrh10" Then
Zelle.Interior.ColorIndex = 52
End If
If Zelle.Value = ">AAVrh10" Then
Zelle.Interior.TintAndShade = 0.3
End If
If Zelle.Value = ">AAVpo.1" Then
Zelle.Interior.ColorIndex = 10
End If
If Zelle.Value = ">AAVpo.1" Then
Zelle.Interior.TintAndShade = 0.1
End If
If Zelle.Value = ">AAV12" Then
Zelle.Interior.ColorIndex = 43
End If
If Zelle.Value = ">AAV12" Then
Zelle.Interior.TintAndShade = 0.1
End If
If Zelle.Value = ">AAV4O" Then
Zelle.Interior.ColorIndex = 44
End If
If Zelle.Value = ">AAV4O" Then
Zelle.Interior.TintAndShade = 0.6
End If
If Zelle.Value = ">AAV5O" Then
Zelle.Interior.ColorIndex = 41
End If
If Zelle.Value = ">AAV5O" Then
Zelle.Interior.TintAndShade = 0.6
End If
If Zelle.Value = "indet" Then
Zelle.Interior.ColorIndex = 15
End If
If Zelle.Value = "indet" Then
Zelle.Interior.TintAndShade = 0.2

Protocol step
End If
If Zelle.Value = "MUT" Then
Zelle.Interior.ColorIndex = 2
End If
Next Zelle
End Sub

Table 2: CellProfiler protocol for the microscopy picture quantification of nuclei and cells positively stained with the B1- or A20-antibody.

Protocol step	Setting
CellProfiler	Version,2014-07-23T17:45:00 6c2d896
Images	
Filter images?:	No filtering
Metadata	
Extract metadata?:	No
NamesAndTypes	
Assign a name to:	Images matching rules
Select the rule criteria:	and (file does contain "dapi.tif"")
Name to assign these images:	DNA
Select the image type:	Grayscale image
Set intensity range from:	Image metadata
Select the rule criteria:	and (file does contain "TxR.tif"")
Name to assign these images:	TexasRed
Select the image type:	Grayscale image
Set intensity range from:	Image metadata
Crop	
Select the input image:	DNA
Name the output image:	CropBlue
Select the cropping shape:	Rectangle
Select the cropping method:	Coordinates
Apply which cycleś cropping pattern?:	First
Left and right rectangle positions:	1, 200
Top and bottom rectangle positions:	1, 200
Remove empty rows and columns?:	Edges
IdentifyPrimaryObjects:	
Select the input image:	DNA
Name the primary objects to be identified:	Nuclei
Typical diameter of objects, in pixel units (Min,Max):	11,40
Discard objects outside the diameter range?:	Yes
Try to merge too small objects with nearby larger objects?:	No
Discard objects touching the border of the image?:	Yes
Threshold strategy:	Global
Thresholding method:	Background
Select the smoothing method for thresholding:	Automatic

Protocol step	Setting
Threshold correction factor:	1
Lower and upper bounds on threshold:	0,1
Method to distinguish clumped objects:	Intensity
Method to draw dividing lines between clumped objects:	Intensity
Automatically calculate size of smoothing filter for declumping?:	Yes
Automatically calculate minimum allowed distance between local maxima?:	Yes
Speed up by using lower-resolution image to find local maxima?:	Yes
Retain outlines of the identified objects?:	Yes
Name the outline image:	NucOutlines
Fill holes in identified objects?:	After both thresholding and declumping
Handling of objects if excessive number of objects identified:	Continue
IdentifyPrimaryObjects:	
Select the input image:	TexasRed
Name the primary objects to be identified:	B1_or_A20_positive_cells
Typical diameter of objects, in pixel units (Min,Max):	11,80
Discard objects outside the diameter range?:	Yes
Try to merge too small objects with nearby larger objects?:	No
Discard objects touching the border of the image?:	Yes
Threshold strategy:	Global
Thresholding method:	Background
Select the smoothing method for thresholding:	Automatic
Threshold correction factor:	1
Lower and upper bounds on threshold:	0,1
Method to distinguish clumped objects:	Intensity
Method to draw dividing lines between clumped objects:	Intensity
Automatically calculate size of smoothing filter for declumping?:	Yes
Automatically calculate minimum allowed distance between local maxima?:	Yes
Speed up by using lower-resolution image to find local maxima?:	Yes
Retain outlines of the identified objects?:	No
Name the outline image:	PrimaryOutlines
Fill holes in identified objects?:	After both thresholding and declumping
Handling of objects if excessive number of objects identified:	Continue
MeasureObjectIntensity:	
Select an image to measure:	TexasRed
Select objects to measure:	B1_or_A20_positive_cells
ExportToSpreadsheet:	
Select the column delimiter:	Comma ("","")
Output file location:	Default Input Folder sub-folder
Add a prefix to file names?:	No
Overwrite without warning?:	Yes
Add image metadata columns to your object data file?:	No
Limit output to a size that is allowed in Excel?:	No
Representation of Nan/Inf:	NaN

Protocol step	Setting
Calculate the per-image mean values for object measurements?:	Yes
Calculate the per-image median values for object measurements?:	Yes
Calculate the per-image standard deviation values for object measurements?:	Yes
Create a GenePattern GCT file?:	No
Export all measurement types?:	Yes

Table 3: CellProfiler protocol for the microscopy picture quantification of cells positively stained with A20- and anti-AAP2-antibodies.

Protocol step	Setting
CellProfiler	Version,2014-07-23T17:45:00 6c2d896
Images	
Filter images?:	No filtering
Metadata	
Extract metadata?:	No
NamesAndTypes	
Assign a name to:	Images matching rules
Select the rule criteria:	and (file does contain "dapi.tif"")
Name to assign these images:	DNA
Select the image type:	Grayscale image
Set intensity range from:	Image bit-depth
Select the rule criteria:	and (file does contain "AF488.tif"")
Name to assign these images:	Green
Select the image type:	Grayscale image
Set intensity range from:	Image metadata
Select the rule criteria:	and (file does contain "AF568.tif"")
Name to assign these images:	Red
Select the image type:	Grayscale image
Set intensity range from:	Image metadata
CorrectIlluminationCalculate	
Select the input image:	Red
Name the output image:	IllumRed
Select how the illumination function is calculated:	Background
Block size:	60
Rescale the illumination function?:	Yes
Calculate function for each image individually, or based on all images?:	Each
Smoothing method:	No smoothing
Retain the averaged image?:	No
Retain the dilated image?:	No
CorrectIlluminationApply:	
Select the input image:	Red

Protocol step	Setting
Name the output image:	CorrRed
Select the illumination function:	IllumRed
Select how the illumination function is applied:	Divide
IdentifyPrimaryObjects:	
Select the input image:	DNA
Name the primary objects to be identified:	Nuclei
Typical diameter of objects, in pixel units (Min,Max):	11,40
Discard objects outside the diameter range?:	Yes
Try to merge too small objects with nearby larger objects?:	No
Discard objects touching the border of the image?:	Yes
Threshold strategy:	Global
Thresholding method:	Background
Select the smoothing method for thresholding:	Automatic
Threshold correction factor:	1
Lower and upper bounds on threshold:	0,1
Method to distinguish clumped objects:	Intensity
Method to draw dividing lines between clumped objects:	Intensity
Automatically calculate size of smoothing filter for declumping?:	Yes
Automatically calculate minimum allowed distance between local maxima?:	Yes
Speed up by using lower-resolution image to find local maxima?:	No
Retain outlines of the identified objects?:	Yes
Name the outline image:	NucOutlines
Fill holes in identified objects?:	After both thresholding and declumping
Handling of objects if excessive number of objects identified:	Continue
IdentifyPrimaryObjects:	
Select the input image:	Green
Name the primary objects to be identified:	B1_or_A20_positive_cells
Typical diameter of objects, in pixel units (Min,Max):	10,50
Discard objects outside the diameter range?:	Yes
Try to merge too small objects with nearby larger objects?:	No
Discard objects touching the border of the image?:	Yes
Threshold strategy:	Global
Thresholding method:	Background
Select the smoothing method for thresholding:	Automatic
Threshold correction factor:	1
Lower and upper bounds on threshold:	0,1
Method to distinguish clumped objects:	Shape
Method to draw dividing lines between clumped objects:	Intensity
Automatically calculate size of smoothing filter for declumping?:	Yes
Automatically calculate minimum allowed distance between local maxima?:	Yes
Speed up by using lower-resolution image to find local maxima?:	Yes
Retain outlines of the identified objects?:	No
Name the outline image:	PrimaryOutlines
Fill holes in identified objects?:	Never

Protocol step	Setting
Handling of objects if excessive number of objects identified:	Continue
IdentifyPrimaryObjects:	
Select the input image:	CorrRed
Name the primary objects to be identified:	AAP
Typical diameter of objects, in pixel units (Min,Max):	10,50
Discard objects outside the diameter range?:	Yes
Try to merge too small objects with nearby larger objects?:	No
Discard objects touching the border of the image?:	Yes
Threshold strategy:	Global
Thresholding method:	Background
Select the smoothing method for thresholding:	Automatic
Threshold correction factor:	1.0
Lower and upper bounds on threshold:	0.0,1.0
Method to distinguish clumped objects:	Shape
Method to draw dividing lines between clumped objects:	Intensity
Automatically calculate size of smoothing filter for declumping?:	Yes
Automatically calculate minimum allowed distance between local maxima?:	Yes
Speed up by using lower-resolution image to find local maxima?:	Yes
Retain outlines of the identified objects?:	No
Fill holes in identified objects?:	Never
Handling of objects if excessive number of objects identified:	Continue
ExportToSpreadsheet:	
Select the column delimiter:	Comma ("","")
Output file location:	Default Input Folder sub-folder
Add a prefix to file names?:	No
Overwrite without warning?:	Yes
Add image metadata columns to your object data file?:	No
Limit output to a size that is allowed in Excel?:	No
Representation of Nan/Inf:	NaN
Calculate the per-image mean values for object measurements?:	No
Calculate the per-image median values for object measurements?:	No
Calculate the per-image standard deviation values for object measurements?:	No
Create a GenePattern GCT file?:	No
Export all measurement types?:	Yes

Table 4: CellProfiler protocol for the microscopy picture quantification of cells positively stained with B1- and anti-AAP2-antibodies.

Protocol step	Setting
CellProfiler	Version,2014-07-23T17:45:00 6c2d896
Images	
Filter images?:	No filtering
Metadata	

Protocol step	Setting
Extract metadata?:	No
NamesAndTypes	
Assign a name to:	Images matching rules
Select the rule criteria:	and (file does contain "dapi.tif"")
Name to assign these images:	DNA
Select the image type:	Grayscale image
Set intensity range from:	Image bit-depth
Select the rule criteria:	and (file does contain "AF488.tif"")
Name to assign these images:	Green
Select the image type:	Grayscale image
Set intensity range from:	Image metadata
Select the rule criteria:	and (file does contain "AF568.tif"")
Name to assign these images:	Red
Select the image type:	Grayscale image
Set intensity range from:	Image metadata
CorrectIlluminationCalculate	
Select the input image:	Red
Name the output image:	IllumRed
Select how the illumination function is calculated:	Background
Block size:	60
Rescale the illumination function?:	Yes
Calculate function for each image individually, or based on all images?:	Each
Smoothing method:	No smoothing
Retain the averaged image?:	No
Retain the dilated image?:	No
CorrectIlluminationApply:	
Select the input image:	Red
Name the output image:	CorrRed
Select the illumination function:	IllumRed
Select how the illumination function is applied:	Divide
IdentifyPrimaryObjects:	
Select the input image:	DNA
Name the primary objects to be identified:	Nuclei
Typical diameter of objects, in pixel units (Min,Max):	11,40
Discard objects outside the diameter range?:	Yes
Try to merge too small objects with nearby larger objects?:	No
Discard objects touching the border of the image?:	Yes
Threshold strategy:	Global
Thresholding method:	Background
Select the smoothing method for thresholding:	Automatic
Threshold correction factor:	1
Lower and upper bounds on threshold:	0,1
Method to distinguish clumped objects:	Intensity
Method to draw dividing lines between clumped objects:	Intensity

Protocol step	Setting
Automatically calculate size of smoothing filter for declumping?:	Yes
Automatically calculate minimum allowed distance between local maxima?	:Yes
Speed up by using lower-resolution image to find local maxima?:	No
Retain outlines of the identified objects?:	Yes
Name the outline image:	NucOutlines
Fill holes in identified objects?:	After both thresholding and declumping
Handling of objects if excessive number of objects identified:	Continue
IdentifyPrimaryObjects:	
Select the input image:	Green
Name the primary objects to be identified:	B1_or_A20_positive_cells
Typical diameter of objects, in pixel units (Min,Max):	10,50
Discard objects outside the diameter range?:	Yes
Try to merge too small objects with nearby larger objects?:	No
Discard objects touching the border of the image?:	Yes
Threshold strategy:	Global
Thresholding method:	Background
Select the smoothing method for thresholding:	Automatic
Threshold correction factor:	1
Lower and upper bounds on threshold:	0,1
Method to distinguish clumped objects:	Shape
Method to draw dividing lines between clumped objects:	Intensity
Automatically calculate size of smoothing filter for declumping?:	Yes
Automatically calculate minimum allowed distance between local maxima?:	Yes
Speed up by using lower-resolution image to find local maxima?:	Yes
Retain outlines of the identified objects?:	No
Name the outline image:	PrimaryOutlines
Fill holes in identified objects?:	Never
Handling of objects if excessive number of objects identified:	Continue
IdentifyPrimaryObjects:	
Select the input image:	CorrRed
Name the primary objects to be identified:	AAP
Typical diameter of objects, in pixel units (Min,Max):	10,50
Discard objects outside the diameter range?:	Yes
Try to merge too small objects with nearby larger objects?:	No
Discard objects touching the border of the image?:	Yes
Threshold strategy:	Per object
Thresholding method:	MoG
Masking objects:	B1_or_A20_positive_cells
Approximate fraction of image covered by objects?:	0.01
Select the smoothing method for thresholding:	Automatic
Threshold correction factor:	1.0
Lower and upper bounds on threshold:	0.0,1.0
Method to distinguish clumped objects:	Shape
Method to draw dividing lines between clumped objects:	Intensity

Protocol step	Setting
Automatically calculate size of smoothing filter for declumping?:	Yes
Automatically calculate minimum allowed distance between local maxima?:	Yes
Speed up by using lower-resolution image to find local maxima?:	Yes
Retain outlines of the identified objects?:	No
Fill holes in identified objects?:	Never
Handling of objects if excessive number of objects identified:	Continue
ExportToSpreadsheet:	
Select the column delimiter:	Comma ("","")
Output file location:	Default Input Folder sub-folder
Add a prefix to file names?:	No
Overwrite without warning?:	Yes
Add image metadata columns to your object data file?:	No
Limit output to a size that is allowed in Excel?:	No
Representation of Nan/Inf:	NaN
Calculate the per-image mean values for object measurements?:	No
Calculate the per-image median values for object measurements?:	No
Calculate the per-image standard deviation values for object measurements?:	No
Create a GenePattern GCT file?:	No
Export all measurement types?:	Yes

Table 5: Microsoft excel macro for flow cytometry data extraction

Protocol step
Sub macro_FACS
Dim i As Integer
Dim j As Integer
Sheets.Add
ActiveSheet.Name = "result"
For i = 1 To 12
For j = 1 To 8
Worksheets("result").Cells(j, i).Value = Worksheets("Result 1").Cells(9 + ((i - 1) + (j - 1) * 12) * 7, 8)
Next
Next
Sheets.Add
ActiveSheet.Name = "result_positive cells"
For i = 1 To 96
'For j = 1 To 8
Worksheets("result_positive cells").Cells(i, 1).Value = Worksheets("Result 1").Cells(9 + ((i - 1)) * 7, 8)
'Next
Next
Sheets.Add
ActiveSheet.Name = "result_x-mean"
For i = 1 To 96
For j = 1 To 8

Protocol step
Worksheets("result_x-mean").Cells(i, 1).Value = Worksheets("Result 1").Cells(9 + ((i - 1)) * 7, 10)
Next
Next
Sheets.Add
ActiveSheet.Name = "result_x-median"
For i = 1 To 96
For j = 1 To 8
Worksheets("result_x-median").Cells(i, 1).Value = Worksheets("Result 1").Cells(9 + ((i - 1)) * 7, 9)
Next
Next
End Sub

Danksagung

An dieser Stelle möchte ich als Erstes meinem Betreuer Dr. Dirk Grimm dafür danken, mir die Arbeit an diesem Projekt, in seinem Labor und generell im spannenden Gebiet der Gentherapie ermöglicht zu haben. Durch sein Engagement war es mir möglich, nicht nur mein Thema erfolgreich zu bearbeiten, sondern auch Einblicke in RNAi und TuDs, HBV, HCV sowie CRISPR/Cas zu erlangen bzw. an entsprechenden Projekten essentiell mitzuwirken. Seine Unterstützung, sein Ideenreichtum sowie sein exzellentes Hintergrundwissen im Bereich der AAV-basierten Gentherapie haben erheblich zum Voranschreiten der vorliegenden Arbeit beigetragen. Gleichzeitig konnte ich meine Ideen frei verwirklichen und an Konferenzen teilnehmen, ohne mir Sorgen um die entsprechende Finanzierung machen zu müssen. Zudem bin ich sehr dankbar für die hilfreichen Kommentare während der Anfertigung dieser Arbeit.
Des Weiteren danke ich Prof. Martin Müller für seine Bereitschaft meine Arbeit als Erstgutachter zu betreuen sowie Prof. Oliver Fackler für seine Teilnahme an meinen jährlichen TAC-Meetings. Beide haben mit ihren Vorschlägen zur Realisierung meines Projektes beigetragen. Einen zusätzlichen Dank an Frauke Melchior, die sich bereit erklärt hat, den Vorsitz über mein Prüfungskomitee zu übernehmen.
Außerdem danke ich Boehringer Ingelheim für die Finanzierung der Sequenz-Optimierung der AAV4 sowie AAV5 Kapside, ohne die diese Arbeit nicht hätte begonnen werden können.
Bei meiner Graduiertenschule HBIGS (insbesondere Rolf Lutz und Sandra Martini) möchte ich mich für die Hilfe und Unterstützung während meiner Doktorandenzeit sowie die Bereitstellung der exzellenten Weiterbildungsmöglichkeiten bedanken.
Ein sehr großes „Dankeschön“ geht an alle (alten und neuen) Mitglieder der AG Grimm, die den Laboralltag um so vieles bereichert und erleichtert haben. Die Stimmung sowie der Zusammenhalt im Labor sind einzigartig und konnten mir in schwierigen experimentellen Phasen immer Unterstützung geben. In diesem Zusammenhang möchte ich besonders den Labormitgliedern der „ersten Stunde“ danken: Ellen Wiedtke für die Einführung und weitere Hilfe während meiner Virusproduktionen sowie unsere geteilte Vernarrtheit in Katzen; Nina Schürmann für jeglichen Beistand in sämtlichen Situationen und die Beantwortung aller Fragen; Stefan Mockenhaupt für die Einführung in RNAi und die Möglichkeit sein Projekt abschließen zu dürfen sowie Eike Kienle für die Einweihung in AAV-Shuffling und die Unterstützung zu Beginn meiner Arbeit. Weiterhin danke ich speziell Elena Senis Herrero für die wunderbare Zeit im Labor, die Ermunterungen und Hilfe bei Fragen & Problemen, die gemeinsamen & unvergesslichen Reisen nach Madrid sowie Südafrika und für alle spanisch-deutschen (kulinarischen) Erlebnisse! Anne-Kathrin Herrmann, Franziska Hentzschel und nochmals Nina Schürmann schulde ich neben der unterhaltsamen und inspirierenden Gesellschaft im Labor besonderen Dank für die lustigen Spieleabende. Dominik Niopek danke ich für seinen Beistand in allen Lebenslagen über seine aktive Zeit in unserem Labor hinaus. Neben allen weiteren Labormitgliedern möchte ich noch Julia Fakhiri, Jonas Weinmann, Ahmed Boumezbeur, Ma-

rio Lederle, Silke Uhrig-Schmidt, Chronis Fatouros, Kathrin Tegeler, Benjamin Kachel und Florian Schmidt erwähnen, die essentiell zur tollen Arbeitsatmosphäre beigetragen haben. Außerdem bin ich meinen ehemaligen Praktikanten Volker Klöss, Sina Stäble, Katarina Schmidt, Sarah Klinnert sowie Jannik Traut zu Dank verpflichtet. Insbesondere gegen Ende meiner Arbeit leisteten sie mir wertvolle Unterstützung bei der Generierung von wichtigen Daten und konnten dabei hoffentlich nützliches Wissen für ihr weiteres Studium erwerben.
Außerhalb des Labors gilt mein Dank Christian Bender für die unermüdliche und ehrgeizige Weiterentwicklung von Salanto, ohne die die vorliegende Arbeit in dieser Form nicht existieren würde. Katharina Höhn danke ich für den Versuch, ungereinigte AAVs im Elektronenmikroskop darzustellen, auch wenn diese Ergebnisse hier leider nicht gezeigt werden. Weiterhin haben mir Volkan Sakin und Andreas Werner wertvolle Hinweise zum Nachweis post-translationaler Modifikationen gegeben, auch wenn leider keiner davon bisher zum Erfolg geführt hat. Andrea Imle danke ich für die Unterstützung aus dem „Virologie-Hauptlabor“ und die Gabe verschiedenster Antikörper-Aliquote.
Zudem möchte ich Monika, Ivonne und Peter für die dauerhafte Entlastung unseres Laboralltags danken – insbesondere für ihre Bereitschaft, wenn wieder kurzfristig etwas autoklaviert werden musste. Weiterhin verdient Daniel Brown meinen Dank für die Lösung aller IT-Fragen und Probleme, die im Laufe der Jahre aufgetreten sind.
Mein tiefster Dank aber gilt meinem Freund René für die größtmögliche Unterstützung während meiner ganzen Zeit als PhD-Studentin. Du hast mir immer wieder Halt und Kraft gegeben sowie jede Situation mit mir durchgestanden. Danke für die kontinuierliche Motivation und Liebe. Ich freue mich auf unsere gemeinsame Zukunft!
Außerdem bin ich meiner Familie unendlich dankbar, die mir nicht nur meine Doktorarbeit, sondern auch mein Bachelor- und Masterstudium ermöglicht hat. Danke, Mama und Papa, für eure Liebe, Hilfe und den Glauben an mich!