

Charakterisierung der *in vivo*-Selektion und  
Optimierung von auf Adeno-assoziierten Viren  
exprimierten randomisierten  
Peptidbibliotheken

Dissertation

zur Erlangung des Doktorgrades im Fachbereich Chemie  
der Fakultät für Mathematik, Informatik und Naturwissenschaften  
der Universität Hamburg

vorgelegt von

Melanie Lampe

Hamburg

2016



Gutachter:

Prof. Dr. med. Martin Trepel

Prof. Dr. med. Elke Oetjen

Tag der Disputation:

14.10.2016

Tag der Druckfreigabe:

28.10.2016

Die vorliegende Arbeit wurde im Zeitraum von Januar 2012 bis zum März 2016 in der Arbeitsgruppe Receptortargeting von Prof. Dr. Martin Trepel am Universitätsklinikum Hamburg-Eppendorf durchgeführt.

*„Nur wenige wissen, wie viel man wissen muss, um zu wissen, wie wenig man weiß.“*

*- Werner Heisenberg -*

# Inhaltsverzeichnis

<b>I Abkürzungsverzeichnis .....</b>	<b>i</b>
<b>II Abbildungsverzeichnis .....</b>	<b>iii</b>
<b>III Tabellenverzeichnis .....</b>	<b>iv</b>
<b>Zusammenfassung.....</b>	<b>1</b>
<b>Abstract .....</b>	<b>3</b>
<b>1 EINLEITUNG.....</b>	<b>5</b>
1.1 Gentherapie und Gentherapievektoren.....	5
1.2 Adeno-assoziierte Viren (AAV) .....	5
1.2.1 Das AAV-Genom .....	6
1.2.2 AAV-Kapsidstruktur und zelluläre Rezeptoren.....	7
1.2.3 Transduktion und Replikationszyklus.....	8
1.3 AAV als Gentherapievektor .....	10
1.3.1 Klinische Studien .....	10
1.3.2 Produktion von rekombinanten AAV (rAAV) für gentherapeutische Ansätze .....	10
1.3.3 AAV2-Tropismus.....	11
1.3.4 Gewebespezifische Promotoren und AAV-Serotypen .....	11
1.3.5 Modifikation des AAV-Kapsids .....	12
1.3.6 Randomisierte AAV-Peptidbibliotheken ( <i>random AAV display peptide libraries</i> ) .....	13
1.3.7 <i>In vitro</i> - und <i>in vivo</i> -Selektion von randomisierten AAV-Peptidbibliotheken .....	15
1.3.8 Die Niere und das Pankreas als gentherapeutisches Ziel.....	17
1.4 Next-Generation Sequencing zur Auswertung der Selektion von AAV-Peptidbibliotheken.....	17
1.4.1 DNS-Sequenzierung der ersten, zweiten und dritten Generation .....	17
1.4.2 Illumina® Next-Generation Sequencing .....	18
1.4.3 Next-Generation Sequencing zur Untersuchung von Peptidbibliotheken.....	19
1.5 Zielsetzung der Arbeit .....	21
<b>2 MATERIAL .....</b>	<b>22</b>
2.1 Laborgeräte .....	22
2.2 Molekularbiologische Kits .....	23
2.3 Reagenzien .....	23
2.4 Enzyme .....	23
2.5 Antikörper.....	24
2.6 Chemikalien .....	24
2.7 DNS-Marker .....	24
2.8 Verbrauchsmaterialien .....	24
2.9 Lösungen .....	24

---

2.9.1 Standardpuffer .....	24
2.10 Medien .....	25
2.10.1 Medien zur Kultivierung von prokaryotischen Zellen .....	25
2.10.2 Medien zur Kultivierung von eukaryotischen Zellen .....	26
2.11 Zellen .....	26
2.11.1 Prokaryotische Zellen ( <i>Escherichia coli</i> ) .....	26
2.11.2 Eukaryotische Zelllinien .....	26
2.12 Antibiotika und Zusätze .....	27
2.13 Oligonukleotide .....	27
2.14 Mäuse .....	28
2.15 Viren .....	29
2.16 Rekombinante AAV2 .....	29
2.17 Plasmide .....	29
2.18 Software und Datenbanken .....	30
2.19 Unternehmen .....	30
<b>3 METHODEN .....</b>	<b>31</b>
3.1 Molekularbiologische Methoden .....	31
3.1.1 Isolierung von Plasmid-DNS .....	31
3.1.2 DNS-Quantifizierung .....	32
3.1.3 Enzymatischer Verdau von DNS mittels Restriktionsendonukleasen .....	32
3.1.4 Agarose Gelelektrophorese .....	32
3.1.5 DNS-Gelextraktion .....	33
3.1.6 DNS-Sequenzierung .....	33
3.1.7 Ligation von DNS-Fragmenten .....	33
3.1.8 DNS-Präzipitation mit Ethanol .....	34
3.1.9 Polymerase-Kettenreaktion ( <i>polymerase chain reaction</i> , PCR) .....	34
3.1.9.1 PCR zur Amplifikation von viraler DNS .....	34
3.1.9.2 Kolonie-PCR (cPCR) .....	35
3.1.9.3 Quantitative Echtzeit PCR (q-PCR) .....	36
3.1.10 Next-Generation Sequencing (NGS) .....	37
3.1.10.1 DNS-Isolierung .....	38
3.1.10.2 Amplifikation der viralen das randomisierte Insert beinhaltenden DNS - erste PCR ..	38
3.1.10.3 Anfügen der Linker - zweite PCR .....	38
3.1.10.4 Anfügen der Illumina®-Adaptersequenzen mit Barcode - dritte PCR .....	39
3.1.10.5 Bioinformatische Auswertung der NGS-Daten .....	40

3.1.10.6 Berechnung und Bewertung von Peptidanreicherung und -spezifität zur Identifizierung von potenziell gewebespezifischen AAV-Vektoren .....	40
3.2 Mikrobiologische Methoden .....	42
3.2.1 Kultivierung von <i>E.coli</i> .....	42
3.2.1.1 Suspensionskultur .....	42
3.2.1.2 Kultivierung auf Agarplatten .....	42
3.2.1.3 Lagerung in Glycerin .....	42
3.2.2 Herstellung von kompetenten <i>E. coli</i> .....	42
3.2.2.1 Herstellung von chemisch-kompetenten Bakterien (Calciumchlorid-Methode) .....	42
3.2.2.2 Herstellung von elektro-kompetenten Bakterien .....	43
3.2.3 Transformation in <i>E. coli</i> .....	43
3.2.3.1 Transformation mittels Hitzeschock.....	43
3.2.3.2 Transformation mittels Elektroporation .....	43
3.2.4 Bestimmung der Transformationseffizienz .....	44
3.3 Zellbiologische Methoden .....	45
3.3.1 Kultivierungsbedingungen.....	45
3.3.2 Auftauen und Einfrieren von eukaryotischen Zellen.....	45
3.3.3 Passage von adhärennten Zellen.....	45
3.3.4 Bestimmung der Zellzahl .....	45
3.4 Virologische Methoden und AAV-Selektion .....	46
3.4.1 <i>In vitro</i> -Selektion von randomisierten Peptidbibliotheken.....	46
3.4.1.1 Infektion mit AAV und Adenovirus.....	46
3.4.1.2 Ernte der AAV .....	46
3.4.1.3 Isolierung von viraler DNS zur NGS-Probenvorbereitung .....	47
3.4.2 <i>In vivo</i> -Selektion von randomisierten AAV-Peptidbibliotheken.....	47
3.4.2.1 Tiere.....	47
3.4.2.2 Virusinjektion und -zirkulation .....	47
3.4.2.3 Organentnahme .....	47
3.4.2.4 Isolierung von Langerhans-Inseln aus dem murinen Pankreas.....	48
3.4.2.5 DNS-Isolierung.....	48
3.4.2.6 Amplifikation der die randomisierten Inserts beinhaltenden viralen DNS .....	49
3.4.2.7 Verdau der die randomisierten Inserts beinhaltenden viralen DNS .....	49
3.4.2.8 Restriktion des AAV2-Genom codierenden Plasmids pMT202-6.....	50
3.4.2.9 Ligation der Insert tragenden DNS in das Plasmid pMT202-6.....	51
3.4.2.10 Elektroporation der sekundären Plasmidbibliothek in elektrokompetente DH5α .....	51



---

3.4.2.11 Produktion der AAV2 Transfer-shuttles .....	52
3.4.2.12 Zellernte und Isolierung der rAAV .....	53
3.4.2.13 Reinigung der rAAV mittel Iodixanol Dichtegradientenzentrifugation .....	53
3.4.2.14 Bestimmung des replikativen Titers .....	55
3.4.2.15 Produktion von sekundären AAV2-Peptidbibliotheken .....	56
3.4.2.16 Beenden einer Selektion .....	56
3.4.3 Produktion von Kapsid-modifizierten rAAV2-Vektoren mit Reporter gen.....	56
3.4.3.1. Herstellung von Plasmiden mit Insertion im AAV2- <i>cap</i> .....	57
3.4.3.2 Großproduktion von selektierten Kapsid-modifizierten rAAV2-Einzelklonen mit Reporter gen in HEK 293T Zellen.....	57
3.4.4 Evaluation von selektierten Kapsidvarianten.....	58
3.4.4.1 Evaluation von <i>in vitro</i> selektierten Kapsidvarianten mittels Luciferaseassay.....	58
3.4.4.2 Analyse der Vektorverteilung in der Maus ( <i>homing</i> ) zur Evaluation der <i>in vivo</i> selektierten Kapsidvariante.....	58
3.4.5 Statistik.....	58
<b>4 ERGEBNISSE.....</b>	<b>59</b>
4.1 <i>In vivo</i> -Selektion von randomisierten AAV2-NNK <sub>7</sub> -Peptidbibliotheken .....	59
4.1.1 <i>In vivo</i> -Selektion von randomisierten AAV2-NNK <sub>7</sub> -Peptidbibliotheken in Niere und Pankreas von Mäusen .....	59
4.1.1.1 Generierung von sekundären AAV2-NNK <sub>7</sub> -Peptidbibliotheken zur Selektion in Niere und Pankreas .....	59
4.1.1.2 Anreicherung von Viruspartikeln im Laufe der Selektionen .....	60
4.1.2 Selektion von AAV2-Peptidbibliotheken in Langerhans-Inseln des murinen Pankreas .....	61
4.1.2.1 Isolierung von Langerhans-Inseln aus dem murinen Pankreas.....	61
4.1.2.2 Generierung von sekundären AAV2-NNK <sub>7</sub> -Peptidbibliotheken zur Selektion in Langerhans-Inseln.....	61
4.2 Analyse des <i>in vivo</i> -Selektionsverlaufs mittels Next-Generation Sequencing (NGS).....	62
4.2.1 Probenherstellung für die NGS-Analyse.....	62
4.2.2 Analyse der Tiefensequenzierung der AAV-Bibliotheken während der <i>in vivo</i> -Selektion – Möglichkeiten der Auswertung.....	63
4.2.2.1 Anreicherung von Peptidsequenzen während der Selektionen.....	64
4.2.2.2 Sequenzmotivanalyse der zehn häufigsten Peptidsequenzen.....	66
4.2.2.3 Vergleich der häufigsten Peptidsequenzen aus Zielgewebe und jeweiligen Kontrollorganen.....	69

4.2.2.4 Analyse von Peptidanreicherung- und Spezifität zur Bewertung der identifizierten Sequenzmotive .....	70
4.2.2.5 Evaluierung eines exemplarischen selektierten AAV-Klons .....	76
4.3 Analyse und Vergleich von AAV2-NNK <sub>7</sub> -Plasmidbibliothek und Virusbibliotheken .....	78
4.3.1 NGS-Daten und Vergleich der AAV2-NNK <sub>7</sub> -Plasmidbibliothek und den daraus generierten Virusbibliotheken hinsichtlich der Peptiddiversität .....	78
4.3.2 Vergleich und Sequenzmotivanalyse der zehn häufigsten Peptidsequenzen.....	79
4.3.3 Häufigkeit der 20 Aminosäuren in den randomisierten Positionen der unselektierten NNK <sub>7</sub> -Bibliotheken .....	80
4.3.4 Erwartete und tatsächliche Aminosäurehäufigkeit der NNK <sub>7</sub> -Bibliotheken .....	82
4.4 Einführung der neuen Peptidbibliothek AAV2-Trimer <sub>6</sub> -(-C).....	85
4.4.1 AAV2-Trimer <sub>6</sub> -(-C) im Vergleich zur AAV2-NNK <sub>7</sub> -Peptidbibliothek .....	85
4.4.2 Vergleich der Peptiddiversität der AAV2-Trimer <sub>6</sub> -(-C)-Plasmidbibliothek und der daraus generierten Virusbibliothek .....	86
4.4.3 Erwartete und tatsächliche Aminosäurehäufigkeit in den 2T <sub>6</sub> -Bibliotheken.....	87
4.4.4 Sequenzmotivanalyse der zehn häufigsten Peptidsequenzen .....	89
4.4.5 Häufigkeit der 19 Aminosäuren an den randomisierten Positionen der unselektierten AAV2-Trimer <sub>6</sub> -(-C)-Bibliothek .....	89
4.4.6 <i>In vitro</i> -Selektion zum direkten Vergleich der AAV2-Trimer <sub>6</sub> -(-C)-Peptidbibliothek mit der AAV2-NNK <sub>7</sub> -Peptidbibliothek .....	90
4.4.6.1 Anreicherung der zehn häufigsten Peptidsequenzen der letzten Selektionsrunde.....	92
4.4.6.2 Evaluation der <i>in vitro</i> selektierten AAV2-Kapside.....	94
<b>5 DISKUSSION .....</b>	<b>96</b>
5.1 Überblick .....	96
5.2 Übertragung der Selektionsmethode auf neue Ziele <i>in vivo</i> .....	97
5.2.1 Die Niere als Zielorgan.....	97
5.2.2 Das Pankreas und die Langerhans-Inseln als Zielgewebe .....	98
5.2.3 Retrospektive Untersuchung der AAV-Peptidbankselektion im Gehirn .....	100
5.3 Next-Generation Sequencing zur Untersuchung von AAV-Selektionen.....	100
5.4 Anreicherung von Peptidsequenzen während einer Selektion.....	102
5.5 Bestimmung und Bewertung von Sequenzmotiven und potenziell spezifischen Klonen .....	102
5.6 Vergleich der Peptidbibliotheken AAV2-NNK <sub>7</sub> - und AAV2-Trimer <sub>6</sub> -(-C) hinsichtlich Diversität und Abdeckung.....	104
5.7 Vergleich der Plasmidbibliotheken AAV2-NNK <sub>7</sub> und AAV2-Trimer <sub>6</sub> -(-C).....	106
5.8 Vergleich der nicht-selektierten Virusbibliotheken AAV2-NNK <sub>7</sub> und AAV2-Trimer <sub>6</sub> -(-C).....	108

---

5.9 Die Sequenzmotive „NXXRXX( <sup>E</sup> / <sub>X</sub> )“ und „XXDXT <sup>R</sup> / <sub>X</sub> X“ .....	109
5.10 Praktischer Vergleich der Peptidbibliotheken AAV2-NNK <sub>7</sub> und AAV2-Trimer <sub>6</sub> - (-C) mittels <i>in vitro</i> -Selektion .....	111
5.11 Ausblick.....	112
<b>6 Literaturverzeichnis .....</b>	<b>114</b>
<b>7 Anhang.....</b>	<b>128</b>
7.1 Anhang zu den Ergebnissen zu 4.3.4.....	128
7.2 Auflistung der verwendeten Gefahrenstoffe nach GHS (global harmonisiertes System).....	130
7.3 Publikationen.....	133
<b>Danksagung .....</b>	<b>134</b>
<b>Eidesstattliche Versicherung:.....</b>	<b>135</b>
<b>Änderungsvermerk.....</b>	<b>136</b>

## I Abkürzungsverzeichnis

A	Adenosin / Alanin
AAV	Adeno-associated Virus
Ad	Adenovirus
AMP	Ampicillin
ANOVA	Varianzanalyse
As	Aminosäure
bp	Basenpaar(e)
c	Konzentration
C	Cytidin/Cystein
CaCl <sub>2</sub>	Calciumchlorid
Cap	<i>capsid protein</i> , Kapsidprotein
cm	Zentimeter
CMV	Cytomegalovirus
ddH <sub>2</sub> O	Doppelt-destilliertes Wasser
D	Asparaginsäure
Da	Dalton
DMEM	Dulbecco's modified Eagle's Minimal Essential Medium
DMSO	Dimethylsulfoxid
dNTP(s)	Deoxyribonukleotid(e)
E	Glutaminsäure
<i>E. coli</i>	<i>Escherichia coli</i>
F	Phenylalanin / Farad
FBS	<i>fetal bovine serum</i> , Fötales Kälberserum
<i>Fgf-4</i>	<i>fibroblast growth factor 4</i> , Fibroblasten Wachstumsfaktor 4
FGFR	<i>fibroblast growth factor receptor</i> , Fibroblasten Wachstumsfaktor Rezeptor
for	<i>forward</i> , vorwärts
g	Gramm
G	Guanosine/Glycin
gP	Genomische Partikel
H	Histidin
HBSS	Hank's Balanced Salt Solution
HCl	Chlorwasserstoff
HEK 293T	<i>human embryonic kidney 293 with the simian virus 40 large T antigen</i> , Humane embryonale Nierenzellen mit Simianvirus 40 large T-Antigen
HEPES	2-(4-(2-Hydroxyethyl)-1-piperazinyl)-ethansulfonsäure
HGFR	<i>hepatocyte growth factor receptor</i> , Hepatozyten Wachstumsfaktor Rezeptor
HSPG	Heparansulfat-Proteoglycan
I	Isoleucin
i.v.	intravenös
ITR	Inverted Terminal Repeat, invertierte terminale Wiederholungen
IVC	<i>individual ventilated cage</i> , einzeln belüftete Käfige
K	Lysin
K <sub>2</sub> HPO <sub>4</sub>	Dikaliumhydrogenphosphat
kb	Kilobase
KCl	Kaliumchlorid
KH <sub>2</sub> PO	Kaliumdihydrogenphosphat
l	Liter
L	Leucin
LB	Lysogeny broth
μ	Mikro
M	molar (= mol/l)
m	Masse
mM	millimolar
M	Methionin
mg	Milligramm

MgCl <sub>2</sub>	Magnesiumchlorid
MgCl <sub>2</sub> 6H <sub>2</sub> O	Magnesiumchloridhexahydrat
min	Minute(n)
MOI	<i>multiplicity of infection</i> , Multiplizität der Infektion
n	Anzahl
N	Adenosin (A), Cytidin (C), Guanosin (G) oder Thymidin (T)/Asparagin
Na <sub>2</sub> HPO <sub>4</sub> 2H <sub>2</sub> O	Dinatriumhydrogenphosphatdihydrat
NaCl	Natriumchlorid
NGS	Next-Generation Sequencing
Ω	Ohm
OD	Optische Dichte
ORF	<i>open reading frame</i> , Offener Leserahmen
P	Prolin
PBS	<i>phosphate buffered saline</i> , Phosphatgepufferte Lösung
PBS-MK	Phosphatgepufferte Lösung mit Magnesium und Kalium
PCR	<i>polymerase chain reaction</i> , Polymerase-Kettenreaktion
PEG	Polyethylenglycol
pfu	<i>plaque forming units</i> , Plaque formierende Einheiten
Q	Glutamin
q-PCR	<i>quantitative polymerase chain reaction</i> , quantitative Polymerase Kettenreaktion
R	Arginin
rAAV	recombinantes Adeno-assoziiertes Virus
RBE	<i>Rep binding element</i> , Rep bindendes Element
rev	<i>reverse</i> , rückwärts
Rep	<i>replication protein</i> , Protein für die Replikation
rep E	Replikative Einheiten
rpm	<i>rounds per minute</i> , Runden per Minute
RPMI	Roswell Park Memorial Institute
RT	Raumtemperatur
S	Serin
SOB	super optimal broth
SOC	super optimal broth with catabolite repression
SV40	Simianvirus 40
T	Thymidin / Threonin
TAE	Tris-Acetat-EDTA-Puffer, tris(hydroxymethyl)aminomethan-acetat-ethylendiamin-tetraessigsäure
TB	terrific broth
TE (Tris-EDTA)	Tris(hydroxymethyl)aminomethanethylendiaminetetraessigsäure
Tm	<i>melting temperature</i> , Schmelztemperatur
Tris	Tris(hydroxymethyl)aminomethan
trs	terminal resolution site
U	Unit(s), Einheit(en)
UKE	Universitätsklinikum Hamburg-Eppendorf
UV	Ultraviolett
V	Valin / Volt
v/v	Volumen pro Volumen
vG	Virale Genome
vol.	Volumen
VP	Virionprotein
W	Tryptophan
w/v	<i>weight per volume</i> , Gewicht pro Volumen
x g	Vielfaches der Erdbeschleunigung
Y	Tyrosin

## II Abbildungsverzeichnis

Abbildung 1: Schematische Darstellung des AAV2-Genoms und der transkribierten prä-mRNS. ....	6
Abbildung 2: Schematische Darstellung des AAV2-Kapsids und Sekundärstruktur einer VP1-Untereinheit.....	7
Abbildung 3: Schematische Darstellung des AAV-DNS-Replikationsmodells.....	9
Abbildung 4: Genom des AAV2-Wildtypen im Vergleich zu dem eines rekombinanten AAV-Vektors. ....	11
Abbildung 5: Schematische Darstellung des AAV2-Kapsids mit Peptidinsertion an Position R588 (VP1) und Teilsequenz des <i>cap</i> -Gens mit modifizierter Insertionsstelle. ....	13
Abbildung 6: Schema zur Produktion von AAV-Peptidbibliotheken mittels Transfer-Shuttle-Produktion. ....	14
Abbildung 7: Prinzip der Selektion von randomisierten AAV-Peptidbibliotheken.....	16
Abbildung 8: Schematische Darstellung der Clusterbildung beim NGS nach Illumina®.....	19
Abbildung 9: Schematische Darstellung der NGS-Probenvorbereitung zur Untersuchung von AAV- Peptidbibliotheken.....	37
Abbildung 10: Beipielsequenz des finalen PCR-Produktes der NGS-Probe.....	40
Abbildung 11: Anreicherung von viralen Partikeln während der Selektion.....	60
Abbildung 12: Expandiertes Pankreas und isolierte Langerhans-Inseln umgeben von exokrinen Zellen.....	61
Abbildung 13: Analyse der PCR-Produkte der NGS-Proben mittels Agarose-Gelelektrophorese.....	63
Abbildung 14: Untersuchung der Peptidanreicherung während der Selektionen.....	65
Abbildung 15: Untersuchung der Peptidanreicherung unter Berücksichtigung der Peptidhäufigkeiten. ....	66
Abbildung 16: Bewertung der Peptidanreicherung und Peptidspezifität mit Hilfe der <i>GE</i> -, <i>GS</i> - und <i>C</i> -Scores. ...	76
Abbildung 17: Relative Anreicherung der Peptidsequenz „GDDATRD“ während der Selektion im Pankreas. ....	76
Abbildung 18: Vergleich der Vektorverteilungen des rAAV2-GDDATRD mit dem Kontrollvektor mit AAV2- Wildtypkapsid. ....	77
Abbildung 19: Vergleich der AAV2-NNK7-Plasmidbibliothek und daraus generierten Virusbibliotheken bezüglich der Anreicherung von Peptidsequenzen.....	79
Abbildung 20: Häufigkeit der 20 Aminosäuren an den randomisierten Positionen der unselektierten Plasmidbibliothek im Vergleich zu den drei Virusbibliotheken.....	82
Abbildung 21: Verteilung der Aminosäuren der Peptidsequenzen der Plasmidbibliothek im Vergleich zum NNK- Codierungsschema begründeten Referenzwert. ....	83
Abbildung 22: Codonhäufigkeiten und Abweichungen vom Erwartungswert (AAV2-NNK <sub>7</sub> -Bibliotheken). ....	84
Abbildung 23: Vergleich der AAV2-Trimer <sub>6</sub> -(-C)-Plasmidbibliothek und daraus generierten Virusbibliotheken bezüglich der Anreicherung von Peptidsequenzen.....	86
Abbildung 24: Codonhäufigkeiten und Abweichungen vom Erwartungswert (AAV2-Trimer <sub>6</sub> -(-C)-Bibliotheken).88	88
Abbildung 25: WebLogo basierte Analyse der Plasmidbibliothek im Vergleich zur Virusbibliothek. ....	90
Abbildung 26: Untersuchung der Peptidanreicherung während der Selektionen.....	92
Abbildung 27: Verlauf der häufigsten zehn Peptidsequenzen der letzten Selektionsrunde während der Selektion.....	93
Abbildung 28: Bestimmung der Transduktion der selektierten und ausgewählten Kapsidvarianten. ....	95
Abbildung 29: Codonhäufigkeit und Abweichungen vom Erwartungswert (AAV2-NNK <sub>7</sub> -Virusbibliothek 2). ....	128
Abbildung 30: Codonhäufigkeit und Abweichungen vom Erwartungswert (AAV2-NNK <sub>7</sub> -Virusbibliothek 3). ....	129

## III Tabellenverzeichnis

Tabelle 1: Laborgeräte .....	22
Tabelle 2: Molekularbiologische Kits.....	23
Tabelle 3: Reagenzien .....	23
Tabelle 4: Enzyme .....	23
Tabelle 5: Antikörper.....	24
Tabelle 6: Prokaryotische Zellen .....	26
Tabelle 7: Eukaryotische Zellen.....	26
Tabelle 8: Antibiotika und Zusätze.....	27
Tabelle 9: Oligonukleotide .....	27
Tabelle 10: Barcodesequenzen. ....	28
Tabelle 11: Mäuse .....	28
Tabelle 12: Viren .....	29
Tabelle 13: Rekombinante AAV2.....	29
Tabelle 14: Plasmide .....	29
Tabelle 15: Software und Datenbanken.....	30
Tabelle 16: Unternehmen .....	30
Tabelle 17: Zusammenfassung wichtiger Kenngrößen der AAV2-Peptidbibliothek-Selektion in Niere und Pankreas.....	60
Tabelle 18: Übersicht über die Selektionsrunden der Selektion in den Langerhans-Inseln.....	61
Tabelle 19: Zusammenfassung wichtiger Kenngrößen der AAV2-Peptidbibliothek-Selektion in Langerhans-Inseln.....	62
Tabelle 20: Zusammenfassung der NGS-Daten für die AAV2-Peptidbibliothek-Selektionen in Niere, Pankreas, Langerhans-Inseln und Gehirn.....	64
Tabelle 21: Sequenzmotivanalyse der zehn häufigsten Peptidsequenzen der Selektionen in Niere, Pankreas, Langerhans-Inseln und Gehirn.....	68
Tabelle 22: Vergleich der häufigsten Peptidsequenzen in Ziel- und Kontrollgeweben. ....	69
Tabelle 23: Zusammenfassung der NGS-Daten der AAV2-NNK <sub>7</sub> -Plasmid- und Virusbibliotheken. ....	78
Tabelle 24: Peptidsequenzanalyse und Vergleich der AAV2-NNK <sub>7</sub> -Plasmid- und Virusbibliotheken.....	80
Tabelle 25: Vergleich der verschieden codierten Peptidbibliotheken AAV2-Trimer <sub>6</sub> -(C) und AAV2-NNK <sub>7</sub> .....	85
Tabelle 26: Zusammenfassung der NGS-Daten der AAV2-Trimer <sub>6</sub> -(C) Plasmid- und Virusbibliothek. ....	86
Tabelle 27: Peptidsequenzanalyse und Vergleich der AAV2-Trimer <sub>6</sub> -Plasmid- und Virusbibliotheken.....	89
Tabelle 28: Zusammenfassung der NGS-Daten der <i>in vitro</i> -Selektionen.....	91
Tabelle 29: Auflistung der verwendeten Gefahrenstoffe nach GHS.....	130
Tabelle 30: Legende zu den aufgelisteten Piktogrammen.....	131





## Zusammenfassung

Vor 50 Jahren wurde das Adeno-assoziierte Virus (AAV) entdeckt, seit über 30 Jahren wird an der Verwendung des Virus als Gentherapievektor geforscht und seit 13 Jahren werden modifizierte Kapsidvarianten des Vektors im Rahmen von AAV-Peptidbibliotheken hauptsächlich *in vitro* aber auch *in vivo* selektiert, um deren natürlichen Tropismus neu auszurichten. AAV-Peptidbibliotheken bestehen aus einer Vielzahl an Kapsid-modifizierten Vektoren, die sich durch die gezielte Insertion von im Virusgenom codierten und an der Kapsidoberfläche exprimierten randomisierten Peptiden voneinander unterscheiden. Trotz vieler Untersuchungen und Erkenntnisse rund um die Selektion dieser Bibliotheken, bei der sich für den jeweiligen Anwendungszweck bevorzugte Kapsidvarianten anreichern, gleicht besonders der Prozess der *in vivo*-Selektion und die Bestimmung der tatsächlichen Diversität der Peptidbibliotheken immer noch einer Art „Black-Box“. So war eine qualitative Auswertung der Peptidinserts lange Zeit nur über die Sanger-Sequenzierung zufällig ausgewählter Einzelklone möglich. Je nach Probenumfang können so zwar die häufigsten Sequenzen jeder Selektionsrunde identifiziert werden; eine sichere Abschätzung der Diversität und der Anreicherung der Sequenzen ist auf diese Weise, aufgrund des hohen Arbeitsaufwandes, allerdings unrealistisch und die Wahrscheinlichkeit, einen potenziell gewebespezifischen Klon zu finden ist eher gering. Durch die Entwicklung von automatisierten Sequenzierungsmethoden der nächsten Generation, des Next-Generation Sequencings (NGS), ist es nun möglich, einen wesentlich tieferen Einblick in die Zusammensetzung nicht-selektierter wie auch selektierter Peptidbibliotheken zu gewinnen und Aussagen über die Diversität, Anreicherung und Spezifität einzelner Sequenzen oder Sequenzmotive zu machen. Des Weiteren kann die Qualität einer Peptidbibliothek bereits vor ihrer Selektion bestimmt und die Suche nach einem gewebespezifischen AAV-Klon wesentlich vereinfacht werden.

Im ersten Teil dieser Arbeit wurden AAV2-Peptidbibliotheken mit NNK-codierten (N = jedes der vier möglichen Nukleotide; K = Guanin oder Thymin) sieben Aminosäure langen Peptidinsertionen (NNK<sub>7</sub>-Peptidbibliothek) in den Zielgeweben Niere, Pankreas und Langerhans-Inseln, nach systemischer Injektion in Mäusen selektiert, mittels NGS analysiert und mit retrospektiv erhobenen NGS-Daten einer erfolgreichen *in vivo*-Selektion im Gehirn im Hinblick auf Diversität und Anreicherung verglichen. Hierbei wurde für die Selektion im Gehirn die stärkste Anreicherung bestimmt. Neben der Verifizierung des Sequenzmotives der Selektion im Gehirn („XDGXXWX“), welches bereits durch die Sanger-Sequenzierung von zehn Einzelklonen in Vorarbeiten gesehen wurde, wurden zwei weitere Sequenzmotive identifiziert. Während sich das Motiv aus der Selektion in der Niere („NXXRXX<sup>E</sup>/X“) als Gewebe-unspezifisch herausstellte, wurde bei den Selektionen in Pankreas und Langerhans-Inseln das gleiche Sequenzmotiv („XXDXT<sup>R</sup>/X“) selektiert und stark angereichert. Zusätzlich konnte durch die Anwendung von Bewertungskriterien zur Berechnung von Anreicherung und Spezifität von

Sequenzmotiven und einzelnen Sequenzen im Falle der Selektion im Pankreas ein das selektierte Motiv tragender Klon mit einer zum Zielorgan verschobenen Vektorverteilung erfolgreich identifiziert werden.

Im Zuge der Suche nach dem Ursprung des unspezifischen und dennoch sehr auffallenden Sequenzmotivs „NXXRXX<sup>E</sup>/<sub>x</sub>“ wurden im zweiten Teil der Arbeit die Plasmidbibliothek sowie die unselektierte Virusbibliothek mittels NGS analysiert. Neben der erwarteten Anreicherung bestimmter Aminosäuren aufgrund des NNK-Codierungsschemas der Peptide, wurde in der Plasmidbibliothek eine systematische Ungleichverteilung bestimmter Codons und in der Virusbibliothek eine verhältnismäßig starke Anreicherung des Motivs „NXXRXX“ festgestellt.

Zum Zweck einer besseren initialen Gleichverteilung der Codons wurde im dritten Teil der Arbeit eine neue von unserer Arbeitsgruppe generierte AAV2-Trimer<sub>6</sub>-(-C)-Peptidbibliothek, mit Trimer-codierten und zur Erreichung einer höheren Abdeckung nur sechs Aminosäure langen Peptidinsertionen, auf Diversität untersucht und im Vergleich zur NNK-codierten Bibliothek *in vitro* selektiert. Trotz einer wesentlich verbesserten Gleichverteilung der Codons der Plasmidbibliothek, wurde in der Virusbibliothek das bekannte unspezifische, nur um eine Aminosäure verkürzte, Sequenzmotiv „NXXRXX“ angereichert. Bei diesem Motiv handelt es sich folglich um ein initial in der Virusbibliothek angereichertes Motiv, das einen strukturellen Vorteil für das AAV2-Kapsid bedeuten könnte. Schließlich zeigte der *in vitro* Vergleich der Trimer-basierten und der NNK-codierten Bibliothek zwar keinen generellen aber einen vom Selektionsziel abhängigen Vorteil entweder der Trimer- oder der NNK-codierten Bibliothek. Damit konnte gezeigt werden, dass sich auch die neue AAV2-Peptidbibliothek zur Selektion von Tropismus-modifizierten Kapsidvarianten und zur Erweiterung des Zielspektrums von AAV2-Peptidbibliothek-Selektionen eignet.

## Abstract

Adeno-associated virus (AAV) was discovered 50 years ago, its applicability as gene therapy vector has been investigated for over 30 years and capsid-modified variants of this virus have been selected in the context of AAV display peptide libraries to redirect the viral tropism for 13 years. AAV display peptide libraries consist of a high number of capsid-modified vectors that differ in genome-coded insertions of randomized peptides within the viral capsid.

Particles with desired properties can be isolated from such AAV display peptide libraries by screening them in subsequent selection rounds *in vitro* or *in vivo*. However, especially the selection process of *in vivo* screenings which is associated by a change of the library composition and a decreasing peptide diversity can still be considered a „black-box“.

For a long time, sanger-sequencing of a limited number of randomly chosen library clones was the only possibility to analyze the enriched capsid variants. Depending on the sample number, this type of analysis enabled the identification of the most frequent peptide inserts of each selection round. A good estimation of peptide diversity and enrichment, however, is quite unrealistic with only few analyzed clones and the possibility to find a potential tissue-specific variant is low. The development of automated sequencing methods of the next generation (Next-Generation Sequencing = NGS) provides deeper insights into the composition of primary and preselected peptide libraries and allows to estimate peptide diversity, as well as enrichment and specificity of single sequences or sequence motifs. Furthermore, the quality of a peptide library can be determined before using it for *in vitro* or *in vivo* screenings and the search for a tissue specific AAV clone can be simplified.

In the first part of this work, AAV2 display peptide libraries containing NNK-coded seven amino acid peptide insertions (N = each of the four possible nucleotides; K = guanine or thymine), were screened in the target tissues kidney, pancreas and islets of Langerhans after systemic injection into mice. After several selection rounds, the libraries were analyzed via NGS and compared with the retrospectively NGS-analyzed data of a successful *in vivo* screening in the brain, regarding diversity and enrichment. Furthermore, the existence of the brain-enriched sequence motif (“XDGXXWX”), which was seen already during Sanger sequencing of ten single clones previously, was verified and two additional sequence motifs were identified. The sequence motif selected in the kidney (“NXXRXX<sup>E</sup>/<sub>X</sub>”) turned out to be non-tissue-specific, while the second sequence motif (“XXDXT<sup>R</sup>/<sub>X</sub>X”) was highly enriched during the screening in pancreas and the islets of Langerhans. Furthermore, by applying assessment criteria to calculate the enrichment and specificity of sequence motifs and single sequences, a single pancreas-enriched clone could be identified.

In order to assess the origin of the non-specific but striking sequence motif “NXXRXX<sup>E</sup>/<sub>X</sub>”, the plasmid library and the non-selected virus library were analyzed via NGS in the second part of this work. In

addition to the expected enrichment of distinct amino acids due to the NNK coding scheme, a systematic unequal distribution of distinct codons in the plasmid library and a relatively strong enrichment of the motif “NXXRXXX” in the virus library was identified.

To ensure a more equal distribution of amino acids within the peptide library’s inserts, our group generated a new AAV2 display peptide library with trimer-based nucleotide inserts. In order to enhance the coverage, the length of the peptide insert was reduced to six amino acids (i.e. six nucleotide trimers). In the third part of this study, this new library, named AAV2 Trimer<sub>6</sub> (-C) display peptide library, was assessed regarding its diversity and selected *in vitro* in comparison to the NNK-coded library. Despite an improved equal distribution of the plasmid library’s codons, the non-specific sequence motif “NXXRXX” was enriched in the virus library, being only one amino acid shorter than the non-specific motif observed in the NNK-library.

Consequently, this motif seems to be initially enriched in the virus library, and seems to confer a structural advantage for the AAV2 capsid. Finally, the *in vitro* comparison of the trimer-based and the NNK-coded library showed that each library can be superior to the other one, depending on the target cells. Consequently, the selection of tropism-modified capsid variants from the new peptide library is feasible and furthermore, the target spectrum for the selection of AAV2 display peptide libraries can be extended with this library.

# 1 EINLEITUNG

## 1.1 Gentherapie und Gentherapievektoren

Das allgemeine Prinzip der Gentherapie ist der Transfer von Nukleinsäuren in eukaryotische Zellen mit dem Ziel der Veränderung der Genexpression von therapeutisch relevanten Genen zur Linderung oder gar Heilung von Krankheiten. Die größte Herausforderung stellt hierbei die Entwicklung von sicheren und effizienten Vektorsystemen dar<sup>1,2</sup>. In den letzten Jahren wurden viele Methoden für den Transfer von Nukleinsäuren entwickelt und untersucht. Neben Vektoren aus Lipiden<sup>3</sup>, Peptiden<sup>4</sup>, Polysacchariden<sup>1</sup> oder Nanopartikeln<sup>5</sup> stellen Viren ein weiteres Vektorsystem zum Transport von genetischer Information dar. Virale Vektoren weisen im Gegensatz zu nicht-viralen Vektoren den Vorteil stärkerer Transduktionseffizienzen und länger anhaltender Genexpression auf<sup>6</sup>.

Im Laufe der Zeit wurden verschiedene virale Vektorsysteme entwickelt. Neben Retroviren, Herpes-Simplex-Viren und Adeno-assoziierten Viren ist das Adenovirus ein vor allem in der Vergangenheit häufig eingesetztes Vektorsystem<sup>7</sup>. So wurden im Laufe der Jahre viele klinische Studien zum Beispiel zur Behandlung des schweren kombinierten Immundefektes aufgrund von Adenosindesaminase-Mangel (*severe combined immunodeficiency, ADA-SCID*)<sup>8</sup>, der zystischen Fibrose<sup>9</sup>, der Hämophilie B<sup>10</sup>, der Duchenne-Muskeldystrophie<sup>7</sup>, von Parkinson<sup>11</sup> oder verschiedenen Krebsarten<sup>7</sup> mit viralen Vektorsystemen durchgeführt. Während Retroviren stabil in das Wirtsgenom integrieren und eine lang anhaltende Genexpression vermitteln können, weisen Adenoviren eine hohe und transiente Genexpression auf<sup>6</sup>. Allerdings bergen Retroviren das Risiko der Insertionsmutagenese und Adenoviren sind sehr immunogen<sup>6</sup>.

Das Adeno-assoziierte Virus (AAV) eignet sich sehr gut als Basis viraler Vektoren für die Gentherapie. Im Gegensatz zu Adenoviren gilt AAV nach wie vor als weitestgehend nicht-pathogen und nur sehr gering immunogen<sup>12</sup>. Des Weiteren können AAV verschiedenste Zelltypen und Gewebe effizient transduzieren und zeigen langanhaltende Genexpression<sup>6</sup>.

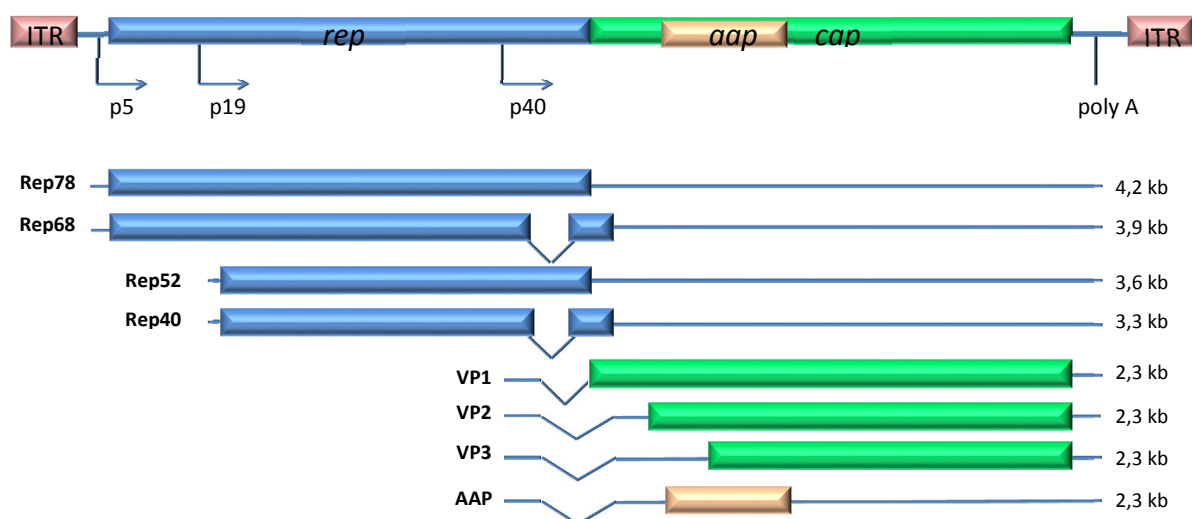
## 1.2 Adeno-assoziierte Viren (AAV)

Adeno-assoziierte Viren wurden erstmals 1965 als Kontaminationen von adenoviralen Isolaten beschrieben<sup>13-15</sup>. AAV sind kleine, unbehüllte und Helfervirus-abhängige DNS-Viren und gehören zur Familie der Parvoviren<sup>16</sup>, genauer zur Gattung der Dependoviren<sup>17</sup>. Bis heute sind zwölf natürliche Serotypen aus Geweben von Menschen und anderen Wirbeltieren isoliert worden<sup>18-21</sup> und mehr als 100 verschiedene Varianten beschrieben<sup>22,23</sup>. Der Serotyp 2 ist der am besten charakterisierte<sup>12</sup> und weist mit ca. 80 %<sup>17</sup> die höchste Seroprävalenz in der Bevölkerung auf<sup>25</sup>.

### 1.2.1 Das AAV-Genom

Das Genom des AAVs besteht aus ca. 4,7 kb einzelsträngiger DNS<sup>26–28</sup>, codiert für drei offene Leserahmen (*open reading frames, ORFs*) *rep*, *aap* und *cap*<sup>28,29</sup>, und ist von 145 Nukleotiden langen invertierten terminalen Wiederholungen (*Inverted Terminal Repeats, ITRs*) flankiert<sup>30</sup>. Die ITRs besitzen 125 Nukleotid lange palindromische Nukleotidsequenzen<sup>31</sup>, die T-förmige Haarnadelstrukturen ausbilden<sup>30</sup> und für die Replikation und Verpackung des Genoms sowie für die Integration in das Wirtsgenom benötigt werden<sup>17</sup>.

Insgesamt besitzt das Genom drei Promotoren (P5, P19 und P40)<sup>32,33</sup> und codiert für acht zum Teil durch alternatives Spleißen translatierte Proteine. Hierbei codiert *rep* für vier nicht-strukturelle Proteine Rep78 (78 kDa), Rep68 (68 kDa), Rep52 (52 kDa) und Rep40 (40 kDa)<sup>34–36</sup>. Alle vier Rep-Proteine zeigen Helicase- und ATPase-Aktivität<sup>35–38</sup>. Während die beiden größeren Proteine Rep78 und Rep68 unter anderem für die Virus-Replikation, die Regulierung der Transkription und die spezifische Integration in das humane Genom benötigt werden, wird den kleineren Proteinen Rep52 und Rep40 eine Rolle bei der Verpackung der viralen DNS in das Kapsid zugeschrieben<sup>39–42</sup>. Der offene Leserahmen *cap* codiert für drei Strukturproteine (VP1 (87 kDa), VP2 (72 kDa) und VP3 (62 kDa))<sup>17,43,44</sup>, die das virale Kapsid bilden. Die Proteine VP2 und VP3 sind hierbei Spleißvarianten des Strukturproteins VP1. *aap* codiert für das *assembly-activating protein* (AAP (23 kDa)), für welches eine Funktion beim Zusammenbau des viralen Kapsids beschrieben wurde<sup>29</sup>. Eine schematische Darstellung des AAV2-Genoms ist in **Abbildung 1** dargestellt.

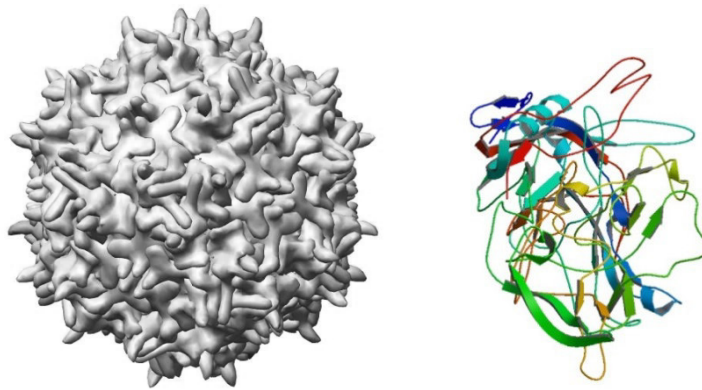


**Abbildung 1: Schematische Darstellung des AAV2-Genoms und der transkribierten prä-mRNAs.**

Das AAV2-Genom besteht aus den offenen Leserahmen *rep*, *aap* und *cap*, und ist von ITRs flankiert. Die Transkription der Genprodukte Rep78 und Rep68 steht unter der Kontrolle des p5 Promotors, während die Transkription für Rep52 und Rep40 durch den p19 Promotor reguliert wird. Die Genprodukte Rep68 und Rep40 sind jeweils Spleißvarianten von Rep78 und Rep52. *Cap* codiert für die strukturellen Kapsidproteine VP1, VP2 und VP3. Hierbei stellen die Proteine VP2 und VP3 Spleißvarianten des Strukturproteins VP1 dar. Der offene Leserahmen *aap* codiert für das *assembly-activating protein* (AAP). Die Transkription der mRNAs für die Strukturproteine wird durch den p40 Promotor reguliert. Die generierten RNA Transkripte sind als dünne Linien dargestellt. Die offenen Dreiecke stellen Introns dar. Nach Büning et al.<sup>45</sup>, modifiziert.

### 1.2.2 AAV-Kapsidstruktur und zelluläre Rezeptoren

Das AAV-Kapsid ist ikosaedrisch und hat einen Durchmesser von ungefähr 22 nm<sup>17</sup>. Es besteht aus 60 Untereinheiten der Proteine VP1, VP2 und VP3 (molares Verhältnis von 1:1:10)<sup>46</sup>. Die atomare Struktur der verschiedenen Serotypen (AAV1, 2, 4, 6, 7, 8, und 9)<sup>47-53</sup> wurden mittels Kristallographie, bzw. im Fall des AAV5 mittels Kryo-Elektronen Mikroskopie gelöst<sup>54</sup>. Die Hauptstruktur des Kapsids wird durch ein konserviertes, acht-strängiges, antiparalleles  $\beta$ -barrel-Motiv mit großen Schleifen (*loops*) zwischen den  $\beta$ -Strängen gebildet<sup>12</sup>. Die Schleifen benachbarter Proteinuntereinheiten bilden Serotyp-charakteristische Ausstülpungen um die Dreifachachse und werden als *threefold spikes* bezeichnet<sup>19,24</sup>. Die meisten Interaktionen zwischen dem Kapsid und zellulären Rezeptoren finden über diese Schleifen statt. Die Struktur des AAV2-Kapsids ist in **Abbildung 2** schematisch dargestellt.



**Abbildung 2: Schematische Darstellung des AAV2-Kapsids und Sekundärstruktur einer VP1-Untereinheit.**  
([www.rcsb.org](http://www.rcsb.org) PDB ID 1LP3, Xie et al.<sup>48</sup>)

Heparansulfat-Proteoglykane (HSPG) stellen die primären zellulären Rezeptoren für AAV2 dar<sup>55</sup>. Hierbei bildet ein Cluster, bestehend aus fünf Aminosäuren (R484, R487, K532, R585 und R588), die entsprechende Bindungsstelle<sup>56,57</sup>. Insertionen an der Position R588 beeinflussen die Bindung an HSPG und können zu einer Neuausrichtung von AAV2 Vektoren führen<sup>58-60</sup>.

HSPG wird auf den Zellen vieler verschiedener Zelltypen präsentiert. Eine erfolgreiche Bindung und Internalisierung des Virus benötigen allerdings zusätzlich zu HSPG weitere Co-Rezeptoren. Als solche wurden Integrin  $\alpha\beta 5$ <sup>61</sup>, Integrin  $\alpha 5\beta 1$ <sup>62</sup>, der Fibroblasten Wachstumsfaktor Rezeptor 1 (*fibroblast growth factor receptor 1, FGFR1*)<sup>63</sup>, der Hepatozyten Wachstumsfaktor Rezeptor 1 (*hepatocyte growth factor receptor 1, HGFR1*)<sup>64</sup>, und der 37/67-kDa Laminin Rezeptor (LamR)<sup>65</sup> identifiziert.

### 1.2.3 Transduktion und Replikationszyklus

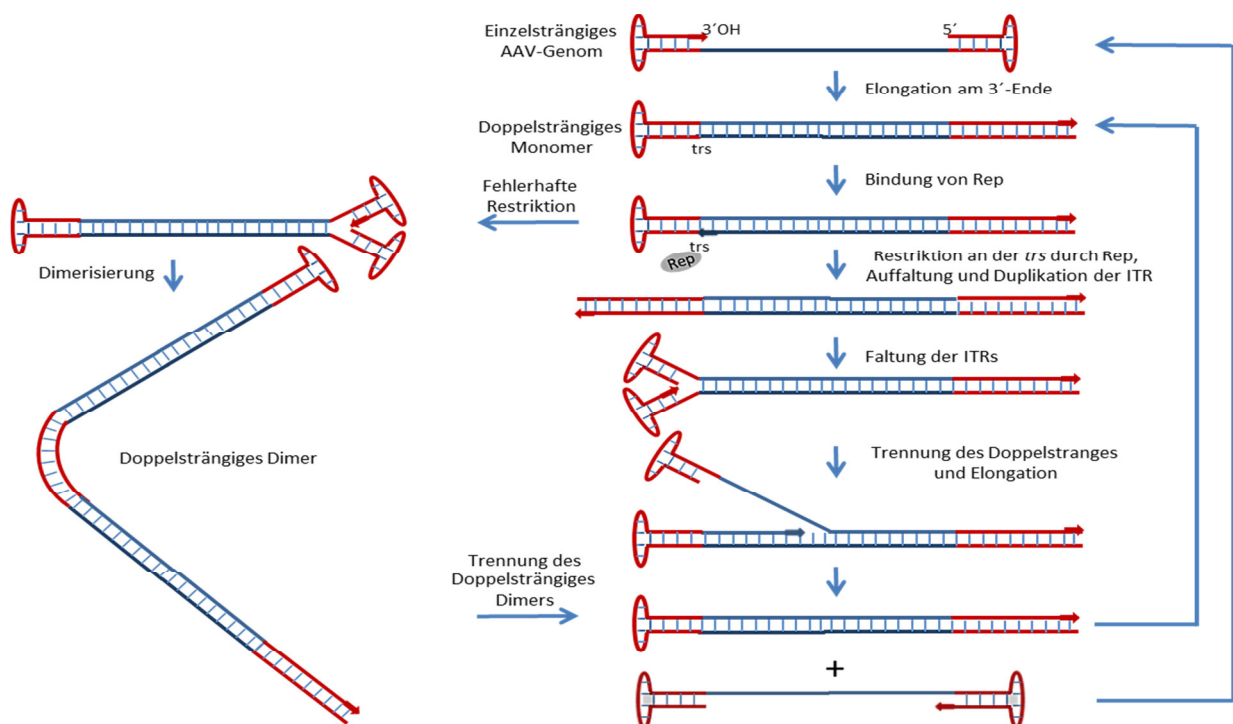
Nach der Bindung an zelluläre Rezeptoren wird AAV über Endozytose mittels *clathrin-coated pits*<sup>66</sup> Caveolae<sup>67</sup> oder Makropinozytose<sup>67</sup> internalisiert. Zusätzlich wurden alternative Zell-spezifische Internalisierungswege<sup>68</sup>, sowie die Möglichkeit der Transzytose<sup>69</sup> beschrieben. Im frühen Endosom wird das Virus an Mikrotubuli und Mikrofilamenten entlang in Richtung Zellkern transportiert<sup>70</sup>. Virale Partikel wurden in Endosomen, Lysosomen, perinukleären Recycling-Endosomen und im Golgi-Apparat gefunden<sup>12,23,67</sup>. Wie für die Internalisierung und den Transport der AAV zum Zellkern sind auch bezüglich der Freilassung des Virus aus dem Endosomen und bezüglich des Transportes durch die Kernmembran verschiedene potenzielle Wege beschrieben worden. Die N-terminalen Domänen des VP1 Proteins beinhalten die für die Freilassung aus dem Endosomen und für die Transduktion notwendige Phospholipase A2 (PLA2)<sup>71-73</sup>, sowie drei basische Cluster (BC1, 2 und 3), die vermutlich als Kernlokalisierungssignal (*nuclear localization signal, NLS*) für das Kapsid fungieren<sup>67,74</sup>. Bisherige Beobachtungen lassen darauf schließen, dass das Virus durch einen Konformationswechsel der N-Termini der Proteine VP1 und VP2, welche dabei exponiert werden, aus dem Endosom ins Cytoplasma freigelassen wird<sup>67,75</sup>. Weitere Studien legen nahe, dass für die Freilassung aus dem Endosom und/oder für spätere Schritte der Transduktion Cathepsin B und L notwendig sind<sup>76</sup>. Nach der Freilassung in das Cytoplasma akkumuliert das Virus perinukleär<sup>66,74,77,78</sup> und transloziert über den nukleären Poren Komplex (*nuclear pore complex, NPC*) in den Zellkern<sup>66</sup>, wo die virale DNS entlassen wird<sup>79</sup>.

Ohne die Gegenwart von sogenannten Helferviren wie Adenoviren<sup>13,80</sup> oder Herpes-Simplex Viren<sup>81</sup>, verweilen AAV intrazellulär in einem latenten Status. In diesem Status kann das Virus Rep-abhängig an einer bestimmten Stelle des Chromosomen 19 (19q13.4-qter, „AAVS1“) in das humane Genom integrieren<sup>82,83</sup> oder als extra-chromosomales Episom persistieren<sup>84</sup>. Unspezifische Integrationen sind sehr selten<sup>85</sup>. Die Co-Infektion mit einem Helfervirus ermöglicht den Übergang vom latenten in den lytischen Zyklus, wobei die virale DNS repliziert und neue Viren produziert werden<sup>42</sup>. Bezüglich des Adenovirus als Helfervirus stellen die Proteine E1A, E1B, E2A, E4orf6 und VA die unterstützenden Faktoren dar<sup>42,86-88</sup>. Dabei scheint keines der Faktoren direkt an der Replikation beteiligt zu sein<sup>89</sup>. Zusätzlich wurde beobachtet, dass die Bestrahlung oder die Behandlung von Zellen mit genotoxischen Reagenzien ebenfalls die Replikation der AAV-DNS, auch in Abwesenheit eines Helfervirus, ermöglichen<sup>90</sup>.

Die virale Replikation von Parvoviren und somit auch die von AAV findet nach einem sogenannten *rolling-hairpin*-Mechanismus statt<sup>16,91</sup>. Nachdem die einzelsträngige virale DNS in den Zellkern entlassen wurde, wird sie in doppelsträngige DNS umgeschrieben; hierbei werden virale Rep-Proteine wie auch die zelluläre Replikationsmaschinerie benötigt<sup>92</sup>. Bei der Replikation des AAV-Genoms spielen die ITRs eine besondere Rolle. Die 3'-terminale ITR stellt eine Hydroxylgruppe zur



Verfügung und fungiert als Primer und Replikationsursprung. Hier bindet die zelluläre Polymerase und kopiert die DNS bis zum 5'-Ende. Im nächsten Schritt binden die AAV Proteine Rep78 und/oder Rep68, die neben ihrer Helicase-Aktivität Ort- und Strang-spezifische Endonuklease-Funktion<sup>35</sup> besitzen, an dem Rep-bindenden Element (*Rep Binding Element, RBE*)<sup>93</sup> der 3'-terminalen ITR und schneiden an der gegenüberliegenden terminalen Schnittstelle (*terminal resolution site, trs*). An das hierbei entstehende freie 3'-Ende bindet erneut die Polymerase und kopiert die nun einzelsträngige ITR zu einer Doppelsträngigen. Hierdurch entsteht ein linearer doppelsträngiger DNS-Strang, die Haarnadelstrukturen der ITRs formen sich neu und das freiliegende 3'-Ende wird verlängert. Schließlich werden die Stränge in ein doppelsträngiges Monomer und ein einzelsträngiges AAV-Genom getrennt. Bei einer fehlerhaften Restriktion an der *trs* kommt es zur Dimerisierung und Entstehung eines doppelsträngigen Dimers, welches an der ITR der Symmetrieachse wieder getrennt werden kann (**Abbildung 3**)<sup>17</sup>. Die Kapside werden in den Nukleoli zusammengebaut und anschließend in das Nukleoplasma transportiert<sup>94,95</sup>. Im Nukleoplasma wird die virale DNS mit Hilfe der Proteine Rep52 und Rep40 in das Kapsid gebracht<sup>39,41,96</sup>. Hierbei werden Genome mit positiver und negativer Polarität zu gleichen Teilen verpackt<sup>97</sup>. Schließlich wird das neue Virus mit der Lyse der Wirtszelle freigelassen.



**Abbildung 3: Schematische Darstellung des AAV-DNS-Replikationsmodells.**

Durch Verlängerung des freien 3'-Endes wird das einzelsträngige AAV-Genom zu einem doppelsträngigen Monomer mit einem geschlossenen Ende. Hierbei fungiert die 3'-terminale ITR als Replikationsursprung und Primer. Virale Rep Proteine binden an der dem 3'-Ende gegenüberliegenden *trs* und ermöglichen somit die Kopie der ITR. Die Haarnadelstrukturen der ITRs bilden sich neu und durch die Verlängerung des freien 3'-Endes wird der DNS-Doppelstrang in einen Einzelstrang (neues AAV-Genom) und ein als Replikationsmatrize fungierendes doppelsträngiges Monomer getrennt. Wird dieses nicht an der *trs* getrennt, formieren sich die ITRs an dem offenen Ende zu Haarnadelstrukturen. Durch das freie 3'-Ende wird eine Dimerisierung ermöglicht und ein doppelsträngiges Dimer entsteht (Kopf-an-Kopf oder Ende-an-Ende-Orientierung). Dieses kann an der ITR der Symmetrieachse wieder zu einem doppelsträngigen Monomer getrennt werden. Nach Gonçalves<sup>17</sup>, modifiziert.

### 1.3 AAV als Gentherapievektor

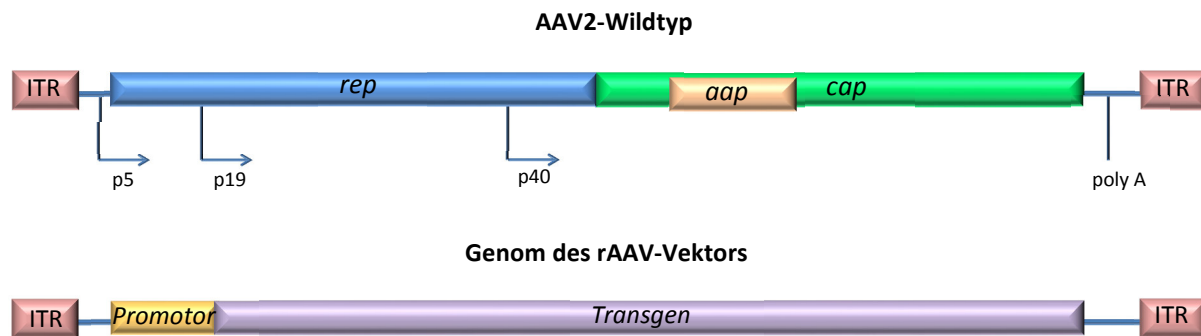
#### 1.3.1 Klinische Studien

Aufgrund ihrer bereits beschriebenen Eigenschaften sind AAV-basierte Vektoren sehr gut für gentherapeutische Ansätze geeignet. In den 1980er-Jahren wurden die ersten AAV-Vektoren beschrieben<sup>98-100</sup>. Ungefähr ein Jahrzehnt später wurde die erste klinische Studie mit einem AAV-Vektor zur Behandlung der zystischen Fibrose im Menschen durchgeführt<sup>9,101,102</sup>. Seitdem wurden 137 klinische Studien mit AAV durchgeführt oder laufen noch<sup>7</sup>. Die meisten therapeutischen Ansätze betreffen monogenetische Erkrankungen wie Hämophilie B, zystischen Fibrose, oder Duchenne Muskeldystrophie oder sind Ansätze zur Behandlung von Krebs<sup>12</sup>. Neun Jahre nachdem in China das erste Gentherapeutikum Gendicin (rekombinantes Adenovirus mit Tumorsuppressor-Gen p53) gegen Tumore im Kopf-Hals -Bereich zugelassen wurde, wurde im Jahre 2012 das erste Gentherapeutikum in Europa zugelassen<sup>103,104</sup>. Alipogen Triprovoc (Glybera, AAV1-LPLS447X) ist eine Ser447X-Variante der Lipoproteinlipase in AAV1 und wird intramuskulär gegen die Lipoproteinlipase-Defizienz eingesetzt<sup>105</sup>.

#### 1.3.2 Produktion von rekombinanten AAV (rAAV) für gentherapeutische Ansätze

Eine wesentliche Struktur des AAV Genoms sind die ITRs. Sie sind die Voraussetzung für die Replikation des Genoms und die Verpackung dessen in das Kapsid<sup>106</sup>. Da Rep- und Cap-Proteine in *trans* bereitgestellt werden können, kann bei der Herstellung von rAAV theoretisch jedes beliebige Gen, das zwischen die ITRs kloniert wurde, in rAAV-Kapside verpackt werden (**Abbildung 4**). Hierbei sind die Gene in ihrer Größe auf die des natürlichen AAV Genoms limitiert.

Für die Produktion von rAAV werden drei Plasmide in Zellen einer passenden Zelllinie co-transfiziert<sup>107</sup>. Diese drei Plasmide codieren für 1.) das jeweilige Transgen, welches durch die ITRs flankiert ist, 2.) die *rep* und *cap* Sequenzen ohne ITRs und 3.) die für die Produktion notwendigen adenoviralen Helfergene (E2A, E4, VA)<sup>108,109</sup>. Die so produzierten rAAV tragen das Gen von Interesse anstatt des AAV Genoms und können z.B. mittels Dichtegradientenzentrifugation oder Affinitätschromatographie gereinigt werden.



**Abbildung 4: Genom des AAV2-Wildtypen im Vergleich zu dem eines rekombinanten AAV-Vektors.**

Die *rep/cap*-Genkassette des AAV2-Wildtypen ist in dem rekombinanten AAV-Vektor durch das Transgen und den zugehörigen Promotor ausgetauscht und somit in *cis* vorhanden. Die genetische Information für die Strukturproteine wird bei der Produktion der rAAV-Vektoren in *trans* bereitgestellt.

### 1.3.3 AAV2-Tropismus

Das AAV2 besitzt einen breiten und gleichzeitig limitierten nativen Tropismus. Während z.B. in Hepatozyten, Skelettmuskeln, im Herzen, in Epithelzellen der Atemwege, Zellen des Gehirns und in einigen Krebszelllinien ein lang anhaltender Gentransfer gezeigt werden konnte<sup>12</sup>, sind manche Zellen wie embryonale Stammzellen<sup>110</sup> und endotheliale Zellen<sup>111</sup> wenig oder gar nicht permissiv für AAV2. Die Verwendung von AAV2-Vektoren für gentherapeutische Ansätze ist somit zum einen auf transduzierbare Zelltypen limitiert und zum anderen nach systemischer Applikation *in vivo* unspezifisch. Folglich ist die Veränderung des unspezifischen AAV2-Tropismus zu einem spezifischen neuausgerichteten Tropismus von großem Interesse.

### 1.3.4 Gewebespezifische Promotoren und AAV-Serotypen

Eine Möglichkeit der Vektorausrichtung ist die Verwendung von gewebespezifischen Promotoren<sup>112</sup>. Da es nicht für alle Zelltypen spezifische Promotoren gibt und nicht alle Zelltypen permissiv für AAV sind, ist dieser Ansatz aber limitiert<sup>12</sup>. Eine weitere Möglichkeit den natürlichen Tropismus von AAV zu verändern, ist der Einsatz von unterschiedlichen AAV-Serotypen. Da verschiedene Serotypen in unterschiedlichen Geweben variierende Transduktionseffizienzen zeigen<sup>113</sup>, ist es möglich, den Tropismus von AAV2 durch die Verwendung von verschiedenen und neu isolierten Serotypen, sowie durch die Pseudotypisierung von Serotypen (Produktion von Viruspartikeln deren Genom und Kapsid von verschiedenen Serotypen stammen) zu erweitern. Allerdings gibt es, ähnlich wie bei den Promotoren, weniger Serotypen als Zelltypen<sup>12</sup>. Die Herstellung von chimären oder von sogenannten „Mosaik-Vektoren“ ist eine weitere Möglichkeit, den Tropismus von AAV-Vektoren zu verändern. Hierbei werden im Falle der chimären Vektoren Kapsidomänen oder -sequenzen ausgetauscht und im Falle der Mosaik-Vektoren ganze Kapsiduntereinheiten von verschiedenen Serotypen oder

Kapsidmutanten und somit ihre zugehörigen vorteilhaften Eigenschaften kombiniert<sup>12</sup>. Allerdings ist die Reproduzierbarkeit der Stöchiometrie der Kapsiduntereinheiten sehr schwierig und die Wahrscheinlichkeit der Existenz von Antikörpern gegen mindestens einen der Serotypen relativ hoch<sup>12,114</sup>. Zur Umgehung dieser Probleme wurden Strategien, die die direkte Modifizierung der Kapsidstruktur beinhalten, entwickelt.

### 1.3.5 Modifikation des AAV-Kapsids

AAV-Kapside können indirekt und direkt modifiziert werden. Eine indirekte Methode besteht darin, rezeptorbindende Liganden wie bi-spezifische  $F(ab'\gamma)_2$ -Antikörper an das Kapsid zu konjugieren<sup>115</sup>. Eine weitere Möglichkeit ist die Verwendung von Avidin-gekoppelten Fusionsproteinen in Kombination mit biotinylierten Kapsiden<sup>12,116</sup>. Im Gegensatz zum direkten Ansatz kann die Verwendung von Konjugaten *in vivo* zu einer erhöhten Immunogenität und Instabilität des Vektors führen<sup>12</sup>.

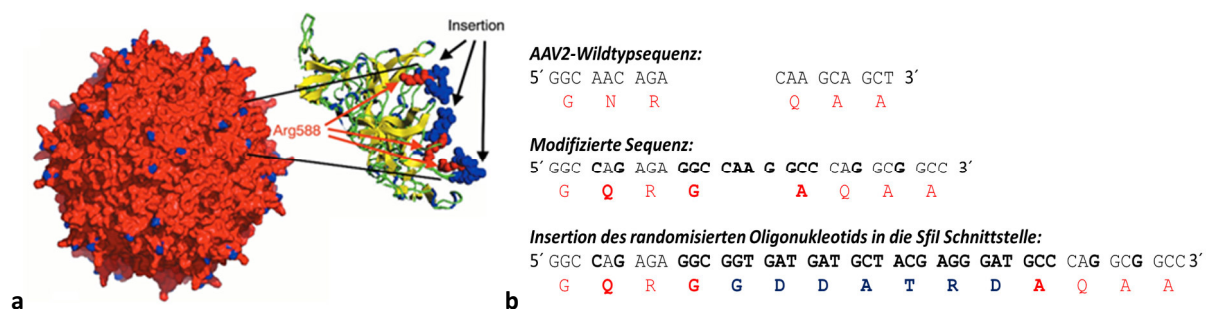
Durch das Einfügen einer für einen Liganden codierenden Aminosäuresequenz in das Kapsid wird dieses auf direktem Wege modifiziert. Der erste erfolgreiche Versuch, AAV-Vektoren auf ursprünglich nicht permissive Zellen auszurichten war die Insertion eines 14 Aminosäuren langen Peptids mit sogenanntem RGD-Motiv in die Aminosäureposition 587 des *cap*-Gens<sup>12,58</sup>. Aus verschiedenen Gründen eignet sich diese Position besonders gut für Peptidinsertionen. So konnte gezeigt werden, dass die Insertion von 14 Aminosäuren keinen negativen Einfluss auf die Verpackung des viralen Genoms oder die Struktur des Kapsids hat<sup>58</sup>. Des Weiteren ist die Aminosäure 587 Teil der "threefold spikes". Aufgrund der Kapsidstruktur sind sie häufig (60-mal) und gut exponiert auf der Virusoberfläche vorhanden und eignen sich somit sehr gut für die Interaktion mit zellulären Rezeptoren. Außerdem konnte gezeigt werden, dass Peptidinsertionen in der Nähe der Aminosäuren 585/587 (Teile des Heparin-Bindungsmotivs) potenziell mit der Bindung an den zellulären Rezeptor HSPG interferieren<sup>56</sup>, was zu einer Inhibierung des natürlichen Tropismus führen und für eine Neuausrichtung von Vorteil sein kann. Des Weiteren reduzieren Modifikationen der Position 587 den neutralisierenden Effekt von humanen Antikörpern<sup>117</sup>. Zusätzlich zu der Position 587 wurden noch weitere Positionen, an denen Peptidinsertionen toleriert werden und eventuell den Tropismus von AAV2 verändern können, identifiziert<sup>118</sup>.

Um auch ohne Kenntnis des entsprechenden Rezeptors Peptidliganden identifizieren zu können, wurden Phagen-Bibliotheken (*phage display libraries*) mittels *in vivo* Phage Display selektiert<sup>119-121</sup>. Die selektierten Peptidliganden wurden in AAV-Kapside kloniert und konnten, zum Teil Zellen, die die entsprechenden Rezeptoren trugen, erfolgreich erreichen<sup>12</sup>. Dennoch hat diese Vorgehensweise ihre Limitierung, da die selektierten Peptide nun im Kontext des AAV-Kapsids vorliegen und die neue strukturelle Umgebung einen sterischen Einfluss auf die Konformation des eingefügten Peptides

haben kann. Hierdurch wiederum kann es zu einer Veränderung der ursprünglichen Spezifität und Affinität des Peptides kommen und der dieses Peptid präsentierende Vektor hätte folglich nicht die gleichen Bindungseigenschaften wie das freie im Phagenkontext selektierte Peptid<sup>122</sup>. Zur Umgehung dieses Problems wurde zur Identifikation von spezifisch transduzierenden Vektoren eine Selektionsmethode, die auf der Modifikation des Vektorkapsids selber beruht, entwickelt<sup>122–124</sup>.

### 1.3.6 Randomisierte AAV-Peptidbibliotheken (*random AAV display peptide libraries*)

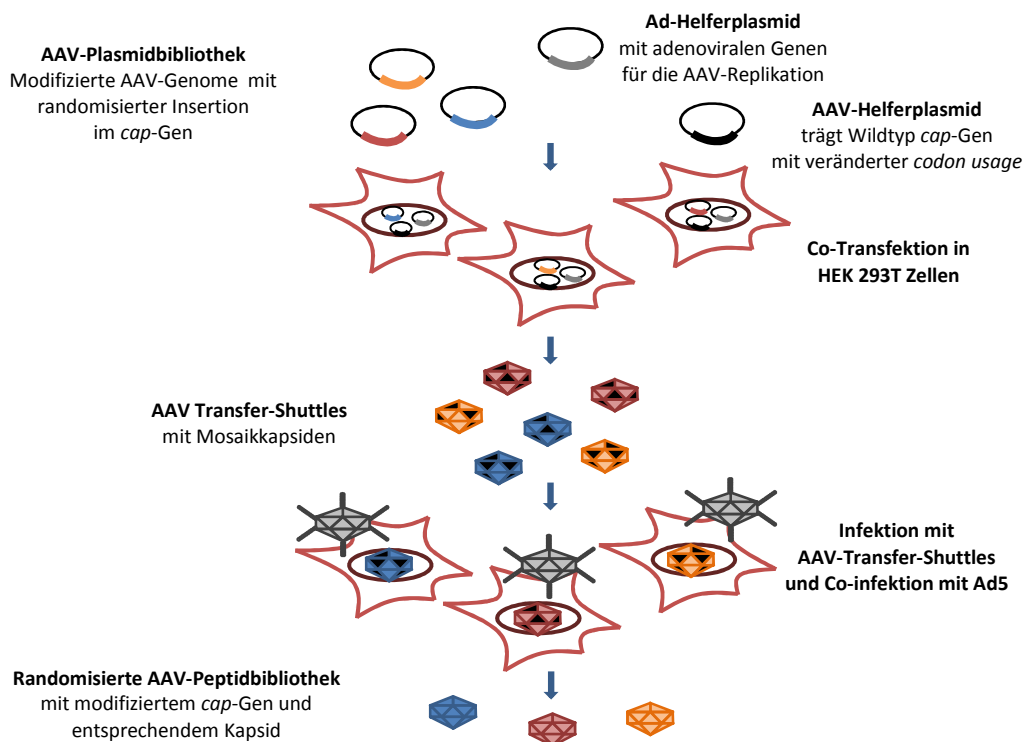
2003 entwickelten Perabo et al.<sup>124</sup> und parallel dazu Müller et al.<sup>122</sup> zwei ähnliche und dennoch leicht variierende Methoden zu Herstellung und Selektion von randomisierten AAV-Peptidbibliotheken<sup>122,124</sup>. Beide Methoden beruhen auf der Insertion von randomisierten sieben Aminosäuren langen Peptiden in das AAV-Kapsid. Die von Müller et al. generierte Kapsidmodifikation, auf dessen die in dieser Arbeit verwendeten Peptidbibliotheken beruhen, ist in **Abbildung 5** dargestellt.



**Abbildung 5:** Schematische Darstellung des AAV2-Kapsids mit Peptidinsertion an Position R588 (VP1) und Teilsequenz des *cap*-Gens mit modifizierter Insertionsstelle. **a)** Struktur des AAV2 Kapsids und molekulare Struktur der Insertionsstelle R588 mit integriertem sieben Aminosäure langem Peptid. Das integrierte Peptid (blau) ist 60 mal über das gesamte Kapsid gut exponiert in den "threefold-spikes" angeordnet. (Übernommen mit Erlaubnis der Macmillan Publishers Ltd: [*Nat. Biotechnol.*], Müller et al.<sup>122</sup>, ©(2003).) **b)** Teilsequenz des *cap*-Gens mit Insertionsstelle. Gezeigt sind die AAV2-Wildtypsequenz, die zur Generierung der *SfiI*-Schnittstellen modifizierte Sequenz mit Platzhalter (generiert eine Leserasterverschiebung, wenn der Platzhalter nicht durch das Insert getauscht wird) zur Verhinderung der Produktion von Kapsiden ohne Peptidinsertion und die modifizierte Sequenz mit Insertion. Schwarz: DNS-Sequenz; schwarz hervorgehoben: Modifizierte DNS-Sequenz; rot und blau: Aminosäuresequenz im Ein-Buchstaben-Code; rot hervorgehoben: Modifizierte Aminosäuresequenz; blau hervorgehoben: Peptidinsertion. Nach Müller et al.<sup>122</sup>, modifiziert.

Neben der sich um eine Aminosäure unterscheidenden Insertionsstelle, N587<sup>124</sup> und R588<sup>122</sup> des *cap*-Gens, wurden die AAV-Peptidbibliotheken einmal auf direktem und einmal auf indirektem Wege hergestellt. Der von Perabo et al. verfolgte direkte Weg besteht aus zwei Schritten 1.) der Generierung der Plasmidbibliothek und 2.) der Produktion der Virusbibliothek mittels Co-Transfektion von a) Plasmiden aus der Bibliothek und b) einem adenoviralem Helferplasmid in einem molaren Verhältnis von 1:1<sup>124</sup>. Bei dieser Vorgehensweise könnte es aber passieren, dass während der Transfektion mehr als ein Plasmid aus der Bibliothek dieselbe Zelle transfiziert, so dass chimäre Kapside mit verschiedenen modifizierten Cap Untereinheiten, von denen nur eines im Genom codiert

ist, produziert würden<sup>122</sup>. Aus diesem Grund wurde von Müller et al. der indirekte Weg zur Herstellung von AAV Peptidbanken eingeführt. Dieser Ansatz impliziert 1.) die Generierung einer Plasmidbibliothek mit randomisierten Oligonukleotidinsertionen, 2.) die Herstellung von sogenannten Transfer-Shuttles mit Mosaik-Kapsiden und 3.) die Produktion der Virusbibliothek (**Abbildung 6**). Die Herstellung von Transfer-Shuttles erfolgt durch Co-Transfektion von a) Plasmiden aus der Bibliothek, b) dem Helferplasmid, welches das Wildtyp *cap*-Gen ohne ITRs codiert und c) dem adenoviralen Helferplasmid. Hierbei entstehen Mosaik-Kapside, die aus einer Mischung aus Wildtyp- und modifizierten Cap-Proteinuntereinheiten bestehen. Für die anschließende Produktion der Virusbibliothek werden im dritten Schritt HEK 293T Zellen mit den Transfer-Shuttles in einer möglichst niedrigen MOI zur Verhinderung von Mehrfachinfektionen und anschließend mit Adenoviren zur Amplifikation der AAV co-infiziert<sup>122</sup>. Auf diese Weise werden AAV-Peptidbibliotheken mit einer Diversität von bis zu  $10^8$  verschiedenen Kapsidvarianten und Titern von  $1 \times 10^{12}$  viralen Genomen hergestellt. Da es bei der Produktion von Transfer-Shuttles zu homologen Rekombinationen zwischen dem modifiziertem *cap*-Gen und dem Wildtyp *cap*-Gen kommen kann, verbesserten Waterkamp et al.<sup>125</sup> 2006 die Qualität der Peptidbibliotheken, indem das Wildtyp *cap*-Gen gegen ein Codon-modifiziertes *cap*-Gen, das die homologe Rekombination verhindert, ausgetauscht wurde<sup>125</sup>.



**Abbildung 6: Schema zur Produktion von AAV-Peptidbibliotheken mittels Transfer-Shuttle-Produktion.**

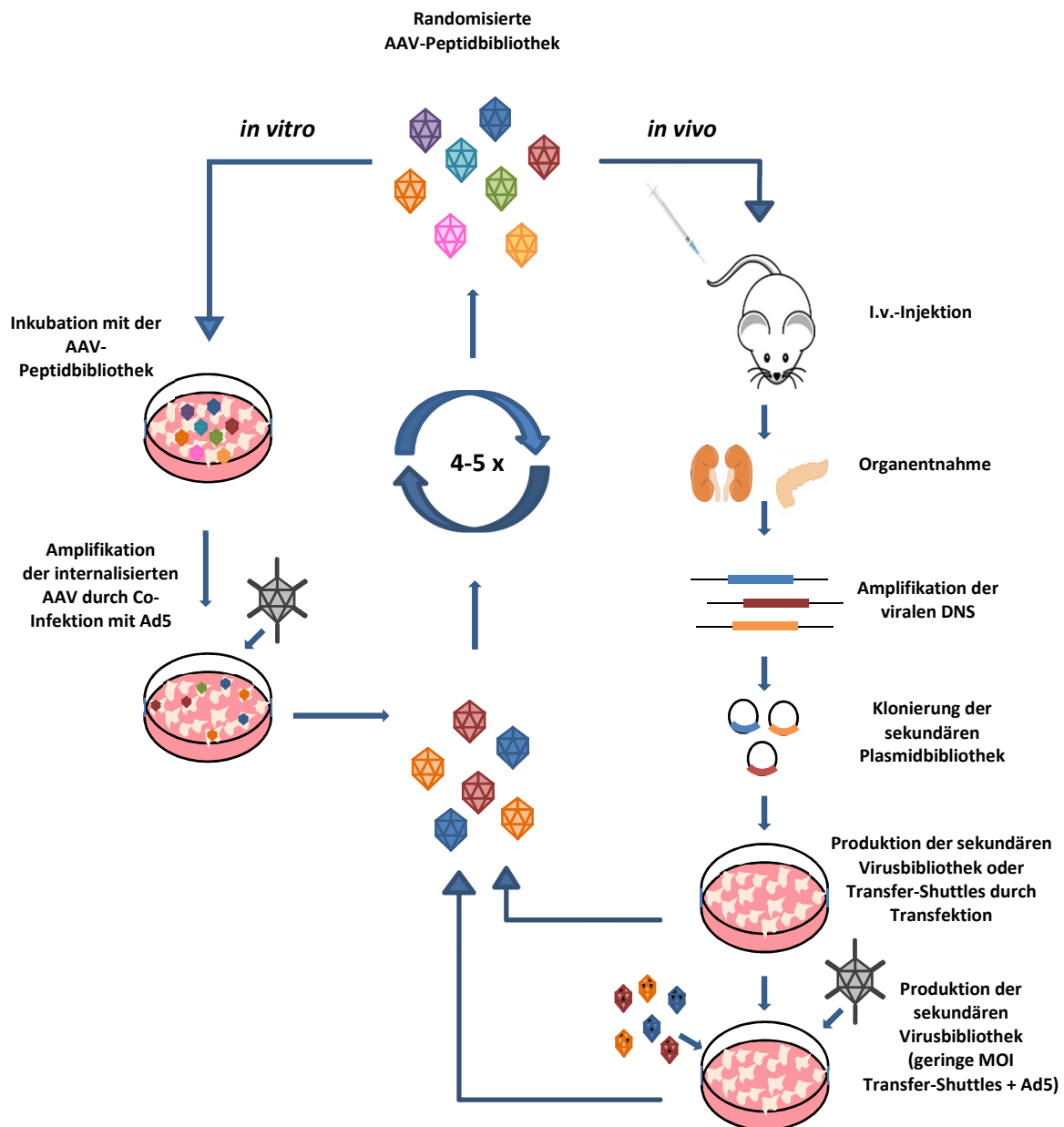
Zur Herstellung von AAV-Peptidbibliotheken werden zunächst Transfer-Shuttles mit Mosaik-Kapsiden mittels Tripel-Transfektion von a) Plasmiden der Plasmidbibliothek, b) dem Helferplasmid mit *codon usage* modifiziertem Wildtyp *cap* Gen ohne ITRs und c) dem adenoviralen Helferplasmid hergestellt. Anschließend werden zur Produktion der finalen Virusbibliothek HEK 293T Zellen mit den Transfer-Shuttles in geringer MOI infiziert und danach mit Ad5 co-infiziert. Nach Müller et al.<sup>122</sup> und Waterkamp et al.<sup>125</sup>, modifiziert.

### 1.3.7 *In vitro*- und *in vivo*-Selektion von randomisierten AAV-Peptidbibliotheken

Zur Identifikation von spezifischen AAV-Vektoren können AAV-Peptidbibliotheken *in vitro* und *in vivo* selektiert werden. Bei der *in vitro*-Selektion werden die Zielzellen zunächst mit der AAV-Bibliothek und später mit Adenovirus co-infiziert. Erfolgreich infizierte AAV können in Gegenwart des Adenovirus replizieren und reichern sich hierbei mit jeder Selektionsrunde an, während AAV, die nicht vollständig transduzieren, nach und nach aus der Bibliothek verschwinden (**Abbildung 7**). Auf diese Weise konnten Kapsid-modifizierte AAV in verschiedensten Zellen *in vitro* selektiert werden<sup>12,122,124,126–128</sup>. Neben der Insertion von sieben Aminosäuren langen Peptiden wurden auch Peptidbibliotheken mit fünf, zwölf, 19 und 26 Aminosäuren langen Peptidinsertionen in der Kapsidposition 588 generiert und *in vitro* auf zwei Zelltypen, primäre HCAEC (*Human Coronary Artery Endothelial Cells*) und H9C2 (Cardiomyoblasten aus der Ratte), selektiert<sup>129</sup>. Die hierbei selektierten Klone wurden auf Transduktionseffizienz und -Selektivität in 44 humanen Tumorzelllinien getestet. Hierbei zeigten die verschiedenen Vektoren ähnliche Transduktionseffizienzen aber unterschiedliche Selektivitäten, was darauf schließen lässt, dass die Insertion von verschieden langen Peptiden zu unterschiedlichen Zellrezeptorspezifitäten der entsprechenden Vektoren führen kann<sup>129</sup>.

Die *in vivo*-Verwendung von *in vitro* selektierten Vektoren ist aber aus verschiedenen Gründen begrenzt. Zum einen unterscheidet sich die Rezeptorausstattung der Zellen aufgrund der Kultivierung oder der Tatsache, dass es sich um Zelllinien handelt oftmals von der *in vivo* Situation<sup>130</sup>, zum anderen findet bei der *in vitro*-Selektion keine Negativselektion statt. Zusätzlich können Vektoren *in vivo* unspezifisch durch die Bindung an die extrazelluläre Matrix oder durch das mononukleäre Phagozytensystem und das retikuloendotheliale System in der Milz, Leber und Niere weggefangen werden<sup>131</sup>.

Im Gegensatz zu einer *in vitro*-Selektion ist eine *in vivo*-Selektion wesentlich aufwendiger. Nach der Injektion und Zirkulation der Virusbibliothek in der Maus wird die virale das Insert beinhaltende DNS aus den entsprechenden Geweben isoliert und mittels PCR amplifiziert, um eine Superinfektion mit Adenovirus zu umgehen. Anschließend wird das amplifizierte Insert zur Herstellung der Plasmidbibliothek in Plasmide kloniert. Die Plasmidbibliothek wird dann entweder direkt mittels Transfektion oder indirekt über den Schritt der Produktion von Transfer-Shuttles zur Produktion der sekundären Virusbibliothek verwendet (**Abbildung 7**)<sup>122</sup>. AAV-Vektoren wurden in der Maus z.B. in der Lunge<sup>108</sup>, dem Herzen<sup>132</sup> und in Brusttumoren<sup>133</sup> selektiert. Hierbei zeigte jedoch nur der im Herzen selektierte Vektor Spezifität. Erst kürzlich gelang es unserer Arbeitsgruppe, zwei AAV2-Vektoren zu selektieren, die das Endothel des Gehirns bzw. der Lunge spezifisch transduzieren<sup>134,135</sup>.



**Abbildung 7: Prinzip der Selektion von randomisierten AAV-Peptidbibliotheken.**

**Links:** *In vitro*-Selektion einer randomisierten AAV-Peptidbibliothek. Nach der Inkubation der Zellen mit der AAV-Peptidbibliothek werden internalisierte AAV durch Co-Infektion mit Ad5 amplifiziert. Die hierbei produzierte sekundäre Virusbibliothek wird für die nächste Selektionsrunde eingesetzt. **Rechts:** *In vivo*-Selektion einer randomisierten AAV-Peptidbibliothek. Die AAV-Peptidbibliothek wird intravenös in die Schwanzvene von Mäusen injiziert. Nach der Zirkulationszeit (hier zwei Tage) werden die zu untersuchenden Gewebe entnommen. Aus der isolierten Gesamt-DNS wird die virale DNS per PCR amplifiziert und in Plasmide kloniert. Mit der hierbei entstehenden Plasmidbibliothek wird die sekundäre Virusbibliothek entweder direkt nur über Transfektion oder indirekt über die Produktion von Transfer-Shuttles und anschließender Zellinfektion neuer Zellen in geringer MOI und Co-Infektion mit Ad5 produziert. (Bilder der Maus und der Organe aus [www.openclipart.org](http://www.openclipart.org)).



### 1.3.8 Die Niere und das Pankreas als gentherapeutisches Ziel

Für viele Krankheiten der Niere sind gentherapeutische Behandlungsstrategien denkbar. Hierzu gehören u.a. die diabetische Nephropathie, das Alport-Syndrom, die Zystenniere und vererbte Funktionsstörungen der Tubuli<sup>136</sup>. Doch die komplexe Anatomie, eine Vielzahl verschiedener Zelltypen und der Zugang zu diesen aufgrund ihrer Lokalisation oder des Fehlens einer adäquaten Blutversorgung machen bestimmte Kompartimente der Niere zu einem sehr schwierig erreichbaren Ziel und die Anzahl gentherapeutischer Studien sind eher limitiert<sup>137–140</sup>.

Neben der Niere ist auch das Pankreas aus gentherapeutischer Sicht ein sehr interessantes Zielorgan. Mit dem Ziel das Pankreas, insbesondere die Langerhans-Inseln, effizient und möglichst spezifisch zu transduzieren, wurden verschiedene AAV Serotypen *ex vivo* und *in vivo* getestet<sup>141–143</sup>. Wang et al.<sup>143</sup> untersuchten neben der Transduktion verschiedener Serotypen (1, 2, 5, 6 und 8) unterschiedliche Wege der Vektorverabreichung *in vivo*. Hierfür wurden rekombinante dsAAV mit GFP als Reportergen produziert und die Genexpression in den Langerhans-Inseln nach interperitonealer, intraductaler und intravenöser Injektion verglichen<sup>143</sup>. Sie beobachteten, dass nach intraperitonealer (AAV1, 2, 5, 6 und 8) und intraductaler (AAV2, 6 und 8) Injektion hauptsächlich die peripheren Zellen der Langerhans-Inseln Genexpression zeigten. Nach intravenöser Injektion von AAV6 und AAV8 und zusätzlicher transients Blockade der Leber hingegen wurden die meisten Inseln gleichmäßig transduziert<sup>143</sup>. Folglich ist eine intravaskuläre Injektion die beste Variante die Langerhans-Inseln gleichmäßig zu transduzieren. Der Einsatz eines gewebespezifischen Promotors sorgte zwar für eine spezifische Genexpression, ohne die transiente Blockade der Leber konnte aber kaum eine Genexpression in den Langerhans-Inseln detektiert werden, da das Virus sich im Pankreas nicht anreicherte<sup>143</sup>. Ein Vektor, der sich von sich aus nach systemischer Injektion im Pankreas anreichert wäre somit von Vorteil.

## 1.4 Next-Generation Sequencing zur Auswertung der Selektion von AAV-Peptidbibliotheken

### 1.4.1 DNS-Sequenzierung der ersten, zweiten und dritten Generation

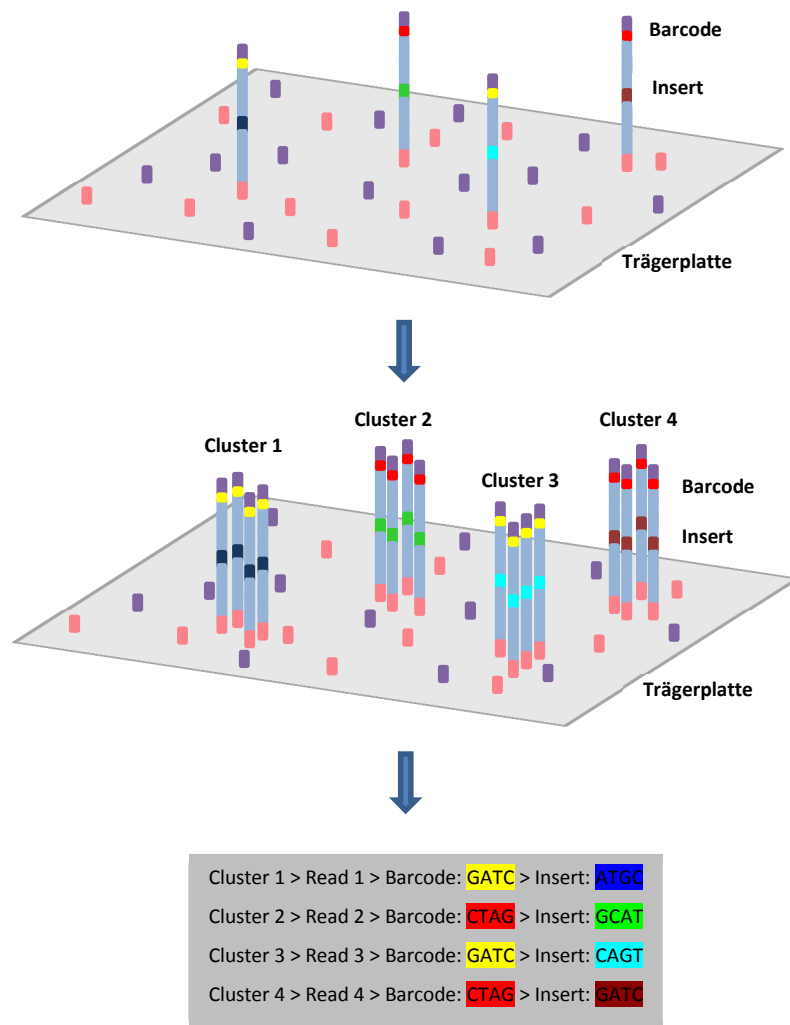
In den 1970-er Jahren wurde zur Sequenzierung von DNS-Sequenzen neben der Fragmentierungsmethode von Maxam und Gilbert<sup>144</sup> die Didesoxymethode oder auch Kettenabbruch-Synthese nach Sanger<sup>145</sup> entwickelt und blieb lange Zeit die Methode der Wahl. Im Laufe der Jahre und im Rahmen des humanen Genomprojektes wurde allerdings die Nachfrage nach effizienteren Sequenzierungsmethoden stärker. Ende der 1980-er Jahre wurden aufbauend auf der Sanger-Methode die ersten automatisierten Sequenzierungsgeräte, Geräte der „ersten Generation“, eingeführt und trieben das humane Genomprojekt, das 1990 startete, voran<sup>146</sup>. Das gesamte Projekt war auf 15 Jahre mit einer Förderung von drei Milliarden US Dollar ausgelegt; im Jahre 2004 wurde

das komplette humane Genom publiziert. Aufgrund der hohen Kosten und des enormen Zeitaufwandes schrieb das NHGRI (*National Human Genome Research Institute*) im selben Jahr ein Förderungsprogramm aus, mit dem Ziel innerhalb von zehn Jahren den Preis für die Sequenzierung eines humanen Genoms auf 1.000 US Dollar zu reduzieren<sup>146</sup>. Bereits ein Jahr später revolutionierten Sequenzierungsgeräte der „nächsten Generation“ den Markt. Mittels massiver paralleler Tiefensequenzierung war es nun möglich bis zu einer Gigabase pro Lauf zu sequenzieren. Damit war das Zeitalter des *Next-Generation Sequencings* (NGS) angebrochen. Verschiedene Sequenzierungsmethoden und -geräte wurden entwickelt. Durch weitere technische Verbesserungen ist es seit 2014 sogar möglich 1,8 Terabasen pro Lauf zu sequenzieren. Sowie die Effizienz erhöht wurde, verringerte sich der Preis; nun war es möglich 45 humane Genome an nur einem Tag für 1.000 Dollar pro Genom zu sequenzieren und das Ziel des NHGRI war somit erreicht (<http://www.illumina.com>).

### 1.4.2 Illumina® Next-Generation Sequencing

Im Laufe der Zeit haben sich verschiedene Sequenzierungsmethoden und Anbieter solcher etabliert. Aufgrund der vielfältigen Anwendungsmöglichkeiten, dem hohen Durchsatz und den geringsten Kosten pro Base stellt Illumina® den derzeitigen Marktführer in Sachen NGS dar<sup>146</sup>.

Das Prinzip der Illumina® Sequenzierung besteht grob betrachtet aus vier Schritten, 1.) der Herstellung der Bibliothek, 2.) der Generierung von sogenannten Clustern, 3.) der Sequenzierung und 4.) der Datenanalyse. Je nach Anwendung ändert sich die Art der Probenvorbereitung bzw. der Datenanalyse, während der Vorgang der Sequenzierung im Wesentlichen gleich bleibt. Im Allgemeinen wird bei der Herstellung der Bibliothek die zu untersuchende DNS oder cDNS mit sogenannten Barcode-beinhaltenen Adaptern versehen. Die hierbei entstehenden Fragmente werden im zweiten Schritt auf eine Trägerplatte, die auf ihrer Oberfläche zu den Adaptern der Bibliothek komplementäre Oligonukleotide präsentiert, gegeben und mittels Brückensynthese amplifiziert. Auf diese Weise entsteht für jede gebundene Sequenz ein ganzes Sequenz-Cluster (**Abbildung 8**). Im Anschluss werden die DNS-Fragmente durch Sequenzierung während der Synthese (*sequencing by synthesis*) analysiert. Hierbei werden fluoreszenz-markierte Nukleotide in den wachsenden DNS-Strang eingebaut und die jeweiligen Basen anhand der Emissionswellenlänge des Fluoreszenzfarbstoffes der Sequenzen eines Clusters nach jedem Zyklus identifiziert. Schließlich werden die Daten je nach Anwendung bioinformatisch durch den Vergleich mit Referenzsequenzen ausgewertet (<http://www.illumina.com>).



**Abbildung 8: Schematische Darstellung der Clusterbildung beim NGS nach Illumina®.** Nach der Bindung der DNS-Fragmente an die immobilisierten Adapter der Trägerplatte, werden sie in mehreren Schritten mittels Brückensynthese amplifiziert. Hierbei entstehen Cluster gleicher Sequenzen. Die Sequenzierung erfolgt während der Synthese (*sequencing by synthesis*). Nach der Sequenzierung werden die Sequenzen mit den jeweiligen Referenzsequenzen verglichen und können anhand der Barcodes korrekt zugeordnet werden.

### 1.4.3 Next-Generation Sequencing zur Untersuchung von Peptidbibliotheken

Die wichtigste Kenngröße einer Peptidbibliothek ist ihre Peptiddiversität. Während die Peptiddiversität einer primären (unselektierten) Peptidbibliothek so groß wie möglich sein sollte, um sämtliche mögliche Peptidvarianten abzudecken, sollte sich die Diversität während einer Selektion aufgrund der Anreicherung von potenziell spezifischen Klonen deutlich verringern. Im Umkehrschluss ist es möglich, anhand einer sich verändernden Peptiddiversität die Stärke der Anreicherung und somit den Erfolg einer Selektion zu bestimmen.

Während die Diversität einer Plasmidbibliothek anhand der Anzahl der transformierten Bakterienklone bestimmt werden kann, konnte die Diversität einer Virusbibliothek vor der

Möglichkeit der Tiefensequenzierung nur durch die stichprobenartige Sequenzierung von zufällig gewählten Einzelklonen abgeschätzt werden. Mit der Entwicklung des NGS ist es zum einen möglich, die Diversität von Peptidbibliotheken zu bestimmen, zum anderen ermöglicht es einen wesentlich tieferen Einblick in die Dynamik der Selektion einer Peptidbibliothek. Zusätzlich kann mit Hilfe dieser Methode die Suche und Auswahl potenziell gewebespezifischer Klone vereinfacht werden<sup>135</sup>.

Während der Gehirndothel-spezifische Vektor „AAV-BR1“<sup>134</sup> und der Lungenendothel-spezifische Klon konventionell durch die Analyse in einer zufälligen kleinen Stichprobe sequenzierter Klone nach *in vivo*-Selektion der AAV-Peptidbibliothek gefunden wurde, wurde letzterer zusätzlich mittels der Methode des Next-Generation Sequencings identifiziert<sup>135,147</sup>. Hierfür wurden Bewertungskriterien zur Bestimmung der Anreicherung bestimmter Sequenzen während der Selektion und ihrer Spezifität gegenüber nicht-Zielorganen formuliert. Anhand der Berechnung eines Anreicherungs-Scores (*Enrichment Score*), eines Spezifitäts-Scores (*General Specificity Score*) und eines kombinierten Scores (*Combined Scores*) wurde dann der Klon identifiziert, der die höchste Wahrscheinlichkeit besitzt spezifisch zu sein<sup>135</sup>.

## 1.5 Zielsetzung der Arbeit

Die Selektion von AAV-Peptidbibliotheken stellt eine äußerst vielversprechende Methode zur Identifikation von gewebespezifischen Gentherapievektoren dar. Erst kürzlich wurden von unserer Gruppe zwei gewebespezifische Vektoren *in vivo* selektiert<sup>134,135</sup> und weitere Selektionen sind geplant. Trotz der intensiven Erforschung des AAV als Gentherapievektor und obwohl seit über zehn Jahren Kapsid-modifizierte AAV in Form von AAV-Peptidbibliotheken *in vitro* und *in vivo* selektiert werden, besteht bis heute weder eine genaue Kenntnis über die tatsächliche Diversität der verwendeten primären und sekundären Peptidbibliotheken, noch wurde die Anreicherung bestimmter Peptidsequenzen oder Sequenzmotive systematisch untersucht.

Das Ziel dieser Arbeit war es, mit Hilfe des NGS einen tieferen und besseren Einblick in die Komplexität der Selektion von AAV-Peptidbibliotheken zu erhalten. Hierzu sollte gezielt die Anreicherung und Veränderung von Peptidsequenzen und Sequenzmotiven während der *in vivo*-Selektion von AAV2-Peptidbibliotheken nach systemischer Injektion in die Maus untersucht werden. Aufgrund der erfolgreichen Selektion eines Gehirndothel- und eines Lungenendothel-spezifischen AAV-Vektors sollte gleichzeitig die Übertragbarkeit der Selektionsmethode auf die Zielgewebe Niere und Pankreas, sowie die funktionellen endokrinen Einheiten des Pankreas, den Langerhans-Inseln, getestet werden. Zusätzlich sollten die sekundären Peptidbibliotheken der publizierten Selektion im Gehirn retrospektiv mittels NGS analysiert und mit den Ergebnissen der neuen Selektionen verglichen werden.

Des Weiteren sollte eine neue von uns generierte AAV2-Peptidbibliothek, die AAV2-Trimer<sub>6</sub>-(-C)-Peptidbibliothek, auf Anwendbarkeit *in vitro* getestet und mit unserer bisher verwendeten AAV2-NNK<sub>7</sub>-Peptidbibliothek im Hinblick auf Diversität und Abdeckung verglichen werden.

## 2 MATERIAL

### 2.1 Laborgeräte

Die in dieser Arbeit verwendeten Laborgeräte sind in **Tabelle 1** aufgelistet.

**Tabelle 1: Laborgeräte**

Geräte	Produktname	Hersteller/Vertreiber
Agarosegelelektrophoresekammer	Sub-Cell® GT, Wide Mini Sub-Cell® GT	Biorad
Anästhesiesystem	XGI-8 Gas Anesthesia System	Xenogen
Autoklav	Varioklav®	HP Medizintechnik
Digitale Mikroskopkamera	DP-25	Olympus
Elektroporationssystem	Gene Pulser Xcell™ Microbial System	Biorad
Fluoreszenzmikroskop	Axiovert 40 CFL	Carl Zeiss
Gefrierschrank (-20 °C)	Biomedical Freezer	Sanyo Biomedical
Gefrierschrank (-80 °C)	V.I.P.™ Series	Sanyo Biomedical
Geldokumentationssystem	Gene Genius Bio Imaging System	Syngene
Gelelektrophoresenetzteile	PowerPac™ Basic	Biorad
Gewebehomogenisator	Precellys®24	Precellys®/Bertin technologies
Kühltischzentrifuge	centrifuge 5417 R	Eppendorf AG
Kühlzentrifuge	Avanti® J-E, Rotors: JA-14, JA-20.1	Beckman Coulter
Lichtmikroskop	CK 2	Olympus
Luminometer	Mithras LB 940	Berthold Technologies
Magnetrührer	Hotplate stirrer Model L-81	A. Hartenstein
Microbiological incubator platform	Innova 2000	New Brunswick Scientific
Mikrobiologischer Inkubator	Kelvitron® t	Heraeus Instruments
Mikroplatten Photometer	Sunrise™	Tecan
pH-Meter	Microprocessor pH Meter	WTW
Pipetten	Eppendorf Research®plus	Eppendorf AG
Präzisionswaage	Lab Style 3002	Mettler Toledo
Proteinelektrophoresekammer	Mini-PROTEAN® Tetra Cell	Biorad
Real-time PCR cycler	LightCycler® Nano	Roche Diagnostics
Schüttelinkubator	Incubator Shaker	New Brunswick Scientific
Spectrophotometer	NanoDrop 2000c	Thermo Fisher Scientific
Sterilbank	HeraSafe HS 15	Heraeus Instruments
Thermoblöcke	Thermoblock	Peqlab
Thermocycler	Mastercycler® personal	Eppendorf AG
Tischschüttler	MaxQ 4000	Thermo Fisher Scientific
Tube toppler	Quick-Seal Cordless Tube Topper	Beckman Coulter
Ultrazentrifuge	Optima™ LE-80K, Rotor: Type 70.1	Beckman Coulter
UV-Lichttransilluminator	FLUO_LINK	BIOBLOCK SCIENTIFIC
Vortexer	peqTwist	Peqlab
Wasserbad	Thermomix ME	B. Braun
Zellkulturinkubator	HERAcell 240	Thermo Fisher Scientific
Zentrifuge, groß	Centrifuge 5810	Eppendorf AG
Zentrifuge, klein	Centrifuge 5424	Eppendorf AG

## 2.2 Molekularbiologische Kits

Verwendete Molekularbiologische Kits sind in **Tabelle 2** aufgelistet.

**Tabelle 2: Molekularbiologische Kits**

Kit	Hersteller
DNeasy® Blood & Tissue Kit (250)	Qiagen
NucleoBond® PC 100	Macherey-Nagel
NucleoBond® Xtra Maxi	Macherey-Nagel
PeqGold Gel Extraction Kit	Peqlab
PeqGold Plasmid Mini Kit	Peqlab
Qiaquick® PCR Purification Kit (250)	Qiagen

## 2.3 Reagenzien

In dieser Arbeit verwendeten Reagenzien sind in **Tabelle 3** aufgelistet.

**Tabelle 3: Reagenzien**

Reagenz	Hersteller/Vertreiber
6x Orange DNS Loading Dye	Thermo Scientific
Ethidiumbromid (10 mg/ml)	Sigma-Aldrich®
Luciferase Assay Reagent (LAR)	Promega
OptiPrep™ Dichtegradient Medium (60% w/v iodixanol)	Axis-Shield
PolyFect	Qiagen
Roti®-Safe Ladepuffer	Carl Roth®

## 2.4 Enzyme

Die in dieser Arbeit verwendeten Enzyme sind in **Tabelle 4** gelistet.

**Tabelle 4: Enzyme**

Enzym	Hersteller/Vertreiber
Benzonase® Nuklease ( $\geq 250$ U/ $\mu$ l)	Sigma-Aldrich®
Kollagenase P (>1.5 U/mg nach Wunsch)	Roche
FastStart Essential Green Mastermix (SYBR Green)	Roche
GoTaq® DNS Polymerase (5 U/ $\mu$ l)	Promega
HotStar Taq® DNS Polymerase (5 U/ $\mu$ l)	Qiagen
Phusion High Fidelity DNS Polymerase (2,000 U/ml)	New England Biolabs
Proteinase K	Qiagen
Restriktionsenzyme (divers)	Fermentas und New England Biolabs
RNase A (100 mg/ml)	Qiagen
T4 DNS Ligase (400,000 U/ml)	New England Biolabs
Trypsin/EDTA 10x Lösung	PAA Laboratories

## 2.5 Antikörper

Verwendete Antikörper sind in **Tabelle 5** beschrieben.

**Tabelle 5: Antikörper**

Bezeichnung	Beschreibung	Quelle
AK76.3	Monoklonaler Maus anti-AAV Rep, Überstand von Hybridoma Zellen	Prof. Dr. Jürgen Kleinschmidt (DKFZ, Heidelberg, Deutschland) <sup>148</sup> ,
sc-2010	Polyklonaler Ziege anti-Maus IgG, FITC-konjugiert	Santa Cruz

## 2.6 Chemikalien

Die in dieser Arbeit verwendeten Chemikalien wurden von den folgenden Herstellern bezogen: Affymetrix, BD Bioscience, Biorad, Carl Roth, Dako, Fermentas, Fluka, GE-Healthcare, Invitrogen, Merck, Polysciences, Serva, Sigma-Aldrich, Roche, Vetequip.

## 2.7 DNS-Marker

Für die Bestimmung der DNS-Fragmentgrößen mittels Gelelektrophorese wurden der Quick Load® 100 bp DNA Ladder (New England Biolabs) oder der Quick Load® 1 kp DNA Ladder (New England Biolabs) verwendet.

## 2.8 Verbrauchsmaterialien

Die in dieser Arbeit verwendeten Verbrauchsmaterialien wurden von den folgenden Herstellern bezogen: Beckman Coulter, Biorad, Braun, Eppendorf, Greiner, Miltenyi Biotec, Pierce, Sarstedt, Vetequip, Whatman.

## 2.9 Lösungen

### 2.9.1 Standardpuffer

PBS (*phosphate buffered saline*, Phosphat gepufferte Lösung), pH 7.4

Inhaltsstoff	Formel	Konzentration, 10x	Molarität (1x)
Natriumchlorid	NaCl	80 g/l	137 mM
Kaliumchlorid	KCl	2 g/l	2,68 mM
Dinatriumhydrogenphosphatdihydrat	Na <sub>2</sub> HPO <sub>4</sub> – 2H <sub>2</sub> O	17,8 g/l	10,0 mM
Kaliumdihydrogenphosphat	KH <sub>2</sub> PO <sub>4</sub>	2,4 g/l	1,76 mM

Eingestellt auf pH 7,4 und autoklaviert. Mit <sub>d</sub>H<sub>2</sub>O verdünnt.



PBS-MK (PBS mit Magnesium [MgCl<sub>2</sub>] und Kalium [KCl]), pH 7.4

Inhaltsstoff	Formel	Konzentration, 10x	Molarität (1x)
Natriumchlorid	NaCl	80 g/l	137 mM
Kaliumchlorid	KCl	2 g/l	2,68 mM
Dinatriumhydrogenphosphatdihydrat	Na <sub>2</sub> HPO <sub>4</sub> – 2H <sub>2</sub> O	17,8 g/l	10,0 mM
Kaliumdihydrogenphosphat	KH <sub>2</sub> PO <sub>4</sub>	2,4 g/l	1,76 mM
Magnesiumchlorid, 1M	MgCl <sub>2</sub>	0,1 % (v/v)	0,1 mM
Kaliumchlorid	KCl	186 mg/l	0,25 mM

Mit einem 0,2 Sterilfilter sterilfiltriert.

TAE (Tris-Acetate-EDTA)

Inhaltsstoff	Formel	Konzentration, 50x	Molarität (1x)
Tris	NH <sub>2</sub> C(CH <sub>2</sub> OH) <sub>3</sub>	242 g/l	40 mM
Na <sub>2</sub> EDTA (0.5 M, pH8)	C <sub>10</sub> H <sub>14</sub> N <sub>2</sub> Na <sub>2</sub> O <sub>8</sub>	10 % (v/v)	1 mM
Acetic acid	C <sub>2</sub> H <sub>4</sub> O <sub>2</sub>	5,71 % (v/v)	20 M

Mit ddH<sub>2</sub>O verdünnt.

## 2.10 Medien

### 2.10.1 Medien zur Kultivierung von prokaryotischen Zellen

LB (*lysogeny broth*), pH 7

Inhaltsstoff	Konzentration
Bacto Typtone	1 % (w/v)
Bacto Yeast extract	0,5 % (w/v)
Natriumchlorid	1 % (w/v)

Eingestellt auf pH 7 und autoklaviert.

TB (*terrific broth*), pH 7

Inhaltsstoff	Konzentration
Terrific Broth	5,04 1 % (w/v)
Glycerin	0,4 % (v/v)

Eingestellt auf pH 7 und autoklaviert.

SOB (*super optimal broth*), pH 7,4

Inhaltsstoff	Konzentration
Bacto Typtone	2 % (w/v)
Bacto Yeast extract	0,5 % (w/v)
Natriumchlorid	0,05 % (w/v)
Kaliumchlorid	2,5 mM

Eingestellt auf pH 7,4 und autoklaviert.

SOC (*super optimal broth with catabolite repression*), pH 7

Inhaltsstoff	Konzentration
SOB	
Magnesiumchlorid*	16 mM
Glucose*	20 mM

\*Sterilfiltriert.

## 2.10.2 Medien zur Kultivierung von eukaryotischen Zellen

### Zellkulturmedien

Medium	Abkürzung	Hersteller
Dulbecco's Modified Eagle Medium	DMEM	Gibco® Life Technologies™
RPMI Medium 1640	RPMI	Gibco® Life Technologies™

### Einfriermedium

Inhaltsstoff	Konzentration
Jeweiliges Medium	45 % (v/v)
FBS	45 % (v/v)
DMSO	10 % (v/v)

## 2.11 Zellen

### 2.11.1 Prokaryotische Zellen (*Escherichia coli*)

Verwendete Bakterienstämme sind in **Tabelle 6** beschrieben.

**Tabelle 6: Prokaryotische Zellen**

Bezeichnung	Beschreibung des Genotyps	Referenz
DH5α™	F <sup>-</sup> φ80d <i>lacZ</i> ΔM15 ( <i>lacZ</i> YA-argF) U169 <i>deoR recA1 endA1 hsdR17</i> ( <i>r<sub>K</sub><sup>-</sup>, m<sub>K</sub><sup>+</sup></i> ), <i>phoA supE44 λ<sup>-</sup> thi-1 gyrA69 relA1</i>	Referenz <sup>149</sup>
XL1-Blue	<i>recA1 endA1 gyrA96 thi-1 hsdR17 sup44 relA1 lac</i> [F' <i>proAB lacI<sup>q</sup>ΔM15 Tn10</i> (Tet <sup>r</sup> )]	Referenz <sup>150</sup>

### 2.11.2 Eukaryotische Zelllinien

Eukaryotische Zelllinien sind in **Tabelle 7** beschrieben. Alle Zellkulturmedien wurden mit 10 % (v/v) fötalem Kälberserum (FBS) und 1 % (v/v) Penicillin / Streptomycin versetzt.

**Tabelle 7: Eukaryotische Zellen**

Name	Beschreibung	Medium	Quelle/ Referenz
<b>AAV2 Produktion</b>			
HEK 293T	Humane embryonale Nierenzellen, tragen die Temperatursensitive Mutante des SV-40 large T-Antigens ( <i>tsA1609neo</i> ) und die Gene E1a, E1b und E4 des Adenovirus type 5	DMEM	Referenz <sup>151,152</sup>
<b>AAV2 in vitro-Selektion</b>			
4T1	Murine Tumorzelllinie der Brustdrüse einer Balb/c Maus.	RPMI	Referenz <sup>153</sup>
HT-29	Humane Tumorzelllinie aus dem Darm.	RPMI	Referenz <sup>154</sup>
MCF-7	Humane Tumorzelllinie aus der Brustdrüse.	DMEM	Referenz <sup>155</sup>
Paca5061	Humane Tumorzelllinie aus dem Pankreas.	RPMI	Prof. Dr. med. Udo Schumacher, UKE Hamburg, <sup>156</sup>

## 2.12 Antibiotika und Zusätze

In **Tabelle 8** sind verwendete Antibiotika und Zusätze beschrieben.

**Tabelle 8: Antibiotika und Zusätze**

Antibiotikum	Abkürzung	Konzentration	Konzentration (verwendet)
<b>Prokaryotische Zellkultur</b>			
Ampicillin	AMP	50 mg/ml	150 mg/l
<b>Eukaryotische Zellkultur</b>			
Penicillin	PEN	5,000 U/ml	50 U/ml
Streptomycin	STREP	5,000 µg/ml	50 mg/l

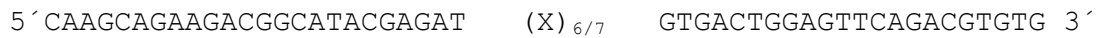
## 2.13 Oligonukleotide

Alle in dieser Arbeit verwendeten Oligonukleotide wurden von Metabion synthetisiert. In **Tabelle 9** sind sie aufgelistet.

**Tabelle 9: Oligonukleotide**

Bezeichnung	DNS-Sequenz [5´- 3´]	Anzahl der Basen	Tm [°C]	GC-Gehalt [%]	Referenz
<b>Primer für die Amplifizierung der Peptidinsertion</b>					
AAV2 3+	GGTTTCATCTTTGGGAAG	19	55	47,4	Referenz <sup>133</sup>
AAV2 3-	CTCCTCCACAGATTTCATCA	21	59	47,6	Referenz <sup>133</sup>
AAV2 4+	ATGGCAAGCCACAAGGACGATG	22	64	54,5	Referenz <sup>133</sup>
AAV2 4-	CGTGGAGTACTGTGTGATGAAG	22	62	50	Referenz <sup>133</sup>
<b>Primer für die AAV2 Quantifizierung</b>					
AAV2 1+	GCAGTATGGTTCTGTATCTACCAACC	26	66	46,2	Referenz <sup>133</sup>
AAV2 1-	GCCTGGAAGAACGCCCTTGTGTG	22	66	59,1	Referenz <sup>133</sup>
CMV fwd	GGGACTTTCCTACTTGCA	19	57	52,6	Referenz <sup>157</sup>
CMV rev	GCGGAGTTGTTACGACAT	19	57	52,6	Referenz <sup>157</sup>
<b>Primer für die NGS-Probenvorbereitung</b>					
NGS1FwdA	ACACTCTTCCCTACACGACGCTCTTCC GATCTGCTCCAGAGAGGCCAGAGAG	53	89	56,6	Dr. Timo Sieber
NGS1RevA	TGACTGGAGTTCAGACGTGTGCTCTTC CGATCTATGAGCATCTGCGGTGGCCGC CTG	57	90	57,9	Dr. Timo Sieber
NGS1FwdB	ACACTCTTCCCTACACGACGCTCTTCC GATCTTAGCTCCAGAGAGGCCAGAGA G	55	89	54,5	Dr. Timo Sieber
NGS1RevB	TGACTGGAGTTCAGACGTGTGCTCTTC CGATCTCATCATCTGCGGTGGCCGCT G	55	90	58,2	Dr. Timo Sieber
NGS1FwdC	ACACTCTTCCCTACACGACGCTCTTCC GATCTATCTCTCCAGAGAGGCCAGAGA G	56	88	53,6	Dr. Timo Sieber
NGS1RevC	TGACTGGAGTTCAGACGTGTGCTCTTC CGATCTTCATCTGCGGTGGCCGCTG	53	90	58,5	Dr. Timo Sieber
NGS2Fwd	AATGATACGGCGACCACCGAGATCTAC ACTCTTCCCTACACGAC	45	82	51,1	Dr. Timo Sieber

Die DNS-Sequenz der reversen die Barcodes beinhaltenden Oligonukleotide (Adaptorsequenzen von Illumina® konstruiert) ist die folgende:



In **Tabelle 10** sind die verwendeten Barcodesequenzen aufgeführt.

**Tabelle 10: Barcodesequenzen.**

Barcodenummer	Barcode DNS- Sequenz [5' - 3']	Anzahl der Basen	Tm [°C]	GC [%]	Referenz
5	CACTGT	52	87	51.9	Illumina®
6	ATTGGC	52	87	51.9	Illumina®
7	GATCTG	52	87	51.9	Illumina®
9	CTGATC	52	87	51.9	Illumina®
12	TACAAG	52	86	50	Illumina®
19	TTTCAC	52	86	50	Illumina®
132	TCAAGGC	53	88	52.8	Illumina®
133	GACTTAT	53	86	49.1	Illumina®
134	GTAGGGT	53	88	52.8	Illumina®
135	AGATTCA	53	86	49.1	Illumina®
136	CCCCATT	53	88	52.8	Illumina®
137	ACGAGGA	53	88	52.8	Illumina®
138	CCATCCA	53	88	52.8	Illumina®
139	CAATTTT	53	85	47.2	Illumina®
140	TACCCTT	53	87	50.9	Illumina®
141	ATGTAGA	53	86	49.1	Illumina®
142	ATCTACC	53	87	50.9	Illumina®
143	CAGTCCC	53	88	54.7	Illumina®
144	AAACGGA	53	87	50.9	Illumina®
145	TTTAGGA	53	86	49.1	Illumina®
146	GGGAGAA	53	88	52.8	Illumina®
147	CGTAATA	53	86	49.1	Illumina®
148	GAATATC	53	86	49.1	Illumina®
149	TACACAA	53	86	49.1	Illumina®
150	CAACCAT	53	87	50.9	Illumina®
151	GGTTAGG	53	88	52.8	Illumina®
152	GAGTCGA	53	88	52.8	Illumina®
153	TAGCGCT	53	88	52.8	Illumina®
154	GGTATGA	53	87	50.9	Illumina®
155	TACGTGC	53	88	52.8	Illumina®
156	TAATTGA	53	85	47.2	Illumina®
157	AGAACAA	53	86	49.1	Illumina®
158	CAGCACA	53	88	52.8	Illumina®
159	CAGAGAG	53	88	52.8	Illumina®

## 2.14 Mäuse

Der verwendete Mausstamm ist in **Tabelle 11** beschrieben.

**Tabelle 11: Mäuse**

Mausstamm	Beschreibung	Referenz
FVB/N	Immunokompetenter Inzucht-Albino-Stamm, geboren in der UKE Versuchstierhaltung	Referenz <sup>158</sup>

## 2.15 Viren

Die für diese Arbeit verwendeten Viren sind in **Tabelle 12** aufgeführt.

**Tabelle 12: Viren**

Virus	Beschreibung	Quelle
Ad5	Humans Wildtyp Adenovirus 5	Laboratoire de Thérapie Génique (Nantes France)
MAAdV Typ 2	Murines Adenovirus Typ 2	PD Dr. Silvio Hemmi, Institut für Molecular Life Science, Universität Zürich, Schweiz

## 2.16 Rekombinante AAV2

In **Tabelle 13** sind rekombinante AAV aufgelistet.

**Tabelle 13: Rekombinante AAV2**

Bezeichnung	Beschreibung	Referenz
AAV2-Transfer-Shuttles	AAV2 Transfer-Shuttles, tragen das AAV2 <i>rep</i> des pMT202-6 und das modifizierte AAV2 <i>cap</i> von pMT202-6 und pVP3cm, Mosaikkapside mit Codon modifizierten Wildtyp Untereinheiten und durch die Insertion von randomisierten (X) <sub>7</sub> Peptiden an R588 modifizierten Untereinheiten ( <b>Tabelle 14</b> ).	Referenz <sup>122</sup>
AAV2-NNK <sub>7</sub> - Peptidbibliothek	AAV2-Peptidbibliothek mit NNK-codierten sieben Aminosäuren langen Peptidinsertionen, tragen das AAV2 <i>rep</i> des pMT202-6 und das modifizierte AAV2 <i>cap</i> von pMT202-6 ( <b>Tabelle 14</b> ).	Referenz <sup>122</sup>
AAV2-Trimer <sub>6</sub> -(-C)- Peptidbibliothek	AAV2-Peptidbibliothek mit Trimer-codierten sechs Aminosäuren langen Peptidinsertionen, tragen das AAV2 <i>rep</i> des pMT202-6 und das modifizierte AAV2 <i>cap</i> von pMT202-6 ( <b>Tabelle 14</b> ).	Diese Arbeit
rAAV2-Wildtyp-CMV-LUC	rAAV2 mit Luciferase unter CMV Promotor als Transgen und Wildtyp Kapsid ( <b>Tabelle 14</b> ).	Diese Arbeit
rAAV2-GDDATRD-CMV-LUC	rAAV2 mit Luciferase unter CMV Promotor als Transgen und GDDATRD Peptidinsertion an R588 ( <b>Tabelle 14</b> ).	Diese Arbeit

## 2.17 Plasmide

Alle verwendeten Plasmide sind in **Tabelle 14** beschrieben.

**Tabelle 14: Plasmide**

Name	Beschreibung	Referenz
pMT202-6	Plasmid codiert für das AAV2-Genom mit ITRs, <i>cap</i> R588Ins, Plasmid zur Produktion von AAV2-Peptidbibliotheken ( <b>Tabelle 13</b> ).	Referenz <sup>133</sup>
pXX <sub>6</sub>	Helferplasmid mit den adenoviralen Genen E1A, E1B, E2A, E4-orf6, VA, zur Produktion von AAV2- Peptidbibliotheken und rAAV2 ( <b>Tabelle 13</b> ).	Referenz <sup>159</sup>
pXX <sub>2</sub> 187	Plasmid codiert für das AAV2-Genom ohne ITRs, <i>cap</i> R588Ins, Plasmid zur Produktion von rAAV2 ( <b>Tabelle 13</b> ).	Referenz <sup>126</sup>
pXX <sub>2</sub> 187 GDDATRD	Plasmid codiert für das AAV2-Genom ohne ITRs, enthält Insertion für Peptid „GDDATRD“ im <i>cap</i> an R588, Plasmid zur Produktion von rAAV2-GDDATRD-CMV-Luc ( <b>Tabelle 13</b> )	Diese Arbeit
pVP3cm	Plasmid codiert für das Codon modifizierte AAV2 <i>cap</i> vp3, zur Produktion von Transfer-Shuttles ( <b>Tabelle 13</b> ).	Referenz <sup>125</sup>
pAAV CMV tetO <sub>2</sub> Luc	Plasmid enthält die AAV2 Expressionskassette, Luciferase unter CMV Promotor als Reporter gen zwischen den ITRs, zur Produktion von rAAV2 mit Luciferase als Transgen ( <b>Tabelle 13</b> ).	Dr. Stefan Michelfelder

## 2.18 Software und Datenbanken

In **Tabelle 15** sind die für diese Arbeit verwendeten Software und Datenbanken aufgelistet.

**Tabelle 15: Software und Datenbanken**

Bezeichnung	Anwendung	Quelle
Acrobat® 7.0	Generierung von PDF-Dokumenten	Adobe®
ApE 2.0 – A plasmid Editor (M. Wayne Davis)	<i>In silico</i> DNS-Analyse	<a href="http://biologylabs.utah.edu/30orgensen/wayned/ape/">http://biologylabs.utah.edu/30orgensen/wayned/ape/</a>
Cell^A	Mikroskopie-Dokumentation und -Analyse	Olympus
CLC Workbench 6	<i>In silico</i> DNS-Analyse	CLCbio A Qiagen®Company
Excel 2010	Datenanalyse	Microsoft
Gene Snap	Dokumentation von Agarosegelen	Syngene
Nanodrop 2000	Messung von DNS-Konzentrationen	Thermo Fisher Scientific
PeLiCa	Berechnung von Peptidbibliotheken	<a href="http://www.pelica.org/">http://www.pelica.org/</a>
pepDB	Analyse von NGS-Daten	<a href="http://pepdb.zbh.uni-hamburg.de">http://pepdb.zbh.uni-hamburg.de</a>
Power Point 2010	Bearbeitung von Abbildungen	Microsoft
Prism 3.0	Statistik und grafische Bearbeitung	GraphPad
PubMed	Literaturrecherche	<a href="http://www.ncbi.nlm.nih.gov/pubmed">http://www.ncbi.nlm.nih.gov/pubmed</a>
Word 2010	Textbearbeitung	Microsoft

## 2.19 Unternehmen

Material und Geräte wurden von den folgenden in **Tabelle 16** aufgeführten Unternehmen bezogen.

**Tabelle 16: Unternehmen**

Unternehmen	Sitz des Unternehmens	Unternehmen	Sitz des Unternehmens
Adobe Systems	Munich, Deutschland	Lonza	Cologne, Deutschland
Abcam®	Cambridge, Großbritannien	Macherey-Nagel	Dueren, Deutschland
Axis Shield	Dundee, Schottland	Merck	Darmstadt, Deutschland
BD Bioscience	Heidelberg, Deutschland	Metabion	Planegg/Steinkirchen, Deutschland
Beckman Coulter	Krefeld, Deutschland	Mettler Toledo	Hamburg, Deutschland
Berthold Technologies	Bad Wildbad, Deutschland	Microsoft	Unterschleissheim, Deutschland
Biometra	Goettingen, Deutschland	New England Biolabs (NEB)	Frankfurt a. M., Deutschland
Biorad	Munich, Deutschland	Olympus	Hamburg, Deutschland
Braun	Melsungen, Deutschland	PAA Laboratories	Cölbe, Deutschland
Caliper Lifesciences	Mainz, Deutschland	Peqlab	Erlangen, Deutschland
Carl Roth	Karlsruhe, Deutschland	Pierce/Thermo Scientific	Schwerte, Deutschland
Carl Zeiss	Jena, Deutschland	Qiagen	Hilden, Deutschland
Eppendorf	Hamburg, Deutschland	Roche	Mannheim, Deutschland
Fermentas	Sankt Leon-Rot, Deutschland	Santa Cruz	Heidelberg, Deutschland
Fluka	Taufkirchen, Deutschland	Sanyo Biomedical	Etten-Leur, Niederlande
GE-Healthcare	Munich, Deutschland	Sarstedt	Nuembrecht, Deutschland
Gibco®/ Invitrogen	Darmstadt, Deutschland	Seqlab	Goettingen, Deutschland
Gilson	Mettmenstetten, Schweiz	Serva	Heidelberg, Deutschland
GraphPad/Statcon	Witzenhausen, Deutschland	Sigma-Aldrich®	Taufkirchen, Deutschland
Greiner Bio-One	Frickenhausen, Deutschland	Syngene	Cambridge, Großbritannien
HP Medizintechnik	Oberschleissheim, Deutschland	Tecan Group	Crailsheim, Deutschland
Heraeus Instruments	Hanau, Deutschland	Thermo Fisher Scientific	Schwerte, Deutschland
Invitrogen	Darmstadt, Deutschland	Vetequip/VWR	Darmstadt, Deutschland
Leica	Wetzlar, Deutschland		

## 3 METHODEN

### 3.1 Molekularbiologische Methoden

#### 3.1.1 Isolierung von Plasmid-DNS

Für die Amplifizierung von Plasmid-DNS wurden über Nacht-Kulturen unterschiedlicher Volumina mit transformierten Bakterien oder Bakterien aus Glycerolstocks angesetzt. Nach Erreichen einer bestimmten Dichte wurde die Plasmid-DNS isoliert. Für die Isolierung der Plasmid-DNS von kleinen Kulturen (5 ml, "Mini-Präp") wurde das peqGOLD Miniprep Kit I von peqlab verwendet, für größere Kulturen (200 ml, "Midi-Präp") das NucleoBond® PC Kit von Macherey-Nagel. Plasmid-DNS von noch größeren Kulturen (800 ml, "Maxi-Präp") wurde mit Hilfe des NucleoBond® Xtra Maxi Kit ebenfalls von Macherey-Nagel isoliert. Hierbei wurde nach den Angaben des Herstellers verfahren.

Zur Isolierung der Plasmid-DNS aus kleinen Kulturen wurden alle Zentrifugationsschritte bei Raumtemperatur (RT) durchgeführt. Zunächst wurden die Bakterien für 1 min bei 10.000 x g zentrifugiert und das Bakterienpellet in 250 µl Resuspensions-Puffer resuspendiert. Um die Bakterien zu lysieren, wurden 250 µl Lyse-Puffer hinzugegeben und maximal 5 min bei RT inkubiert. Anschließend wurden 350 µl Neutralisations-Puffer zugegeben und das Lysat für 10 min mit 10.000 x g zentrifugiert. Der Überstand wurde auf eine Silica-Membran-Säule gegeben und die Säulen wurden für 1 min bei 10.000 x g zentrifugiert. Danach wurden die Säulen mit 750 µl Wasch-Puffer durch Zentrifugation für 1 min mit 10.000 x g gewaschen. Durch einen weiteren Zentrifugationsschritt wurde die Säule für 2 min bei 10.000 x g zentrifugiert und damit getrocknet. Schließlich wurde die DNS mit 30 - 50 µl Elutions-Puffer und Zentrifugation für 1 min mit 5.000 x g eluiert.

Für die Isolierung von Plasmid-DNS aus größeren Kulturen wurden diese für 10 min mit 6.000 x g bei 4°C zentrifugiert und das Pellet in einem von der Pelletgröße abhängigen Volumen an Resuspensions-Puffer resuspendiert. Danach wurde zur Lyse der Bakterien die gleiche Menge an Lyse-Puffer hinzugegeben und für maximal 5 min bei RT inkubiert. Anschließend wurde Neutralisations-Puffer mit dem gleichen Volumen wie für die ersten Puffer zugegeben und das Lysat für 10 min mit 10.000 x g zentrifugiert. Nach der Equilibrierung der Säule mit Equilibrierungs-Puffer wurde der Überstand auf die Säule geladen. Nachdem der komplette Überstand durch die Säule gelaufen war, wurde die Säule mit Wasch-Puffer gewaschen und die DNS mit 5 ml Elutions-Puffer im Falle einer Midi-Präp und mit 25 ml bei einer Maxi-Präp eluiert. Aufgrund des großen Volumens wurde die DNS anschließend mit dem 0,8-fachen Volumen an Isopropanol präzipitiert. Hierfür wurde die DNS nach der Zugabe des Isopropanols für 30 min mit 15.000 x g bei 4°C zentrifugiert. Das Pellet wurde mit 2 ml 70 % (v/v) Ethanol gewaschen und danach für 5 min mit 15.000 x g bei RT zentrifugiert. Schließlich wurde das

DNS-Pellet an der Luft getrocknet und dem gewünschten Volumen an Elutions-Puffer oder  $ddH_2O$  gelöst.

### 3.1.2 DNS-Quantifizierung

Zur Bestimmung der DNS-Konzentration und zur Bestimmung der Reinheit der DNS wurde das Spektrophotometer NanoDrop 2000c von Thermo Fisher Scientific verwendet. Bei einem Absorptions-Verhältnis 260 nm/280 nm bei ungefähr 1,8 und dem Verhältnis 260 nm/230 nm bei einem Wert von ungefähr 2, wurde die DNS als rein von Protein und anderen Kontaminationen, die bei diesen Wellenlängen absorbieren, bezeichnet.

### 3.1.3 Enzymatischer Verdau von DNS mittels Restriktionsendonukleasen

Um Plasmid-DNS auf Korrektheit der Schnittstellen zu testen und für Klonierungen wurden Restriktionsendonukleasen von New England Biolabs oder *Fast Digest* Enzyme von Thermo Scientific, wenn nicht anders angegeben, nach folgendem Protokoll verwendet.

Komponente	Menge
DNS	0,5 – 1 $\mu$ g
Restriktionsendonuklease	1 $\mu$ l per $\mu$ g DNS
Empfohlener Puffer, 10x	1/10 vol.
BSA, 10x (wenn empfohlen)	1/10 vol.
$ddH_2O$	ad 1 vol.

Inkubationszeiten wurden nach Herstellerangaben durchgeführt.

### 3.1.4 Agarose Gelelektrophorese

Zur Überprüfung von PCR-Produkten und DNS-Fragmenten nach Restriktion wurden Agarose-Gelelektrophoresen durchgeführt. Hierfür wurden Sub-Cell® GT oder Wide Mini Sub-Cell® GT Kammern von Biorad und das PowerPac™ basic ebenfalls von Biorad verwendet.

Für die Auftrennung von DNS-Fragmenten über 1.000 bp wurden 1 % Agarosegele und für DNS-Fragmente unter 1.000 bp wurden 2 % Agarosegele verwendet. Hierfür wurden im Falle eines 1 % Gels 1 g Agarose von Invitrogen™ in 100 ml TAE-Puffer gelöst und aufgekocht bis es komplett geschmolzen war. Zu 50 ml geschmolzener Agarose wurden 5 - 10  $\mu$ l Ethidiumbromid von Sigma-Aldrich® oder 10  $\mu$ l Roti®-Safe GelStain von Carl Roth® zugegeben und das Gemisch in einem Gelschlitten von Biorad gegossen. Nach der Härtung des Gels wurden 10  $\mu$ l des entsprechenden DNS-Markers (1 kb DNS-Marker für Fragmente über 1.000 bp und 100 bp DNS-Marker für Fragmente unter 1.000 bp) und die vorher mit Ladepuffer vorbereiteten DNS-Proben auf das Gel geladen.



Abhängig von der Fragmentgröße und der Dichte des Gels wurden Spannungen von 80 V bis 120 V für ca. 1 h, bzw. bis die Fragmente gut aufgetrennt waren, angelegt.

Die Detektion und Dokumentation wurde unter UV-Licht mit Hilfe des Geldokumentationssystems Gene Genius Bio Imaging System und der Software Gene Snap von Syngene durchgeführt.

### 3.1.5 DNS-Gelextraktion

DNS-Gelextraktionen wurden mit dem PeqGOLD Gel Extraction Kit von Peqlab entsprechend der Anleitung des Herstellers durchgeführt. Nach dem Ausschneiden des jeweiligen DNS-Fragmentes, wurde das Gelstück gewogen und in 100 µl PB Bindungspuffer pro 100 mg Gel bei 55 - 65 °C und zwischenzeitlichem vortexen geschmolzen. Das geschmolzene Gel wurde anschließend auf eine Säule überführt und bei 10.000 x g für 1 min bei RT zentrifugiert. Zum Waschen der DNS wurden 750 µl CG Waschpuffer auf die Säule gegeben und diese erneut unter denselben Bedingungen zentrifugiert. Dieser Schritt wurde wiederholt. Zum Trocknen der Säule wurde diese erneut zentrifugiert. Und schließlich wurde die DNS in 30 - 50 µl Elutions-Puffer oder ddH<sub>2</sub>O mittels Zentrifugation bei 5.000 x g für 1 min bei RT eluiert.

### 3.1.6 DNS-Sequenzierung

Die Sequenzierung von Plasmid-DNS wurde von der Firma SeqLab durchgeführt. Hierfür wurden 1,2 µg DNS mit 10 pmol des jeweiligen Primers gemischt und auf ein Endvolumen von 15 µl gebracht. Die *in silico*-Analyse der Sequenzen erfolgte mit den Programmen ApE 2.0 und CLC Workbench (**Tabelle 15**).

### 3.1.7 Ligation von DNS-Fragmenten

Für die Ligation von DNS-Fragmenten wurde die T4 DNS Ligase von New England Biolabs® verwendet. Die Menge des Inserts für ein molares Verhältnis von 1:1 wurde nach der folgenden Formel berechnet:

$$\text{Menge des Inserts [ng]} = \frac{\text{Länge des Inserts [kb]}}{\text{Länge des Vektors [kb]}} * \text{Menge des Vektors [ng]}$$

Für Standardligationen wurde ein molares Verhältnis von Vektor zu Insert von 1:3 verwendet. Die Ligationen wurden wie folgt angesetzt:

Komponente	Menge
Vektor-DNS	50 ng
Insert-DNS	X
ddH <sub>2</sub> O	ad 20 µl

Die DNS wurde gemischt, für 2 min bei 65 °C inkubiert und anschließend für 5 min auf Eis abgekühlt. Anschließend wurde der T4 DNS Ligase-Puffer und die T4 DNS Ligase ergänzt.

Komponente	Volumen [µl]
T4 DNS Ligase (400 U/µl, New England Biolabs)	2,5
10 x T4 DNS Ligase-Puffer	2,5

Schließlich wurden die Reaktionen für 16 h bei 16°C inkubiert. Für jede Ligation wurde eine Religationskontrolle, bei der ddH<sub>2</sub>O anstatt von Insert-DNS verwendet wurde, mitgeführt.

### 3.1.8 DNS-Präzipitation mit Ethanol

Um DNS zu konzentrieren und von Proteinen und Salzen zu reinigen, wurde sie mit Ethanol wie folgt präzipitiert.

Komponente	Volumen
DNS	1 vol.
Natriumacetat (3M)	1/10 vol.
Ethanol (>99 % (v/v))	2 vol.
Glycogen	1 µl*

\*Für kleine DNS-Fragmente empfohlen.

Die Mischung wurde gut gemischt und für 30 min bis 2 h bei -20°C oder über Nacht inkubiert. Nach der Präzipitation wurde die DNS für 30 min bei 20.000 x g und 4°C zentrifugiert. Anschließend wurde die DNS mit eiskaltem 70 % (v/v) Ethanol gewaschen und für 8 min bei 20.000 x g und 4°C zentrifugiert. Der Überstand wurde verworfen und das DNS-Pellet wurde für ca. 10 min bei RT getrocknet. Schließlich wurde das getrocknete DNS-Pellet in ddH<sub>2</sub>O oder dem gewünschten Puffer resuspendiert.

### 3.1.9 Polymerase-Kettenreaktion (*polymerase chain reaction, PCR*)

#### 3.1.9.1 PCR zur Amplifikation von viraler DNS

Während der Selektion und auch für die NGS-Probenherstellung wurde die DNS zunächst aus den Zellen (3.4.1.3) oder dem jeweiligen Gewebe isoliert (3.4.2.4). Von dieser Gesamt-DNS wurde die virale DNS des *cap*-Gens, die die randomisierten Peptide enthielt mittels PCR amplifiziert. Hierfür wurden 10 Reaktionen mit jeweils 1 µg DNS nach dem folgenden Schema vorbereitet.

Komponente	Volumen [µl]	Endkonzentration
DNS 1 µg	X	20 ng/µl
10 x Hot StarTaq® Reaktionspuffer	5,00	1x
dNTPs (Fermentas, 10mM)	1,00	0,2 mM
Primer + (20µM)	1,00	0,4 µM
Primer - (20µM)	1,00	0,4 µM
Hot StarTaq® Polymerase (Qiagen 5 U/µl)	0,25	0,025 U/µl
ddH <sub>2</sub> O	ad 50	

Das folgende PCR-Protokoll wurde verwendet.

Schritt	Temperatur [°C]	Zeit [min]
Initiale Denaturierung	95	15
Denaturierung	94	1
Primerhybridisierung	61	1
Elongation	72	1
Finale Elongation	72	10
	4	∞

} 36x

Nach der PCR wurden die Proben vereint und mit Ethanol gefällt (3.1.8).

### 3.1.9.2 Kolonie-PCR (cPCR)

Nach der Ligation von DNS-Fragmenten in Plasmide wurden diese in Bakterien transformiert, auf Agar-Platten ausgestrichen und am nächsten Tag Einzelklone gepickt. Zur Überprüfung, ob die Ligation erfolgreich war wurden Kolonie-PCRs durchgeführt. Hierfür wurden Einzelklone gepickt und in 20 µl ddH<sub>2</sub>O verdünnt. Von dieser Bakteriensuspension wurde ein 1 µl für die PCR eingesetzt. Die Ansätze wurden nach dem folgenden Protokoll angesetzt.

Komponente	Volumen [µl]	Endkonzentration
Bakterienlösung	1	
5 x GoTaq <sup>®</sup> Reaktionspuffer	4	1x
dNTPs (Fermentas, 10mM)	0,25	0,125 mM
Primer + (5µM)	1	0,5 µM
Primer - (5µM)	1	0,5 µM
GoTaq <sup>®</sup> Polymerase (Promega, 5 U/µl)	0,1	0,025 U/µl
ddH <sub>2</sub> O	ad 20	

Das folgende PCR-Protokoll wurde verwendet.

Schritt	Temperatur [°C]	Zeit [s]
Initiale Denaturierung	94	180
Denaturierung	94	30
Primerhybridisierung	61	30
Elongation	72	30
Finale Elongation	72	300
	4	∞

} 29x

In dieser Arbeit wurden Kolonie-PCRs nur nach der Ligation von *sticky-ends* durchgeführt. Da hierbei nur eine Orientierung des Inserts möglich war, konnten richtig ligierte Klone von religierten Klonen mittels Gelelektrophorese und anhand ihrer Fragmentgröße voneinander unterschieden werden (3.1.4). Bakterien, die richtig ligierte Plasmid trugen wurden weiter inkubiert (3.2.1.1) und die Plasmid-DNS wie beschrieben isoliert (3.1.1).

### 3.1.9.3 Quantitative Echtzeit PCR (q-PCR)

Die Bestimmung des genomischen Titers von AAV-Peptidbibliotheken, Vektorproduktionen und viraler DNS in Mausgeweben erfolgte mittels q-PCR. Hierfür wurden der FastStart Essential Green Mastermix (SYBR Green) und der LightCycler® Nano von Roche verwendet. Alle q-PCR-Reaktionen wurden wie folgt durchgeführt.

Komponente	Volumen [μl]	Endkonzentration
DNS-Lösung/Viruslösung	2	
Primer + (5μM)	1	0,25 μM
Primer - (5μM)	1	0,25 μM
FastStart Essential Green Mastermix (SYBR Green)	10	
ddH <sub>2</sub> O	ad 20	

Alle hier verwendeten Primer wurden so generiert, dass ein ähnliches PCR-Programm verwendet werden konnte. Lediglich die Temperaturen für die Primerhybridisierung wurden jeweils angepasst. Das folgende Programm entspricht dem Standardprogramm.

Schritt	Temperatur [°C]	Zeit [s]	Anstieg [°C/s]
Initiale Denaturierung	95	600	5
Denaturierung	95	30	5
Primerhybridisierung	X	30	5
Elongation	72	30	5
Schmelzkurve initial	60	20	4
Schmelzkurve final	95	20	0,1

Zur Bestimmung der Anzahl an Virusgenomen wurden Standards hergestellt. Für jeden Standard wurde ein Plasmid, das die gleiche DNS-Sequenz beinhaltet bis zu einer Konzentration von  $1 \times 10^{10}$  Kopien pro μl verdünnt. Hierfür wurde zunächst die Masse des Plasmids nach folgender Formel berechnet.

$$m_p[g] = n_{bp} * 1.096 * 10^{-21} g/bp$$

$m_p$  = Masse (Plasmid);  $n_{bp}$  = Anzahl an Basenpaaren

Danach wurde die Masse von  $1 \times 10^{10}$  Plasmiden berechnet, die entsprechende Konzentration eingestellt und durch mehrfache spektrofotometrische Messungen bestätigt. Anschließend wurde diese Verdünnung sukzessive verdünnt und eine Standardkurve mit definierten Anzahlen an Kopien hergestellt ( $1 \times 10^7$ ;  $1 \times 10^6$ ;  $1 \times 10^5$ ;  $1 \times 10^4$ ;  $1 \times 10^3$  Genomkopien).

Aufgrund hoher Virustiter wurden die Viruslösungen vor der Messung in 1:10 Verdünnungsschritten verdünnt. Im Falle von gereinigtem Virus wurden die 1:1.000 und die 1:10.000 Verdünnungen gemessen; Virus aus Zelllysate wurde 1:10.000 und 1:100.000 verdünnt. Zur Bestimmung der Virustiter wurde die folgende Formel angewandt

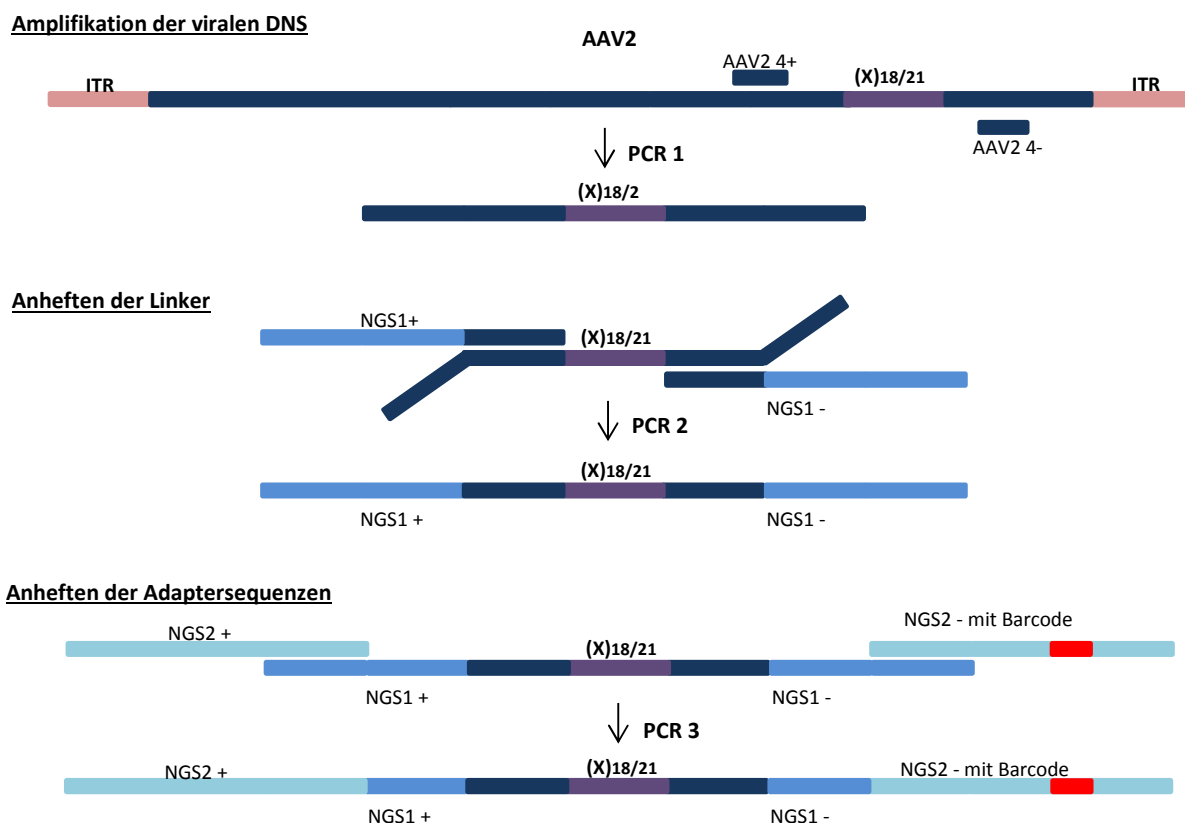
$$\text{Virale Genome pro } \mu\text{l} \left[ \frac{\text{vg}}{\mu\text{l}} \right] = \frac{(X1 * \text{Verdünnung 1}) + (X2 * \text{Verdünnung 2})}{2 \mu\text{l}}$$

X = Anzahl der gemessenen viralen Genome

### 3.1.10 Next-Generation Sequencing (NGS)

*Protokoll zur Probenvorbereitung zur NGS-Sequenzierung von Peptidbibliotheken nach Dr. Timo Sieber*

Die hier beschriebene NGS-Probenvorbereitung kann in vier Teile eingeteilt werden. Zunächst wird Gesamt-DNS aus den untersuchten Geweben isoliert. Anschließend wird die virale DNS amplifiziert. An die amplifizierte virale DNS wird ein Linker gefügt, der die Brücke zwischen der viralen DNS und den universellen Illumina®-Adaptersequenzen darstellt. Schließlich werden diese Adaptersequenzen, die zur Probenidentifikation einen Barcode besitzen, an die Linker gefügt (**Abbildung 9**). Da das NGS eine Tiefensequenzierungsmethode ist, sind Kontaminationen während der gesamten Probenvorbereitung strengstens zu vermeiden.



**Abbildung 9: Schematische Darstellung der NGS-Probenvorbereitung zur Untersuchung von AAV-Peptidbibliotheken.**

Die Vorbereitung der NGS-Proben zur Untersuchung von AAV-Peptidbibliotheken besteht aus drei Schritten, die per PCR erfolgen. Im ersten Schritt werden die randomisierten Oligonukleotidinsertionen aus der isolierten Gesamt-DNS spezifisch amplifiziert. Um diese Sequenzen mit den universellen Adaptersequenzen verbinden zu können, werden anschließend Verbindungssequenzen und im letzten Schritt die Adaptersequenzen angehängt.

### 3.1.10.1 DNS-Isolierung

DNS-Isolierungen aus Zellen und Geweben für Selektionsrunden sind in den Kapiteln **3.4.1.3** und **3.4.2.5** beschrieben. Für die Bestimmung der Vektorverteilung in der Maus wurde die DNS aus einem Gewebestück mit der Größe eines halben Mausherzens isoliert.

### 3.1.10.2 Amplifikation der viralen das randomisierte Insert beinhaltenden DNS - erste PCR

Für die Amplifikation der viralen DNS aus Gewebe- oder Zelllysaten, wurden die das Insert beinhaltenden Sequenzen mittels PCR und den Primern AAV2 4+ und AAV2 4- (**2.13**) amplifiziert. Hierfür wurden 10 Reaktionen mit jeweils 1 µg DNS nach dem unter **3.1.9.1** beschriebenen Protokoll vorbereitet.

Um zu verifizieren, dass die PCR funktioniert hat und dass die PCR-Produkte die richtige Größe besaßen, wurden 5 µl-Proben auf ein 2 % Agarosegel geladen und mittels Gelelektrophorese untersucht (**3.1.4**). Nach der Gelelektrophorese wurden die PCR-Produkte aller Ansätze vereint und mittels Ethanol präzipitiert (**3.1.8**). Aufgrund des hohen Kontaminationsrisikos wurde hier auf eine Gelextraktion verzichtet. Da bei der Präzipitation auch die genomische DNS gefällt wird, konnte eine Konzentrationsbestimmung nicht über die spektrofotometrische Messung mittels NanoDrop erfolgen. Um die DNS-Konzentration dennoch zu bestimmen, wurde die Bandenstärke der Proben im Agarosegel mit der des Markers verglichen und die Konzentration abgeschätzt. Hierfür wurde eine definierte Menge des Markers auf das Gel geladen. Für die zweite PCR wurde die Konzentration der Proben schließlich auf 5 ng pro µl eingestellt.

### 3.1.10.3 Anfügen der Linker - zweite PCR

Mit der zweiten PCR wurden die Linker zwischen der AAV2-Sequenzen und den Illumina®-Adaptersequenzen angefügt. Um zu verhindern, dass die Diversität der Peptidbibliothek durch die Verdünnung der Proben eingeschränkt wird, sollte für diese genug DNS eingesetzt werden.

*Berechnung zur Aufrechterhaltung der Diversität:*

Länge des PCR-Produkts der ersten PCR: 486 bp

1 bp wiegt ungefähr 660 Da.

1 Da  $\triangleq$   $1,6605 \times 10^{-27}$  kg =  $1,6605 \times 10^{-15}$  ng.

Das PCR-Produkt der ersten PCR wiegt:

$$486 \text{ bp} \times 660 \text{ Da} \times 1,6605 \times 10^{-15} \text{ ng} = \underline{5,3262 \times 10^{-10} \text{ ng}}$$

→ 5 ng beinhalten  $9,387555855 \times 10^9$  PCR-Produkte.

Die initiale AAV2-Peptidbibliothek hat auf Plasmid-Ebene eine theoretische Diversität von  $2 \times 10^8$ .

→ 47-fache Abdeckung bei der Verwendung von 5 ng DNS.

Für jede Probe wurden zwei Ansätze je 5 ng DNS angesetzt. Aufgrund der geringeren Fehlerrate von  $4,4 \times 10^{-7}$  wurde für diese PCR die Phusion® High-Fidelity Polymerase von NEB verwendet. Da die finalen Sequenzen sich nur in den Inserts unterschieden und ansonsten komplett identisch wären, kann es sein, dass das Sequenziergerät ins „Stottern“ kommt. Um dies zu verhindern, beinhalten die Linker kleine Sequenzänderungen (**Tabelle 9 und Abbildung 10**). Die PCR wurde nach dem folgenden Protokoll durchgeführt.

Komponente	Volumen [ $\mu$ l]	Endkonzentration
DNS 5 ng	1	0,25 ng/ $\mu$ l
5 x Phusion® HF Puffer	4,00	1x
dNTPs (Fermentas, 10mM)	0,40	0,2 mM
NGS1 Fwd (10 $\mu$ M) ( <b>2.16</b> )	1,00	0,5 mM
NGS1 Rev (10 $\mu$ M) ( <b>2.16</b> )	1,00	0,5 mM
Phusion® High-Fidelity DNS Polymerase (2 U/ $\mu$ l, NEB)	0,25	0,025 U/ $\mu$ l
ddH <sub>2</sub> O	ad 20	

Programm:

Schritt	Temperatur [°C]	Zeit [s]
Initiale Denaturierung	98	30
Denaturierung	98	10
Primerhybridisierung	72	20
Elongation	72	10
	4	$\infty$

} 28x

#### 3.1.10.4 Anfügen der Illumina®-Adaptersequenzen mit Barcode - dritte PCR

Vor der letzten PCR, bei der die Adaptersequenzen angefügt werden, wurde die DNS-Konzentration der amplifizierten PCR-Produkte mittels Agarose-Gelelektrophorese und Vergleich der Bandenstärke mit dem Marker bestimmt (**3.1.10.2**). Für die PCR wurden pro Probe zwei Reaktionen mit jeweils 1 ng angesetzt. Während der PCR wurde ein universelles Oligonukleoid an das 5'-Ende und ein den Barcode beinhaltendes Oligonukleotid an das 3'-Ende gefügt. Die Reaktionen wurden wie folgt angesetzt. Das PCR-Programm entspricht dem der zweiten PCR (**3.1.10.3**).

Komponente	Volumen [ $\mu$ l]	Endkonzentration
DNS 1 ng	1	0,05 ng/ $\mu$ l
5 x Phusion® HF Puffer	4,00	1x
dNTPs (Fermentas, 10mM)	0,40	0,2 mM
NGS2 Fwd (10 $\mu$ M) ( <b>2.16</b> )	1,00	0,5 mM
NGS2 Rev (10 $\mu$ M) ( <b>2.16</b> )	1,00	0,5 mM
Phusion® High-Fidelity DNS Polymerase (2 U/ $\mu$ l, NEB)	0,25	0,025 U/ $\mu$ l
ddH <sub>2</sub> O	12,35	

Die ca. 190 bp langen PCR-Produkte (**Abbildung 10**) wurden anschließend mittels Agarose-Gelelektrophorese über ein langes 3 % Agarosegel von Primern gereinigt (**3.1.5**) und mit Ethanol präzipitiert (**3.1.8**).

**Beispielsequenz des finalen PCR-Produktes:**

```
5´ AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGCTCCAGAGAGGCCAGAGAGGC (X)18/21GCCCA
GGCGGCCACCGCAGATGCTCATAGATCGGAAGAGCACACGTCTGAACTCCAGTCACTACAGTGAATCTCGTATGCCGTCTTCTGCTTG 3´
```

**Abbildung 10: Beispielsequenz des finalen PCR-Produktes der NGS-Probe.** Hellgrau: Im Linker zusätzlich eingefügte Nukleotide. Schwarz: Barcodesequenz.

Schließlich wurde die DNS-Konzentration mittels spektrofotometrischer Messung am NanoDrop gemessen (**3.1.2**) und die Proben zu gleichen Teilen vereinigt. Die genaue Bestimmung der DNS-Konzentration mit einem Qubit Fluorometer, die Sequenzierung der Proben mit dem MiSeq Sequencing System von Illumina® und die bioinformatische Datenanalyse wurde von der NGS-Facility des Heinrich-Pette-Institutes durchgeführt.

### 3.1.10.5 Bioinformatische Auswertung der NGS-Daten

Die bioinformatische Auswertung wurde von Malik Alawi des Heinrich-Pette-Institutes durchgeführt. Um Sequenzierungsfehler weitestgehend auszuschließen, wurden für die bioinformatische Auswertung der NGS-Daten verschiedene Qualitätsfilter angewendet. So musste eine Sequenz neben der geforderten Länge des Inserts auch dem jeweiligen Codierungsschema entsprechen, die flankierenden Erkennungssequenzen (CCAGAGAGGC und GCCCAGGCGG) mussten exakt mit der Referenzsequenz übereinstimmen und in der Zielsequenz durfte kein nicht eindeutig identifizierbares Nukleotid vorhanden sein. Des Weiteren wurden Sequenzen, die sich in nur einem Nukleotid von einer 100-fach häufigeren Sequenz unterschieden und Sequenzen, die seltener als 100-fach beobachtet wurden, von der Analyse ausgeschlossen. Sequenzen, die Stoppcodons beinhalten, wurden ebenfalls nicht analysiert. Sequenzen, die nicht den Anforderungen entsprachen wurden herausgefiltert und separat ausgegeben.

### 3.1.10.6 Berechnung und Bewertung von Peptidanreicherung und -spezifität zur Identifizierung von potenziell gewebespezifischen AAV-Vektoren

Für die Bestimmung eines potenziell gewebespezifischen AAV-Vektors wurden drei Bewertungskriterien herangezogen – Peptidanreicherung (*General Enrichment Score, GE-Score*), Peptidspezifität (*General Specificity Score*) und die Kombination aus beidem (*Combined Score, C-Score*). Zu deren Berechnung wurden die Peptidsequenzen der letzten Selektionsrunde nach ihrer Häufigkeit geordnet und ihre relativen Häufigkeiten in allen Selektionsrunden und Kontrollorganen



verglichen. Für die Berechnung der durchschnittlichen Peptidanreicherung über die gesamte Selektion, wurden zunächst die Anreicherungen für jede Selektionsrunde, durch Vergleich der relativen Häufigkeiten einer Peptidsequenz in einer Selektionsrunde zur relativen Häufigkeit derselben Peptidsequenz aus der vorherigen Selektionsrunde, berechnet (*E-Score*). Diese Werte wurden dann zur Berechnung des *GE-Scores* multipliziert. Zur Berechnung der *GS-Scores* wurden die relativen Häufigkeit der Peptidsequenzen im Zielgewebe mit denen in den Kontrollorganen verglichen (*S-Score*) und anschließend zur Berechnung des *GS-Scores* multipliziert. Der kombinierte *C-Score* schließlich vereint Anreicherung und Spezifität der jeweiligen Peptidsequenzen, durch Multiplikation von *GE-Score* und *GS-Score*. Für die Berechnung der *Scores* zur Bestimmung von Anreicherung und Spezifität wurde die folgende Formel verwendet.

$$GE - \text{ oder } GS - \text{ Score} = (N)^x; N = 1 - \frac{1}{\left(\frac{Rz}{Ry} + 1\right)}$$

*Rz* = Relative Häufigkeit in der letzten Selektionsrunde; relative Häufigkeit im Zielgewebe

*Ry* = Relative Häufigkeit in der vorherigen Selektionsrunde; relative Häufigkeit in den Kontrollorganen

*x* = Anzahl der Selektionsrunden; Anzahl der Kontrollorgane

Mit Hilfe dieser Formel wurden die Verhältnisse der Häufigkeiten berechnet und in einen Wert zwischen 0 und 1 konvertiert. Hierbei bedeutet ein Wert nahe 1 eine starke Anreicherung bzw. eine hohe Spezifität. In manchen Fällen wurde eine Peptidsequenz in der unselektierten, ersten oder zweiten Selektionsrunde nicht gefunden, weil die Sequenzierertiefe nicht alle möglichen Sequenzen abdeckte. Eine Peptidsequenz, die in einer späteren Selektionsrunde gefunden wurde, muss theoretisch aber auch in den vorherigen Selektionsrunden vorhanden gewesen sein. In solchen Fällen wurde als Häufigkeit dieses Peptides die Hälfte der Häufigkeit einer DNS-Sequenz berechnet, um den Unterschied zu einer tatsächlich einmal sequenzierten Peptidsequenz zu wahren.

Grundsätzlich wurde davon ausgegangen, dass die Peptidsequenz mit dem höchsten Wert die höchste Wahrscheinlichkeit für Gewebespezifität besitzt. Zur Orientierung und Einordnung eines Wertes wurden zu jedem Bewertungskriterium Orientierungswerte berechnet. Die Diagramme zur Darstellung der Anreicherung besitzen eine Orientierungslinie auf der Höhe des Wertes bei dem die Peptide im Durchschnitt in den Selektionsrunden gleich häufig wären. Ein Wert über dieser Linie bedeutet, dass sich das jeweilige AAV-Partikel im Durchschnitt während der Selektion angereichert hat. Für die Diagramme zur Berechnung der Spezifität zeigt die Orientierungslinie einen Wert an dem eine Peptidsequenz im Durchschnitt gleichhäufig in dem untersuchten Zielgewebe wie in den Kontrollorganen zu finden war. Peptidsequenzen mit Werten über dieser Linie besitzen durchschnittlich eine höhere relative Häufigkeit im Zielgewebe als in den Kontrollorganen. Bezüglich des *C-Scores* bedeutet ein Wert auf Höhe der Orientierungslinie eine durchschnittliche Verdopplung der relativen Häufigkeit über die Selektion und eine Halbierung der relativen Häufigkeit dieser Peptidsequenz in den Kontrollorganen im Vergleich zum Zielorgan.

## 3.2 Mikrobiologische Methoden

### 3.2.1 Kultivierung von *E.coli*

#### 3.2.1.1 Suspensionskultur

Zur Kultivierung von *E.coli* in Suspensionskultur wurden Einzelklone von Agar-Platten mit entsprechendem Antibiotikum oder aus mit Glycerin eingefrorenen Kulturen in 5 ml Medium bei 37°C und 200 rpm über Nacht kultiviert bis sie eine hohe Dichte hatten. Bakterien, die mit einem ITR tragenden Plasmid transformiert wurden, wurden für ein langsames Wachstum bei 30°C und 175 rpm kultiviert. Für alle Suspensionskulturen wurde LB- oder TB-Medium (**2.10.1**) mit Ampicillin (150 µg/ml) verwendet (**2.12**).

#### 3.2.1.2 Kultivierung auf Agarplatten

Nach der Transformation (**3.2.3**) wurden die Bakterien auf Agar-Platten mit Ampicillin (150 µg/ml) (**2.12**) ausgestrichen und in einem Brutschrank über Nacht bei 37°C inkubiert. Am nächsten Tag wurden Einzelklone zur amplifizierenden Kultivierung gepickt und die Agar-Platten bis zu vier Wochen bei 4°C im Kühlschrank gelagert.

#### 3.2.1.3 Lagerung in Glycerin

Für eine längere Lagerung wurden transformierte Bakterien aus Suspensionskulturen vorsichtig mit Glycerin in einer finalen Konzentration von 50 % Glycerin gemischt und bei -80°C gelagert.

### 3.2.2 Herstellung von kompetenten *E. coli*

Zur Herstellung von kompetenten Bakterien wurden alle Schritte bei 4°C durchgeführt. Alle Puffer, Flaschen und Röhrchen wurden vorgekühlt und die Bakterien wurden auf Eis resuspendiert.

#### 3.2.2.1 Herstellung von chemisch-kompetenten Bakterien (Calciumchlorid-Methode)

Die Bakterien wurden zunächst in einer Vorkultur von 30 ml LB-Medium (**2.10.1**) ohne Antibiotikum bei 37°C und 200 rpm über Nacht kultiviert. Am nächsten Tag wurden 200 ml LB-Medium mit den Bakterien beimpft und bis zu einer OD<sub>600</sub> von 0,5 bis 0,6 weiter kultiviert.

Zur Pellettierung wurden die Bakterien für 5 min bei 3.000 x g und 4°C zentrifugiert. Das Pellet wurde in 100 mM eiskaltem CaCl<sub>2</sub> resuspendiert. Nach einer 30-minütigen Inkubation auf Eis wurden die Bakterien erneut zentrifugiert (5 min, 3.000 x g, 4°C) und anschließend in 15 ml eiskaltem 20 % (v/v) Glycerin in 50 mM CaCl<sub>2</sub>-Lösung resuspendiert. Die Bakterien wurden schließlich in vorgekühlte Eppendorf-Reaktionsgefäße aliquotiert, in flüssigem Stickstoff überführt und bei -80°C gelagert.

### 3.2.2.2 Herstellung von elektro-kompetenten Bakterien

Zur Herstellung von elektrokompenten Bakterien wurde zunächst eine Vorkultur der jeweiligen Bakterien in 50 ml LB-Medium (**2.10.1**) bei 37°C und 200 rpm über Nacht inkubiert.

Am nächsten Tag wurden 500 ml SOB-Medium (**2.10.1**) mit 0,5 ml der Vorkultur beimpft und bis zu einer Dichte von OD<sub>600</sub> von 0,75 – 0,8 weiter inkubiert. Nach Erreichen der Dichte wurde die Bakterienkultur für 15 min auf Eis gestellt und danach für 15 min bei 600 x g und 4°C zentrifugiert. Die Bakterien wurden anschließend in 100 ml eiskaltem 1 mM HEPES resuspendiert und erneut zentrifugiert (15 min, 600 x g, 4 °C). Nachdem dieser Waschschrift wiederholt wurde, wurden die Bakterien in 10 ml eiskaltem 10 % (v/v) Glycerin resuspendiert und erneut zentrifugiert (5 min, 3.000 x g, 4 °C). Schließlich wurden die Bakterien in 2,5 ml eiskaltem 10 % (v/v) Glycerin resuspendiert, aliquotiert, sofort in flüssigen Stickstoff überführt und bei -80 °C gelagert.

### 3.2.3 Transformation in *E. coli*

Um Plasmid-DNS zu amplifizieren, wurden Bakterien mit den entsprechenden Plasmiden mittels Hitzeschock oder Elektroporation transformiert.

#### 3.2.3.1 Transformation mittels Hitzeschock

Für die Transformation mittels Hitzeschock wurden ungefähr 50 ng Plasmid-DNS in einem Volumen von 1–5 µl zu 50 µl Bakterien gegeben, vorsichtig gemischt und für 20 min auf Eis gestellt. Für den Hitzeschock wurden die Bakterien für 45 s in ein 42 °C warmes Wasserbad überführt. Anschließend wurden die Bakterien für 2 min auf Eis gestellt und danach in 150 µl vorgewärmten SOC-Medium (**2.10.1**) resuspendiert. Zur Regeneration wurden die Bakterien für 1 h bei 37°C und 200 rpm inkubiert und schließlich auf Agar-Platten mit dem entsprechenden Antibiotikum ausgestrichen und über Nacht inkubiert (**3.2.1.2**).

#### 3.2.3.2 Transformation mittels Elektroporation

Da hohe Salzkonzentrationen die Elektroporation verhindern können, wurden nur zuvor mit Ethanol präzipitierte (**3.1.8**) Plasmide mittels Elektroporation transformiert. Hierfür wurden zunächst ca. 20 ng Plasmid-DNS in einem Volumen von 1-5 µl ddH<sub>2</sub>O und 50 µl Bakterien in eine vorgekühlte Küvette (1 mm, Biorad) überführt. Für den elektrischen Puls wurde der Gene Pulser Xcell™ Microbial System von Biorad (**2.1**) und die folgenden Parameter verwendet:

Kapazität: 25 µF  
Widerstand: 200 Ω  
Spannung: 1,8 kV

Anschließend wurden die Bakterien sofort in vorgewärmtes SOC-Medium (**2.10.1**) aufgenommen und zur Regeneration für 1 h bei 37°C und 200 rpm inkubiert. Schließlich wurden sie auf Agar-Platten mit dem entsprechenden Antibiotikum ausgestrichen und über Nacht inkubiert (**3.2.1.2**).

#### **3.2.4 Bestimmung der Transformationseffizienz**

Für die Herstellung einer AAV-Peptidbibliothek ist die Transformationseffizienz besonders wichtig. Sie stellt einen limitierenden Faktor der Diversität der Bibliothek dar. Da es auch während einer Selektion wichtig ist die Diversität der Plasmidbibliothek nicht künstlich zu limitieren, wurden auch die Transformationseffizienzen der sekundären Plasmidbibliotheken wie folgt bestimmt.

$$\text{Transformationseffizienz [Transformanten}/\mu\text{g DNS]} = \frac{n}{\text{DNS } [\mu\text{g}]} * d$$

n = Anzahl der Kolonien, d = Verdünnung

### 3.3 Zellbiologische Methoden

#### 3.3.1 Kultivierungsbedingungen

Die Arbeit mit eukaryotischen Zellen wurde unter sterilen Bedingungen in einer Sterilbank der Sicherheitsklasse S2 durchgeführt (2.1). Die Zellen wurden in Zellkulturinkubatoren unter Standardbedingungen bei 37°C mit 5 % CO<sub>2</sub> inkubiert.

#### 3.3.2 Auftauen und Einfrieren von eukaryotischen Zellen

Im Gegensatz zum Einfrieren von Zellen sollte das Auftauen so schnell wie möglich passieren. Hierfür wurden die eingefrorenen Zellen aus dem flüssigen Stickstoff direkt für 1 min in einem 37 °C warmen Wasserbad aufgetaut. Nach der Zugabe von ca. 10 ml Medium mit Zusätzen wurden die Zellen 5 min mit 200 x g zentrifugiert und somit gewaschen. Das DMSO beinhaltende Medium wurde verworfen und die Zellen wurden in frischem Medium mit Zusätzen vorsichtig resuspendiert und in Zellkulturflaschen überführt. Am nächsten Tag wurde das Medium gegen frisches Medium gewechselt. Zum Einfrieren von eukaryotischen Zellen wurden diese zunächst gezählt und dann eine definierte Menge für 5 min mit 200 x g zentrifugiert. Das Zellpellet wurde anschließend in Einfriermedium resuspendiert und in Cryopreservationsröhrchen überführt. Diese wurden dann in mit Isopropanol gefüllten Einfrierbehälter gestellt und auf -80 °C langsam heruntergekühlt. Die Langzeitlagerung fand in flüssigem Stickstoff statt.

#### 3.3.3 Passage von adhärenenten Zellen

Bei einer Konfluenz von 80 bis 90 % wurden die Zellen passagiert. Hierfür wurden sie mit 10 ml PBS gewaschen und anschließend mit Trypsin/EDTA für ungefähr 5 min bei 37°C im Zellkulturinkubator inkubiert. Danach wurden sie je nach ihrer Verdopplungszeit in einem geeigneten Verhältnis gesplittet und in frischem Medium neu ausgebracht.

#### 3.3.4 Bestimmung der Zellzahl

Zur Bestimmung der Zellzahl wurden die Zellen mit PBS gewaschen, mit Trypsin/EDTA von der Zellkulturschale oder -flasche gelöst und in frischem Medium mit Zusätzen aufgenommen. Danach wurden sie in einem Verhältnis von 1:10 in 0.2 % Trypanblau in PBS verdünnt. Von der Verdünnung wurden 10 µl in eine Zählkammer überführt, die lebenden Zellen in den vier großen Quadranten ausgezählt und die Zellzahl pro ml wurde nach der folgenden Formel berechnet.

$$c \left[ \frac{\text{Zellen}}{\text{ml}} \right] = \frac{n}{4} * d * 10.000$$

c= Konzentration der Zellen, n= Anzahl der gezählten Zellen, d= Verdünnung

### 3.4 Virologische Methoden und AAV-Selektion

#### 3.4.1 *In vitro*-Selektion von randomisierten Peptidbibliotheken

Nach Müller et al.<sup>122</sup>, modifiziert

Das generelle Protokoll einer *in vitro*-Selektion lautet wie folgt. Zunächst werden die Zellen von Interesse mit der AAV-Peptidbibliothek und zur Ermöglichung der Virusamplifikation mit Adenovirus co-infiziert. Bei beginnender Zelllyse werden die AAV aus den Zellen und dem Zellkulturüberstand geerntet und nach Abnahme einer Probe für die spätere Auswertung für die nächste Selektionsrunde eingesetzt. Nach maximal fünf Selektionsrunden wird die Selektion beendet.

##### 3.4.1.1 Infektion mit AAV und Adenovirus

Am Tag vor der Virusinfektion wurden  $5 \times 10^5$  Zellen pro 6-Well in einer 6-Well-Platte ausgesät. Am nächsten Tag wurden die Zellen mit der AAV-Peptidbibliothek mit einer MOI 1.000 vG/Zelle in 1 ml Medium mit Zusätzen für 6 bis 8 h inkubiert. Nach der Inkubation wurden die Zellen zweimal mit PBS gewaschen und mit Ad5 oder MAdV2 mit einer MOI 5 Plaque formierenden Einheiten pro Zelle (pfu/Zelle) in 750 µl Medium ohne Zusätze für 1 h inkubiert. Schließlich wurde der Zellkulturüberstand auf 2 ml mit Medium mit Zusätzen aufgefüllt und die infizierten Zellen für weitere 48 bis 72 h, bis ein cytopathischer Effekt bei 50 % der Zellen zu sehen war, bei 37°C inkubiert.

##### 3.4.1.2 Ernte der AAV

Nach der Inkubation wurden die AAV aus den Zellen und dem Zellkulturüberstand geerntet. Hierfür wurden die Zellen und der Überstand gesammelt und zunächst die Ad5 45 min bei 55°C in einem Wasserbad inaktiviert. Danach wurden die Zellen für 5 min bei 200 x g und Raumtemperatur (RT) zentrifugiert und getrennt vom Überstand aufgearbeitet.

Für die Isolierung von AAV aus Zellen wurden das Zellpellet in 100 µl PBS resuspendiert und die Zellen durch frier/tau-Zyklen aufgebrochen. Hierfür wurde das Zelllysate in Trockeneis mit 96 % (v/v) Ethanol für ca. 5 min eingefroren und danach in einem Wasserbad bei 37°C für 5 min aufgetaut. Der Zyklus wurde insgesamt dreimal durchgeführt. Anschließend wurde das Lysat 5 min mit 350 x g und RT zentrifugiert und der Überstand abgenommen.

Die AAV aus dem Zellkulturüberstand wurden mittels Säulen (Vivaspin® 20 mit 10.000 MWCO von Sartorius) und Zentrifugation für 10 min mit 3.000 x g bei RT in 400 µl konzentriert. AAV aus den Zellen und dem Überstand wurden vereinigt und eine Probe von 50 µl für die Auswertung der viralen DNS wurde genommen. Zum Start einer neuen Selektionsrunde wurden erneut Zellen ausgesät und mit den restlichen 450 µl der Viruslösung infiziert (**3.4.1.1**).

### 3.4.1.3 Isolierung von viraler DNS zur NGS-Probenvorbereitung

Zur Isolierung von viraler DNS zur NGS-Probenvorbereitung wurden 25 µl der zuvor genommenen Probe (3.4.1.3) mit 5 µl Proteinase K von Qiagen (2.4) für 10 min bei 56 °C inkubiert. Zur Inaktivierung der Proteinase K wurden die Proben für 15 min bei 95 °C inkubiert. Schließlich wurde die virale DNS aus 0,5 µl der Proben zur NGS-Probenvorbereitung verwendet (3.1.10).

### 3.4.2 *In vivo*-Selektion von randomisierten AAV-Peptidbibliotheken

Nach Michelfelder et al.<sup>133</sup>, modifiziert

#### 3.4.2.1 Tiere

Für die *in vivo*-Selektionen wurden 2 bis 3 Monate alte FVB (PymT (+/T)) Mäuse verwendet.

Die Tiere wurden im Einklang mit dem deutschen Tierschutzgesetz gehalten. Die Tierversuche des Projektes „Selektion gewebespezifischer Gentherapievektoren mit auf adeno-assoziierten Viren exprimierten randomisierten Peptidbibliotheken“, G42/09 wurden von der Behörde für Soziales, Familie, Gesundheit und Verbraucherschutz – Lebensmittelsicherheit und Veterinärwesen, Hamburg, DE genehmigt. Die Tiere wurden in der Forschungstierhaltung des UKEs in einzeln belüfteten Käfigen (*individual ventilated cages, IVCs*) mit einem Tag-/Nachtzyklus von 12 h gehalten.

#### 3.4.2.2 Virusinjektion und -zirkulation

Für die Virusinjektionen wurden die Mäuse mittels Inhalationsnarkose mit einem Gemisch aus 2,5 % (v/v) Isofluran und 97,5 % (v/v) Sauerstoff narkotisiert. Für die erste Selektionsrunde wurden  $1 \times 10^{11}$  virale Genome in 150 - 200 µl vorgewärmten HBSS intravenös in die Schwanzvene injiziert. Weitere Selektionsrunden wurden mit  $1 \times 10^{10}$  bis  $1 \times 10^{11}$  viralen Genomen gestartet. Die Zirkulationszeit der Viren betrug 48 h.

#### 3.4.2.3 Organentnahme

Nach einer Zirkulationszeit von 48 h wurden die Mäuse unter Inhalationsnarkose (3.4.2.3) mittels zervikaler Dislokation (Genickbruch) getötet. Anschließend wurden die Ziel- und Kontrollorgane (Niere, Pankreas, Herz, Lunge, Leber, Skelettmuskel Milz und Gehirn) entnommen, in flüssigen Stickstoff transferiert und bei -80°C gelagert.

#### **3.4.2.4 Isolierung von Langerhans-Inseln aus dem murinen Pankreas**

*Nach Lacy und Kostianovsky<sup>160</sup>, modifiziert*

Zur Isolierung von Langerhans-Inseln gibt es verschiedene Methoden. Dabei ist den meisten Methoden die Dissoziation der Inseln aus dem Verband des exokrinen Gewebes mittels Kollagenase gemein. Für die hiesige Selektion wurden für jede Selektionsrunde die Langerhans-Inseln von vier männlichen Mäusen isoliert.

Hierfür wurde jede einzelne Maus zunächst mittels Inhalationsnarkose (**3.4.2.2**) narkotisiert und anschließend mittels zervikaler Dislokation getötet (**3.4.2.3**). Danach wurde so schnell wie möglich die Bauchhöhle geöffnet und das Pankreas weitestgehend freigelegt. Mit einem Faden wurde der Gallengang am Eingang zum Dünndarm abgeschnürt. Für die Injektionen der Kollagenase P von Roche mit einer Konzentration von 1 mg/ml (**2.4**) wurde der Gallengang etwas gestrafft. Nach der Injektion von 2 bis 3 ml der Kollagenaselösung wurde das Pankreas entnommen. Für den Fall, dass das Organ nach der Injektion nicht vollständig expandiert war, wurde die Kollagenase nach der Organentnahme an verschiedenen Stellen des Organs in einer Petrischale injiziert.

Das Gewebe wurde anschließend für ca. 15 min in einem Wasserbad bei 37°C und zwischenzeitlichem Schütteln inkubiert. Danach wurden 25 ml eiskaltes HBSS zugegeben und das Lysat für ca. 10 s stark geschüttelt. Da die Inseln wesentlich größer als die Zellen des exokrinen Gewebes sind, wurden sie danach mittels Sedimentation voneinander getrennt. Hierfür wurden die Lysate für ca. 5 bis 10 min auf Eis gestellt. Danach wurde der Überstand abgenommen und das sedimentierte Lysat vorsichtig mit neuem Puffer resuspendiert. Auf diese Weise wurden bis zu drei Waschschriffe durchgeführt. Nach dem letzten Waschschriff wurden die Zellen vorsichtig in ca. 5 ml eiskaltem HBSS resuspendiert und unter dem Lichtmikroskop mit einer Eppendorf-Pipette per Hand gepickt.

#### **3.4.2.5 DNS-Isolierung**

##### *3.4.2.5.1 DNS-Isolierung aus Organen*

Zur Isolierung der DNS wurden ganze Organe im Falle der Selektionen und im Falle der DNS-Isolierung für die Vektorverteilung ein Gewebestück der Größe eines halben Mäuseherzens mit einem Gewebe-Homogenisator von Peqlab (**2.1**) homogenisiert. Für die Homogenisierung ganzer Organe wurden diese in ca. 50 mg schwere Stücke zerteilt.

Die Gewebestücke wurden in Röhrchen mit Keramikügelchen und ATL Lyse-Puffer (180 µl pro 25 mg Gewebe) des DNeasy<sup>®</sup> Blood & Tissue Kits von Qiagen überführt und gemäß der folgenden Einstellungen homogenisiert.



Gewebe	Bedingungen für die Homogenisierung
Niere, Herz, Leber, Milz	2 x 20 s, 5.500 rpm, 10 s Pause
Pankreas und Gehirn	1 x 20 s, 5.500 rpm
Skelettmuskel	2 x 30 s, 6.000 rpm, 20 s Pause

Nach der Homogenisierung wurde den Anweisungen des Protokolls für DNS-Isolierung aus Geweben des DNeasy® Blood & Tissue Kit von Qiagen gefolgt. Hierfür wurden die Gewebelysate mit Proteinase K (20 µl pro 25 mg Gewebe) für 1 bis 3 h bei 56 °C inkubiert, bis das Lysat homogen war. Danach wurde die zelluläre RNS mittels 4 µl RNase A (100 mg/ml) pro 25 mg Gewebe bei RT verdaut. Die Lysate wurden mit dem Vortexer für 15 s geschüttelt und Puffer AL (200 µl pro 25 mg Gewebe) wurden hinzugegeben. Nach der Zugabe von 100 % (v/v) Ethanol (200 µl pro 25 mg Gewebe) wurden die Lysate auf Silica-Membran-Säulchen transferiert und zentrifugiert (6.000 x g, 1 min, RT). Anschließend wurde die DNS mit 500 µl Puffer AW1 und Zentrifugation (6.000 x g, 1 min, RT) gewaschen. Danach erfolgte ein zweiter Waschschrift mit 500 µl Puffer AW2 und Zentrifugation (20.000 x g, 3 min, RT). Durch diesen Schritt wurde die Säule gleichzeitig getrocknet und die DNS wurde anschließend zweimal mit 100 µl AE-Puffer oder ddH<sub>2</sub>O mittels Zentrifugation (6.000 x g, 1 min, RT) eluiert. Schließlich wurde die DNS einer Probe vereinigt und wenn nötig mit Ethanol präzipitiert (**3.1.8**).

#### 3.4.2.7.2 DNS-Isolierung aus Langerhans-Inseln

Für die Isolierung der DNS aus Langerhans-Inseln wurden die isolierten Inseln von vier Mäusen in zwei Röhrchen mit Keramikkügelchen und ATL-Puffer überführt. Die Zellen eines Röhrchens wurden wie 25 mg Gewebe behandelt und nach dem DNeasy® Blood & Tissue Kit von Qiagen aufgearbeitet (**3.4.2.5.1**).

#### 3.4.2.6 Amplifikation der die randomisierten Inserts beinhaltenden viralen DNS

Die virale DNS wurde mittels PCR und den Primern AAV2 +/-4 (**2.13**) in 10 bis 20 Ansätzen amplifiziert (**3.1.9.1**). Um die PCR-Produkte auf ihre richtige Größe zu testen wurden 5 µl von drei zufällig gewählten Ansätzen und 5 µl der Vereinigung aller Ansätze auf einem 2 % Agarosegel mittels Gelelektrophorese analysiert (**3.1.4**). Anschließend wurde die amplifizierte DNS mit Ethanol präzipitiert (**3.1.8**).

#### 3.4.2.7 Verdau der die randomisierten Inserts beinhaltenden viralen DNS

Zur Herstellung einer sekundären Plasmidbibliothek müssen zunächst die randomisierten Inserts in das Plasmid pMT202-6 (**2.17**) ligiert werden. Um die Insert-codierende DNS mit den Schnittstellen

des Plasmids kompatibel zu machen wurden die ca. 400 bp langen gefällten PCR-Produkte mit dem Restriktionsenzym *BglI* nach dem folgenden Protokoll geschnitten.

Komponente	Volumen [ $\mu$ l]
PCR-Produkt (40 $\mu$ g)	X
10 x Puffer 3 (NEB) oder 10 x FD Puffer (Thermo Scientific)	10
<i>BglI</i>	10
ddH <sub>2</sub> O	ad 100

Die Reaktion wurde für 4 h bei 37°C (mit Enzym von NEB) oder für 1 h bei 37°C (mit Enzym von Thermo Scientific) in einem Thermocycler oder Heizblock inkubiert.

Nach der Restriktion hatten die das Insert beinhaltenden DNS-Fragmente eine Länge von 32 bp. Da die Silica-Membran-Säule des PCR Purification Kits von Qiagen Oligonukleotide mit einer Länge von bis zu 40 bp durchlassen, während größere Fragmente in der Säule gebunden werden, wurden diese dafür verwendet die gewünschten Fragmente von den längeren nicht gewünschten Fragmenten zu trennen. Hierfür wurde der Restriktionsverdau mit der 5-fachen Menge an Bindungspuffer gemischt und auf eine Silica-Membran-Säule überführt und zentrifugiert (13.000 rpm, 1 min, RT). Die 32 bp langen DNS-Fragmente befanden sich nun im Durchfluss, wurden mittels Ethanol Präzipitation (**3.1.8**) gefällt und in 40  $\mu$ l Elutionspuffer (1 mM Tris-HCl, pH 8.5) resuspendiert. Die mit *BglI* geschnittenen Fragmente waren kompatibel mit den *SfiI* geschnittenen Plasmid pMT202-6.

#### 3.4.2.8 Restriktion des AAV2-Genom codierenden Plasmids pMT202-6

Für die spätere Ligation des mit *BglI* geschnittenen DNS-Inserts wurde das das AAV2-Genom codierende Plasmid pMT202-6 (**2.17**) mit dem Restriktionsenzym *SfiI* nach folgendem Protokoll verdaut.

Komponente	Volumen [ $\mu$ l]
Plasmid-DNS (20 $\mu$ g)	X
10 x Puffer 2 (NEB) oder 10 x FD Puffer (Thermo Scientific)	10
100 x BSA (bei NEB-Enzym)	1
<i>SfiI</i>	10
ddH <sub>2</sub> O	ad 100

Der Restriktionsverdau wurde für 4 h bei 50°C in einem Thermocycler oder Heizblock inkubiert. Das geschnittene Plasmid wurde nach dem Verdau mit dem Qiagen PCR Purification Kit nach Herstellerangaben gereinigt und in 50  $\mu$ l Elutions-Puffer (1 mM Tris-HCl, pH 8.5) eluiert.

### 3.4.2.9 Ligation der Insert tragenden DNS in das Plasmid pMT202-6

Zur Ligation des mit *Bgl*I geschnittenen Inserts in das mit *Sfi*I geschnittenen Plasmids pMT202-6 wurden 500 - 1.000 ng Plasmid-DNS verwendet. Da es zur Erhaltung der Peptiddiversität, besonders in der ersten Selektionsrunde, wichtig ist eine möglichst hohe Transformationseffizienz zu erreichen, wurden für die Ligation verschiedene molare Verhältnisse getestet. Die Menge des Inserts für ein molares 1:1 Verhältnis von Vektor zu Insert wurde nach der folgenden Formel berechnet.

$$\text{Menge des Inserts [ng]} = \frac{\text{Länge des Inserts [kb]}}{\text{Länge des Vektors [kb]}} * \text{Menge des Vektors [ng]}$$

Generell wurden molare Verhältnisse von Vektor zu Insert von 1:30, 1:100 und 1:150 nach dem folgenden Protokoll getestet.

Komponente	Menge
Plasmid-DNS	500 ng
Insert-DNS	X
ddH <sub>2</sub> O	ad 20 µl

Die DNS wurde gemischt, für 2 min bei 65°C inkubiert und für 5 min auf Eis gestellt. Danach wurde die T4 DNS Ligase und T4 DNS Ligase-Puffer hinzugefügt.

Komponente	Volumen [µl]
T4 DNS Ligase (400 U/µl, NEB)	2,5
10 x T4 DNS Ligase-Puffer	2,5

Schließlich wurden die Reaktionen bei 16°C für 16 h inkubiert. Für die Religationskontrolle wurde statt der Insert-DNS ddH<sub>2</sub>O eingesetzt.

### 3.4.2.10 Elektroporation der sekundären Plasmidbibliothek in elektrokompetente DH5α

Nach der Ligation (3.4.2.9) wurden die Plasmide präzipitiert, das Pellet in 30 µl ddH<sub>2</sub>O resuspendiert (3.1.8), in elektrokompetente (3.2.2.2) DH5α (2.11.1) elektroporiert (3.2.3.2) und auf Agar-Platten mit Ampicillin (150 µg/ml) in verschiedenen Verdünnungen (1:10, 1:100 und 1:1.000) ausgestrichen. Wenn die Religationskontrolle weniger als 1 % der Transformanten der Ligationen aufwies und die Transformationseffizienz ungefähr bei 2x10<sup>7</sup> Transformanten lag, wurde der restliche Ligationsansatz in einer Großproduktion elektroporiert. Hierfür wurden der Ligationsansatz zu 500 µl Bakterien gegeben, auf 10 Küvetten (1 mm, Biorad) verteilt, elektroporiert (3.2.3.2) und in insgesamt 20 ml SOC-Medium (2.10.1) aufgenommen. Nach der Regeneration wurde eine kleine Probe zur Bestimmung der Transformationseffizienz genommen und die restliche Bakterienkultur zu 500 ml LB-Medium gegeben und bei 30 °C und 175 rpm über Nacht inkubiert. Nachdem die Kultur eine OD<sub>600</sub> von ca. 2 hatte wurde die Plasmid-DNS mit dem NucleoBond® PC 100 Kit oder dem NucleoBond® Xtra Maxi Kit von Macherey-Nagel isoliert (3.1.1).

### 3.4.2.11 Produktion der AAV2 Transfer-shuttles

Nach Müller et al.<sup>122</sup> und Waterkamp et al.<sup>125</sup>, modifiziert

#### 3.4.2.11.1 Transfektion von HEK 293T Zellen mit PolyFect von Qiagen

Am Tag vor der Transfektion wurden  $1 \times 10^7$  HEK 293T Zellen in fünf 15 cm Zellkulturschalen mit 20 ml DMEM Medium ausgesät. 1 bis 2 h vor der Transfektion wurde das verbrauchte Zellkulturmedium gegen 17 ml frisches DMEM Medium mit Zusätzen getauscht.

Für die Produktion von Transfer-Shuttles mit PolyFect wurde folgendes Schema verwendet (pro Zellkulturschale):

Komponente	Menge / Volumen
AAV2-Plasmidbibliothek (ITRs und <i>rep/cap</i> mit Insert) (2.17)	3 µg*
pVp3cm (Codon-modifiziertes Wildtyp <i>cap</i> ) (2.17)	3 µg*
pXX <sub>6</sub> (E1A, E1B, E2A, E4-orf6, VA) (2.17)	6 µg
Serumfreies Medium	ad 450 µl

\* Verhältnis 1:1

Der DNS-Mix und 120 µl PolyFect von Qiagen wurden gemischt und gemäß den Herstellerangaben für 5 bis 10 min bei RT inkubiert. Schließlich wurden 670 µl des DNS-PolyFect-Mix mit 1 ml Medium ohne Zusätze pro Zellkulturschale gemischt und auf die Zellen gegeben. Am nächsten Tag wurde das Medium gegen frisches Medium mit Zusätzen getauscht und die Zellen wurden für 48 h bei 37 °C inkubiert (3.3.1).

#### 3.4.2.11.2 Transfektion von HEK 293T Zellen mit PEI (Polyethylenimine von Polyplus)

Am Tag vor der Transfektion wurden  $1 \times 10^7$  HEK 293T Zellen in fünf 15 cm Zellkulturschalen mit 20 ml DMEM Medium ausgesät. 1 bis 2 h vor der Transfektion wurde das verbrauchte Zellkulturmedium gegen 17 ml frisches DMEM Medium ohne Antibiotika getauscht.

Für die Transfektion wurden zu transfizierende Plasmide nach folgendem Schema gemischt (pro Zellkulturschale):

Komponente	Menge / Volumen
AAV2-Plasmidbibliothek (ITRs und <i>rep/cap</i> mit Insert) (2.17)	55 µg*
pVp3cm (Codon-modifiziertes Wildtyp <i>cap</i> ) (2.17)	55 µg*
pXX <sub>6</sub> (E1A, E1B, E2A, E4-orf6, VA) (2.17)	110 µg
NaCl (300 mM)	1.000 µl
ddH <sub>2</sub> O	ad 2.000 µl

\* Verhältnis 1:1

Der PEI-Mix wurde wie folgt hergestellt:

Komponente	Volumen [ $\mu$ l]
PEI (333 $\mu$ g/ml)	456
NaCl (300 mM)	1.000
ddH <sub>2</sub> O	544

Der Plasmid-Mix und der PEI-Mix wurden separat pipettiert, dann vereinigt und für 10 min bei RT inkubiert. Danach wurde der Plasmid-PEI-Mix auf die Zellen gegeben und nach einer Inkubationszeit von 4 bis 8 h bei 37°C wurde das Medium gewechselt. Anschließend wurden die Zellen 72 h bei 37°C inkubiert.

#### 3.4.2.12 Zellernte und Isolierung der rAAV

Drei Tage nach der Transfektion wurden die Zellen geerntet und die rAAV isoliert. Zur Herstellung von Zelllysaten wurde der Zellkulturüberstand abgenommen und die Zellen in 10 ml PBS mit einem Zellschaber von der Zellkulturflasche gelöst. Die Zellsuspension wurde danach für 5 min mit 200 x g bei RT zentrifugiert, der Überstand verworfen und das Zellpellet bis zu einem Volumen von 4,5 ml mit PBS-MK (PBS mit MgCl<sub>2</sub> und KCl) aufgefüllt und darin resuspendiert. Um die Zellen zu lysieren, wurden drei frier/tau- Zyklen durchgeführt. Hierfür wurde das Zelllysate in Trockeneis mit 96 % (v/v) Ethanol für 10 min und eingefroren und danach in einem Wasserbad bei 37°C für 10 min aufgetaut. Der Zyklus wurde insgesamt dreimal durchgeführt. Falls das Lysat noch nicht homogen war, wurde es mittels einer Spritze homogenisiert (23 Gauge). Zum Entfernen der freien DNS wurde das Lysat mit Benzonase Endonuklease (50 U/ml) (2.4) für 30 min bei 37°C inkubiert. Schließlich wurde das Zelllysate für 15 min bei 3.000 x g und RT zentrifugiert und der Überstand mit den rAAV abgenommen.

#### 3.4.2.13 Reinigung der rAAV mittel Iodixanol Dichtegradientenzentrifugation

*Nach Hermens et al.*<sup>161</sup>

Nach der Isolierung der rAAV wurden diese mittels Iodixanol (OptiPrep™ von Axis-Shield) und Dichtegradientenzentrifugation gereinigt und konzentriert. Iodixanol eignet sich sehr gut zur Herstellung von Lösungen mit definierten unterschiedlichen Dichten. Zur Herstellung der Phasen eines diskontinuierlichen Gradienten wurden vier Grundlösungen mit verschiedenen Konzentrationen an Iodixanol (15, 25, 40 und 54 % (v/v)) wie folgt angesetzt.

*15 % Iodixanol-(OptiPrep)-Lösung (75 ml):*

Komponente	Volumen [ml]
NaCl (5M)	20
10 x PBS-MK	10
ddH <sub>2</sub> O	45

*25 % Iodixanol-(OptiPrep)-Lösung (60,75 ml):*

Komponente	Volumen [ml]
10 x PBS-MK	10
Phenolrot-Lösung	0,75
ddH <sub>2</sub> O	50

*40 % Iodixanol-(OptiPrep)-Lösung (33,33 ml):*

Komponente	Volumen [ml]
10 x PBS-MK	10
ddH <sub>2</sub> O	23,33

*54 % Iodixanol-(OptiPrep)-Lösung (10,75 ml):*

Komponente	Volumen [ml]
10 x PBS-MK	10
Phenolrot-Lösung	0,75

Der Gradient wurde nach dem folgenden Schema hergestellt.

Phase [%]	Puffervolumen [ml]	Iodixanolvolumen [ml]	Gesamtvolumen [ml]	Verwendetes Volumen [ml]
15	2,495	0,83	3,325	3
25	1,415	0,97	2,385	2
40	0,665	1,33	1,995	1,66
54	0,215	1,8	2,015	1,66

Zur Herstellung der Gradienten wurden zunächst 4,5 ml Lysat in 13,5 ml Quick-Seal™ Ultra-Clear™ Zentrifugenröhrchen überführt. Anschließend wurden die einzelnen Phasen langsam und vorsichtig, beginnend mit der Phase mit der geringsten Dichte, sukzessive unter die vorherige gespritzt. Hierbei wurde die Entstehung von Luftblasen vermieden. Das Phenolrot diente zur Überprüfung der Intaktheit der einzelnen Phasen und Phasengrenzen. Zur Vermeidung von Luftblasen wurde das Röhrchen bis zur Hälfte des Halses mit 1 x PBS-MK aufgefüllt und mit einem *Tube Topper* (Beckman) verschlossen. Die Röhrchen wurden gewogen und mit Parafilm austariert bis die Differenz weniger als 0,05 g betrug. Anschließend wurden die Gradienten in einer Ultrazentrifuge von Beckman für 70 min bei 350.000 x g (58.000 rpm in einem 70.1 Ti Rotor von Beckman) und 18 °C mit aktiver Bremse zentrifugiert. Aufgrund der Größe und der Masse von AAV konzentrierten sich hierbei die Viren in der

40 % Phase. Um an die Phase zu gelangen wurde das Röhrchen zum Druckausgleich zunächst am oberen Teil mit einer Nadel (23 G) angestochen. Anschließend wurde das Röhrchen mit einer Spritze (23 G-Nadel) auf Höhe knapp unter der 40 % Phase angestochen und die Phase abgezogen. Nachdem eine Probe zur Bestimmung des Titers genommen wurde, wurde die Lösung aliquotiert und bei -80°C gelagert.

#### **3.4.2.14 Bestimmung des replikativen Titers**

Vor der Produktion der sekundären Virusbibliothek wurden der replikative und/oder der genomische Titer der Transfer-Shuttles bestimmt. Die Bestimmung des replikativen Titers erfolgte mit einem Replikationsassays über den Nachweis des Rep Proteins mittels Antikörper. Auf der Grundlage der Anzahl der das Rep Protein exprimierenden Zellen wurde der replikative Titer bestimmt.

Hierfür wurden zunächst  $2,5 \times 10^4$  HEK 293T Zellen in 100  $\mu$ l DMEM Medium mit Zusätzen in einer Poly-L-Lysin beschichteten 96-Well-Platte ausgebracht. Am nächsten Tag, wenn die Zellen eine Konfluenz von 70-80 % erreicht hatten, wurden sie mit verschiedenen Konzentrationen der Transfer-Shuttles infiziert. Zu diesem Zweck wurden die Transfer-Shuttles in 100  $\mu$ l DMEM Medium mit Zusätzen 1:5 verdünnt und vorsichtig mit dem Medium des ersten Wells gemischt. Hier beginnend wurden in der ersten Reihe der 96-Well-Platte sukzessive 1:2 Verdünnungen hergestellt. Beginnend mit dem Well der Reihe wurden dann 1:10 Verdünnungen für jede Spalte hergestellt. Nach 1 h wurden die Zellen mit humanem Ad5 mit einer MOI von 5 pfu/Zelle in 50  $\mu$ l DMEM ohne Zusätze überinfiziert. Die Zellen wurden nach 24 h mit 200  $\mu$ l PBS gewaschen, mit 50  $\mu$ l/Well 3,5 % (v/v) Paraformaldehyd in PBS fixiert, mit 50  $\mu$ l/Well 0,2 % (v/v) Triton in PBS permeabilisiert und mit 50  $\mu$ l/Well 3 % (v/v) BSA und 0,05 % (v/v) Tween in PBS für 1 h blockiert. Für die Inkubation mit dem Anti-Rep-Primärantikörper AK76.3 (**2.5**) wurde dieser 1:2 in 3 % (v/v) BSA, 0,05 % (v/v) Tween in PBS verdünnt und zu 50  $\mu$ l/Well auf die Zellen gegeben. Anschließend wurden die Zellen mit 200  $\mu$ l PBS gewaschen und dann mit 50  $\mu$ l/Well mit dem sekundären Ganz-Anti-Maus-FITC-Antikörper (**2.5**) von Santa Cruz in einer 1:200 Verdünnung in 3 % BSA, 0,05 Tween in PBS für 1 h bei RT oder über Nacht bei 4°C inkubiert. Schließlich wurden die FITC exprimierenden Zellen als AAV positive Zellen unter dem Fluoreszenzmikroskop gezählt und der replikative Titer unter Berücksichtigung der Verdünnungen bestimmt. Bei dem Assay wurden zwei Negativkontrollen mitgeführt. Hierbei war eine Probe ohne Ad5 und eine ohne den Primärantikörper.

Da bei dieser Methode der Titerbestimmung z.T. Schwankungen auftraten, wurde in diesen Fällen zusätzlich der genomische Titer mittels q-PCR (**3.1.9.3**) bestimmt.

#### **3.4.2.15 Produktion von sekundären AAV2-Peptidbibliotheken**

Zur Produktion von sekundären AAV-Peptidbibliotheken wurden HEK 293T Zellen mit AAV Transfer-Shuttles mit einer geringen MOI und humanem Ad5 co-infiziert.

Einen Tag vor der Infektion wurden fünf 15 cm Zellkulturschalen mit  $1 \times 10^7$  Zellen pro Schale vorbereitet. Am Tag der Infektion wurde das Zellkulturmedium verworfen und die Zellen mit den AAV Transfer-Shuttles mit einer MOI von 0,1 oder 1 pfu/Zelle in 5 ml DMEM Medium ohne Zusätze bei 37°C inkubiert. Nach 30 min wurden die Zellen mit Ad5 mit einer MOI von 5 pfu/Zelle in 2,5 ml DMEM Medium ohne Zusätze co-infiziert. Nach weiteren 30 min wurden 12,5 ml DMEM Medium mit Zusätzen hinzugegeben. Die Zellen wurden ungefähr 72 h bei 37°C inkubiert bis ein cytophatischer Effekt bei ca. 50 % der Zellen zu sehen war.

Aufgrund der Zugabe von Ad5 fand während der Inkubation eine Zelllyse statt. Zur Isolierung der AAV wurden Zellen und Zellkulturüberstand in 50ml Röhrchen überführt und zur Inaktivierung des Ad5 für 45 min bei 55 °C im Wasserbad inkubiert. Danach wurde der Zellkulturüberstand mittels Zentrifugation (3.000 x g, 15 min, RT) von den toten Zellen und Zelltrümmern getrennt, in Vivaspins® von Sartorius überführt und mittels Zentrifugation bei 3.000 x g auf ein Volumen von ungefähr 4,5 ml konzentriert. Schließlich wurde das Konzentrat mittels Iodixanol Dichtegradientenzentrifugation gereinigt (**3.4.2.13**).

#### **3.4.2.16 Beenden einer Selektion**

Zur Überprüfung der Anreicherung von Insert-Sequenzen im kleinen Maßstab wurde die DNS von fünf bis zehn Klonen nach jeder Selektionsrunde sequenziert. Wurde hierbei die Anreicherung eines Sequenzmotivs oder einer Sequenz festgestellt, wurde die Selektion vor Erreichen der fünften Selektionsrunde beendet. So wurde im Falle der Selektion im Pankreas die Selektion nach der vierten Selektionsrunde, aufgrund der Anreicherung eines Sequenzmotivs, und die Selektion in den Langerhans-Inseln nach Runde drei, aufgrund der Anreicherung einer Peptidsequenz, beendet.

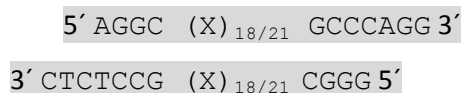
#### **3.4.3 Produktion von Kapsid-modifizierten rAAV2-Vektoren mit Reporter gen**

Zur Evaluation selektierter AAV-Klone wurden Kapsid-modifizierte rAAV-Vektoren mit dem Reportertransgen Luciferase unter CMV Promotor produziert (**1.3.2**). Hierfür wurden neben dem Ad5-Helferplasmid ein Plasmid, das für das Transgen zwischen den ITRs codiert und ein Plasmid, das die Rep und die Cap-modifizierten Proteine codiert benötigt. Während das für das Transgen codierende Plasmid bereits kloniert war und für jeden zu produzierenden Klon verwendet werden konnte, musste das Rep und Cap codierende Plasmid für jeden zu evaluierenden Klon neu kloniert werden.



### 3.4.3.1. Herstellung von Plasmiden mit Insertion im AAV2-*cap*

Hierfür wurden Oligonukleotide, die die jeweilige Insertsequenz und die *Bgl*I Schnittstelle codierten, kreiert und kommerziell von Metabion bezogen.



Für die Hybridisierung der Oligonukleotide wurden diese auf eine Konzentration von 200  $\mu$ M in 1 x T4 DNS Ligase-Puffer verdünnt und mit 10  $\mu$ l des Puffers wie folgt gemischt.

Komponente	Volumen [ $\mu$ l]	Endkonzentration
Oligonukleotid for (200 $\mu$ M)	5	50 $\mu$ M
Oligonukleotid rev (200 $\mu$ M)	5	50 $\mu$ M
Puffer	10	

Die Oligonukleotide wurden für 2 min bei 95°C in einem Thermo-Cycler erhitzt und dann in dem abgeschalteten Cycler gelassen bis die Proben ungefähr 25°C hatten. Die nun doppelsträngigen Oligonukleotide wurden danach über ihre *sticky ends* in einem Vektor- zu Insert-Verhältnis von 1:100 an die Position R588 des AAV2-VP1 in das Plasmid pXX2 187 (2.17) ligiert (3.1.7).

Die Plasmide wurden in *E. coli* XL1 blue transformiert (3.2.3.1), Einzelklone wurden gepickt (3.2.1.2) und mittels Kolonie-PCR auf die korrekte Insertion getestet (3.1.9.2). Positive Einzelklone wurden in einem großen Ansatz amplifiziert und Plasmid-DNS wurde isoliert (3.1.1).

### 3.4.3.2 Großproduktion von selektierten Kapsid-modifizierten rAAV2-Einzelklonen mit Reporter gen in HEK 293T Zellen

Für die Produktion von Kapsidvarianten mit Reportertransgen wurden HEK 293T Zellen mit PolyFect nach dem bereits unter 3.4.2.11 beschriebenen Prinzip transfiziert. Lediglich die zu transfizierenden Plasmide sind andere und deswegen im Folgendem aufgeführt.

Komponente	Menge / Volumen
pAAV CMV tetO <sub>2</sub> Luc (Luciferase zwischen ITRs) (2.17)	3 $\mu$ g*
pXX <sub>2</sub> 187 (AAV2 <i>rep/cap</i> ohne ITRs und mit Insert in <i>cap</i> ) (2.17)	3 $\mu$ g*
pXX <sub>6</sub> (E1A, E1B, E2A, E4-orf6, VA) (2.17)	6 $\mu$ g
Serumfreies Medium	ad 450 $\mu$ l

\*Verhältnis 1:1

Nach der Transfektion wurden die Viren isoliert (3.4.2.12), gereinigt (3.4.2.13), titriert (3.1.9.3) und für weitere Evaluationen *in vitro* 3.4.4.1 oder *in vivo* 3.4.4.2 verwendet.

### 3.4.4 Evaluation von selektierten Kapsidvarianten

#### 3.4.4.1 Evaluation von *in vitro* selektierten Kapsidvarianten mittels Luciferaseassay

Zur Untersuchung der Transduktionseffizienzen der selektierten Klone wurde die Luciferase-expression pro Well mittels Luciferaseassay von Promega bestimmt. Hierfür wurden die jeweiligen Zellen in 96-Well-Platten ausgesät und am nächsten Tag mit den jeweiligen rAAV in Triplikaten mit einer MOI 1.000 vG/Zelle in 200 µl Medium mit Zusätzen für 4 h bei 37°C inkubiert. Danach wurde das Medium gewechselt und die Zellen für 48 h bei 37°C inkubiert. Nach der Inkubationszeit wurde der Zellkulturüberstand abgenommen und 20 µl Reporter Lysis Buffer (RLB) von Promega auf die Zellen gegeben. Die Zellen wurden durch einen frier/tau-Zyklus lysiert und die Luciferaseaktivität wurde mit dem Luminometer (Mithras LB 940 von Berthold Technologies) nach der automatischen Zugabe des Luciferase Assay Reagenz von Promega mit folgenden Einstellungen gemessen.

Parameter:

Volumen:	100 µl pro Well
Geschwindigkeit:	langsam
Verzögerung:	2 s
Messzeit:	10 s

#### 3.4.4.2 Analyse der Vektorverteilung in der Maus (*homing*) zur Evaluation der *in vivo* selektierten Kapsidvariante

Zur Evaluation der *in vivo* selektierten Kapsidvarianten wurden die Vektoren mit Luciferase als Reportergen mit einer Konzentration von  $5 \times 10^{10}$  vG intravenös in die Schwanzvene injiziert (3.4.2.2). Nach zwei Wochen wurden die Tiere getötet und die Organe entnommen (3.4.2.3). Für die Untersuchung der Vektorverteilung wurde die Gesamt-DNS eines Organstückes mit der Größe eines halben Mäuseherzens des Zielorgans und der Kontrollorgane isoliert (3.4.2.4) und auf eine DNS-Konzentration von 20 ng/µl eingestellt (3.1.2). Anschließend wurde die Anzahl der Vektorgenome mittels q-PCR gemessen (3.1.9.3). Und schließlich wurde die Anzahl der Vektorgenome in 100 ng Gesamt-DNS berechnet.

### 3.4.5 Statistik

Die statistische Auswertung der Luciferaseassays (3.4.4.1) und der Vektorverteilung (3.4.4.2) wurde mit einem One-way ANOVA Test des Programms Prism 3.0, mit anschließender Bonferroni-Korrektur, gemacht. Die Werte sind als Mittelwerte mit positiver Standardabweichung gezeigt. Die Signifikanzlevel sind:  $p < 0,05 = *$ ;  $p < 0,01 = **$ ;  $p < 0,001 = ***$ .

## 4 ERGEBNISSE

### 4.1 *In vivo*-Selektion von randomisierten AAV2-NNK<sub>7</sub>-Peptidbibliotheken

Nach der erfolgreichen Selektion eines Gehirn- und eines Lungen-spezifischen AAV2-Vektors in Mäusen<sup>134,135</sup> sollte im Folgenden die Übertragbarkeit der verwendeten AAV-Peptidbibliothek-Selektionsmethode auf neue gentherapeutisch hochrelevante Ziele untersucht werden. Hierfür wurden AAV2-Peptidbibliotheken in der Niere, dem Pankreas und dessen funktionellen Untereinheiten, den Langerhans-Inseln, von Mäusen nach intravenöser Injektion selektiert.

#### 4.1.1 *In vivo*-Selektion von randomisierten AAV2-NNK<sub>7</sub>-Peptidbibliotheken in Niere und Pankreas von Mäusen

##### 4.1.1.1 Generierung von sekundären AAV2-NNK<sub>7</sub>-Peptidbibliotheken zur Selektion in Niere und Pankreas

Die Generierung von sekundären AAV-Peptidbibliotheken während einer *in vivo*-Selektion ist ein mehrstufiger Prozess (1.3.7). Um sicherzustellen, dass die theoretische Diversität der Bibliothek während der Selektion nicht experimentell bedingt limitiert oder verzerrt wird, muss zur Aufrechterhaltung der Diversität jeder Schritt der Generierung von sekundären Peptidbibliotheken quantitativ überprüft werden.

Die Diversität einer AAV-Peptidbibliothek ist durch die Anzahl an transformierten Bakterienklonen bei der Herstellung der Plasmidbibliothek und die Anzahl der Zellen bei der Produktion der Virusbibliothek begrenzt. Die für die *in vivo*-Selektionen verwendete AAV2-NNK<sub>7</sub>-Peptidbibliothek hat eine anhand der Anzahl an Plasmidklonen berechnete theoretische Diversität von  $2 \times 10^8$ . Für die erste Selektionsrunde aller hier durchgeführten Selektionen wurden  $1 \times 10^{11}$  virale Genome (vG) injiziert; dieses entspricht theoretisch einer 500-fachen Abdeckung jedes in der Bibliothek vorhandenen Klons. Trotz der Annahme, dass die Anzahl von Peptidvarianten aufgrund der Anreicherung der viralen Partikel im Zielgewebe von Selektionsrunde zu Selektionsrunde sinkt, wurden für die weiteren Selektionsrunden nie weniger als  $1 \times 10^{10}$  vG (> 50 Kopien jedes viralen Partikels) injiziert, um zu gewährleisten, dass genug virale Partikel das Zielgewebe erreichen und daraus isoliert werden können.

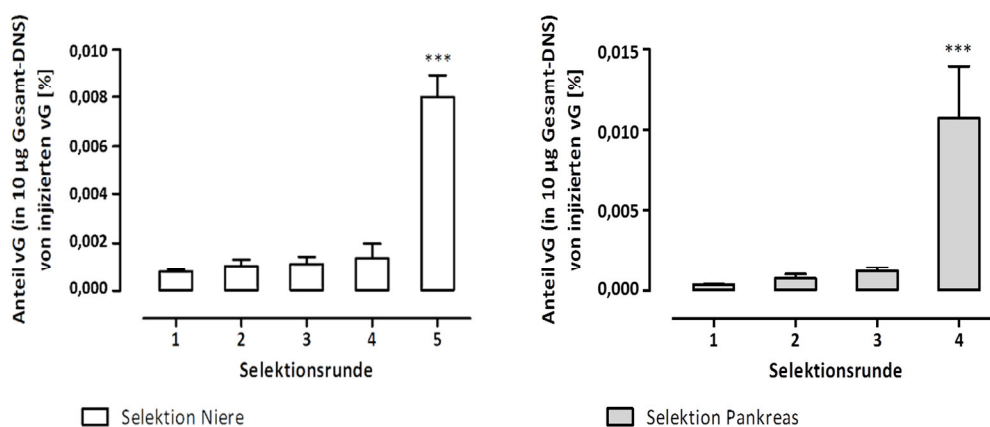
Die für die Qualitätssicherung der Peptidbibliothek relevanten Parameter der Selektion in der Niere und dem Pankreas sind in **Tabelle 17** zusammengefasst.

**Tabelle 17: Zusammenfassung wichtiger Kenngrößen der AAV2-Peptidbibliothek-Selektion in Niere und Pankreas.** Dargestellt sind die Anzahl der injizierten viralen Partikel und der transformierten Klone, sowie die replikativen Titer der Transfer-Shuttles und die genomischen Titer der AAV2-Peptidbibliotheken einzelner Selektionsrunden in Niere und Pankreas.

Selektion und Runde	AAV injiziert [gP]	Plasmidbibliothek [transformierte Klone]	Replikativer Titer der Transfer-Shuttles [rep E/ml]	Genomischer Titer der AAV2-Peptidbibliothek [gP/ml]
<b>Niere</b>				
1	$1 \times 10^{11}$	$8,70 \times 10^6$	$3,00 \times 10^{10}$	$2,40 \times 10^{12}$
2	$1 \times 10^{11}$	$4,34 \times 10^6$	$1,20 \times 10^9$	$5,74 \times 10^{11}$
3	$1 \times 10^{11}$	$3,90 \times 10^6$	$7,42 \times 10^9$	$5,97 \times 10^{12}$
4	$1 \times 10^{11}$	$1,52 \times 10^7$	$2,20 \times 10^9$	$9,97 \times 10^{11}$
5	$1 \times 10^{10}$	$2,43 \times 10^7$	-	-
<b>Pankreas</b>				
1	$1 \times 10^{11}$	$1,49 \times 10^7$	$2,00 \times 10^9$	$1,30 \times 10^{12}$
2	$1 \times 10^{11}$	$7,10 \times 10^6$	$1,17 \times 10^9$	$2,75 \times 10^{12}$
3	$1 \times 10^{11}$	$1,67 \times 10^7$	$6,62 \times 10^8$	$2,21 \times 10^{11}$
4	$1 \times 10^{10}$	$1,65 \times 10^7$	-	-

#### 4.1.1.2 Anreicherung von Viruspartikeln im Laufe der Selektionen

Die positive Selektion von gewebespezifischen viralen Vektoren und die gleichzeitige negative Selektion von unspezifischen Vektoren stellen das Grundprinzip einer erfolgreichen AAV-Peptidbibliothek-Selektion dar. Zur Beurteilung der Anreicherung der viralen Partikel im Zielorgan wurden diese nach jeder Selektionsrunde quantifiziert (**Methoden 3.1.9.3**). **Abbildung 11** zeigt die Anreicherung der viralen Partikel im Zielorgan. Von Runde eins bis vier der Selektion in der Niere und von Runde eins bis drei im Pankreas war ein leichter Anstieg zu verzeichnen; in der jeweils letzten Selektionsrunde wurde dann ein sehr deutlicher Anstieg gemessen. Hier wurden in der Niere in 10 µg aus dem Zielgewebe isolierter Gesamt-DNS 0,008 % der injizierten viralen Partikel gemessen, während dieser Wert im Pankreas 0,011 % betrug.

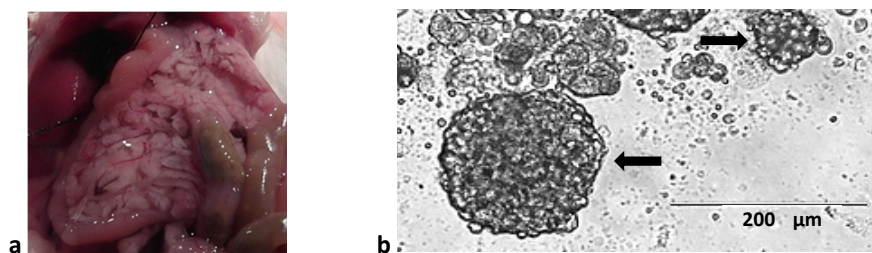


**Abbildung 11: Anreicherung von viralen Partikeln während der Selektion.** Für jede Selektionsrunde wurde Gesamt-DNS aus dem jeweiligen Zielorgan isoliert und die Anzahl der viralen Partikel mittels q-PCR gemessen. Dargestellt ist der prozentuale Anteil an injizierten viralen Partikeln, gemessen in 10 µg Gesamt-DNS. Gezeigt sind die Mittelwerte mit positiver Standardabweichung. Die statistische Auswertung wurde mit einem einseitigen ANOVA Test und anschließender Bonferroni-Korrektur gemacht.  $p < 0,05 = *$ ;  $p < 0,01 = **$ ;  $p < 0,001 = ***$ ; ( $n=3$ ).

#### 4.1.2 Selektion von AAV2-Peptidbibliotheken in Langerhans-Inseln des murinen Pankreas

##### 4.1.2.1 Isolierung von Langerhans-Inseln aus dem murinen Pankreas

Zum Vergleich der AAV-Selektion zwischen Gesamtorgan und einer definierten funktionellen Untereinheit, wurden zusätzlich zum Gesamt-Pankreas AAV-Selektionen in den therapeutisch hochrelevanten endokrinen Langerhans-Inseln durchgeführt (**Methoden 3.4.2.4**). Hierzu wurde das Pankreas der Maus durch Injektion mit Kollagenase expandiert und die Langerhans-Inseln anschließend isoliert (**Abbildung 12**). Die Reinheit der pro Selektionsrunde aus je vier Mäusen isolierten Langerhans-Inseln wurde unter dem Lichtmikroskop auf zwischen 50 % und 70 % geschätzt.



**Abbildung 12: Expandiertes Pankreas und isolierte Langerhans-Inseln umgeben von exokrinen Zellen. a)** Expandiertes Pankreas nach Kollagenase-Injektion. Nach der Tötung der Maus wurde der Bauchraum geöffnet und Kollagenase P in den Gallengang injiziert. **b)** Isolierte Langerhans-Inseln (Pfeile) umgeben von exokrinen Zellen und Zelltrümmern. Phasenkontrastmikroskopie, 20-fach vergrößert.

Die Anzahl der per Hand isolierten Langerhans-Inseln für jede Selektionsrunde und die daraus isolierte Menge an DNS sind in **Tabelle 18** dargestellt. Aufgrund der geringen isolierten DNS-Menge wurde diese in ihrer Gesamtheit für die darauffolgende PCR eingesetzt.

**Tabelle 18: Übersicht über die Selektionsrunden der Selektion in den Langerhans-Inseln.** Übersicht über die Anzahl der isolierten Langerhans-Inseln, der daraus isolierten DNS und die für die PCR eingesetzte DNS-Menge für jede Selektionsrunde.

Selektionsrunde	Anzahl der Langerhans-Inseln	Isolierte DNS nach Fällung [µg]	Für PCR eingesetzte DNS [µg]
1	ca. 120	31,0	31
2	ca. 140	3,8	3,8
3	ca. 120	5,2	5,2

##### 4.1.2.2 Generierung von sekundären AAV2-NNK<sub>7</sub>-Peptidbibliotheken zur Selektion in Langerhans-Inseln

Für die erste Selektionsrunde wurden wie für die Selektionen in Niere und Pankreas  $1 \times 10^{11}$  virale Genome injiziert. Im Laufe der Selektion wurde mangels Ausbeute die Anzahl der injizierten viralen Genome verringert (**Tabelle 19**). Aufgrund von Ungenauigkeiten innerhalb des Tests zur Bestimmung

des replikativen Titers wurde zusätzlich der genomische Titer der Transfer-Shuttles bestimmt. Dieser war hier, wie auch bei anderen Selektionen, ca. 1.000 -fach höher als der replikative Titer (Daten nicht gezeigt). Für die Transfer-Shuttles der zweiten Selektionsrunde wurde der genomische Titer bestimmt und zur Berechnung des replikativen Titers verwendet. Die relevanten Parameter der Selektion in den Langerhans-Inseln sind in **Tabelle 19** dargestellt.

**Tabelle 19: Zusammenfassung wichtiger Kenngrößen der AAV2-Peptidbibliothek-Selektion in Langerhans-Inseln.** Dargestellt sind die Anzahl der injizierten viralen Partikel und der transformierten Klone, sowie die replikativen Titer der Transfer-Shuttles und die genomischen Titer der AAV2-Peptidbibliotheken einzelner Selektionsrunden der AAV2-Peptidbibliothek-Selektionen in Langerhans-Inseln.

Selektions- runde	AAV injiziert [gp]	Plasmidbibliothek [transformierte Klone]	Replikativer Titer der Transfer- Shuttles [rep E/ml]	Genomischer Titer der Transfer-Shuttles [gp/ml]	Genomischer Titer der AAV2- Peptidbibliothek [gp/ml]
1	$1 \times 10^{11}$	$7,20 \times 10^6$	$2,12 \times 10^8$	$5,45 \times 10^{11}$	$4,89 \times 10^{11}$
2	$5 \times 10^{10}$	$3,03 \times 10^6$	-	$5,01 \times 10^{11}$	$1,26 \times 10^{11}$
3	$1 \times 10^{10}$	$4,63 \times 10^6$	-	-	-

## 4.2 Analyse des *in vivo*-Selektionsverlaufs mittels Next-Generation Sequencing (NGS)

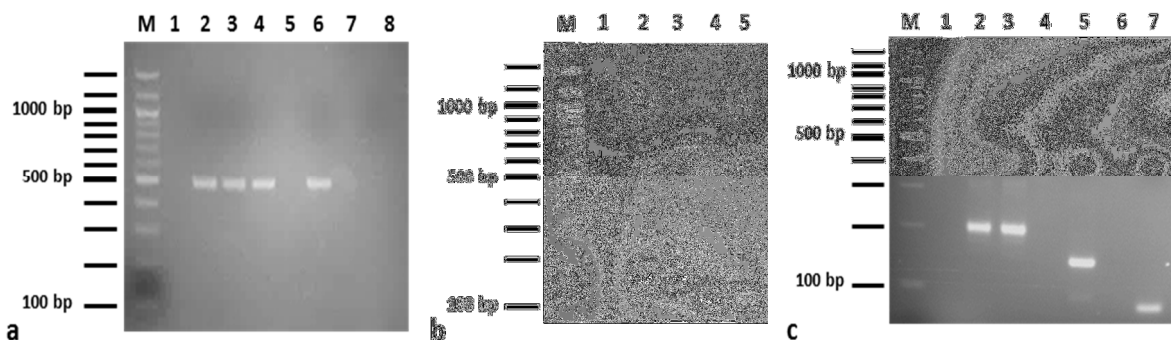
In bisherigen AAV-Peptidbibliothek-Selektionen wurden die DNS-Sequenzen von fünf bis zehn zufällig gewählten Klonen analysiert, um einen gewebespezifischen Vektor zu finden. Hierbei war der Fokus auf die Anreicherung von Peptidsequenzen und Sequenzmotiven gesetzt und angereicherte oder motivtragende Klone wurden als Reportergenvektoren produziert und weiter evaluiert. Diese Evaluierungen zeigten, dass der häufigste Klon nicht zwangsläufig der Klon mit der höchsten Spezifität war, und dass ein gewebespezifischer Klon nicht unbedingt das häufigste Sequenzmotiv trug. Für eine gezieltere Auswahl zu evaluierender Klone und einen tieferen Einblick in die Peptidbibliotheken ist folglich die Sequenzierung von wesentlich mehr als nur zehn Klonen notwendig. Entsprechend wurden die beschriebenen *in vivo*-Selektionen im Folgenden mittels NGS analysiert. Gleichzeitig wurde die erfolgreiche konventionell durchgeführte Selektion im Gehirn<sup>134</sup> retrospektiv mittels NGS ausgewertet und als Positivkontrolle mit den hiesigen Selektionen verglichen.

### 4.2.1 Probenherstellung für die NGS-Analyse

Für die Analyse der Anreicherung von Peptidvarianten während der Selektion wurde für jede Selektionsrunde die virale DNS aus den Zielorganen isoliert und sequenziert, während für die Untersuchung der Spezifität aus der jeweils letzten Selektionsrunde zusätzlich die virale DNS der

Kontrollorgane (Niere, Pankreas, Herz, Lunge, Leber und Skelettmuskel; Selektion Gehirn: Herz, Lunge und Leber) analysiert wurde. Für die Methode der parallelen Tiefensequenzierung und die anschließende Analyse wurde jede zu untersuchende DNS-Probe mit den Illumina® Adapter-Sequenzen versehen (3.1.10). Um unspezifische Amplifikationen auszuschließen und zur Überprüfung der PCR-Produkte und wurden diese während der Probenvorbereitung nach jeder PCR mittels Agarose-Gelelektrophorese (3.1.4) auf ihre korrekte Größe kontrolliert.

In **Abbildung 13** sind exemplarisch die Ergebnisse der PCR-Produkte nach jedem PCR-Schritt gezeigt. Das PCR-Produkt der ersten PCR zur Amplifikation der viralen DNS hatte eine Größe von ca. 480 bp, während das Produkt der zweiten PCR eine Größe von ca. 130 bp hatte. Nachdem die Adapter-Sequenzen angefügt wurden, hatte das finale Produkt eine Länge von ca. 190 bp.



**Abbildung 13: Analyse der PCR-Produkte der NGS-Proben mittels Agarose-Gelelektrophorese. a)** PCR 1 (Amplifikation); 1: leer; 2, 3 und 4: PCR-Produkte (ca. 480 bp); 5: leer; 6: Probe der Mischung aller Ansätze mit gleicher Intensität; 7: leer; 8: Negativkontrolle. **b)** PCR 2 (Anfügen der Linker); 1: leer; 2 und 3: PCR-Produkte (ca. 130 bp); 4 und 5: Negativkontrollen (Primeramplifikationen). **c)** PCR 3 (Anheften der Barcodes); 1: leer; 2 und 3: PCR-Produkte (ca. 190 bp); 4: leer; 5: Negativkontrolle (Primeramplifikationen); 6: leer, 7: Negativkontrolle (Primer). M: 100 bp DNS-Marker (*New England Biolabs*).

#### 4.2.2 Analyse der Tiefensequenzierung der AAV-Bibliotheken während der *in vivo*-Selektion – Möglichkeiten der Auswertung

Um zunächst einen Überblick über die NGS-Daten zu erhalten, ist in **Tabelle 20** die Anzahl der im Rahmen der *in vivo*-Selektion der AAV-Peptidbibliotheken sequenzierten DNS- und Peptidsequenzen aufgelistet. Während sich die Anzahl der gelesenen DNS-Sequenzen (Lesetiefe) nach jeder Selektionsrunde in einer ähnlichen Größenordnung bewegte, verringerte sich die Anzahl der tatsächlich verschiedenen Peptidsequenzen mit jeder Selektionsrunde. Nur im Falle der letzten Selektionsrunde der Selektion in der Niere und im Pankreas wurden vergleichsweise weniger DNS-Sequenzen gelesen. Mit den hier gezeigten Daten wird im Folgenden die Anreicherung von Peptidvarianten genauer untersucht.

**Tabelle 20: Zusammenfassung der NGS-Daten für die AAV2-Peptidbibliothek-Selektionen in Niere, Pankreas, Langerhans-Inseln und Gehirn.** Für jede Selektionsrunde ist die Anzahl der sequenzierten DNS-Sequenzen und die Anzahl der jeweiligen Peptidsequenzen gezeigt. Während die Anzahl der gelesenen DNS-Sequenzen innerhalb einer Selektion für jede Selektionsrunde, mit Ausnahme der Selektion in der Niere (siehe Diskussion), in einer ähnlichen Größenordnung liegt, verringert sich die Anzahl der Peptidsequenzen mit jeder Selektionsrunde.

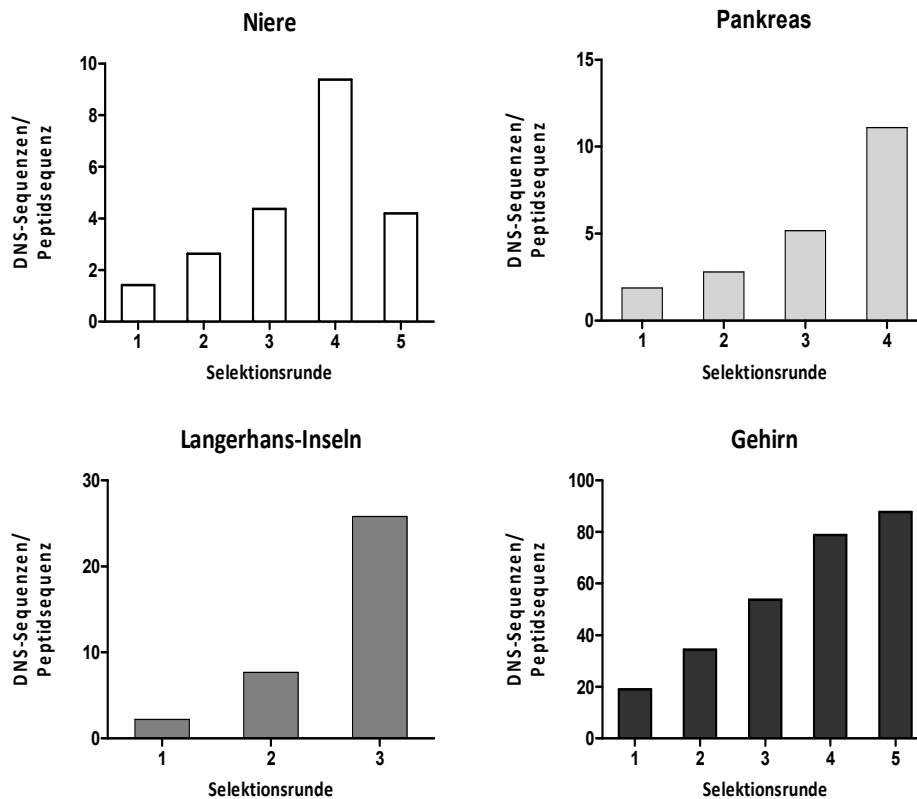
Selektion	Runde	Anzahl DNS-Sequenzen	Anzahl Peptidsequenzen
Niere	1	142318	99999
	2	127886	48437
	3	139220	31834
	4	178619	19013
	5	13257	3159
Pankreas	1	125172	66662
	2	134771	48024
	3	140480	27102
	4	63267	5705
Langerhans- Inseln	1	82780	37431
	2	64079	8322
	3	65370	2533
Gehirn	1	155243	8200
	2	169470	4934
	3	160854	2988
	4	164025	2080
	5	175338	1998

#### 4.2.2.1 Anreicherung von Peptidsequenzen während der Selektionen

Zur Bestimmung der Anreicherung von Peptiden während der Selektionen werden im Folgenden zwei Möglichkeiten vorgestellt. Hierbei wurden im Falle der ersten Variante alle Peptide betrachtet. Für die zweite Auswertungsvariante wurden die Peptide nach ihrer Häufigkeit geordnet und nur die häufigsten Peptide analysiert.

Für die Analyse der Peptidanreicherung unter Berücksichtigung aller sequenzierten Sequenzen wurde für jede Selektionsrunde das Verhältnis von DNS-Sequenzen zu Peptidsequenzen berechnet. **Abbildung 14** zeigt ein steigendes Verhältnis mit jeder Selektionsrunde, außer für die Selektionsrunde fünf der Selektion Niere. Das höchste DNS/Peptid-Verhältnis der für diese Arbeit durchgeführten Selektionen wurde für die Selektion in den Langerhans-Inseln bestimmt. Hier kamen in der letzten Selektionsrunde durchschnittlich ca. 25 DNS-Sequenzen auf eine Peptidsequenz. Eine noch stärkere Anreicherung wurde für die Selektion im Gehirn berechnet. In der letzten Selektionsrunde lag das Verhältnis bei ca. 90 DNS-Sequenzen pro Peptid.

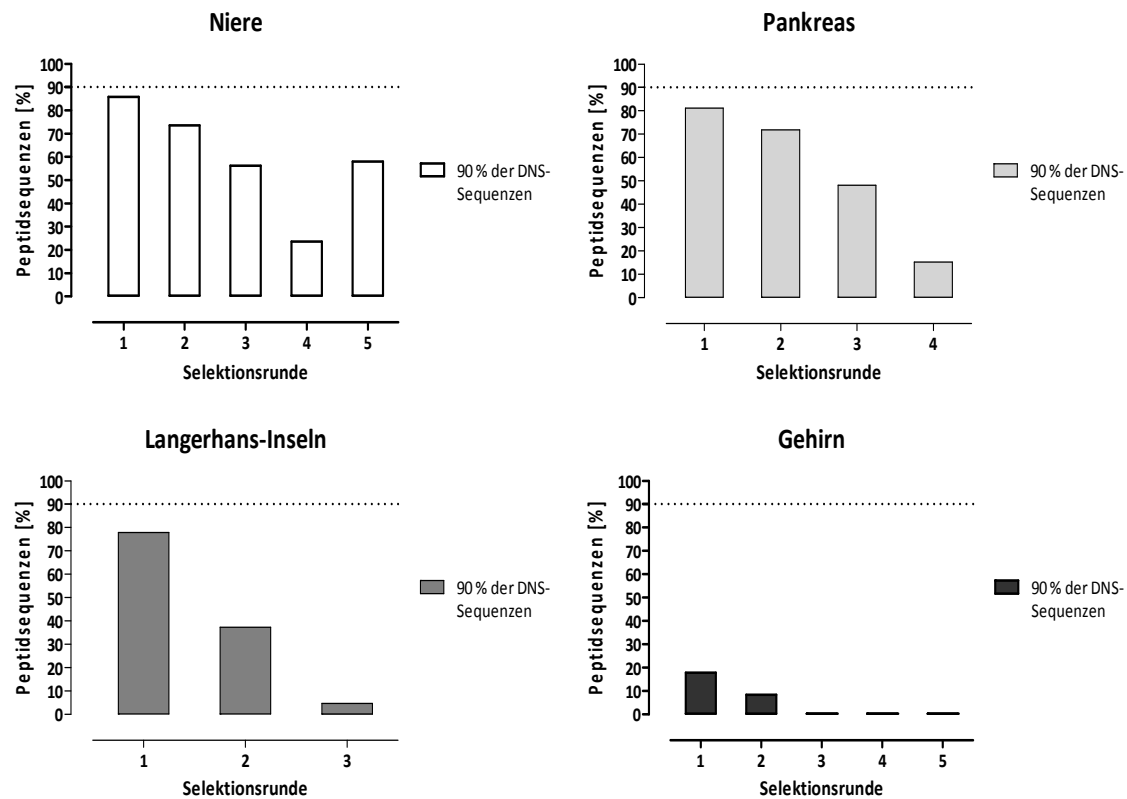




**Abbildung 14: Untersuchung der Peptidanreicherung während der Selektionen.** Für jede Selektionsrunde ist das DNS/Peptid-Verhältnis gezeigt. Hierbei wurden alle Sequenzen berücksichtigt. Für jede Selektion, außer der Selektion in der Niere (siehe Diskussion), ist eine Anreicherung des Verhältnisses mit jeder Selektionsrunde zu sehen.

In **Abbildung 15** wird die Anreicherung der Peptidsequenzen im Zielgewebe noch deutlicher. Für jede Selektionsrunde wurden die Peptidsequenzen ihrer Häufigkeit nach geordnet (von häufig zu selten). Jede Peptidsequenz wird durch eine bestimmte Anzahl von DNS-Sequenzen repräsentiert. Beginnend mit der häufigsten Peptidsequenz wurden die zugehörigen DNS-Sequenzen solange addiert bis 90 % der DNS-Sequenzen einbezogen waren. Jeder Balken repräsentiert somit 90 % der DNS-Sequenzen und die Balkenhöhe zeigt, wie viel Prozent der Peptidsequenzen diese abdecken.

Im Falle der letzten Selektionsrunde der Selektion in der Niere repräsentieren 58 % der Peptidsequenzen 90 % der gelesenen DNS-Sequenzen. Während bei der Selektion im Pankreas nur 15 % der Peptidsequenzen betrachtet werden müssen, um 90 % der DNS-Sequenzen abzudecken. Bezüglich der Selektion in den Langerhans-Inseln sind es 4,5 % der Peptidsequenzen. Die Selektion im Gehirn zeigt jedoch die höchste Anreicherung. Hier decken nur 0,2 % der Peptidsequenzen 90 % der DNS-Sequenzen ab.



**Abbildung 15: Untersuchung der Peptidanreicherung unter Berücksichtigung der Peptidhäufigkeiten.** Jeder Balken repräsentiert 90 % der gelesenen DNS-Sequenzen. Die Balkenhöhe entspricht dem jeweiligen Anteil an Peptidsequenzen.

#### 4.2.2.2 Sequenzmotivanalyse der zehn häufigsten Peptidsequenzen

Zum Vergleich der Selektionsrunden der AAV2-NNK<sub>7</sub>-Peptidbibliothek-Selektionen in verschiedenen Zielgeweben und zur Identifizierung eines Sequenzmotivs wurden die zehn häufigsten Peptidsequenzen aller Selektionen und Selektionsrunden in **Tabelle 21** gegenübergestellt. Als Sequenzmotiv wurde die häufigste Kombination von Aminosäuren in der gleichen Position, die mindestens in drei von zehn Sequenzen zu finden war, definiert.

Die häufigsten Peptidsequenzen der Selektion in der Niere zeigen final zwei Sequenzmotive „NXXRXXE“ und „NDXXRXX“. Ersteres ist bereits in Selektionsrunde zwei und drei zu finden. Besonders auffällig ist die Häufung von Asparagin in der ersten Aminosäureposition jeder Selektionsrunde. Das Peptid „NDAGREV“ ist mit knapp 23 % und 14 % das häufigste Peptid der letzten beiden Selektionsrunden in der Niere.

Das finale Sequenzmotiv der Selektion im Pankreas ist die Sequenz „XXDXT<sup>R</sup>/<sub>X</sub>X“. Dieses Motiv ist bereits in der zweiten Selektionsrunde zu sehen und wird mit jeder Runde häufiger und somit stärker angereichert. Das gleiche Motiv ist auch in den Peptidsequenzen der Selektion in den Langerhans-Inseln bereits in der ersten Selektionsrunde zu finden. Des Weiteren hat sich in dieser Selektion die

Sequenz „EEDRTRA“ bereits in der zweiten Selektionsrunde zu ca. 40 % und in der letzten Selektionsrunde zu ca. 73 % stark angereichert.

In der zweiten Selektionsrunde der Selektion im Gehirn baut sich das Sequenzmotiv „XDGXXWX“ auf. In der dritten Selektionsrunde ist es dominant. Des Weiteren sind Teile hiervon in der als spezifisch befundenen Sequenz „NRGTEWD“ vorhanden<sup>134</sup>. Zusätzlich zu diesem Motiv zeigen die NGS-Daten der Selektion im Gehirn das bereits aus der Selektion in der Niere bekannte Sequenzmotiv „NXXRXXE“. Ein Teil des Motivs („NXXRXXX“) wurde unter den zehn häufigsten Sequenzen in Runde eins, zwei, vier und fünf gefunden, während es in Runde drei zwischenzeitlich von dem Motiv „XDGXXWX“ verdrängt wurde. In Selektionsrunde fünf der Selektion im Gehirn sind folglich zwei Sequenzmotive vorhanden, „XDGXXWX“ und „NXXRXXE“. Erst durch die Betrachtung der in einem Peptid tatsächlich vorkommenden Kombination häufiger Aminosäuren wurde eine Koexistenz von Sequenzmotiven erkannt.

Der Vergleich der hier untersuchten Peptidsequenzen der verschiedenen Selektionen zeigt große Ähnlichkeiten zwischen den jeweiligen ersten Selektionsrunden. Das dominante Asparagin in der ersten Aminosäureposition ist hierbei besonders auffällig. Es ist auch Teil des in der Niere und dem Gehirn selektierten Motivs „NXXRXXE“. Im Falle der Selektionen in Pankreas und den Langerhans-Inseln beginnt sich bereits in der zweiten bzw. ersten Selektionsrunde das Sequenzmotiv „XXDXT<sup>R</sup>/X“ aufzubauen.

**Tabelle 21: Sequenzmotivanalyse der zehn häufigsten Peptidsequenzen der Selektionen in Niere, Pankreas, Langerhans-Inseln und Gehirn.** Dargestellt sind die zehn häufigsten Peptidsequenzen jeder Selektionsrunde der AAV2-NNK<sub>7</sub>-Selektionen in Niere, Pankreas, Langerhans-Inseln und Gehirn und ihr prozentualer Anteil an allen sequenzierten Peptiden. Aminosäuren, die mindestens dreimal an derselben Position vorkommen, sind individuell farbig markiert. Zusätzlich ist das Sequenzmotiv jeder Selektionsrunde dargestellt. Für eine deutlichere Darstellung der Sequenzmotive wurden die Peptide nicht nach ihrer Häufigkeit sortiert.

Gewebe	Runde 1 (%)	Runde 2 (%)	Runde 3 (%)	Runde 4 (%)	Runde 5 (%)
<b>Niere</b>					
	NOVGSWS 0,72	NNVRTSE 7,75	NNVRTSE 6,14	NDAGREV 23,15	NDAGREV 13,92
	NNVRTSE 0,48	NDARREV 0,81	NEARGRE 0,85	NRTAEGY 2,36	NDAGREL 1,38
	NGVKKEG 0,28	NEARGRE 0,34	NDARREV 1,33	NEARGRE 1,45	NDVRRSE 1,51
	NRTAEGY 0,17	NQTSRQD 0,87	NVVRSEA 0,66	NNVRTSE 1,31	NEARGRE 1,55
	NDARREV 0,15	NRTAEGY 2,31	NQTSRQD 1,12	NOVGSWS 1,25	NNVRTSE 1,32
	NVAKSVQ 0,08	ENRTRSE 0,69	NRTAEGY 2,19	NQTSRQD 1,75	EDTTKAG 1,39
	RGDLGLS 0,45	ERDETRS 1,34	ERDETRS 1,58	GQDATRV 4,02	KEGTSWQ 7,96
	GGKERAP 0,12	GQDATRV 0,39	GQDATRV 1,00	ERDETRS 1,88	GVGTSHM 1,24
	VRRPRFW 0,30	RGDLGLS 0,55	ENRTRSE 0,60	KEGTSWQ 5,86	GQDATRV 1,99
	ERDETRS 0,08	ESGLSQS 0,40	KEGTSWQ 0,66	GVGTSHM 1,91	EKLHEGV 1,40
<b>Motiv</b>	NXVXXXX	NXXRXXX	NXXRXXX	-	NXXRXXE NDXXRXX
<b>Pankreas</b>					
	NOVGSWS 2,87	ERDETRS 1,35	ERDETRS 9,16	GQDATRV 14,79	
	NNVRTSE 0,32	EYDATRS 0,66	EYDATRS 3,86	GDATRAN 3,74	
	NGVKKEG 0,20	EIDGTRQ 0,22	EIDGTRQ 2,46	GSDDTQQ 5,13	
	NRTAEGY 0,17	GQDATRV 0,35	GQDATRV 5,56	GQDGTKM 2,04	
	NVAKSVQ 0,13	DTDWTRA 0,32	GDATRAN 1,78	ERDETRS 12,63	
	NDARREV 0,12	NNVRTSE 1,19	NADETRA 1,00	EIDGTRQ 5,72	
	RGDLGLS 2,12	NOVGSWS 0,51	NNVRTSE 0,95	EYDATRS 4,86	
	ERDETRS 0,13	NRTAEGY 0,31	GSDDTQQ 2,23	DGDKTKA 5,56	
	VRRPRFW 0,41	RGDLGLS 0,87	DGDKTKA 1,35	AADGTKA 2,40	
	SDGLAWV 0,15	SDGLAWV 0,26	NRTAEGY 1,37	SEEDTQM 1,42	
<b>Motiv</b>	NXVXXXX	XXDXTRX	XXDXTRX	XXDXTR <sub>R</sub> ,X	
<b>Langerhans-Inseln</b>					
	EEDRTRA 4,02	EEDRTRA 39,79	EEDRTRA 72,51		
	ERDETRS 0,26	ERDETRS 2,83	ERDETRS 5,79		
	DTDWTRA 0,05	GQDATRV 0,54	GQDATRV 1,11		
	GQDATRV 0,05	NNVRTSE 0,79	ENRTRSE 0,49		
	NOVGSWS 12,65	NDARREV 0,38	ESGHGYF 0,24		
	NNVRTSE 0,45	NOVGSWS 9,51	NRTAEGY 1,71		
	NDARREV 0,11	NRTAEGY 2,69	NQTSRQD 0,35		
	NRTAEGY 0,17	NQTSRQD 0,32	NMARQAD 0,26		
	NGVKKEG 0,15	ESGHGYF 2,30	NEVRAHP 0,26		
<b>Motiv</b>	XXDXTRX	XXDXTRX	XXDXTRX		
<b>Gehirn</b>					
	NNVRTSE 7,16	NNVRTSE 30,36	SDGLTWS 49,52	SDGLTWS 56,25	SDGLTWS 62,48
	NVVRSEA 0,49	NEVRAHP 1,55	ADGVQWT 14,95	ADGVQWT 10,67	ADGVQWT 3,62
	NEVRAHP 0,44	NMARQAD 1,06	DDGVSWK 3,15	GDGIAWV 1,67	GDGIAWV 0,88
	NNSRQDT 0,30	NNTRETG 0,89	DDGVQWT 0,53	GDGLTWS 0,71	NNSAARWE 0,32
	NMARQAD 0,30	NQTSRQD 1,20	GDGIAWV 3,03	DDGVSWK 0,60	NNVRTSE 15,16
	NDARREV 1,49	NRCTEWD 1,70	EDCMGWK 0,85	NRCTEWD 0,72	NSVRGGE 0,19
	NGVKKEG 0,33	NGVKKEG 1,54	GDGLTWS 0,64	NNVRTSE 14,35	NEARGRE 0,18
	NRTAEGY 0,59	SDGLTWS 6,06	DDKVTWL 2,77	NMARQAD 4,49	NMARQAD 9,53
	ENRTRSE 0,51	ADGVQWT 4,27	NRCTEWD 6,63	NEVRAHP 2,13	NEVRAHP 2,11
	GNGTKQP 0,37	DDKVTWL 0,91	NNVRTSE 7,59	GNGTKQP 0,58	GNGTKQP 0,44
<b>Motiv</b>	NXXRXXX	NXXRXXX	XDGXXWX	XDGXXWX	XDGXXWX NXXRXXE

#### 4.2.2.3 Vergleich der häufigsten Peptidsequenzen aus Zielgewebe und jeweiligen Kontrollorganen

Die Auswertung der Selektionen mittels NGS erlaubt nicht nur die Untersuchung der Anreicherung bestimmter Peptidsequenzen über die Selektionsrunden; zusätzlich ist es möglich Hinweise zur Spezifität dieser Sequenzen zu bekommen. In **Tabelle 22** sind die häufigsten Peptidsequenzen der AAV2-Selektionen in Niere, Pankreas, Langerhans-Inseln und Gehirn, sowie der jeweiligen Nicht-Zielorgane aufgelistet. Die Peptidsequenz „NDAGREV“ war die häufigste Sequenz der AAV-Peptidbibliothek-Selektion in der Niere. Zusätzlich wurde sie aber auch in dem Kontrollorgan Pankreas am häufigsten gefunden. Die häufigste Peptidsequenz der AAV2-Selektion in der Bauchspeicheldrüse war „GQDATRV“, aber auch in den Kontrollorganen Niere und Lunge wurde diese Sequenz am stärksten angereichert. Die häufigste Peptidsequenz der Selektion in den Langerhans-Inseln „EEDRTRA“ war auch die Sequenz, die am häufigsten in allen untersuchten Kontrollorganen gefunden wurde. Bezüglich der Selektion im Gehirn war die häufigste Sequenz „SDGLTWS“ auch die häufigste der Kontrollorgane Herz und Lunge. Gleichzeitig teilten sich die Kontrollorgane Herz, Lunge und Skelettmuskel der Selektion in der Niere, sowie das Kontrollorgan Herz aus der Selektion Pankreas die gleiche häufigste Peptidsequenz „SDGLAWV“. Des Weiteren war diese Sequenz sehr ähnlich zu der Sequenz „SDGLTWS“ der Selektion im Gehirn. Schließlich stellte die Sequenz „ERDETRS“ die häufigste Sequenz des Kontrollorgans Leber der Selektion in der Niere und der Kontrollorgane Leber und Skelettmuskel der Selektion im Pankreas dar.

Der Vergleich der häufigsten Peptidsequenzen zeigt, dass die im Zielgewebe am häufigsten vorkommenden Kapsidvarianten mindestens in einem Kontrollorgan ebenfalls am häufigsten zu finden waren und somit zumindest anhand der Sequenzanalyse und deren Organverteilung nicht als zielgewebespezifisch eingeordnet werden können. Gleichzeitig zeigen diese Daten das Vorhandensein von Kapsidvarianten, die im Sinne der Zielsetzung zu einem optimierten Gentransfer in bestimmte Gewebe einen Vorteil haben und sich im Zielgewebe, allerdings auch in den Kontrollorganen unspezifisch anreichern. Folglich zeigen diese Untersuchungen, dass es nicht reicht, in der Sequenzanalyse den im Zielgewebe häufigsten AAV-Klon zu identifizieren, um einen tatsächlich zielgewebespezifischen AAV-Vektor zu finden.

**Tabelle 22: Vergleich der häufigsten Peptidsequenzen in Ziel- und Kontrollgeweben.** Gezeigt sind die häufigsten Sequenzen und ihr jeweiliger prozentualer Anteil in den Ziel- und Kontrollgeweben der jeweils letzten Selektionsrunde. Die Sequenz des Zielgewebes jeder Selektion ist hervorgehoben und unterstrichen.

Selektion Gewebe	Niere	Pankreas	Langerhans-Inseln	Gehirn
Niere	<u>NDAGREV</u> 13,92 %	GQDATRV 11,80 %	EEDRTRA 62,12 %	-
Pankreas/ Langerhans-Inseln	NDAGREV 20,61 %	<u>GQDATRV</u> 14,79 %	<u>EEDRTRA</u> 72,51 %	-
Herz	SDGLAWV 19,67 %	SDGLAWV 18,25 %	EEDRTRA 61,40 %	SDGLTWS 65,98 %
Lunge	SDGLAWV 5,50 %	GQDATRV 9,75 %	EEDRTRA 40,86 %	SDGLTWS 69,16 %
Leber	ERDETRS 17,00 %	ERDETRS 12,79 %	EEDRTRA 65,98 %	NNVRTSE 38,91 %
Skelettmuskel	SDGLAWV 16,26 %	ERDETRS 12,31 %	EEDRTRA 66,73 %	-
Gehirn				<u>SDGLTWS</u> 62,48 %

#### 4.2.2.4 Analyse von Peptidanreicherung- und Spezifität zur Bewertung der identifizierten Sequenzmotive

Zur Bewertung der Anreicherung und der Spezifität von Sequenzmotiven und Peptidsequenzen wurden etablierte Bewertungskriterien herangezogen<sup>135</sup>. Hierfür wurden *Scores* (Werte zwischen 0 und 1) zur Einschätzung der Peptidanreicherung (*Enrichment-Score, E-Score*), der Peptidspezifität (*General Specificity-Score, GS-Score*) sowie einer Kombination aus beidem (*Combined-Score, C-Score*) bestimmt. Im Gegensatz zu der Arbeit von Körbelin et al.<sup>135</sup> wurde in der vorliegenden Arbeit zur Einschätzung der Peptidanreicherung ein *GE-Score (General Enrichment-Score)* anstatt eines *E-Scores* berechnet (**3.1.10.6**). Der Vergleich der Werte für Anreicherung und Spezifität ermöglicht es AAV-Klone zu finden, die die höchste Wahrscheinlichkeit besitzen, spezifisch zu sein. Hierbei gilt, je näher der Wert an 1 ist, desto angereicherter bzw. spezifischer ist der entsprechende AAV-Klon. Um die Werte besser einordnen zu können, wurden für jeden *Score* Orientierungswerte (gestrichelte Linien in den Diagrammen) bestimmt. Peptide, deren Werte über dem Orientierungswert des *GE-Scores* liegen haben sich während der Selektion durchschnittlich angereichert; Werte über dem Orientierungswert des *GS-Scores* bedeuten eine durchschnittliche höhere relative Häufigkeit des Peptids im Zielorgan im Vergleich zu den Kontrollorganen. Bezüglich des *C-Scores* bedeuten Werte über der Linie, dass sich das jeweilige Peptid während der Selektion im Zielorgan durchschnittlich mehr als verdoppelt hat, während es in den Kontrollorganen durchschnittlich nur halb so oft vorkam. Im Folgenden sind die berechneten *Scores* für die häufigsten 30 Peptide der Selektionen dargestellt.

Bezüglich der Selektion in der Niere trugen sieben der 30 häufigsten Sequenzen das Motiv „NXXRXX<sup>E</sup>/<sub>x</sub>“ und insgesamt 13 Sequenzen hatten ein Asparagin an erster Position. 26 der 30 häufigsten Sequenzen hatten sich während der Selektion durchschnittlich angereichert; hierbei zeigten die meisten Sequenzen aber nur eine eher schwache Anreicherung. 12 Sequenzen lagen über der Orientierungslinie des *GS-Scores*, davon lagen nur vier Sequenzen deutlich drüber. Sechs der 30 Sequenzen hatten zwar einen Wert über dem der Orientierungslinie des *C-Scores*, insgesamt waren die jeweiligen Werte aber unter 0,1 und somit sehr niedrig (**Abbildung 16a**).

Im Falle der Selektion im Pankreas trugen 23 der 30 häufigsten Sequenzen das selektierte Motiv „XXDXT<sup>R</sup>/<sub>x</sub>“. Alle diese Sequenzen hatten sich durchschnittlich während der Selektion angereichert und alle, bis auf eine, lagen über der Orientierungslinie des *GS-Scores*. Insgesamt neun Sequenzen lagen über der Linie des *C-Scores*, die Sequenz „GDDATRD“ sogar deutlich mit einem Wert über 0,3 (**Abbildung 16b**).

Bezüglich der Selektion in den Langerhans-Inseln hatten 11 Sequenzen das selektierte Motiv „XXDXT<sup>R</sup>/<sub>X</sub>“. Diese Sequenzen hatten sich durchschnittlich angereichert und zehn von ihnen waren im Zielorgan häufiger als in den Kontrollorganen zu finden. Insgesamt lagen fünf Sequenzen über dem Orientierungswert des C-Scores, die jeweiligen Werte waren aber wie schon bei der Selektion in der Niere eher niedrig (**Abbildung 16c**).

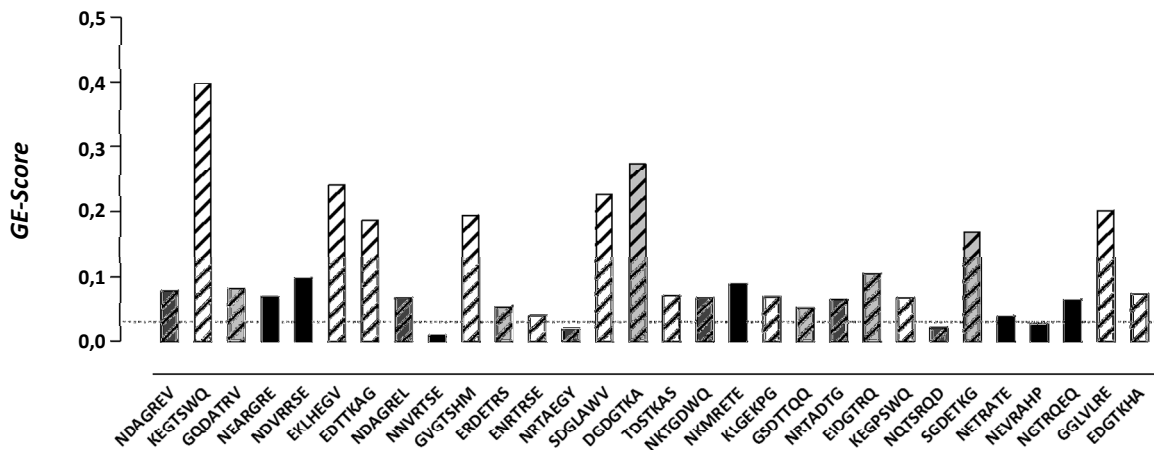
Im Falle der retrospektiv mit NGS analysierten Selektion im Gehirn trugen sechs der 30 häufigsten Sequenzen das selektierte Motiv „XDGXXWX“, 17 Sequenzen das Motiv „NXXRXX<sup>E</sup>/<sub>X</sub>“ und vier zusätzliche Sequenzen ein Asparagin an erster Position. Vier der 11 durchschnittlich angereicherten Sequenzen trugen das in dieser Selektion selektierte Motiv „XDGXXWX“. Alle das Motiv tragenden Sequenzen wurden im Zielorgan Gehirn häufiger gefunden als durchschnittlich in den Kontrollorganen. Die höchste Spezifität allerdings zeigte die Sequenz „NRGTEWD“, die bereits von Körbelin et al. als spezifisch identifiziert wurde. Bezüglich des C-Scores lagen dennoch alle Werte unterhalb der Orientierungslinie (**Abbildung 16d**).

Insgesamt zeigen die Ergebnisse der *GE-Scores* aller Selektionen, dass sämtliche Peptide, die ein Asparagin an erster Stelle besitzen und somit auch jene, die das Motiv „NXXRXX<sup>E</sup>/<sub>X</sub>“ tragen, zwar zu den häufigsten Peptidsequenzen gehören, sich aber im Vergleich zu anderen Peptiden am wenigsten stark oder gar nicht über die Selektionsrunden angereichert haben (**Abbildung 16a, b, c und d GE-Score**). Bei der Betrachtung des *GS-Scores* für die Selektionen in Niere, Pankreas und Langerhans-Inseln fällt auf, dass unter den häufigsten 30 Peptidsequenzen keine Sequenz mit dem Motiv „NXXRXX<sup>E</sup>/<sub>X</sub>“ zu finden ist, die eine höhere relative Häufigkeit im Zielorgan als in den Kontrollorganen aufweist (**Abbildung 16a, b, und c GS-Score**). Nur im Falle der Selektion im Gehirn wurde eine einzige solche Peptidsequenz mit sehr geringer Häufigkeit (0,06 %) gefunden (**Abbildung 16d GS-Score**). Die fehlende Anreicherung (*GE-Score*) und die mangelnde Spezifität (*GS-Score*) für Peptide mit der Sequenz „NXXRXX<sup>E</sup>/<sub>X</sub>“ spiegelten sich auch in den niedrigen Werten des kombinierten *C-Scores* wieder.

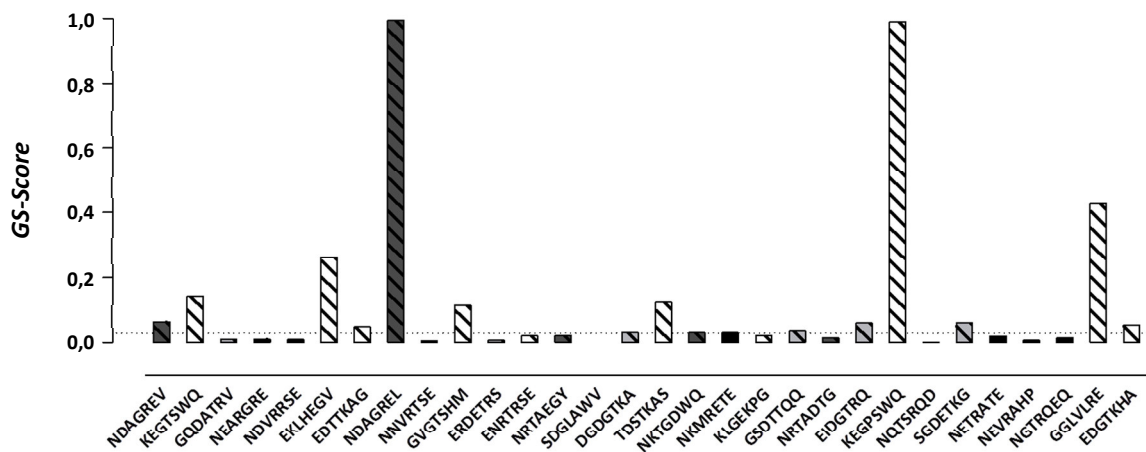
Im Gegensatz dazu wurden Peptide, die das zweite in dieser Arbeit identifizierte Sequenzmotiv „XXDXT<sup>R</sup>/<sub>X</sub>“, selektiert in Pankreas und Langerhans-Inseln, trugen, im Zielorgan stark angereichert und ihre Werte lagen mit einer Ausnahme über dem Orientierungswert des *GS-Scores*.

Zusammenfassend zeigen diese Ergebnisse die Nicht-Spezifität des Sequenzmotives „NXXRXX<sup>E</sup>/<sub>X</sub>“ im untersuchten Kontext; das Motiv „XXDXT<sup>R</sup>/<sub>X</sub>“ hingegen zeigte eine besondere Anreicherung in dem Zielorgan Pankreas und auch in den Langerhans-Inseln. Zusätzlich konnte die Spezifität des Motivs „XDGXXWX“ verifiziert werden.

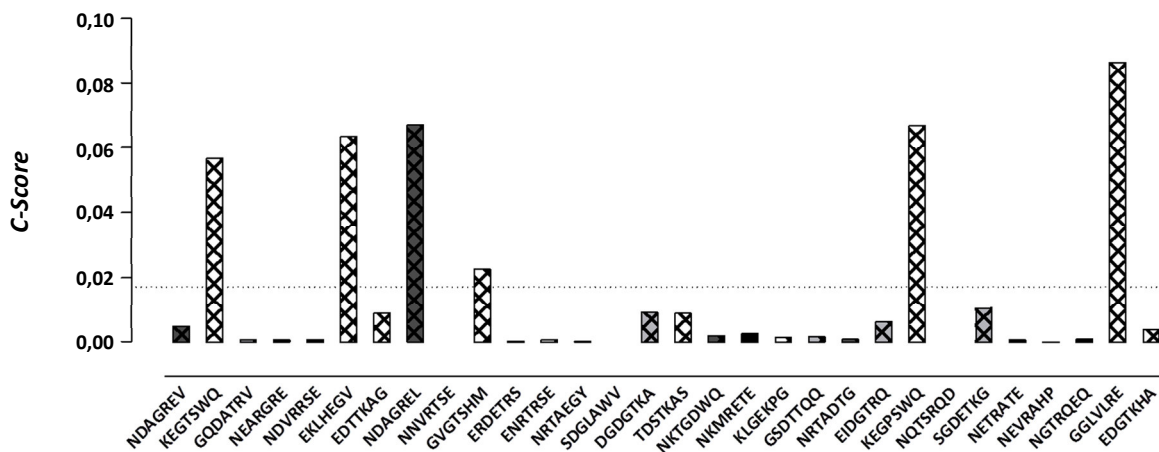
Niere



Niere

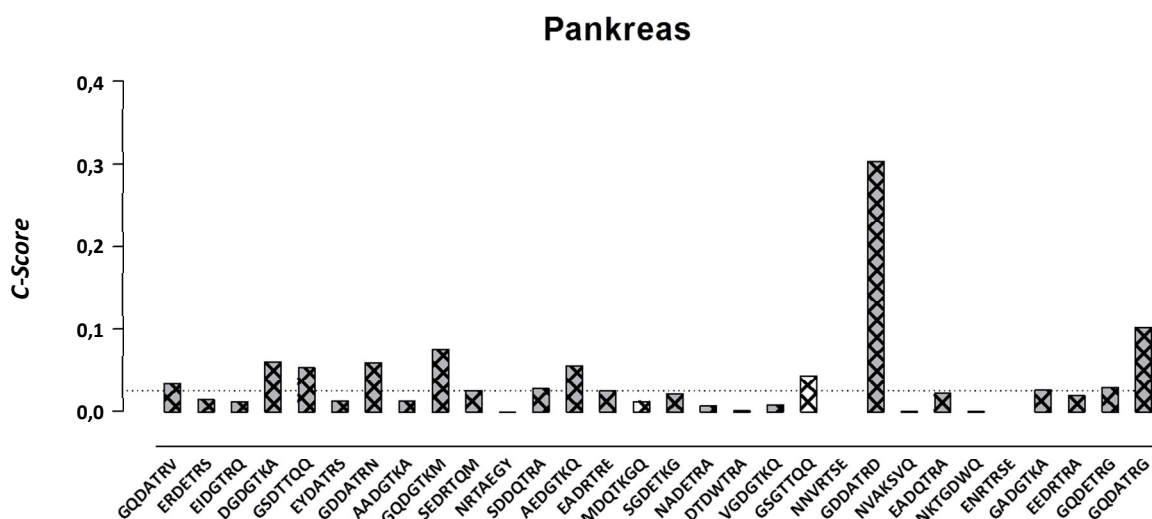
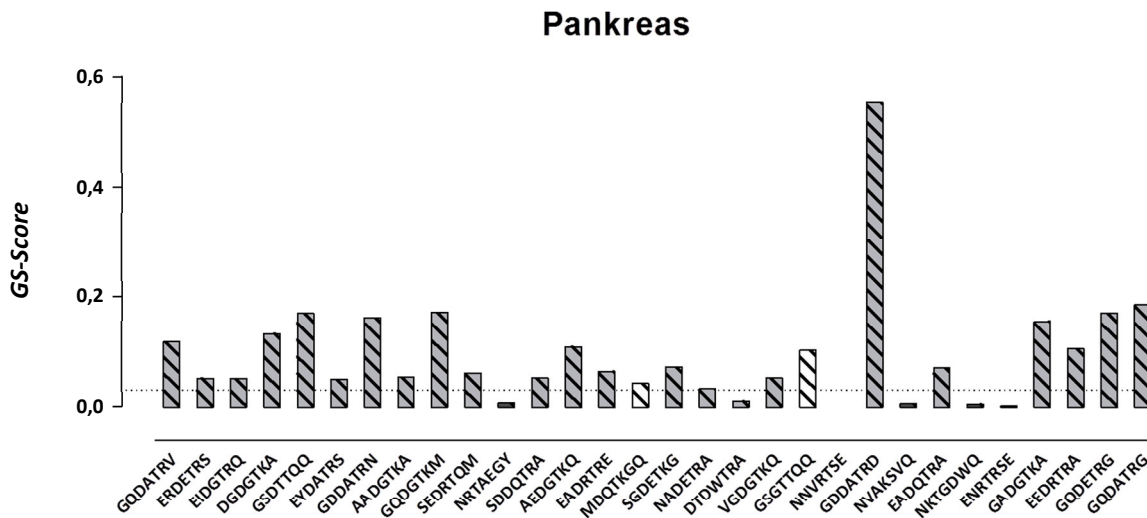
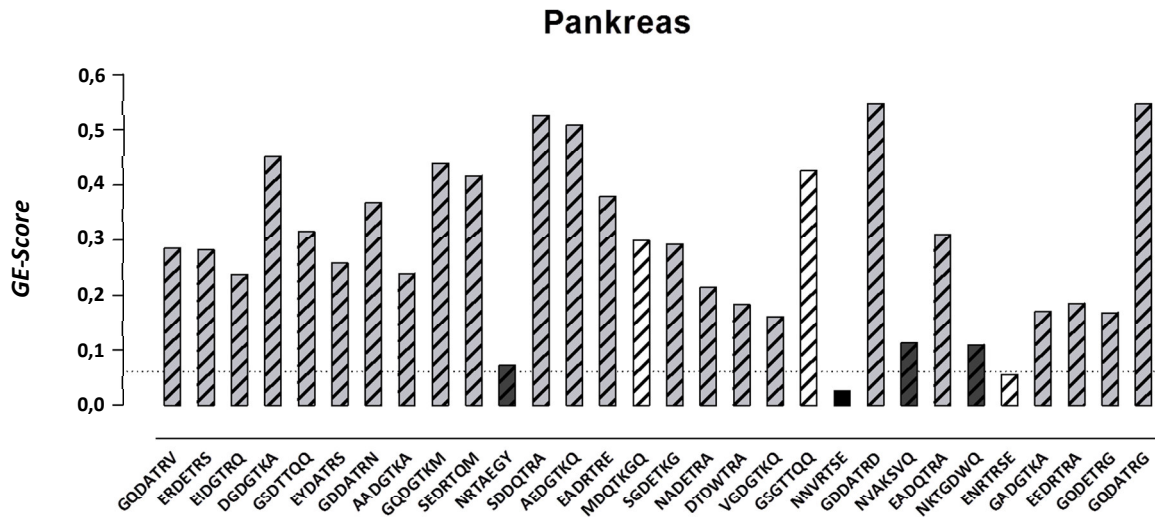


Niere



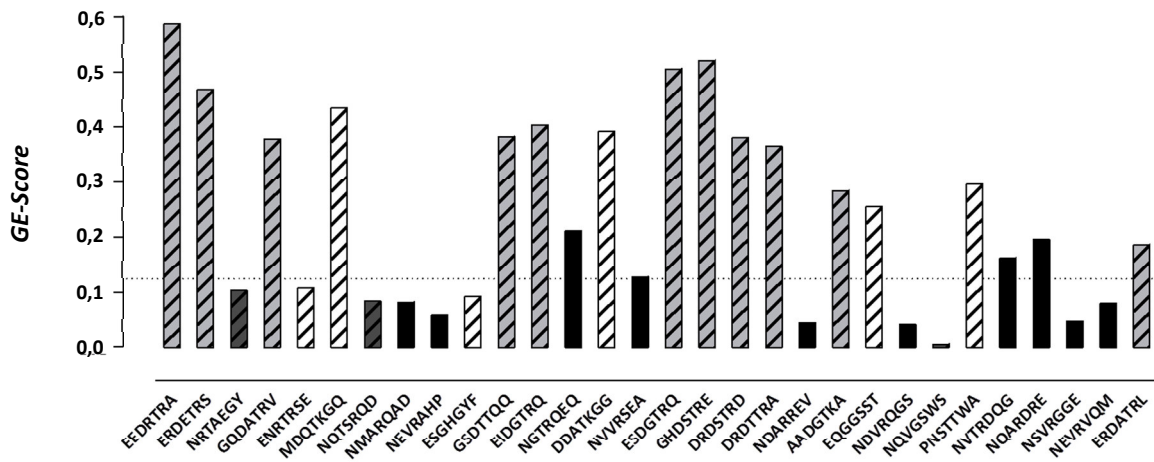
a



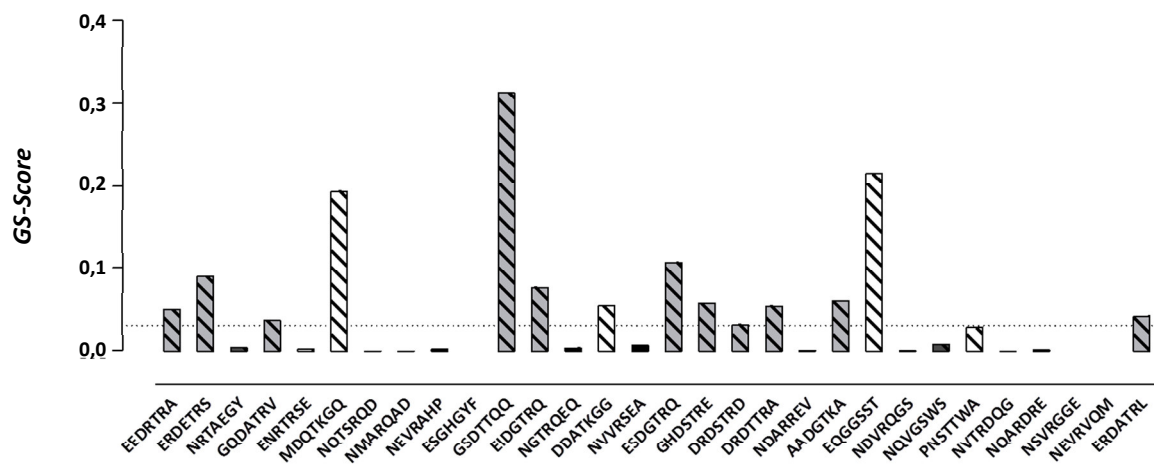


b

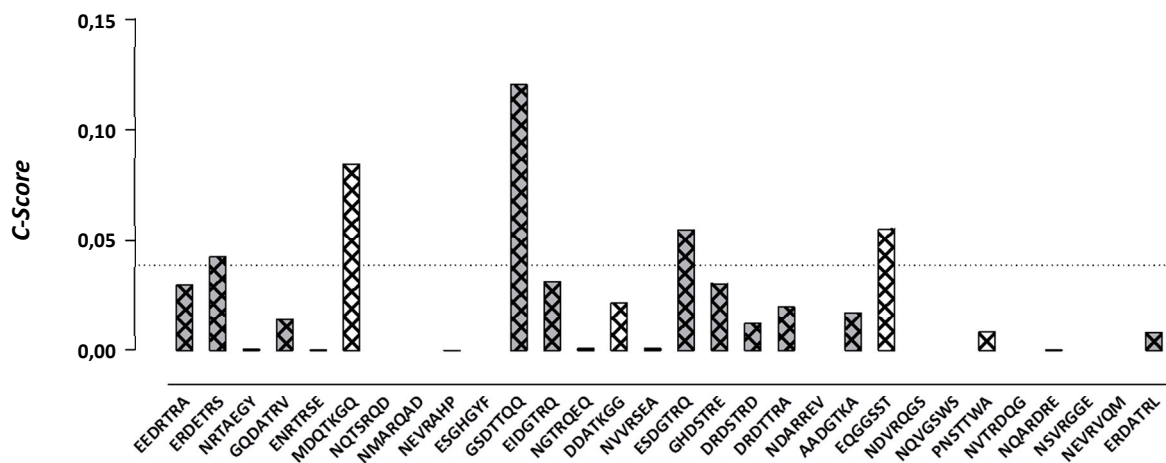
### Langerhans-Inseln



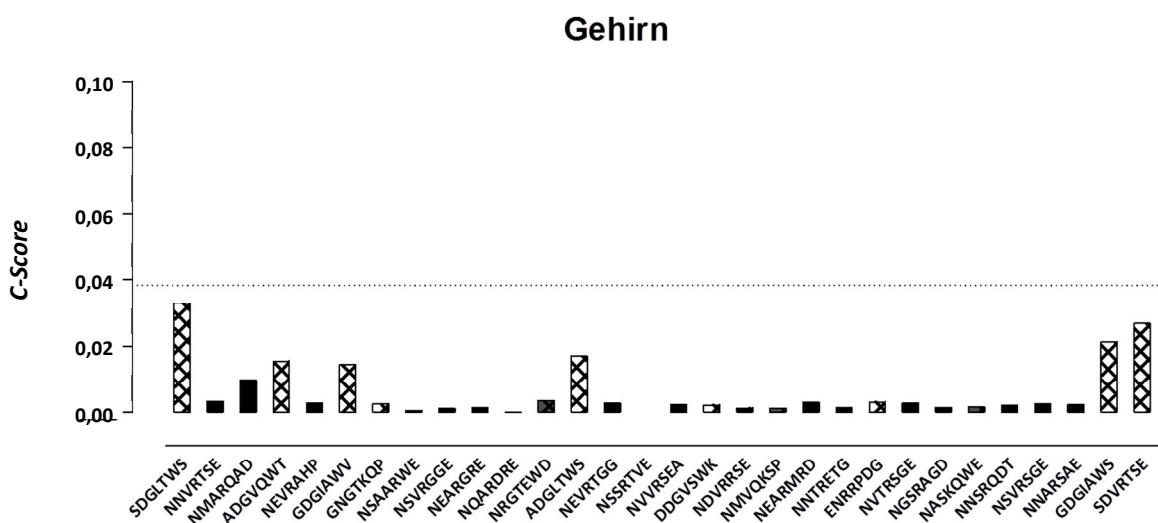
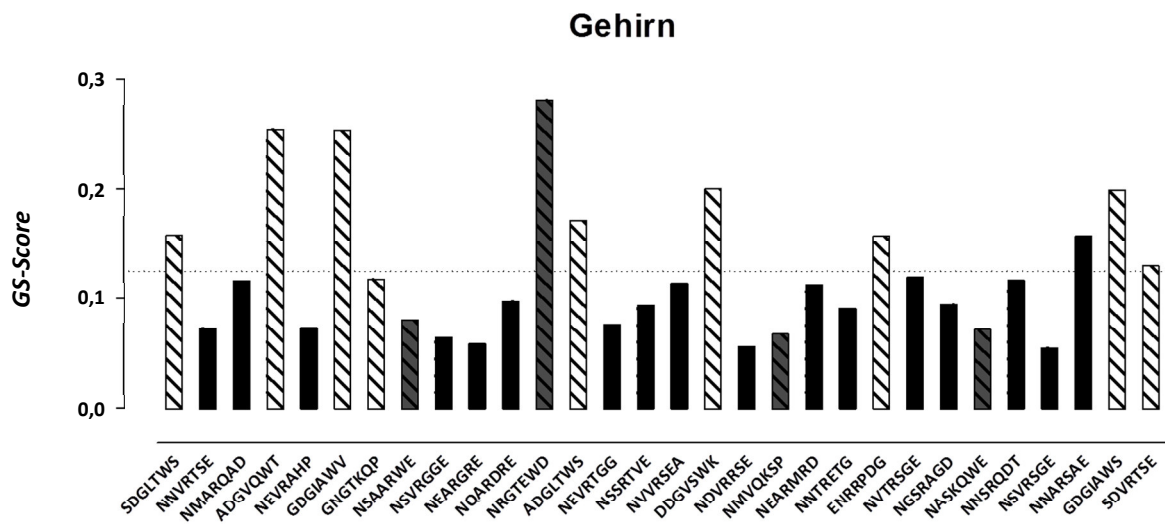
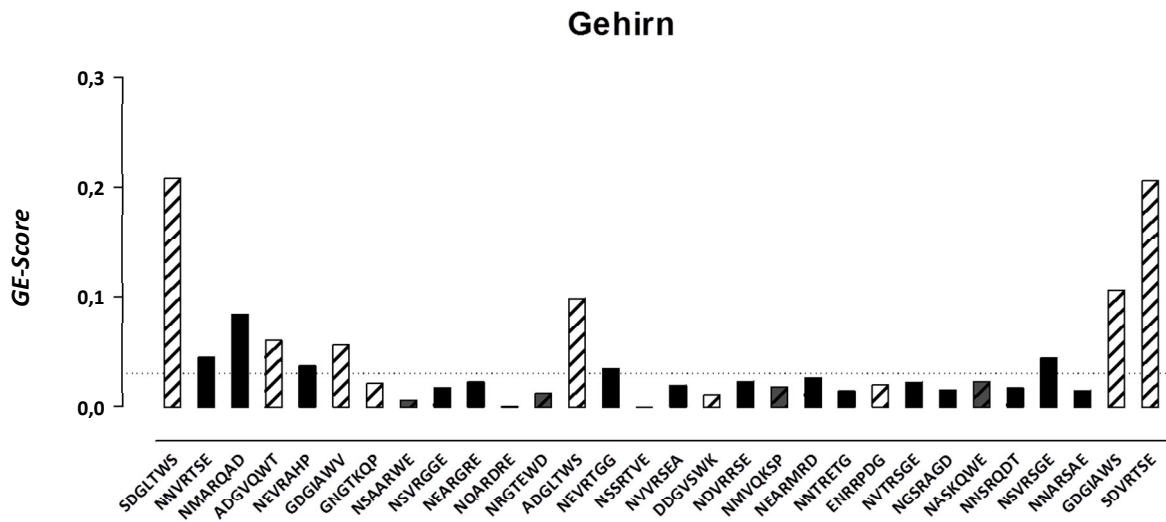
### Langerhans-Inseln



### Langerhans-Inseln



c



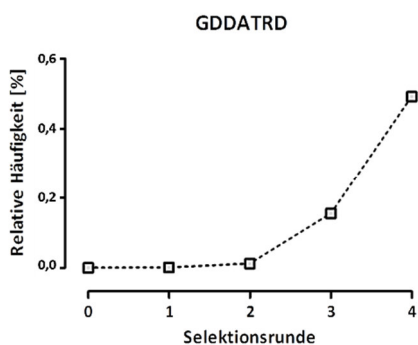
d

**Abbildung 16: Bewertung der Peptidanreicherung und Peptidspezifität mit Hilfe der GE-, GS- und C-Scores.** Für die Auswertung der Selektionen in **a)** der Niere, **b)** im Pankreas, **c)** den Langerhans-Inseln und **d)** dem Gehirn wurden die 30 häufigsten Peptidsequenzen der letzten Selektionsrunde jeder Selektion analysiert. Peptidsequenzen mit den Sequenzmotiven „NXXRXX<sup>E</sup>/X“ (angereichert in Niere und Gehirn, schwarze Balken) und „XXDXT<sup>R</sup>/X“ (angereichert in Pankreas, Langerhans-Inseln, hellgraue Balken) und Peptidsequenzen mit einem Asparagin am Anfang (dunkelgraue Balken) sind markiert. **GE-Score:** Die Orientierungslinie markiert den Wert, der erreicht wird, wenn Peptidsequenzen während der Selektion durchschnittlich eine gleiche relative Häufigkeit haben. Der *GE-Score* ist abhängig von der Anzahl der Selektionsrunden. Für die Selektionen in Niere und Gehirn wurden jeweils fünf, im Pankreas vier und in den Langerhans-Inseln drei Selektionsrunden durchgeführt. **GS-Score:** Die Orientierungslinie zeigt den Wert, an dem Peptidsequenzen im Ziel- und Kontrollgewebe durchschnittlich eine gleiche relative Häufigkeit haben. Der *GS-Score* ist abhängig von der Anzahl der Kontrollorgane. Für die Selektionen in Niere, Pankreas und Langerhans-Inseln wurden jeweils fünf Kontrollorgane untersucht, für die Selektion im Gehirn drei. **C-Score:** Ein Wert auf Höhe der Orientierungslinie bedeutet eine durchschnittliche Verdopplung der relativen Häufigkeit über die Selektion und eine Halbierung der relativen Häufigkeit dieser Peptidsequenz in den Kontrollorganen im Vergleich zum Zielorgan. Der *C-Score* ist abhängig von der Anzahl der Selektionsrunden und Kontrollorgane.

#### 4.2.2.5 Evaluierung eines exemplarischen selektierten AAV-Klons

Bei der Betrachtung der *Scores* der Selektion im Pankreas fiel eine Peptidsequenz besonders auf. Das Peptid „GDDATRD“ zeigte in allen drei Auswertungen den höchsten Wert, bezüglich der Spezifität sogar mit großem Abstand. Zusätzlich trägt es das selektierte Motiv „XXDXT<sup>R</sup>/X“. Aus diesen Gründen wurde die Peptidsequenz „GDDATRD“ und ein AAV-Vektor mit entsprechendem Targetingpeptid im Kapsid im Folgenden näher evaluiert.

In **Abbildung 17** ist der prozentuale Anteil der Peptidsequenz „GDDATRD“ an allen sequenzierten Peptidsequenzen dargestellt. Trotz des relativ geringen Anteils ist eine kontinuierliche Anreicherung dieser Sequenz über die Selektionsrunden festzustellen. Die stärkste Anreicherung fand von der vorletzten zur letzten Selektionsrunde statt.

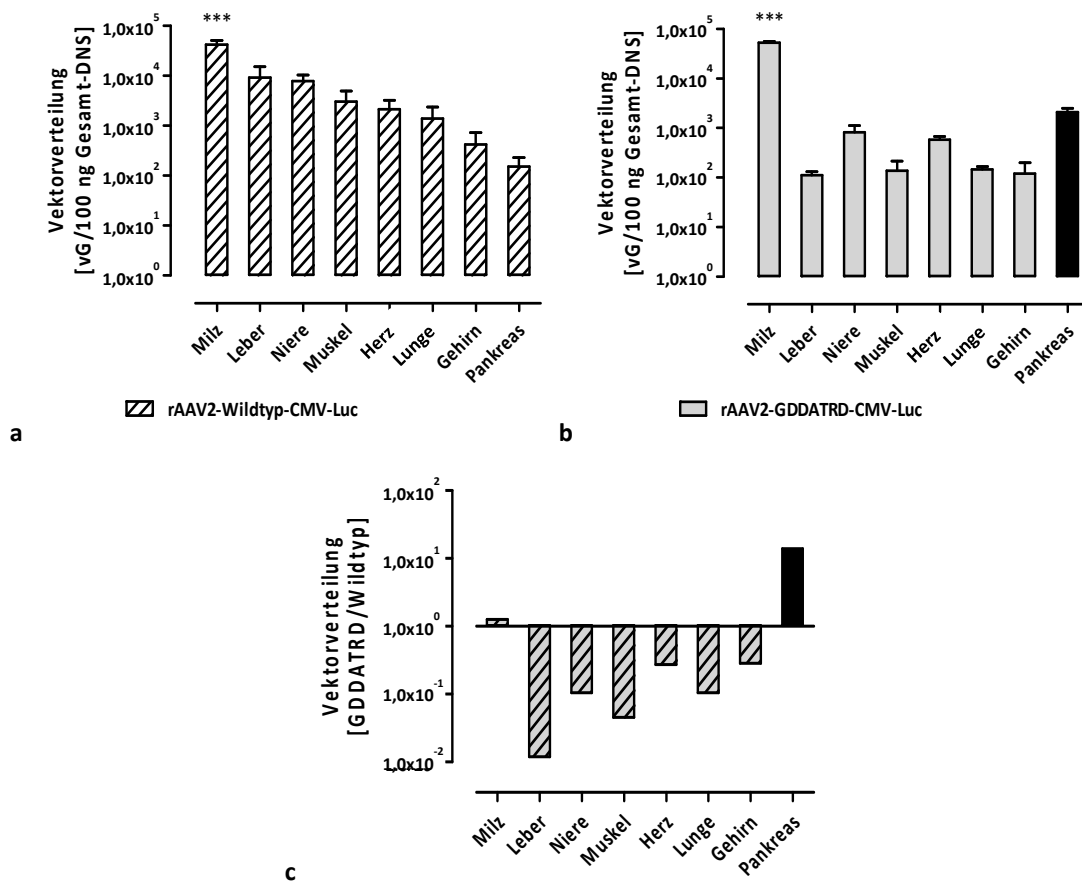


**Abbildung 17: Relative Anreicherung der Peptidsequenz „GDDATRD“ während der Selektion im Pankreas.** Der prozentuale Anteil der Peptidsequenz in jeder Selektionsrunde ist gezeigt.

Aufgrund der bisherigen Ergebnisse wurde ein rekombinanter AAV-Vektor, der die selektierte Peptidsequenz auf seiner Kapsidoberfläche repräsentiert, produziert und die Verteilung des Vektors auf DNS-Ebene in ausgewählten Organen der Maus per q-PCR bestimmt und zu der des Vektors mit Wildtyp-Kapsid verglichen. In **Abbildung 18a** ist die Vektorverteilung von AAV2-Vektoren mit Wildtypkapsid (also ohne Targetingpeptid) in verschiedenen Organen grafisch dargestellt. Die höchste Anzahl von viralen Genomen pro 100 ng Gesamt-DNS wurde in der Milz, der Leber und der Niere bestimmt; die niedrigste hingegen im Pankreas. Die Vektorverteilung des rAAV2-GDDATRD hingegen zeigt eine deutliche Verschiebung der Vektoren; weg von der Leber und hin zum Zielorgan

Pankreas (**Abbildung 18b**). In **Abbildung 18c** wurden beide Vektoren direkt miteinander verglichen. Hier wird die Verschiebung der Vektorverteilung noch deutlicher.

Zusammenfassend konnte mit Hilfe des NGS und der Auswertung der Daten mittels der beschriebenen Bewertungskriterien, zusätzlich zur Bewertung der Spezifität der identifizierten Sequenzmotive, eine AAV2-Kapsidvariante identifiziert werden, die eine zum Zielorgan gelenkte und somit im Vergleich zum Wildtypkapsid veränderte Ausrichtung zeigte.



**Abbildung 18: Vergleich der Vektorverteilungen des rAAV2-GDDATRD mit dem Kontrollvektor mit AAV2-Wildtypkapsid.**

Pro Maus wurden  $5 \times 10^{10}$  genomische Partikel i.v. in die Schwanzvene von jeweils drei FVB-Mäusen injiziert. 14 Tage nach der Injektion wurden die Mäuse getötet, Organe entnommen und Gesamt-DNS isoliert. Die Anzahl viraler Genome pro 100 ng Gesamt-DNS wurde mittels q-PCR bestimmt. Gezeigt sind die Mittelwerte mit positiver Standardabweichung. Die statistische Auswertung wurde mit einem einseitigen ANOVA Test und anschließender Bonferroni-Korrektur gemacht.  $p < 0,05 = *$ ;  $p < 0,01 = **$ ;  $p < 0,001 = ***$ ; (n=3). **a)** rAAV2-Wildtypkapsid, **b)** rAAV2-GDDATRD und **c)** Verhältnis von rAAV2-GDDATRD und rAAV2-Wildtypkapsid. Bezüglich der Vektorverteilung des rAAV2-GDDATRD ist der Balken des Zielorgans Pankreas schwarz hervorgehoben.

### 4.3 Analyse und Vergleich von AAV2-NNK<sub>7</sub>-Plasmidbibliothek und Virusbibliotheken

Im vorherigen Kapitel wurden neben der Verifizierung eines Sequenzmotivs zwei neue Motive identifiziert, wobei sich das Motiv „NXXRXX<sup>E</sup>/X“ als nicht-spezifisch herausstellte. Um den Ursprung dieses so dominanten Sequenzmotivs einschließlich eines möglichen Bias in der Ausgangsbibliothek zu ergründen, wurden im Folgenden die zugrundeliegende primäre Plasmid- und Virusbibliothek, sowie zum Vergleich zwei weitere auf dieser Plasmidbibliothek basierenden Virusbibliothek mittels NGS analysiert.

#### 4.3.1 NGS-Daten und Vergleich der AAV2-NNK<sub>7</sub>-Plasmidbibliothek und den daraus generierten Virusbibliotheken hinsichtlich der Peptiddiversität

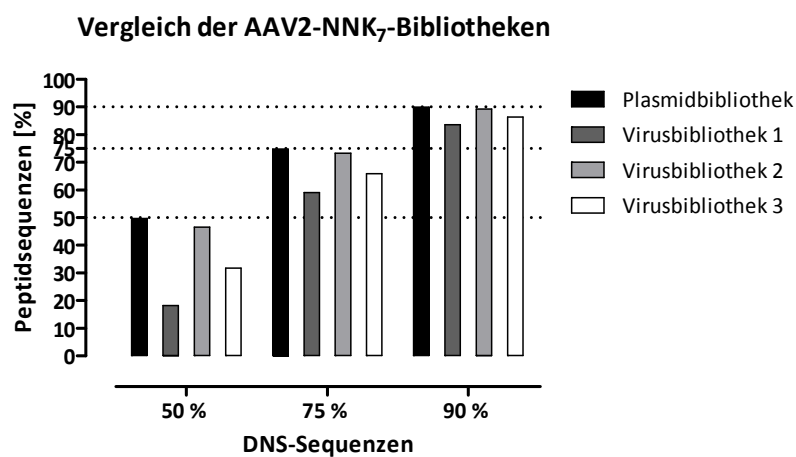
Zur Untersuchung der Diversität der primären Plasmidbibliothek und den drei daraus generierten Virusbibliotheken wurden diese mittels NGS analysiert. **Tabelle 23** zeigt die Anzahl der DNS- und Peptidsequenzen, sowie ihr Verhältnis zueinander. Anhand des DNS/Peptid-Verhältnis kann abgeschätzt werden, wie divers die Bibliotheken sind. Je näher das Verhältnis dem Wert 1 ist, desto diverser ist die Bibliothek, da bei diesem Wert jede gelesene DNS-Sequenz verschieden ist und somit auch ein unterschiedliches Peptid codiert. Hiernach entsprechen die Plasmidbibliothek und die Virusbibliothek 2 nahezu dem Idealwert. Virusbibliothek 1 und 3 hingegen sind mit einem Wert von 1,64 und 1,36 weniger divers.

**Tabelle 23: Zusammenfassung der NGS-Daten der AAV2-NNK<sub>7</sub>-Plasmid- und Virusbibliotheken.**

Bibliothek	Anzahl der DNS-Sequenzen	Anzahl der Peptidsequenzen	Verhältnis DNS- zu Peptidsequenzen
Plasmidbibliothek	98852	98182	1,01
Virusbibliothek 1	225087	137507	1,64
Virusbibliothek 2	75552	70637	1,07
Virusbibliothek 3	550117	403288	1,36

Die Peptidsequenzen und die zugehörigen DNS-Sequenzen wurden ihrer Häufigkeit nach geordnet und anschließend die häufigsten 50 %, 75 % und 90 % der DNS-Sequenzen im Hinblick darauf analysiert, ob sie dem gleichen prozentualen Anteil an Peptidsequenzen entsprechen (**Abbildung 19**). Während 50 % der DNS-Sequenzen der Plasmidbibliothek auch ca. 50 % der Peptidsequenzen entsprechen, decken 50 % der DNS-Sequenzen der Virusbibliothek 1 weniger als 20 % der Peptidsequenzen ab. Dieses entspricht einem theoretischem durchschnittlichem DNS/Peptid-Verhältnis von 2,5, was bedeutet, dass eine Reduktion der ursprünglichen klonalen Diversität

stattgefunden hat. In der Virusbibliothek 2 entsprechen 50 % der DNS-Sequenzen über 45 % der Peptidsequenzen, in der Virusbibliothek 3 sind es etwas über 30 %. Die Betrachtung von 75 % der DNS-Sequenzen zeigt immer noch einen deutlichen Unterschied zwischen den Bibliotheken. Bei Betrachtung von 90 % der DNS-Sequenzen sieht man eine Relativierung der Unterschiede, 90 % der DNS-Sequenzen machen in allen Bibliotheken über 85 % der Peptide aus. Anhand dieser Darstellung wird deutlich, dass im Vergleich zu der Plasmidbibliothek besonders im Falle der Virusbibliotheken 1 und 3 eine Anreicherung von Peptidsequenzen während der Virusproduktion der primär zur Selektion eingesetzten Virusbibliothek stattgefunden hat.



**Abbildung 19: Vergleich der AAV2-NNK<sub>7</sub>-Plasmidbibliothek und daraus generierten Virusbibliotheken bezüglich der Anreicherung von Peptidsequenzen.** Analysiert wurden 50 %, 75 % und 90 % der häufigsten DNS-Sequenzen im Hinblick darauf, wie viel Prozent der tatsächlich vorhandenen unterschiedlichen Peptidsequenzen durch sie repräsentiert werden. Die Orientierungslinien bei 50 %, 75 % und 90 % der Peptidsequenzen stellen den Idealwert (DNS/Peptid- Verhältnis von 1) dar.

#### 4.3.2 Vergleich und Sequenzmotivanalyse der zehn häufigsten Peptidsequenzen

Zur weiteren Analyse der Bibliotheken wurden die zehn häufigsten Peptidsequenzen der AAV2-NNK<sub>7</sub>-Plasmidbibliothek und der daraus generierten Virusbibliotheken miteinander verglichen (**Tabelle 24**). Hierbei zeigte sich eine geringe relative Häufigkeit (<0,0071 %) der häufigsten Peptidsequenzen der Plasmidbibliothek, was die hohe Diversität, die bereits durch das zuvor untersuchte DNS/Peptid-Verhältnis gezeigt wurde, bestätigte. Sieben von zehn Sequenzen der Plasmidbibliothek zeigten eine starke Anreicherung der Aminosäure Glycin. Die häufigste Peptidsequenz bestand sogar nur aus Glycinen.

Der Vergleich der Peptidsequenzen aller drei Virusbibliotheken zeigte eine sehr hohe Ähnlichkeit der Peptidsequenzen. Hierbei unterschieden sich die Sequenzen der Virusbibliothek 1 am meisten, während neun von zehn Sequenzen der Virusbibliotheken 2 und 3 übereinstimmten. Die

Virusbibliothek 1 wurde für die in dieser Arbeit analysierten *in vivo*-Selektionen verwendet. Bei näherer Betrachtung der Virusbibliothek 1 fiel auf, dass alle Peptidsequenzen, die Asparagin in der ersten Aminosäureposition besaßen, auch unter den häufigsten zehn Peptidsequenzen der Selektionen zu finden waren (**Tabelle 21 und Tabelle 24**) und somit nicht herausselektiert wurden. Zusätzlich machte die Sequenz „NNVRTSE“ 1,34 % der sequenzierten Peptidsequenzen der unselektierten Virusbibliothek 1 aus und wurde schließlich als zweithäufigste Sequenz der Selektion im Gehirn identifiziert (**Tabelle 21**).

**Tabelle 24** zeigt des Weiteren, dass alle Virusbibliotheken das Sequenzmotiv „NXXRXXX“ aufwiesen, während dieses Motiv nicht in der Plasmidbibliothek zu finden war. Folglich wurde nicht nur die Selektion bestimmter Peptidsequenzen durch die bereits in der Ausgangsbibliothek vorhandene Anreicherung begünstigt, auch das unspezifische Sequenzmotiv „NXXRXXX“ hat seinen Ursprung in der nicht-selektierten Virusbibliothek.

**Tabelle 24: Peptidsequenzanalyse und Vergleich der AAV2-NNK7-Plasmid- und Virusbibliotheken.** Dargestellt sind die zehn häufigsten Peptidsequenzen jeder Bibliothek und ihre relativen Häufigkeiten. Für jede Position ist die häufigste Aminosäure, die mindestens dreimal zu finden war, farblich markiert. Wenn mehrere Aminosäuren diese Bedingung erfüllten, wurden alle farblich markiert. Zusätzlich wurden die jeweiligen Sequenzmotive angegeben.

Position	Plasmidbibliothek [%]	Virusbibliothek 1 [%]	Virusbibliothek 2 [%]	Virusbibliothek 3 [%]
1	GGGGGGG 0,00708	NNVRTSE 1,34481	NGTRPWE 0,04632	NAVRETS 0,08252
2	NKTATWE 0,00708	NGVKKEG 0,45493	NAARRDE 0,04500	NGVKWDG 0,05853
3	GGGGGEA 0,00506	NDARREV 0,43183	NAVRETS 0,03970	NEARQTG 0,05744
4	EEDRTRA 0,00405	NRTAEGY 0,42872	NDARRDV 0,03838	NGTRPWE 0,05453
5	GKGGGRR 0,00405	GGSGKPG 0,28255	RDGERPG 0,02779	NETRTVQ 0,04944
6	GGGGGGK 0,00405	NEVRAHP 0,21591	NGVKWDG 0,02779	NDARRDV 0,04162
7	GGGGGGV 0,00405	GEWGRAL 0,21191	NETRTVQ 0,02514	NEIRRDG 0,03853
8	EREGRRR 0,00303	GKERAP 0,17282	NEARQTG 0,02382	NGVRGYE 0,03581
9	GGGGGEE 0,00303	NOTSRQD 0,16837	NEIRRDG 0,02382	NAARRDE 0,03453
10	GGGGGEG 0,00303	GERGRME 0,12839	NGVRGYE 0,02382	NDVREPR 0,03217
<b>Motiv</b>	GGGGGGX	NXVXXXX/ NXXRXXX	NXXRXXX	NXXRXXX

#### 4.3.3 Häufigkeit der 20 Aminosäuren in den randomisierten Positionen der unselektierten NNK<sub>7</sub>-Bibliotheken

Bei der bisherigen Analyse der zehn häufigsten Peptidsequenzen wurden Aminosäuren im Kontext ihrer benachbarten Aminosäuren dargestellt und somit die Häufigkeit bestimmter Peptidsequenzen untersucht. Um einen Überblick über die Häufigkeit von bestimmten Aminosäuren an den einzelnen randomisierten Positionen aller sequenzierten Peptidsequenzen der unselektierten Bibliotheken zu erhalten, wurden Sequenzlogos mit der online Software WebLogo generiert. Hierbei wurden jede sequenzierte Peptidsequenz einmal und jede Aminosäureposition der untersuchten Sequenzen



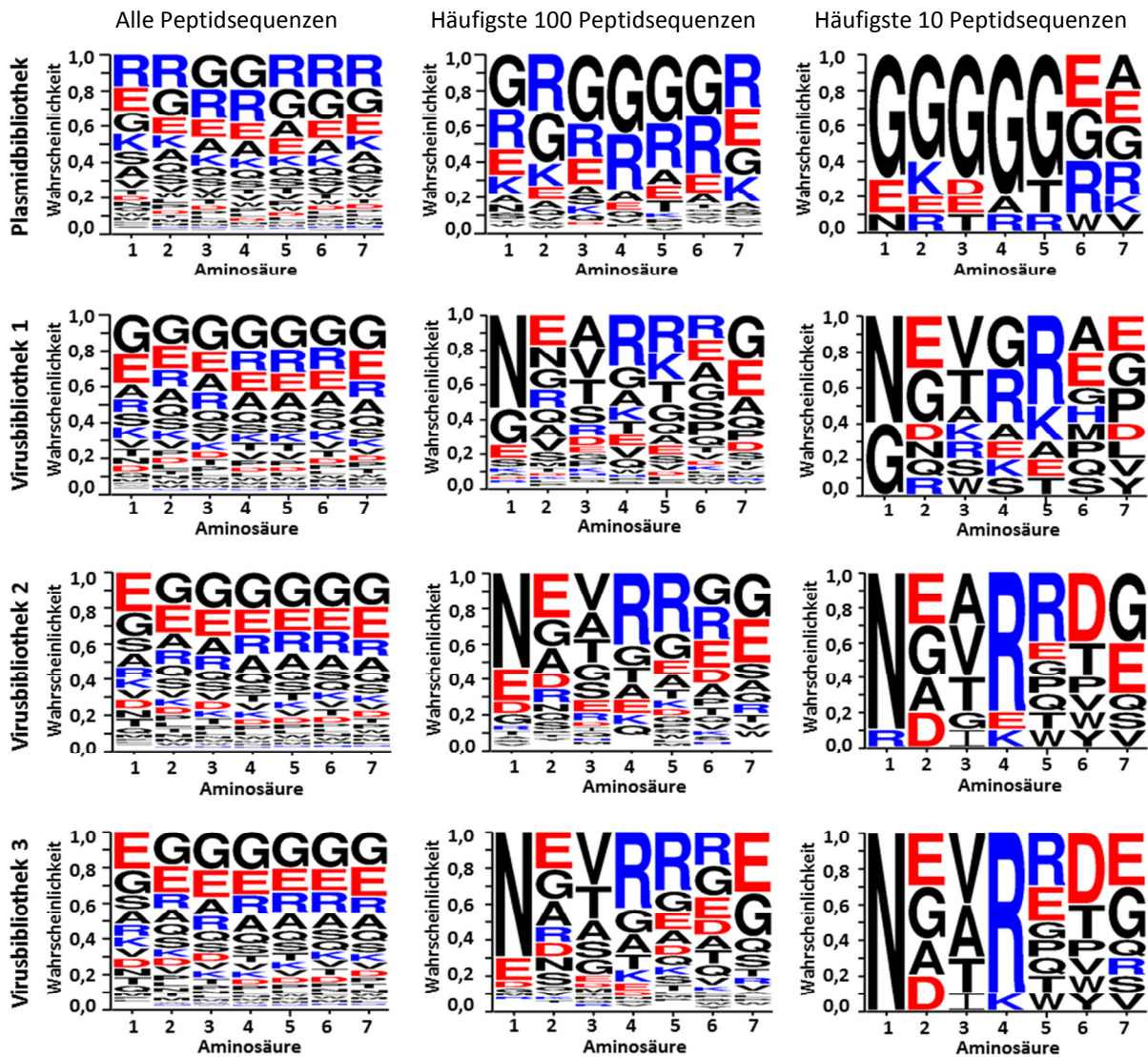
unabhängig von der benachbarten Aminosäure erfasst und die relativen Häufigkeiten der häufigsten Aminosäuren für jede Position dargestellt. **Abbildung 20** zeigt die Sequenzlogos jeweils für alle, die häufigsten 100 und die häufigsten zehn Peptidsequenzen der unselektierten Plasmidbibliothek und der drei daraus generierten Virusbibliotheken.

In der Plasmidbibliothek wurde eine starke Häufung der Aminosäuren Glycin, Arginin und Glutaminsäure beobachtet (in etwa 15-20 % aller Peptide an jeder der sieben Positionen), wobei das Glycin unter den häufigsten 100 und zehn Peptidsequenzen besonders stark überrepräsentiert war (>50 % an Position 1-5 unter den häufigsten zehn Peptiden).

Wie die Plasmidbibliothek zeigten auch die drei daraus generierten Virusbibliotheken eine Häufung der Aminosäuren Glycin, Arginin und Glutaminsäure. Beim Vergleich der 100 häufigsten Peptidsequenzen allerdings waren große Unterschiede zwischen der Plasmidbibliothek und den Virusbibliotheken zu erkennen. Während die Aminosäureverteilung der 100 häufigsten Peptidsequenzen der Plasmidbibliothek der von allen Sequenzen ähnelte, zeigten die Aminosäureverteilungen der 100 häufigsten Sequenzen der Virusbibliotheken große Unterschiede zu den Aminosäureverteilungen aller Peptidsequenzen. Das Asparagin in der ersten Position, war unter den häufigsten 100 Sequenzen der Virusbibliothek in bis zu 70 % der Sequenzen zu finden, während es bei Betrachtung aller Sequenzen mit einer relativen Häufigkeit von unter 10 % kaum auffiel. Zusätzlich waren die Arginine in den Positionen vier und fünf besonders auffallend und in bis zu 40 % der 100 häufigsten Sequenzen zu beobachten. Eine Anreicherung dieser Aminosäuren unter den häufigsten Sequenzen und an bestimmter Position war somit deutlich.

Die hiesigen Ergebnisse bestätigen folglich die Anreicherung des im Rahmen der *in vivo*-Selektionen identifizierten Sequenzmotivs „NXXRXX<sup>X</sup>/<sub>E</sub>“ bereits in den primären Virusbibliotheken, während das Motiv „XXDXT<sup>X</sup>/<sub>R</sub>X“ in diesen nicht angereichert war.

Zusätzlich wird neben dem großen Unterschied zwischen Plasmid- und Virusbibliotheken eine gewisse Ungleichverteilung bestimmter Aminosäuren bereits in der Plasmidbibliothek deutlich.

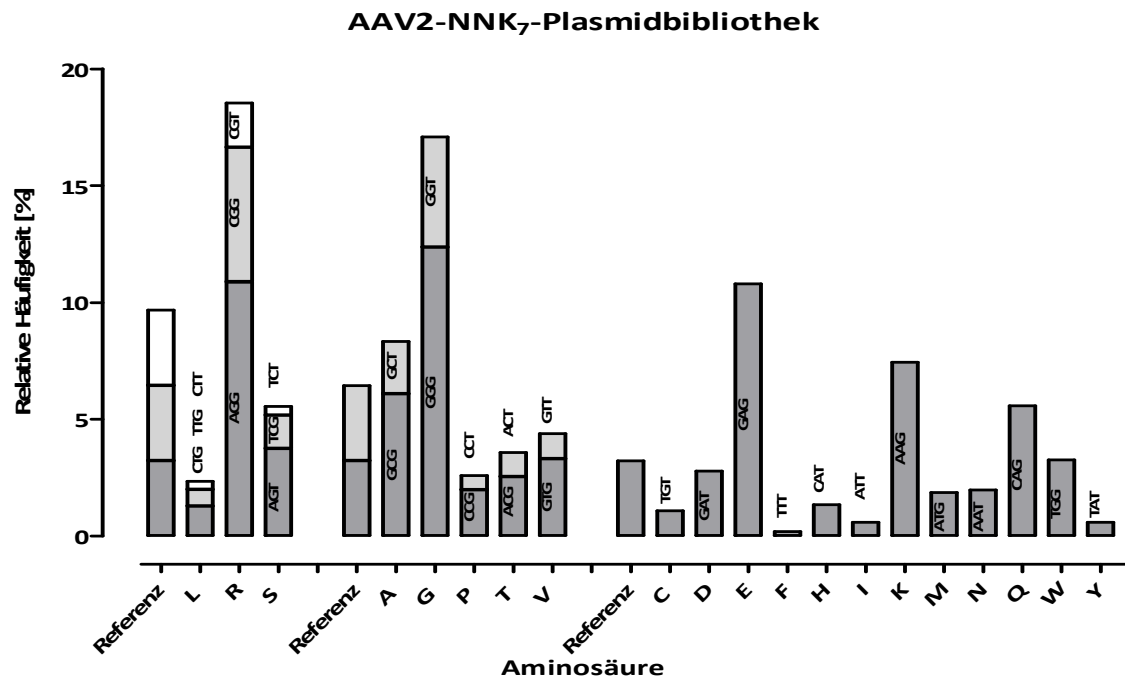


**Abbildung 20:** Häufigkeit der 20 Aminosäuren an den randomisierten Positionen der unselektierten Plasmidbibliothek im Vergleich zu den drei Virusbibliotheken. Für jede Aminosäureposition wurden die häufigsten Aminosäuren ermittelt und deren Häufigkeit grafisch dargestellt. Anhand der Größe der Aminosäuren sind die jeweiligen relativen Häufigkeiten ablesbar. Blau: positiv geladene Aminosäuren (R, H, K); Rot: negativ geladene Aminosäuren (E, D); Schwarz: neutrale Aminosäuren (A, C, F, G, I, L, M, N, P, Q, S, T, V, W). Die Darstellung wurde online generiert: <http://weblogo.threeplusone.com/create.cgi> (Beschriftung der Skalen modifiziert)

#### 4.3.4 Erwartete und tatsächliche Aminosäurehäufigkeit der NNK<sub>7</sub>-Bibliotheken

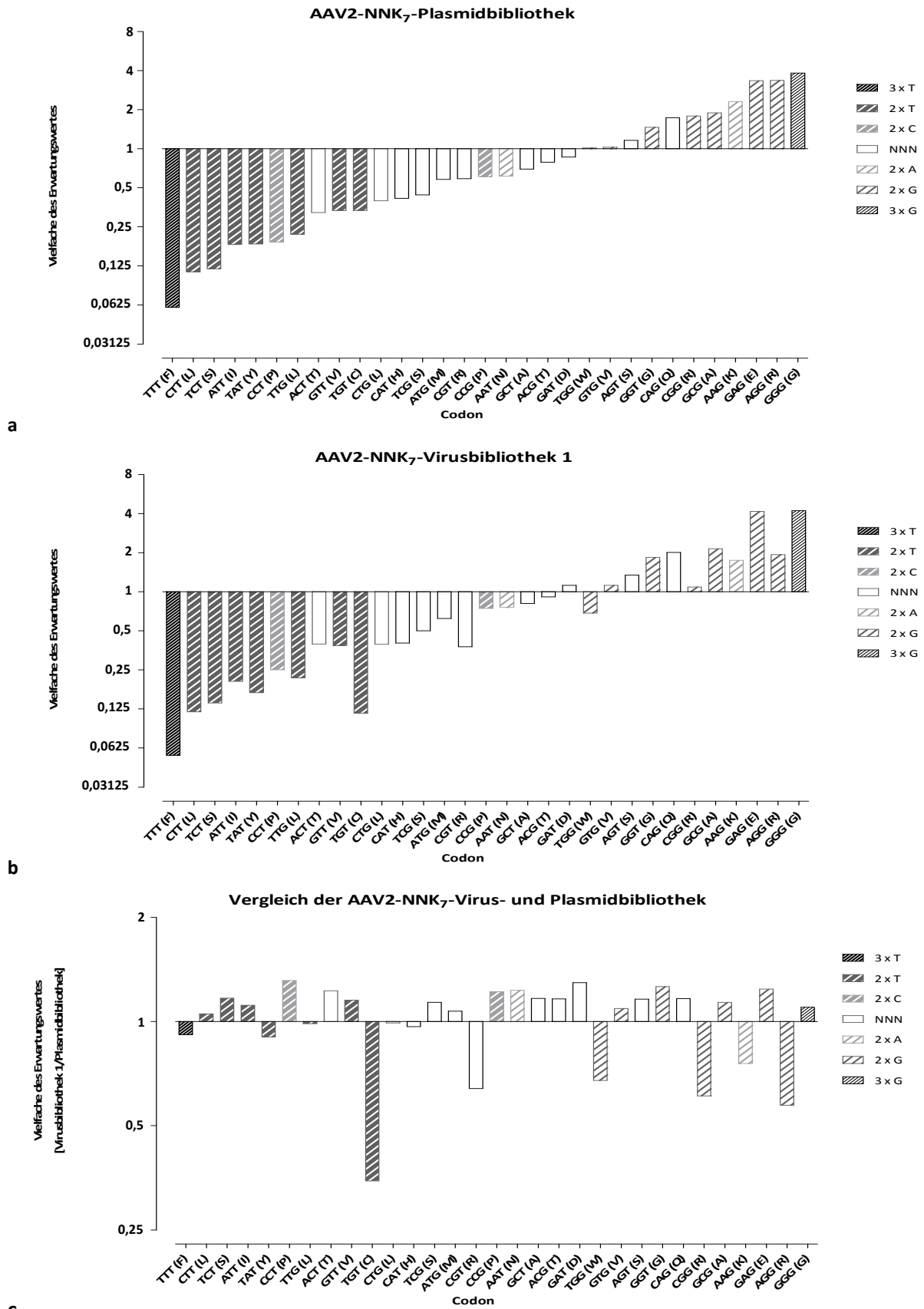
Aufgrund des zur Produktion der Plasmidbibliothek verwendeten NNK-Codierungsschemas sind manche Aminosäuren häufiger präsent als andere. In **Abbildung 21** sind die tatsächliche Verteilung der Aminosäuren der Peptidsequenzen der Plasmidbibliothek und deren theoretische, durch das NNK-Codierungsschema bedingte Referenzwerte dargestellt. Die Aminosäuren Arginin, Glycin, Glutaminsäure, Lysin und Glutamin weisen im Vergleich zu ihrem Referenzwert einen wesentlich höheren prozentualen Anteil auf. Gleichzeitig sind die Anteile der Aminosäuren Leucin, Prolin, Cystein, Phenylalanin, Isoleucin und Tyrosin wesentlich niedriger als unter Einbeziehung des Codierungsschemas erwartet. Dies scheint durch eine starke Ungleichverteilung der Nukleotidcodons

bedingt zu sein. Am Beispiel der Aminosäure Arginin, welche im NNK-Schema durch drei Codons codiert ist (AGG, CGG, CGT), wird diese besonders deutlich. Eines der Codons (AGG) nimmt einen höheren prozentualen Anteil ein als der theoretische Referenzwert für alle drei Codons. Das Gleiche gilt für die überrepräsentierte Aminosäure Glycin und das Codon „GGG“.



**Abbildung 21:** Verteilung der Aminosäuren der Peptidsequenzen der Plasmidbibliothek im Vergleich zum NNK-Codierungsschema begründeten Referenzwert. Der prozentuale Anteil jeder Aminosäure sowie ihrer Codons sind im Vergleich zu ihrem theoretischen Referenzwert grafisch dargestellt.

Die Wahrscheinlichkeit für jedes Codon beträgt  $1/31$  (Stoppcodons wurden im Vorwege von der Analyse ausgeschlossen). Für die Berechnung der Werte in **Abbildung 22** wurde dieser Wert gleich 1 gesetzt. Anschließend wurden die jeweiligen Vielfachen von diesem Erwartungswert für jedes Codon von dem höchsten negativen bis zum höchsten positiven Vielfachen sortiert. Die Ergebnisse zeigen eine klare systematische Präferenz für Guanin- und Adenin-haltige Codons. Cytosin- und Thyminhaltige Codons hingegen sind unterrepräsentiert. Hierbei liegen die Werte zwischen 16-mal unter und viermal über dem Erwartungswert. Anhand dieser Ergebnisse lassen sich die hohen prozentualen Anteile der Aminosäuren Arginin (AGG), Glycin (GGG, GGT) und Glutaminsäure (GAG) aus **Abbildung 21** erklären. Folglich sind die Häufigkeiten der Aminosäuren der Plasmidbibliothek zum einen aufgrund des NNK-Codierungsschemas und zum anderen durch eine systematische Ungleichverteilung der Codons begründet. Des Weiteren wurde durch den direkten Vergleich von Plasmid- und Virusbibliothek eine deutliche Abnahme der Aminosäuren Cystein, Tryptophan, Lysin und Arginin beobachtet (**Abbildung 22c**). Für den hiesigen Vergleich wurde beispielhaft die Virusbibliothek 1 verwendet. Die gleichen Tendenzen sind auch für die Virusbibliothek 2 und 3 beobachtet worden (**Abbildung 29** und **Abbildung 30** im Anhang).



**Abbildung 22: Codonhäufigkeit und Abweichungen vom Erwartungswert (AAV2-NNK<sub>7</sub>-Bibliotheken).** a) AAV2-NNK<sub>7</sub>-Plasmidbibliothek und b) AAV2-NNK<sub>7</sub>-Virusbibliothek 1: Die Vielfachen des Erwartungswertes wurden für jedes Codon bestimmt und von der höchsten negativen Abweichung zur höchsten positiven Abweichung sortiert. c) Vergleich beider Bibliotheken. Durch die Berechnung der Verhältnisse der Vielfachen des Erwartungswertes von Virusbibliothek und Plasmidbibliothek wird die Veränderung durch die Produktion der Virusbibliothek deutlich.

## 4.4 Einführung der neuen Peptidbibliothek AAV2-Trimer<sub>6</sub>-(-C)

### 4.4.1 AAV2-Trimer<sub>6</sub>-(-C) im Vergleich zur AAV2-NNK<sub>7</sub>-Peptidbibliothek

Die oben geschilderten Ergebnisse haben gezeigt, dass aufgrund des NNK-Codierungsschemas und der Ungleichverteilung der Nukleobasen die Codons für manche Aminosäuren bereits in der Plasmidbibliothek überrepräsentiert waren. Des Weiteren enthielten 9,4 % der Sequenzen der bisher verwendeten Plasmidbibliothek nichtfunktionelle Peptide generierende Stoppcodons (Daten nicht gezeigt) und zusätzlich Codons für die Aminosäure Cystein, die bereits während der Viruspartikelproduktion der Peptidbibliothek und auch anschließend während der Selektion negativ selektiert, also abgereichert wurde (generelle Beobachtungen). Es ist daher anzunehmen, dass das Cystein in dieser Position des Kapsids einen strukturellen Nachteil bietet. Zur Verbesserung der Plasmidbibliothek und somit zur Verbesserung der Startbedingungen einer Selektion wurde in unserem Labor (unter Federführung von Dr. Timo Sieber) die AAV2-Trimer<sub>6</sub>-(-C)-, abgekürzt 2T<sub>6</sub>-, Peptidbibliothek generiert. Diese nun erstmals verfügbare Peptidbankvariante wird im Folgenden vor dem Einsatz in Selektionen charakterisiert.

Die bisher verwendete NNK-codierte Peptidbibliothek besteht aus Peptiden mit einer Peptidlänge von sieben Aminosäuren und hat eine theoretische maximale Diversität von  $1,8 \times 10^9$ . Aufgrund der Anzahl von Plasmidklonen von  $2 \times 10^8$  und einer gewissen Redundanz der Peptidvarianten aufgrund des NNK-Schemas wird die mögliche Peptiddiversität begrenzt und liegt bei  $1,3 \times 10^8$ ; die erwartete Abdeckung aller möglichen Peptidsequenzen liegt bei 10,3 % (**Tabelle 25**).

Für die neue 2T<sub>6</sub>-Bibliothek wurden anstatt von 32 NNK-codierten Codons 19 definierte Trimere verwendet und die Peptidlänge wurde auf sechs Aminosäuren gekürzt. Die daraus resultierende maximale Peptiddiversität liegt mit  $4,7 \times 10^7$  unter der Anzahl der Plasmidklone von  $3,6 \times 10^8$  und ist somit nicht von der Anzahl der Plasmidklone begrenzt. Die Peptiddiversität der 2T<sub>6</sub>-Bibliothek ist damit ca. dreimal geringer als die der NNK<sub>7</sub>-Bibliothek. Die mögliche Abdeckung liegt dafür aber bei nahezu 100 % (**Tabelle 25**).

**Tabelle 25: Vergleich der verschieden codierten Peptidbibliotheken AAV2-Trimer<sub>6</sub>-(-C) und AAV2-NNK<sub>7</sub>.**

	AAV2-Trimer <sub>6</sub> -(-C) ( $3,6 \times 10^8$ Plasmidklone)	AAV2-NNK <sub>7</sub> ( $2 \times 10^8$ Plasmidklone)
<b>Codierungsschema</b>	20/20 (definierte Codons)	N = A, T, G, C; K = T/G
<b>Anzahl der Aminosäuren</b>	19	20
<b>Anzahl der Codons</b>	19	32
<b>Peptidlänge</b>	6 As	7 As
<b>Stoppcodons</b>	-	1 (UAG)
<b>Cysteine</b>	-	Ja
<b>Maximale Peptiddiversität</b>	$4,7 \times 10^7$	$1,8 \times 10^9$
<b>Peptiddiversität*</b>	$4,7024 \times 10^7 \pm 1,4919 \times 10^2$	$1,3139 \times 10^8 \pm 6,5451 \times 10^3$
<b>Erwartete Abdeckung*</b>	$0,9995 \pm 3,1711 \times 10^{-6}$	$0,1026 \pm 5,1133 \times 10^{-6}$

\* Berechnet mit PeLiCa ([www.pelica.org](http://www.pelica.org))

#### 4.4.2 Vergleich der Peptiddiversität der AAV2-Trimer<sub>6</sub>-(-C)-Plasmidbibliothek und der daraus generierten Virusbibliothek

Zur Untersuchung der Diversität der neuen 2T<sub>6</sub>-Bibliothek wurden die Plasmid- und die daraus generierte Virusbibliothek mittels NGS sequenziert. Das Verhältnis von DNS-zu Peptidsequenzen war nahezu 1, wobei das der Virusbibliothek wie auch schon bei den NNK<sub>7</sub> Virusbibliotheken etwas höher war (Tabelle 26).

Tabelle 26: Zusammenfassung der NGS-Daten der AAV2-Trimer<sub>6</sub>-(-C) Plasmid- und Virusbibliothek.

Bibliothek	Anzahl der DNS-Sequenzen	Anzahl der Peptidsequenzen	Verhältnis DNS- zu Peptidsequenzen
Plasmidbibliothek	133712	132855	1,01
Virusbibliothek	313508	249172	1,26

Bei der Berechnung des DNS/Peptid-Verhältnis wurden alle Sequenzen einbezogen. Folglich handelt es sich hier um einen Durchschnittswert. Auch wenn dieser Wert nahezu bei 1 liegt, ist eine gewisse relative Anreicherung bestimmter, möglicherweise strukturell günstiger, bzw. Abreicherung anderer, möglicherweise strukturell ungünstiger, Peptidsequenzen nicht ausgeschlossen. Aus diesem Grund wurde in **Abbildung 23** untersucht, ob ein prozentualer Anteil von 50 %, 75 % und 90 % der DNS-Sequenzen dem gleichen prozentualen Anteil an Peptidsequenzen entspricht. In der Plasmidbibliothek wurde wie schon bei der NNK<sub>7</sub>-Plasmidbibliothek nahezu der Idealwert erreicht, während für die Virusbibliothek eine gewisse Anreicherung von Peptiden festzustellen war.

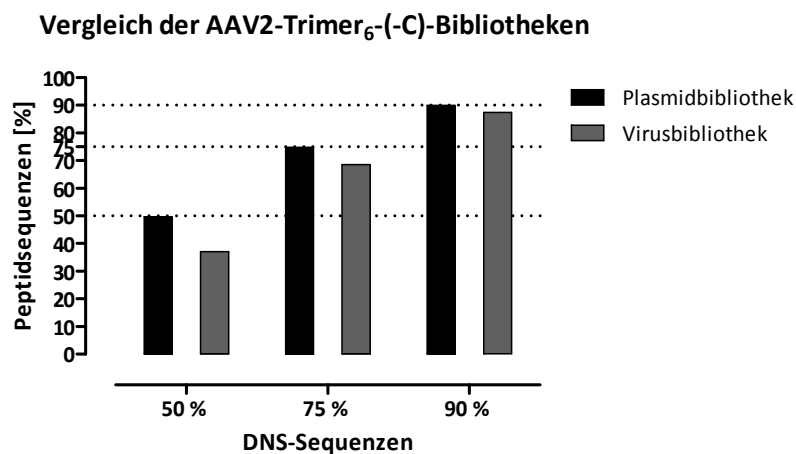


Abbildung 23: Vergleich der AAV2-Trimer<sub>6</sub>-(-C)-Plasmidbibliothek und daraus generierten Virusbibliotheken bezüglich der Anreicherung von Peptidsequenzen. Analysiert wurden 50 %, 75 % und 90 % der häufigsten DNS-Sequenzen im Hinblick darauf wie viel Prozent der Peptidsequenzen durch sie repräsentiert werden. Die Orientierungslinien bei 50 %, 75 % und 90 % der Peptidsequenzen stellen den Idealwert (DNS/Peptid-Verhältnis von 1) dar.

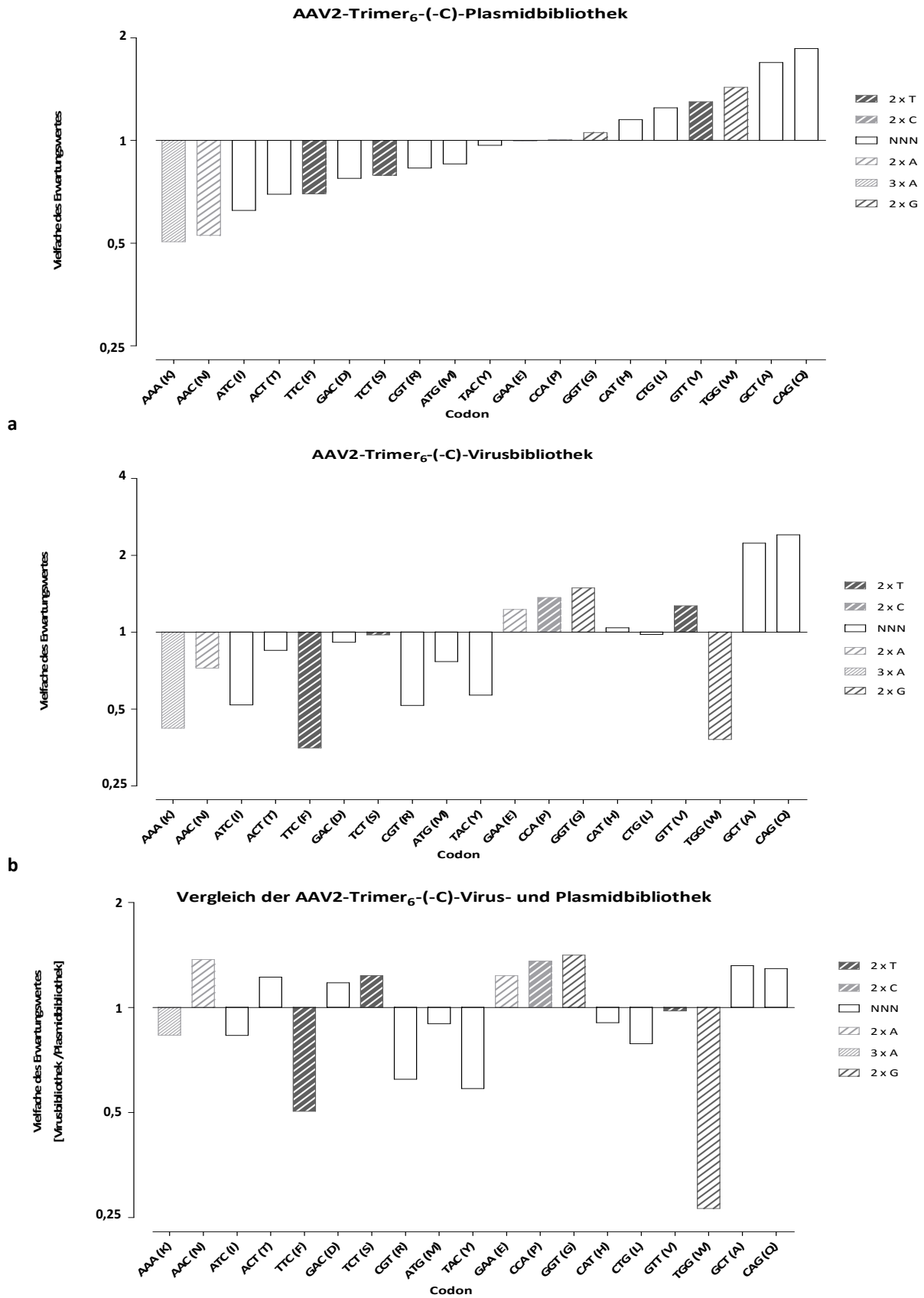
#### 4.4.3 Erwartete und tatsächliche Aminosäurehäufigkeit in den 2T<sub>6</sub>-Bibliotheken

In der NNK<sub>7</sub>-Plasmidbibliothek wurde eine starke relative Anhäufung von Guanin- und Adeninreichen Codons festgestellt, während Cytosin- und Thymin-haltige Codons unterrepräsentiert waren. Um festzustellen, ob solche oder ähnliche Tendenzen auch in der neuen 2T<sub>6</sub>-Plasmidbibliothek existieren, wurden die Vielfachen des Erwartungswertes jedes Codons und somit jeder Aminosäure berechnet. Hierbei liegt die Wahrscheinlichkeit für jede Aminosäure bei 1/19. Dieser Wert wurde gleich 1 gesetzt und die jeweiligen Vielfachen dargestellt.

**Abbildung 24a** zeigt, dass die tatsächlichen Häufigkeiten der Aminosäuren im positiven wie auch im negativen maximal um den Faktor zwei vom Erwartungswert abweichen. Die Häufigkeiten sind somit wesentlich dichter an ihrem Erwartungswert als die der NNK<sub>7</sub>-Plasmidbibliothek, deren Häufigkeiten im negativen bis zu einem Faktor 16 und im positiven bis zu einem Faktor vier abwichen. Des Weiteren konnte bezüglich der 2T<sub>6</sub>-Plasmidbibliothek keine systematische Präferenz für bestimmte Codons oder Nukleotide festgestellt werden.

Für den Vergleich von Plasmidbibliothek und Virusbibliothek wurde letztere auf die gleiche Weise analysiert (**Abbildung 24b**). Im Vergleich zur Plasmidbibliothek zeigte die Virusbibliothek eine klare Abnahme der aromatischen Aminosäuren Tryptophan, Phenylalanin und Tyrosin und zusätzlich für Arginin (**Abbildung 24c**). Die gleichen Aminosäuren und zusätzlich die Aminosäuren Cystein und Lysin zeigten auch beim Vergleich der NNK<sub>7</sub>-Virusbibliothek mit der entsprechenden Plasmidbibliothek die stärkste Abnahme, oder waren bereits in der Plasmidbibliothek unterrepräsentiert (**Abbildung 22**). Die Aminosäure Lysin war bereits in der 2T<sub>6</sub>-Plasmidbibliothek leicht unterrepräsentiert (**Abbildung 24a**). Der größte Zuwachs beim Vergleich der 2T<sub>6</sub>-Bibliotheken war für die Aminosäuren Asparagin und Glycin zu verzeichnen.

Zusammenfassend wurde durch den Einsatz Trimer-codierter Peptidinsertionen nicht nur eine Codierungsschema bedingte Anreicherung bestimmter Aminosäuren verhindert, sondern auch zusätzlich eine initiale, evtl. Synthese bedingte, Ungleichverteilung bestimmter Nukleotide und somit bestimmter Codons auf ein Minimum reduziert. Des Weiteren wurde für die NNK<sub>7</sub>- und die 2T<sub>6</sub>-Bibliotheken eine ähnliche Veränderung, nämlich die Abnahme der aromatischen Aminosäuren Tryptophan, Phenylalanin und Tyrosin, zwischen den Plasmid- und Virusbibliotheken festgestellt.



**c** **Abbildung 24: Codonhäufigkeit und Abweichungen vom Erwartungswert (AAV2-Trimer<sub>6</sub>-(-C)-Bibliotheken).** **a)** AAV2-Trimer<sub>6</sub>-(-C)-Plasmidbibliothek und **b)** AAV2-Trimer<sub>6</sub>-(-C)-Virusbibliothek: Die Vielfachen des Erwartungswerts wurden für jedes Codon bestimmt und von der höchsten negativen Abweichung zur höchsten positiven Abweichung sortiert. **c)** Vergleich beider Bibliotheken. Durch die Berechnung der Verhältnisse der Vielfachen des Erwartungswertes von Virusbibliothek und Plasmidbibliothek wird die Veränderung durch die Produktion der Virusbibliothek deutlich.



#### 4.4.4 Sequenzmotivanalyse der zehn häufigsten Peptidsequenzen

Zusätzlich zur Untersuchung der Codon-Häufigkeiten und deren Abweichungen vom Erwartungswert wurden die zehn häufigsten Peptidsequenzen der AAV2-Trimer<sub>6</sub>-(-C)-Plasmid- und Virusbibliothek verglichen (**Tabelle 27**). Während die Häufigkeiten der einzelnen Peptidsequenzen der Plasmidbibliothek sehr niedrig waren (< 0,003 %), war in der Virusbibliothek bereits eine leichte Anreicherung festzustellen. Das häufigste Peptid „NAARAE“ der Virusbibliothek war mit 0,13 % 43-mal häufiger als das häufigste Peptid der Plasmidbibliothek „ADYAVG“ mit 0,003 %. In der Plasmidbibliothek waren die Aminosäuren Glutamin und Tryptophan überrepräsentiert und die häufigsten Aminosäuren in vier von sechs Aminosäurepositionen. Trotz ihrer Häufigkeit bildeten sie kein Sequenzmotiv per Definition. Im Gegensatz dazu trugen alle der zehn häufigsten Peptidsequenzen der Virusbibliothek ein Asparagin in der ersten Position und in Position vier war eine Häufung von Arginin zu detektieren. Folglich zeichnete sich trotz der besseren initialen Aminosäuregleichverteilung der Peptidinsertionen der Plasmidbibliothek, eine ähnliche Anreicherung der Aminosäure Asparagin an erster Position ab.

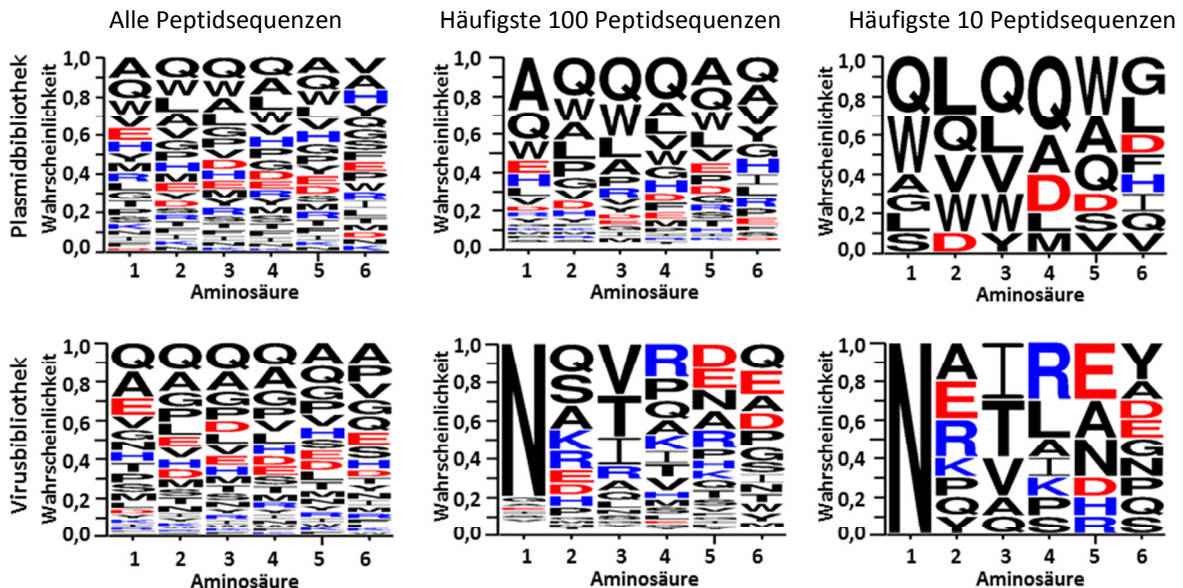
**Tabelle 27: Peptidsequenzanalyse und Vergleich der AAV2-Trimer<sub>6</sub>-Plasmid- und Virusbibliotheken.** Dargestellt sind die zehn häufigsten Peptidsequenzen jeder Bibliothek und ihre relativen Häufigkeiten. Für jede Position ist die häufigste Aminosäure, die mindestens dreimal zu finden war, farbig markiert. Wenn mehrere Aminosäuren diese Bedingung erfüllten, wurden alle farbig markiert. Zusätzlich wurden die jeweiligen Sequenzmotive angegeben.

Position	2T <sub>6</sub> -Plasmidbibliothek [%]	2T <sub>6</sub> -Virusbibliothek [%]
1	ADYAVG 0,00299	NAARAE 0,12854
2	QOVDWI 0,00224	NPQRAA 0,08771
3	QWWLWL 0,00224	NEVLRG 0,06411
4	LVVDDF 0,00224	NRVLDY 0,05007
5	QLQQQV 0,00224	NKIPHD 0,04975
6	WVLOWL 0,00224	NYIIEP 0,04848
7	SWYQQD 0,00224	NRTAFY 0,04465
8	WQLMSG 0,00224	NATREN 0,04401
9	WVQOAH 0,00224	NOTSNS 0,04114
10	GVQAAQ 0,00224	NEIKNQ 0,04082
<b>Motiv</b>	-	NXXRXX; NXTXXX; NXXXEX

#### 4.4.5 Häufigkeit der 19 Aminosäuren an den randomisierten Positionen der unselektierten AAV2-Trimer<sub>6</sub>-(-C)-Bibliothek

Die häufigsten Aminosäuren der Peptidsequenzen in der 2T<sub>6</sub>-Plasmidbibliothek sind Glutamin, Tryptophan, Leucin und Alanin (**Abbildung 25**). Ihr relativer Anteil pro Position liegt bei maximal 10 % (**Abbildung 25**), während der Anteil der häufigsten Aminosäuren bei der NNK<sub>7</sub>-Plasmidbibliothek doppelt so hoch war (**Abbildung 20**). Zusätzlich zu den oben beschriebenen Ergebnissen bezüglich der zehn häufigsten Sequenzen ist damit die höhere Sequenzvielfalt der 2T<sub>6</sub>-Plasmidbibliothek im Vergleich zur NNK<sub>7</sub>-Plasmidbibliothek auch für die Gesamtheit der sequenzierten Sequenzen gezeigt.

Auch die Peptidsequenzen der 2T<sub>6</sub>-Virusbibliothek sind diverser als die der NNK<sub>7</sub>-Virusbibliothek. Bei Betrachtung aller Sequenzen fallen die bereits beschriebene Abreicherung von Tryptophan und die Anreicherung von Prolin und Glycin in der 2T<sub>6</sub>-Virusbibliothek auf. Die häufigsten 100 und die häufigsten zehn Sequenzen der 2T<sub>6</sub>-Virusbibliothek sind immer noch diverser als die der NNK<sub>7</sub>-Virusbibliothek, aber auch hier ist eine starke relative Anreicherung von Asparagin in der ersten Position festzustellen. Zusammen mit der Aminosäure Arginin in Position vier ist auch in dieser Virusbibliothek das, um eine Aminosäure verkürzte, Sequenzmotiv „NXXRXX“ angereichert worden. Zusammenfassend wird mit der Verwendung der neuen 2T<sub>6</sub>-Peptidbibliothek aufgrund der Trimer-Codierung eine initiale Ungleichverteilung und Bevorteilung bestimmter Aminosäuren verhindert. Dennoch scheinen das Motiv „NXXRXX(X)“ tragende Kapsidvarianten und solche, die ein Asparagin an erster Position besitzen, Vorteile bei der Produktion beider Virusbibliotheken (NNK<sub>7</sub> und 2T<sub>6</sub>) zu haben.



**Abbildung 25: WebLogo basierte Analyse der Plasmidbibliothek im Vergleich zur Virusbibliothek.** Für jede Aminosäureposition wurden die häufigsten Aminosäuren ermittelt und deren Häufigkeit grafisch dargestellt. Anhand der Größe der Aminosäuren sind die jeweiligen relativen Häufigkeiten ablesbar. Blau: positiv geladene Aminosäuren (R, H, K); Rot: negativ geladene Aminosäuren (E, D); Schwarz: neutrale Aminosäuren (A, C, F, G, I, L, M, N, P, Q, S, T, V, W). Die Analyse wurde online durchgeführt: <http://weblogo.threeplusone.com/create.cgi> (Beschriftung der Skalen modifiziert)

#### 4.4.6 *In vitro*-Selektion zum direkten Vergleich der AAV2-Trimer<sub>6</sub>-(-C)-Peptidbibliothek mit der AAV2-NNK<sub>7</sub>-Peptidbibliothek

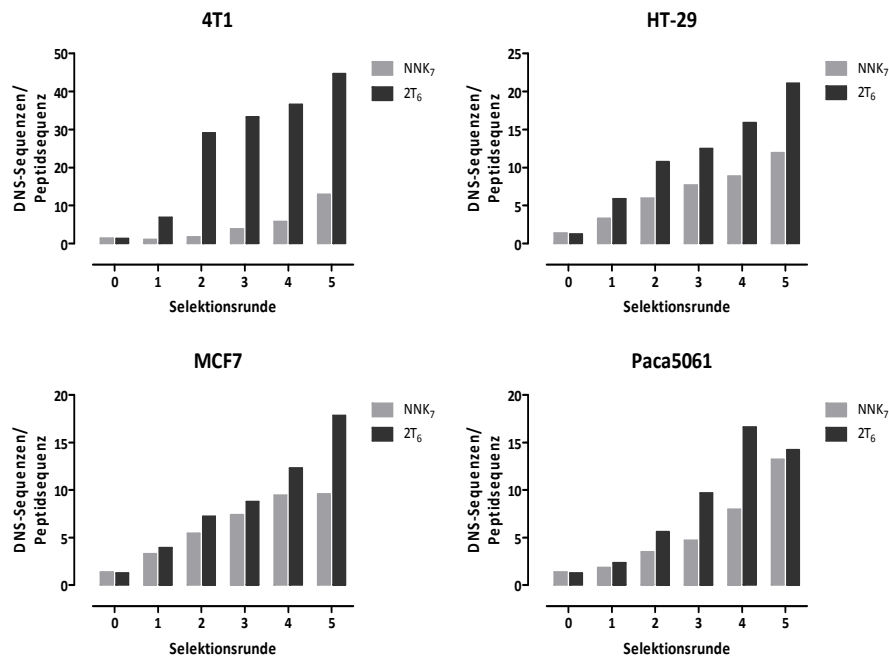
Nachdem die neue 2T<sub>6</sub>-Bibliothek in den vorherigen Abschnitten eingehend charakterisiert wurde, soll sie nun praktisch im Vergleich zur bisher verfügbaren NNK<sub>7</sub>-Bibliothek getestet werden. Da hiermit zunächst untersucht werden soll, ob die neue Peptidbibliothek generell zur Selektion von Tropismus-modifizierten Kapsidvarianten geeignet ist, wurden hierfür *in vitro*- statt *in vivo*-Selektionen mit beiden AAV-Peptidbibliotheken durchgeführt.

Die NNK<sub>7</sub>- und die 2T<sub>6</sub>- Virusbibliothek wurden dafür in vier verschiedenen Zelllinien (4T1, HT-29, MCF7 und Paca5061) über fünf Runden selektiert. Nach jeder Selektionsrunde wurden Proben genommen und diese mittels NGS sequenziert und anschließend analysiert (3.1.10). In **Tabelle 28** sind die Anzahl der sequenzierten DNS-Sequenzen und der zugehörigen Peptidsequenzen für jede Selektion und Selektionsrunde zusammengefasst. Hierbei liegt die Anzahl der sequenzierten DNS-Sequenzen jeder Selektionsrunde in ähnlicher Größenordnung, während die Anzahl der Peptidsequenzen mit jeder Selektionsrunde abnimmt.

**Tabelle 28: Zusammenfassung der NGS-Daten der *in vitro*-Selektionen.**

Selektion	Runde	Anzahl der DNS-Sequenzen	Anzahl der Peptidsequenzen
<b>NNK<sub>7</sub> 4T1</b>	1	75144	71124
	2	81543	48705
	3	77314	19910
	4	52436	9125
	5	86201	6674
<b>NNK<sub>7</sub> HT-29</b>	1	40486	12234
	2	45897	7711
	3	44698	5820
	4	38834	4378
	5	68160	5708
<b>NNK<sub>7</sub> MCF7</b>	1	62925	19281
	2	55591	10186
	3	48162	6524
	4	57590	6094
	4	40603	4245
<b>NNK<sub>7</sub> Paca5061</b>	1	36631	19831
	2	48163	13821
	3	36712	7827
	4	38581	4855
	5	45506	3439
<b>2T<sub>6</sub> 4T1</b>	1	72782	10616
	2	81685	2808
	3	50771	1527
	4	51703	1413
	5	50656	1135
<b>2T<sub>6</sub>HT-29</b>	1	73422	12459
	2	67662	6288
	3	57278	4586
	4	56658	3564
	5	74113	3521
<b>2T<sub>6</sub> MCF7</b>	1	45225	11481
	2	44932	6228
	3	38439	4370
	4	49403	4017
	5	59375	3330
<b>2T<sub>6</sub> Paca5061</b>	1	48528	20714
	2	53010	9448
	3	61672	6367
	4	75302	4537
	5	28966	2035

In **Abbildung 26** wurden die DNS/Peptid-Verhältnisse für die Selektion mit der NNK<sub>7</sub>- und der 2T<sub>6</sub>-Bibliothek verglichen. Beide unselektierten Peptidbibliotheken hatten mit 1,26 im Falle der 2T<sub>6</sub>-Bibliothek und 1,36 bezüglich der NNK<sub>7</sub>-Bibliothek ein ähnliches DNS- zu Peptidsequenzverhältnis und somit eine ähnliche Ausgangsposition. Während der Selektionen zeigte die 2T<sub>6</sub>-Bibliothek jeweils die höchste Anreicherung von Peptidsequenzen in allen vier Zelllinien.

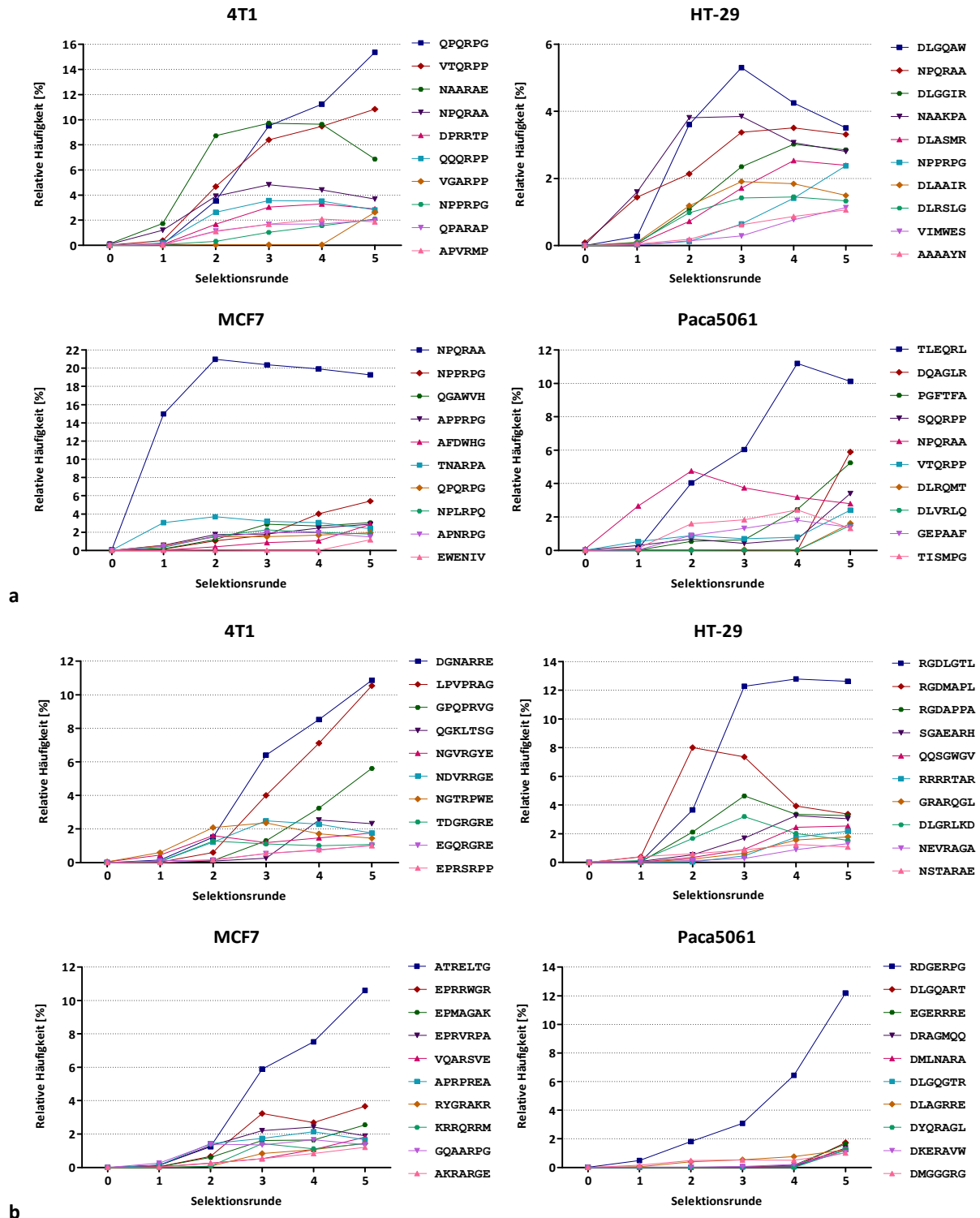


**Abbildung 26: Untersuchung der Peptidanreicherung während der Selektionen.** Für jede Selektion und Selektionsrunde ist das Verhältnis von sequenzierten DNS-Sequenzen zu Peptidsequenzen gezeigt. Hierbei wurden alle Sequenzen berücksichtigt.

#### 4.4.6.1 Anreicherung der zehn häufigsten Peptidsequenzen der letzten Selektionsrunde

Zur Auswahl weiter zu evaluierender Sequenzen wurden die relativen Häufigkeiten der zehn häufigsten Peptidsequenzen der letzten Selektionsrunde und ihr Verlauf während der gesamten Selektion im Folgenden genauer betrachtet. Während der Selektion in den 4T1-Zellen haben sich aus beiden AAV-Bibliotheken jeweils zwei Peptide der häufigsten Sequenzen ab Runde eins am stärksten angereichert und erreichten in Runde fünf einen Anteil von über 10 %. Während bei der Selektion auf HT-29-Zellen ein Einzelklon aus der NNK<sub>7</sub>-Peptidbibliothek klar angereichert wurde, gab es bei der 2T<sub>6</sub>-Bibliothek zwar einen häufigsten Klon, dieser zeigte aber keine klare und kontinuierliche Anreicherung. Im Gegensatz dazu konnte in der Zelllinie MCF-7 aus beiden Peptidbibliotheken jeweils ein häufigster Klon selektiert werden. Der Klon mit der Peptidsequenz „NPQRAA“ aus der 2T<sub>6</sub>-Bibliothek zeigte die stärkste Anreicherung. Dieser Klon wurde auch unter den zehn häufigsten Peptidsequenzen der Selektionen mit den anderen Zelllinien gefunden, wo er sich zwar zunächst anreicherte, aber sich schließlich nicht durchsetzte. Eine solche Beobachtung wurde auch für die mit Asparagin beginnenden Peptidsequenzen „NAARAE“ (4T1) und „NAAKPA“ (HT-29) gemacht. In den

Paca5061-Zellen konnten aus beiden Peptidbibliotheken häufigste Klon identifiziert werden. Hierbei reicherte sich der Klon mit der Sequenz „RDGERPG“ aus der NNK<sub>7</sub>-Peptidbibliothek am stärksten an. Die Anreicherung der zehn häufigsten Peptidsequenzen der letzten Selektionsrunde im Verlauf der Selektion ist in **Abbildung 27** grafisch dargestellt.



**Abbildung 27:** Verlauf der häufigsten zehn Peptidsequenzen der letzten Selektionsrunde während der Selektion. NNK<sub>7</sub>- und 2T<sub>6</sub>-Peptidbibliotheken wurden in vier verschiedenen Zelllinien selektiert (4T1, HT-29, MCF7 und Paca5061). Die zehn häufigsten Peptidsequenzen der letzten Selektionsrunde und ihr prozentualer Anteil an den Sequenzen jeder Selektionsrunde sind gezeigt. **a)** Selektion der AAV2-Trimer<sub>6</sub>(-C)-Peptidbibliothek. **b)** Selektion der AAV2-NNK<sub>7</sub>-Peptidbibliothek.

#### 4.4.6.2 Evaluation der *in vitro* selektierten AAV2-Kapside

Um herauszufinden, ob die selektierten Klone bessere Transduktionseigenschaften als zufällig aus der unselektierten Virusbibliothek gewählte Klone oder Klone der jeweils anderen Peptidbibliothek besitzen, wurde die Transduktion von Reporter-genvektoren mit ausgewählten Kapsidvarianten aus den Peptidbibliotheks-Selektionen untersucht. Hierfür wurde von jeder Selektion der jeweils häufigste Klon der letzten Selektionsrunde bestimmt und rekombinant mit Luciferase als Reporter-gen produziert (3.4.4.1). Falls es unter den zehn häufigsten Klonen welche gab, die von der vierten zur fünften Selektionsrunde eine stärkere Anreicherung zeigten als der häufigste Klon, wurden auch diese untersucht. Zum Vergleich wurden zusätzlich rekombinante Vektoren mit AAV2-Wildtypkapsid und für jede Bibliothek ein zufällig gewählter Klon (NNK<sub>7</sub>: rAAV2-QVWQNAG und 2T<sub>6</sub>: rAAV2-QYGQDP) produziert und untersucht.

Bezüglich der Selektion auf 4T1-Zellen wurden drei Vektoren aus der 2T<sub>6</sub>-Bibliothek (rAAV2-QPQRPG, rAAV2-VTQRPP und rAAV2-VGARPP) und zwei Vektoren aus der NNK<sub>7</sub>-Bibliothek (rAAV2-DGNARRE und rAAV2-LPVPRAG) gegeneinander getestet. Alle selektierten Vektoren zeigten eine stärkere Transduktion als die zufällig aus der unselektierten Bibliothek gewählten Kontrollklone (**Abbildung 28a**). Im Falle der 2T<sub>6</sub>-Bibliothek zeigte der häufigste Vektor auch die stärkste Transduktion. Im Gegensatz dazu war der stärkste Klon der NNK<sub>7</sub>-Bibliothek der mit einer final leicht stärkeren Anreicherung. Der Vektor rAAV2-QPQRPG zeigte als einziger einen zum rAAV2-Wildtypkapsid signifikanten Unterschied.

Im Falle der Selektion auf die HT-29-Zellen wurden zwei Vektoren aus der 2T<sub>6</sub>-Bibliothek (rAAV2-DLGQAW und rAAV2-NPPRPG) und ein Vektor aus der NNK<sub>7</sub>-Bibliothek (rAAV2-RDGLGTL) näher untersucht. Wieder zeigten alle selektierten Vektoren eine stärkere Luciferaseaktivität als die Vektoren mit zufällig ausgewähltem Kapsid aus der unselektierten Virusbibliothek (**Abbildung 28b**). Der Vektor rAAV2-RDGLGTL zeigte mit signifikantem Abstand die stärkste Transduktion.

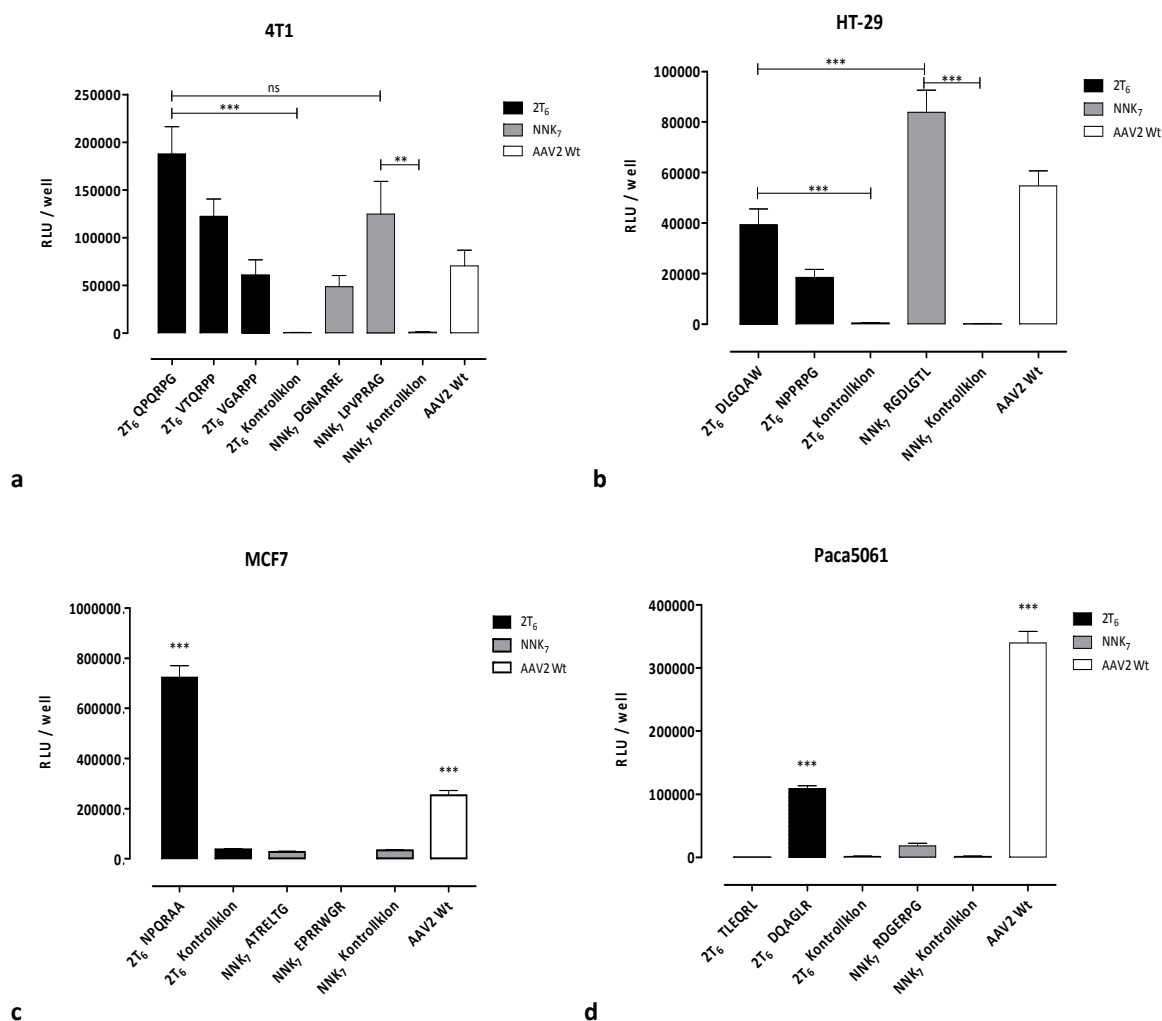
Bei der Analyse der Transduktionseigenschaften der auf MCF7-Zellen selektierten Klone zeigte der Vektor rAAV2-NPQRRA aus der 2T<sub>6</sub>-Bibliothek mit Abstand die stärkste Transduktion (**Abbildung 28c**). Im Gegensatz dazu transduzierten die Vektoren rAAV2-ATRELTG und rAAV2-EPRRWGR der NNK<sub>7</sub>-Bibliothek die MCF7-Zellen schlechter als der Vektor mit rAAV2-Wildtypkapsid und sogar schlechter als der Vektoren mit zufällig ausgewähltem Kapsid.

Im Falle der Selektion auf Paca5061-Zellen zeigte der zweithäufigste Klon der 2T<sub>6</sub>-Bibliothek rAAV2-DQAGLR eine signifikant höhere Expression als der Vektor rAAV2-RDGERPG der NNK<sub>7</sub>-Bibliothek (**Abbildung 28d**). Allerdings war der Vektor mit rAAV2-Wildtypkapsid der Vektor, der die Zellen am besten transduzierte.

Zusammenfassend stammten die am besten transduzierenden Vektoren in einem Fall aus der 2T<sub>6</sub>-Bibliothek (MCF7) und einem anderen aus der NNK<sub>7</sub>-Bibliothek (HT-29). Bezüglich der 4T1-Zellen

waren die Unterschiede der untersuchten Vektoren nicht signifikant und im Falle der Paca5061-Zellen transduzierte der Vektor mit AAV2-Wildtypkapsid am besten.

Somit war zusammenfassend zwar im Vergleich der beiden Bibliotheken kein eindeutiger Vorteil einer der beiden Ausgangs-Konstrukte nachweisbar, sehr wohl aber wurden durch die neue 2T<sub>6</sub>-Bibliothek im Einzelfall Vektoren selektiert, die solchen der herkömmlichen NNK<sub>7</sub>-Bibliothek überlegen waren. Des Weiteren wurde der Beweis erbracht, dass es möglich ist mit der neuen 2T<sub>6</sub>-Peptidbibliothek Tropismus-modifizierte Kapsidvarianten zu selektieren, die eine stärkere Transduktion als zufällig aus der unselektierten Virusbibliothek ausgewählte Kapsidvarianten ermöglichen.



**Abbildung 28: Bestimmung der Transduktion der selektierten und ausgewählten Kapsidvarianten.** Die häufigsten Peptidsequenzen und solche, die sich von der vorletzten zur letzten Selektionsrunde besonders angereichert haben, wurden ausgewählt. AAV-Vektoren, die diese Peptide auf ihrer Kapsidoberfläche repräsentieren, wurden mit Luciferase unter der Kontrolle eines CMV-Promotors als Reportergen produziert. **a)** 4T1-, **b)** HT-29-, **c)** MCF7- und **d)** Paca5061- Zellen wurden mit den jeweiligen Vektoren für 48 h inkubiert. Anschließend wurde die Transduktionseffizienz mittels Luciferase-assay bestimmt. Gezeigt sind die Mittelwerte mit positiver Standardabweichung. Die statistische Auswertung wurde mit einem einseitigen ANOVA Test und anschließender Bonferroni-Korrektur gemacht.  $p < 0,05 = *$ ;  $p < 0,01 = **$ ;  $p < 0,001 = ***$ ; (N=2; n=3). Gezeigt sind die Signifikanzen zwischen dem am stärksten transduzierenden Vektor beider Bibliotheken und zum jeweiligen Kontrollvektor. Zu beachten sind die unterschiedlichen Skalen. Wt: Wildtypkapsid.

## 5 DISKUSSION

### 5.1 Überblick

Stabile, effiziente und zielgerichtete Vektoren, für den Transfer eines therapeutischen Gens in Zielzellen, stellen eine zentrale Komponente bei der Erforschung potenzieller gentherapeutischer Ansätze dar. Die Generierung eines solchen Vektors ist somit eine der wichtigsten und gleichzeitig die am schwierigsten zu erfüllende Aufgabe der gentherapeutischen Forschung. AAV sind stabile Vektoren, die effizient eine Bandbreite verschiedener Zellen transduzieren und eine langanhaltende Expression des therapeutischen Gens vermitteln. Während folglich die Voraussetzungen der Stabilität und der Effizienz erfüllt sind, ist ihr Transduktionsprofil zwar Serotyp-abhängig unerschiedlich, entsprechend des dabei immer noch breiten Zielspektrums aber eher Zell- und Gewebe-unspezifisch. Im Jahr 2003 wurde von zwei Forschungsgruppen (Müller et al.<sup>122</sup>, Perabo et al.<sup>124</sup>) fast gleichzeitig das Prinzip der Selektion von AAV-Peptidbibliotheken und somit die Selektion von Peptiden zur Neuauslenkung des Vektortropismus im Kontext des AAV-Kapsids vorgestellt. Während seitdem eine Vielzahl verschiedener Kapsidvarianten *in vitro* auf verschiedensten Zelltypen selektiert wurde<sup>122,124-126,133,162</sup>, war die Anzahl *in vivo* selektierter Vektoren gering. Im Jahr 2009 konnten von unserer Gruppe Vektoren *in vivo* angereichert werden, die das Zielgewebe zwar stärker als das AAV2-Wildtypkapsid transduzierten, aber dennoch nicht zielgewebespezifisch waren. Ein Jahr später wurde von Ying et al. ein herzspezifischer Vektor selektiert<sup>132,133</sup>.

2013 gelang es unserer Arbeitsgruppe gleich zwei gewebespezifische (Lunge und Gehirn) Vektoren erfolgreich zu selektieren<sup>147</sup> und das große Potenzial der Selektionsmethode zu bestätigen. Nachdem diese Vektoren durch die Evaluation zufällig gewählter Klone identifiziert wurden, wurde 2016 anhand der Selektion in der Lunge eine systematischere Vorgehensweise zur Identifikation eines potenziell gewebespezifischen Vektors mittels der Auswertung und Bewertung von NGS-Daten vorgestellt, anhand derer die Spezifität des identifizierten Klons bestätigt werden konnte<sup>135</sup>.

Dennoch lieferte die Auswertung einer einzigen Selektion mittels NGS-Datenanalyse wenig Information für das Gesamtverständnis der Veränderung einer Peptidbibliothek während einer Selektion. Folglich stellte sich nun aufgrund der erfolgreichen Selektionen zum einen die Frage nach der Übertragbarkeit der Selektions- und Auswertungsmethode auf weitere Zielgewebe. Zum anderen ermöglichte die NGS-Datenanalyse eine umfassende Untersuchung von sekundären und auch primären Peptidbibliotheken und die Ergebnisse lieferten, neben der Idee der Generierung einer Peptidbibliothek mit einer 100-prozentigen Abdeckung der theoretisch möglichen Diversität, einen weiteren Grund für die Optimierung der bisher verwendeten Peptidbibliothek.



## 5.2 Übertragung der Selektionsmethode auf neue Ziele *in vivo*

### 5.2.1 Die Niere als Zielorgan

Trotz einer Vielzahl von Studien zum Einsatz von AAV als Gentherapievektor in verschiedensten Zielorganen ist die Anzahl der Studien bezüglich des Organs Niere eher limitiert und bisher von wenig Erfolg gekrönt. So wurden unterschiedliche Injektionsrouten mit verschiedenen AAV-Serotypen in Mäusen und Ratten getestet. Takeda et al.<sup>163</sup> injizierten die AAV-Serotypen 1 bis 5 über einen intraartiellen Katheter in die äußere Beckenarterie (*Arteria iliaca externa*) von Mäusen und Ratten. Hierbei konnte lediglich in den mit AAV2 injizierten Tieren Genexpression in den tubulären Epithelzellen der Niere nachgewiesen werden<sup>163</sup>. Die direkte intraparenchymale Injektion von AAV2 in die Niere der Maus führte zur Transduktion der Zellen um die Einstichstelle<sup>139</sup>. Die vaskuläre Injektion von AAV2 über die renale Arterie der Niere der Ratte hingegen führte zu einer begrenzten Transduktion proximaler tubulärer Zellen der S<sub>3</sub>-Segmente und Zellen des Sammelrohrs<sup>136</sup>. Und die retrograde Injektion über den Ureter von Mäusen, ebenfalls mit AAV2, führte zur Transduktion von tubulären Zellen der Medulla<sup>164</sup>. Doch bei keinen der Arbeiten konnte Genexpression in den gentherapeutisch betrachtet besonders interessanten Glomeruli oder der Vaskulatur der Niere nachgewiesen werden<sup>140</sup>. Zincarelli et al. untersuchten nach intravenöser Injektion die Serotypen 1 bis 9 auf Vektorverteilung und Genexpression in verschiedenen Organen von Mäusen<sup>113</sup>. Hierbei konnten in der Niere die rekombinanten Viren aller Serotypen auf DNS-Ebene, nicht aber die Expression des Reportergens nachgewiesen werden.

In der vorliegenden Arbeit wurde eine randomisierte AAV2-Virusbibliothek, nach intravenöser Injektion in die Schwanzene, in der Niere von Mäusen selektiert. Hierbei konnte zwar ein Sequenzmotiv („NXXRXX<sup>E</sup>/x“) identifiziert werden; dieses stellte sich aber als nicht gewebespezifisch und bereits in der unselektierten Virusbibliothek angereichert heraus. Während der Selektion wurde eine gewisse Anreicherung von Peptidsequenzen festgestellt. Die Tatsache, dass diese Peptidsequenzen bereits in der initialen unselektierten Virusbibliothek - nicht aber in der Plasmidbibliothek - angereichert waren, könnte auf einen strukturellen Vorteil der entsprechenden AAV2-Kapside hindeuten. Während der Selektion wurden sie dann wahrscheinlich aufgrund ihrer quantitativen Überlegenheit, unspezifischer Bindungen und einer fehlenden starken Gegenselektion von Selektionsrunde zu Selektionsrunde weitergetragen. Ein hoher Grad an unspezifischen Bindungen aufgrund der Beschaffenheit der Gefäße und der Komplexität der Blutversorgung der Niere kann dabei die Ursache für eine fehlende Gegenselektion sein. So tragen die Endothelzellen der Gefäße eine negativ geladene Glykokalyx, an der unspezifische Bindungen möglich wären. Des Weiteren durchläuft das Blut in der Niere zwei Kapillarbetten (glomeruläre Kapillarschlingen und das

peritubuläre Kapillarnetz) mit fenestriertem Endothel (mit und ohne Diaphragma)<sup>165</sup>. Durch das fenestrierte Endothel (Fenster mit einem Durchmesser von 50-100 nm)<sup>165</sup> könnte für AAV der Zugang zur ebenfalls negativ geladenen Basalmembran, auf deren Oberfläche u.a. Integrine und Heparansulfat-Proteoglykane<sup>165</sup>, die den primären zellulären Rezeptor des AAV2-Wildtypkapsids darstellen, zu finden sind, ermöglicht werden.

Neben der Hypothese der unspezifischen Bindungen wäre eine weitere Hypothese, dass das identifizierte unspezifische Sequenzmotiv nicht nur einen strukturellen Vorteil für das Kapsid bedeutet und dadurch initial quantitativ überlegen war, sondern strukturell betrachtet zu einem AAV2-Wildtyp-ähnlichem Kapsid beiträgt und somit ähnliche Stabilität und Bindungseigenschaften aufweist, ohne die Zellen vollständig zu transduzieren. Folglich wäre die Anzahl unspezifisch gebundener bzw. Wildtyp-ähnlicher Kapsidvarianten initial so hoch, dass potenziell spezifische Vektoren quantitativ unterlegen wären und während der Selektion verloren gingen. Des Weiteren zeigt die Vektorverteilung des AAV2-Wildtypkapsids einen vergleichsweise hohen Anteil an viralen Genomen in der Niere (**Abbildung 18a**). Es könnte somit auch sein, dass AAV2-Wildtyp-ähnliche Kapsidvarianten die höchste Wahrscheinlichkeit besitzen Zellen der Niere zu transduzieren.

### 5.2.2 Das Pankreas und die Langerhans-Inseln als Zielgewebe

In dieser Arbeit wurde zum ersten Mal eine randomisierte AAV2-Peptidbibliothek im Pankreas der Maus nach intravenöser Injektion selektiert. Hierbei konnte ein neues Sequenzmotiv und ein Klon mit im Vergleich zum AAV2-Wildtypkapsid veränderten Tropismus selektiert werden. Die Vektorverteilung zeigte eine klare Verschiebung zum Zielorgan Pankreas. Der hohe Anteil an viralen Genomen in der Milz kann durch die unspezifische Aufnahme von AAV durch retikuloendotheliale Zellen begründet sein<sup>166</sup>. Die Tatsache, dass keine Genexpression des Reportergens in der Milz gemessen wurde, bestätigt diese Vermutung (Daten nicht gezeigt). Kontrovers zu den Daten der Vektorverteilung wurde mittels *in vivo*-Bildgebung (*in vivo optical imaging*) die stärkste Genexpression im Herzen und Skelettmuskel gemessen. Die weitere Untersuchung und der Vergleich von Organlysaten mittels Luciferase-assay bestätigten diese Ergebnisse (Daten nicht gezeigt). Da das Selektionsprinzip der *in vivo*-Selektion von AAV-Peptidbibliotheken auf der Anreicherung von Kapsidvarianten im Zielgewebe und nicht auf der Fähigkeit der Kapsidvarianten zur Genexpression beruht, war die Selektion grundsätzlich dennoch erfolgreich, wenn auch die Expression im Zielgewebe noch unbefriedigend war. Der AAV2-Wildtypkapsid-Tropismus wurde geändert und auf das Zielorgan ausgerichtet. Wang et al. erzielten eine gleichmäßige Transduktion der Langerhans-Inseln nach intravenöser Injektion von AAV6 und AAV8 nur bei gleichzeitiger transienter Blockade der Leber<sup>143</sup>. Der selektierte Klon rAAV2-GDDATRD wurde im Vergleich zum AAV2-Wildtypkapsid nach systemischer Injektion von sich aus im Zielorgan Pankreas angereichert ohne dass durch eine

transiente Blockade der Leber die Anreicherung der Vektoren in der Leber verhindert und somit im Zielorgan erzwungen werden musste. Unter Verwendung eines gewebespezifischen Promotors könnte der in dieser Arbeit selektierte Vektor nach systemischer Injektion eine ausreichend starke Genexpression im Pankreas, vielleicht sogar in den Langerhans-Inseln, zeigen.

Die Langerhans-Inseln repräsentieren nur 1 bis 4 % des gesamten Pankreas<sup>167</sup>. Zur Erhöhung des Selektionsdruckes auf die therapeutisch wichtigen funktionellen Untereinheiten des Pankreas wurde parallel zur Selektion im gesamten Pankreas eine AAV-Peptidbibliothek-Selektion in den Langerhans-Inseln durchgeführt. Hierbei ist zu beachten, dass die isolierten Langerhans-Inseln nicht absolut frei von exokrinen Zellen waren. Der Einfachheit halber wurde die Selektion aber Selektion in den Langerhans-Inseln genannt.

Bei der Selektion in den Langerhans-Inseln reicherte sich das gleiche Sequenzmotiv wie bei der Selektion im gesamten Pankreas an. Während sich das Motiv bei der Selektion im Pankreas aber erst in der zweiten Selektionsrunde zu etablieren begann, wurde es bei der Selektion in den Langerhans-Inseln bereits in der ersten Runde sehr deutlich. Diese Tatsache und die Beobachtung, dass die Peptide der Selektion in den Langerhans-Inseln eine z.T. wesentlich stärkere Anreicherung zeigten, sind evtl. auf eine erfolgreiche Erhöhung des Selektionsdrucks zurückzuführen.

Zehn der 30 häufigsten Peptidsequenzen beider Selektionen stimmten überein. Die Sequenz „GSDTTQQ“ schnitt bezüglich des *GS-Scores* und des *C-Scores* bei der Selektion in den Langerhans-Inseln am besten ab. Wie für den rAAV2-GDDATRD aus der Selektion im Pankreas wurde für den Klon mit der Sequenz „GSDTTQQ“ aus der Selektion in den Langerhans-Inseln eine zum Zielorgan verschobene Vektorverteilung identifiziert (Daten nicht gezeigt). Diese war aber, passend zum ebenfalls niedrigeren *GS-Score*, im Vergleich zum im Pankreas selektierten Klon weniger deutlich.

Die Sequenz „GDDATRD“ aus der Selektion im Pankreas wurde ebenfalls in der Selektion in den Langerhans-Inseln unter allen sequenzierten Peptiden der ersten beiden, aber nicht unter denen der dritten Selektionsrunde gefunden (Daten nicht gezeigt). Dies lässt darauf schließen, dass die Sequenz entweder abgereichert oder von anderen Sequenzen verdrängt wurde. Obwohl die Anzahl der Versuchstiere für die Isolierung der Langerhans-Inseln von eins auf vier pro Selektionsrunde gesteigert wurde, konnte z.T. nur relativ wenig DNS isoliert werden. Aus diesem Grund ist es auch möglich, dass der entsprechende AAV-Klon während der Selektion verloren ging. Dennoch konnte auch im Falle der Selektion in den Langerhans-Inseln aus der Gesamtheit an Kapsidvarianten eine Variante selektiert werden, die die höchste Wahrscheinlichkeit besitzt gewebespezifisch zu sein.

Grundsätzlich ist es somit möglich, die AAV-Peptidbibliothek-Selektionsmethode auf andere Organe und funktionelle Untereinheiten zu übertragen. Im Falle der Selektion auf funktionelle

Untereinheiten wurde der Selektionsdruck scheinbar erhöht. Dennoch ist zu beachten, dass aufgrund der geringeren Gewebemenge und somit auch der geringeren Menge an isolierter DNS unter Umständen entsprechend mehr Versuchstiere benötigt werden, da initial selten vorhandene Kapsidvarianten ansonsten während der Selektion verlorengehen können.

### 5.2.3 Retrospektive Untersuchung der AAV-Peptidbankselektion im Gehirn

Zusätzlich zu einem lungenspezifischen Klon konnte in unserer Arbeitsgruppe in Vorarbeiten auch der gehirnspezifische Klon „AAV-BR1“, der die Sequenz „NRGTEWD“ auf seiner Kapsidoberfläche präsentiert, selektiert werden<sup>134</sup>. Der Verlauf der Selektion wurde retrospektiv in der vorliegenden Arbeit mittels NGS analysiert und mit den Daten der anderen Selektionen verglichen. Hierbei zeigte die Selektion im Gehirn die mit Abstand stärkste Peptidanreicherung, das bereits identifizierte Sequenzmotiv „XDGXXWX“ wurde verifiziert und die stark und spezifisch Gehirnspezifikation *in vivo* vermittelnde Peptidsequenz „NRGTEWD“ zeigte den höchsten *GS-Score* und bewies somit erneut den effizienten Nutzen der beschriebenen Bewertungskriterien. In der Arbeit von Körbelin et al.<sup>134</sup> wurde ein die Peptidsequenz „NRGTEWD“ präsentierender Klon in Selektionsrunde drei und vier jeweils einmal unter den zufällig gewählten Klonen gefunden und als zum Sequenzmotiv zugehörig befunden. Die nachträgliche Auswertung der NGS-Daten in der hier vorgelegten Arbeit zeigte, dass dieser Klon zu 6,63 % der sequenzierten Sequenzen in Runde drei, 0,72 % in Runde vier und 0,15 % in Selektionsrunde fünf vorhanden war und sich somit relativ betrachtet sogar abreicherte. Generell ist die Wahrscheinlichkeit, einen so seltenen Klon nur durch die zufällige Auswahl von zehn Bakterienkolonien zu finden, sehr gering und eher zufällig. Trotz der geringen relativen Häufigkeit wäre mittels der Auswertung der NGS-Daten auch in Selektionsrunde fünf der gehirnspezifische Klon schnell und ohne die Evaluation weiterer Klone identifiziert worden.

Des Weiteren wurde das unspezifische Sequenzmotiv „NXXRXX<sup>E</sup>/X“, das ebenfalls in der Niere und den unselektierten Virusbibliotheken angereichert war, in der Selektion im Gehirn vergleichsweise stark angereichert. Die Tatsache, dass die als spezifisch identifizierte Sequenz „NRGTEWD“ ebenfalls ein Asparagin an erster Position trägt, könnte die Erklärung für die Anreicherung des unspezifischen Sequenzmotivs sein. Ein Asparagin an der ersten Position des Peptides scheint bei der Selektion im Gehirn tatsächlich von Vorteil gewesen zu sein.

## 5.3 Next-Generation Sequencing zur Untersuchung von AAV-Selektionen

Zur Identifikation eines potenziell spezifischen Klons wurden in bisherigen Selektionen möglichst viele Klone ausgewählt und deren Genome konventionell sequenziert. Während diese Art der Klonidentifikation in gewissem Maße der Suche nach einer Nadel im Heuhaufen glich, ermöglicht die

Auswertung von AAV-Peptidbibliothek-Selektionen mittels NGS eine wesentlich umfangreichere und systematischere Suche und kann dem Anwender die zeitaufwendige Arbeit der Evaluation vieler Klone ersparen. Wie bei vielen neuen Methoden ergeben sich auch hier Fragen zur richtigen Anwendung. Die Möglichkeit hunderttausende oder gar Millionen von Sequenzen auswerten zu können, bietet dem Anwender einen sehr detaillierten Einblick in Peptidbibliotheken, wirft aber auch Fragen zur Auswahl von auszuwertenden Sequenzen auf. Wie viele DNS-Sequenzen sollten sequenziert werden? Wie viele Peptidsequenzen sollten ausgewertet und verglichen werden? Und wann ist ein Sequenzmotiv oder eine Sequenz potenziell spezifisch?

Generell liefert die Sequenzierung aller möglichen DNS-Sequenzen die maximale Information. Je nach Anwendung ist sie aber eventuell nicht nötig. So ist eine Sequenz, die nach einer Selektion nur einmal unter einer Millionen Sequenzen zu finden ist, höchstwahrscheinlich nicht die gesuchte spezifische Sequenz. Genauso ist eine potenziell spezifische Sequenz, wenn sie einmal in einem Kontrollorgan gefunden wurde, noch lange nicht unspezifisch. Außerdem erfordert eine sehr hohe Anzahl gelesener DNS-Sequenzen (*Reads*) bzw. eine sehr tiefe Sequenzierungstiefe das entsprechende Sequenzierungsgerät und viel Platz auf der Trägerplatte und ist mit entsprechend höheren Kosten verbunden. Im konkreten Fall der in dieser Arbeit mittels *in vivo*-Selektion untersuchten NNK<sub>7</sub>-Peptidbibliothek gehen wir anhand der Anzahl an transformierten Bakterien von einer initialen maximalen Diversität von  $2 \times 10^8$  verschiedener Klone und somit theoretisch möglicher Kapsidvarianten aus. Aber nur ein Teil der modifizierten Kapsidproteine trägt zu funktionellen Kapsiden bei und nur ein Teil der Kapsidvarianten wird positiv selektiert, während im Fall einer *in vivo*-Selektion sogar ein Großteil der Kapsidvarianten durch die anderen Gewebe negativ selektiert bzw. im Zielorgan abgereichert werden sollte. Wie die nachträglich mittels NGS ausgewertete Selektion im Gehirn gezeigt hat, konnte die spezifische Sequenz unter den gegebenen Bedingungen durch die Sequenzierung einer Anzahl von Reads in der Größenordnung von  $10^5$  identifiziert werden.

Während Körbelin et al.<sup>135</sup> aufgrund der hohen Anreicherung mit lediglich 24 Peptidsequenzen 99 % der *Reads* abgedeckt hatten, wären für eine ähnliche Abdeckung im Falle der Selektion in der Niere, bei der die Anreicherung wesentlich geringer war, 3027 Peptidsequenzen auszuwerten gewesen. Von diesen Sequenzen hatten 2939 Peptidsequenzen einen relativen Anteil von unter 0,1 % (Daten nicht gezeigt). Im Zweifelsfall können die beschriebenen Bewertungskriterien zur Identifizierung eines potenziellen spezifischen Klons auch mühelos auf sehr viele Sequenzen angewendet werden. Da sich aber potenziell spezifische Sequenzen während der Selektion anreichern sollten gilt, je geringer die relative Häufigkeit einer Peptidsequenz ist, desto geringer ist die Wahrscheinlichkeit, dass diese Sequenz spezifisch ist. Wie viele Peptidsequenzen ausgewertet werden sollten, ist somit von dem

prozentualen Anteil der Peptide und folglich von der Stärke der Peptidanreicherung abhängig. Aus diesem Grund sollte für jede Selektion individuell festgelegt werden, wie viele Sequenzen näher ausgewertet werden.

#### **5.4 Anreicherung von Peptidsequenzen während einer Selektion**

Zur Beurteilung einer generellen Peptidanreicherung während einer Selektion wurde das Verhältnis DNS-/Peptidsequenz aller Sequenzen für jede Selektionsrunde berechnet. Eine weitere noch deutlichere Möglichkeit, Peptidanreicherungen darzustellen, ist die Auswertung der häufigsten Sequenzen. So wurde für jede Selektionsrunde untersucht, wie viel Prozent der Peptidsequenzen 90 % der DNS-Sequenzen repräsentieren. Da der Zusammenhang zwischen der Anzahl der Reads und der Anzahl der Peptide nicht linear ist, gilt für beide Methoden, dass eine sehr gute Vergleichbarkeit der Werte nur gewährleistet ist, wenn die einzelnen Proben ähnliche Sequenzierungstiefen besitzen und die DNS von ähnlicher Qualität ist. Angenommen, bei einem NGS-Lauf wurden nach Anwendung der Qualitätsfilter  $10^5$  auswertbare DNS-Sequenzen sequenziert und diese entsprächen  $10^4$  Peptidsequenzen. Das entspräche einem DNS-/Peptidsequenz Verhältnis von zehn. Würden von derselben Probe  $10^6$  also zehnmal mehr DNS-Sequenzen gelesen, entsprächen diese nicht automatisch zehnmal mehr verschiedenen Peptidsequenzen, da die zusätzlichen Reads nicht alle unbekannte neue Peptidsequenzen wären. Die bereits bekannten Peptidsequenzen wären ebenfalls und wahrscheinlich mehrfach unter den zusätzlichen Sequenzen. Folglich entsprächen  $10^6$  Reads vielleicht nur  $8 \times 10^4$  statt  $10^5$  Peptidsequenzen und somit einem Verhältnis von 12,5 statt zehn.

Im Falle der Selektion in der Niere stieg das DNS-/Peptidsequenzverhältnis mit jeder Selektionsrunde. In der letzten Selektionsrunde wurde das Verhältnis wieder kleiner, obwohl der prozentuale Anteil an viralen Genomen anstieg. Der Grund für diese Ergebnisse kann somit in der im Vergleich zu den anderen Runden wesentlich niedrigeren Sequenzierungstiefe der letzten Selektionsrunde liegen.

#### **5.5 Bestimmung und Bewertung von Sequenzmotiven und potenziell spezifischen Klonen**

Zur Untersuchung der Etablierung eines Sequenzmotivs oder der Anreicherung eines Klons wurden zunächst von allen Selektionen die zehn häufigsten Peptidsequenzen jeder Selektionsrunde genauer analysiert. Als Sequenzmotiv wurde die häufigste Kombination von Aminosäuren in der gleichen Position, die mindestens in drei von zehn Sequenzen zu finden war, definiert. Auf diese Weise konnten für die Selektion in der Niere („NDXXRXX“ und „NXXRXXE“) und dem Gehirn („XDGXXWX“ und „NXXRXXE“) jeweils zwei und für die Selektion im Pankreas („XXDXT<sup>R</sup>/X“) und den Langerhans-Inseln („XXDXTRX“) das gleiche Sequenzmotiv identifiziert werden. Zum einen fiel hierbei auf, dass

das zweite Motiv der Selektionen in der Niere und dem Gehirn dasselbe war, zum anderen zeigte jede Selektion ein eigenes Motiv und die jeweils häufigsten Sequenzen der Selektionen trugen dieses Motiv oder mindestens Teile davon. Dennoch wurden alle häufigsten Sequenzen theoretisch, durch den Vergleich mit den Sequenzen in den Kontrollorganen, und z.T. auch im Versuch als unspezifisch befunden (Referenz<sup>134,135</sup>, Daten nicht gezeigt). Die Existenz eines Sequenzmotivs scheint somit für die Anreicherung der entsprechenden Vektoren von Vorteil, bedeutet aber nicht automatisch Spezifität.

Zur Untersuchung der Spezifität der identifizierten Sequenzmotive bzw. der selektierten Sequenzen wurden die sequenzierten Peptidsequenzen auf Anreicherung während der Selektion und Spezifität gegenüber Kontrollorganen mittels Berechnung von drei Bewertungskriterien (für Anreicherung, Spezifität und die Kombination aus beidem) untersucht. Mit Hilfe der *Scores* wurden die relativen Häufigkeiten der Peptidsequenzen zueinander ins Verhältnis gesetzt, in einen Wert zwischen 0 und 1 konvertiert und somit besser vergleichbar gemacht. Aufgrund leicht variabler Sequenzierungstiefen ist die absolute Anzahl der viralen Genome im Zielgewebe und in den Kontrollorganen leicht verschieden und nicht direkt vergleichbar. Deshalb wurden zur Berechnung der *Scores* die relativen Häufigkeiten der Peptide miteinander verglichen. Beim Vergleich der relativen Häufigkeiten kann es passieren, dass in einem Organ eine höhere relative Häufigkeit bestimmt wurde als in einem anderen, diese aber nicht einer höheren absoluten Häufigkeit entspricht, weil in dem anderen Organ viel mehr Sequenzen sequenziert wurden. Um diesem entgegenzuwirken und um Rückschlüsse auf die tatsächliche Anzahl viraler Genome ziehen zu können, kann diese vor der Sequenzierung normiert, oder im Nachhinein bestimmt und berechnet werden. Dennoch würden unterschiedliche Sequenzierungstiefen für eine gewisse Variabilität der Werte sorgen. Des Weiteren geht es bei dem Vergleich der *Scores* darum, den Klon zu finden, der die höchste Wahrscheinlichkeit besitzt, spezifisch zu sein und die Reihenfolge der Peptidhäufigkeiten innerhalb einer Probe, bzw. eines Organs bleibt unabhängig davon, ob mit absoluten oder relativen Werten gerechnet wird, die gleiche.

Kürzlich wurde von unsere Arbeitsgruppe ein Lungenendothel-spezifischer Kapsid-modifizierter AAV2-Klon durch *in vivo*-Selektion einer AAV2-Peptidbibliothek selektiert und die Auswertung mittels NGS und den genannten Bewertungskriterien vorgestellt<sup>135</sup>. Während hierbei die Anreicherung von der vorletzten zur letzten Selektionsrunde betrachtet wurde, wurde in der vorliegenden Arbeit die durchschnittliche Anreicherung über die gesamte Selektion berechnet, da es theoretisch sein kann, dass sich ein Klon bereits am Anfang der Selektion besonders anreichert. So hat sich der Gehirnspezifische Klon „AAV-BR1“ in den ersten drei Selektionsrunden angereichert, in den letzten beiden Runden aber nicht.

Einer der Qualitätsfilter bei der Auswertung der NGS-Daten bestand darin, dass eine Sequenz, die sich von einer anderen Sequenz in nur einem Nukleotid unterscheidet und 100-fach seltener vorkam, von der Auswertung ausgeschlossen wurde, weil es sich bei der selteneren Sequenz höchstwahrscheinlich um eine Mutation oder einen Sequenzierungsfehler handelte. Bei der Betrachtung des *GS-Scores* der Selektion in der Niere lagen die Werte für die beiden Sequenzen „NDAGREL“ (Platz 8) und „KEGPSWQ“ (Platz 23) nahezu bei 1. Bei diesen Sequenzen handelte es sich um solche, die sich aufgrund der Änderung nur eines Nukleotids von den häufigeren Sequenzen „NDAGREV“ (Platz 1) bzw. „KEGTSWQ“ (Platz 2) unterschieden. Zusätzlich kamen die selteneren Sequenzen jeweils in keiner weiteren Selektionsrunde oder in der unselektierten Virusbibliothek oder der Plasmidbibliothek vor. Lediglich die Sequenz „NDAGREL“ wurde mit einer relativen Häufigkeit von  $1,31 \times 10^{-5}$  in dem Kontrollorgan Pankreas, in dem ebenfalls die Sequenz „NDAGREV“ die häufigste war, gefunden. In den Sequenzen aller anderen Kontrollorgane wurden weder die Sequenz „NDAGREL“ noch die Sequenz „KEGPSWQ“ gefunden. Aufgrund ihrer Häufigkeit wurden diese Sequenzen nicht durch den beschriebenen Qualitätsfilter herausselektiert, dennoch liegt die Vermutung nahe, dass es sich um eine Mutations- oder Sequenzierungsfehler bedingte Variante der jeweils häufigeren Sequenz handelte.

Die Sequenz „GDDATRD“ der Selektion im Pankreas stach besonders hervor, indem sie von den untersuchten Sequenzen eine der stärksten Anreicherungen und die höchste Spezifität zeigte. Von der häufigeren Sequenz „GQDATRV“ (Platz 1) unterscheidet sich die Sequenz „GDDATRD“ (Platz 22) in zwei Aminosäuren; insgesamt unterscheiden sie sich aber in sieben von 21 Nukleotiden. Zusätzlich ist die Sequenz in allen Selektionsrunden gefunden worden, was darauf schließen lässt, dass es sich um eine „echte“ und nicht um eine durch Sequenzierungsfehler oder Mutationen entstandene Sequenz handelt. Diese Beispiele zeigen, dass die Bewertungskriterien eine wichtige Hilfe zur Identifikation eines potenziell spezifischen Klons geben, aber dennoch kritisch ausgewertet werden sollten.

## **5.6 Vergleich der Peptidbibliotheken AAV2-NNK<sub>7</sub>- und AAV2-Trimer<sub>6</sub>-(-C) hinsichtlich Diversität und Abdeckung**

Die in dieser Arbeit miteinander verglichenen Peptidbibliotheken unterscheiden sich in der Peptidinsertion hinsichtlich des Codierungsschemas und der Peptidlänge und somit in der Höhe der Peptiddiversität und der Abdeckung theoretisch möglicher Peptidsequenzen. Zur Berechnung von Peptiddiversitäten und -abdeckungen für Peptide mit unterschiedlichen Längen und Codierungsschemata haben Sieber et al. ein Programm mit dem Namen PeLiCa (*Peptide Library*



*Calculator*) entwickelt ([www.pelica.org](http://www.pelica.org))<sup>168</sup>. Dieses wurde auch in der hier vorgelegten Arbeit verwendet.

Die meisten Peptide der bisher zur Selektion verwendeten AAV2-Peptidbibliotheken besaßen eine Länge von sieben Aminosäuren<sup>129</sup>. Die theoretisch mögliche maximale Peptiddiversität solcher Peptidbibliotheken liegt bei  $1,8 \times 10^9$ . Da die Anzahl der transformierten Bakterienklone den limitierenden Faktor darstellt, wird die zur maximalen Diversität notwendige Anzahl aber in der Regel nicht erreicht. So besitzen die für diese Arbeit verwendeten AAV2-NNK<sub>7</sub>-Peptidbibliotheken eine Abdeckung aller möglichen Peptidsequenzen von ca. 10 %. Bei diesen Bibliotheken ist es somit möglich, dass eine potenziell spezifische Sequenz von vornherein nicht in der Bibliothek vorhanden ist. Zur Erhöhung der Abdeckung haben die Peptide der hier verwendeten Trimer-codierten Peptidbibliothek eine Länge von sechs Aminosäuren und somit eine theoretische und praktische Peptiddiversität von  $4,7 \times 10^7$  und damit eine Abdeckung von nahezu 100 %. Aufgrund der kürzeren Peptidsequenz besitzt diese Bibliothek aber rechnerisch siebenmal weniger Peptidvarianten. Unter Berücksichtigung des NNK-Codierungsschemas, der damit verbundenen Redundanz und der Präsenz von Stoppcodons in der NNK<sub>7</sub>-Peptidbibliothek beträgt der Unterschied immerhin noch das Dreifache.

Hinter dem Vergleich der beiden Bibliotheken mit entsprechenden Eigenschaften steht die Frage nach der Notwendigkeit die beste Sequenz aller möglichen Sequenzen selektieren zu müssen. Im Falle einer solchen Notwendigkeit, wäre die Verwendung der 2T<sub>6</sub>-Peptidbibliothek von Vorteil, genügt es aber eine zweit- oder sogar drittbeste Sequenz zu finden, bietet die NNK<sub>7</sub>-Peptidbibliothek mehr Möglichkeiten. Die ideale Lösung wäre natürlich eine Peptidbibliothek mit einer hohen Diversität und einer 100-prozentigen Peptidabdeckung; aus technischen Gründen ist das aber bisher nicht möglich.

Um festzustellen, ob der Tropismus von bereits selektierten gewebespezifischen AAV2-Vektoren durch den Austausch einer Aminosäure der Peptidinserts beibehalten wird, wurden in der Dissertation von Körbelin (2013) im Falle des in der Lunge selektierten AAV2 Klon mit der Sequenz „ESGHGYF“ und dem im Gehirn selektierten Klon mit der Sequenz „NRGTEWD“ sämtliche Aminosäuren des Peptidinserts einzeln durch die Aminosäure Alanin ausgetauscht und untersucht<sup>147</sup>. Während in der Sequenz „ESGHGYF“ die ersten beiden Aminosäuren durch Alanin ausgetauscht werden konnten, ohne dass der Tropismus vom Zielorgan Lunge weggelenkt wurde, bewirkte der Aminosäureaustausch in jeder Position der Sequenz „NRGTEWD“ einen veränderten Tropismus. Bezüglich der Sequenz „ESGHGYF“ waren folglich mehrere Sequenzen möglich, im Fall der Sequenz „NRGTEWD“ hingegen durfte keine der Aminosäuren, zumindest nicht durch Alanin, ausgetauscht werden, wenn der Tropismus beibehalten werden sollte.

Folglich wäre zur Findung der Sequenz „ESGHGYF“ eher eine Peptidbibliothek mit einer höheren Peptiddiversität zu empfehlen, denn auch wenn nicht alle Peptidvarianten zur Auswahl stünden, würde höchstwahrscheinlich eine der möglichen passenden Varianten dabei sein. Das Beispiel mit dem die Sequenz „NRGTEWD“ präsentierenden Klon spricht hingegen eher für die Verwendung einer Peptidbibliothek mit einer höheren Peptidabdeckung, denn wenn die entsprechende Sequenz aufgrund einer geringen Abdeckung nicht vorhanden wäre, würde auch eine besonders hohe Diversität nicht zur Findung des Klons beitragen.

Auch die Ergebnisse der hier vorgelegten Arbeit zeigen, dass ein Sequenzmotiv alleine keine Spezifität ausmacht. Viele der im Pankreas selektierten Sequenzen trugen das Sequenzmotiv in Kombination mit verschiedenen Aminosäuren, doch erst die Kombination mit bestimmten Aminosäuren zeigte einen besonders hohen Wert bezüglich der Spezifität. Theoretisch betrachtet sprechen diese Ergebnisse eher für die Verwendung einer Peptidbibliothek mit einer hohen Peptidabdeckung. Dennoch zeigt die Praxis, dass auch mit einer Bibliothek mit vergleichsweise geringer Peptidabdeckung gewebespezifische AAV2-Vektoren selektiert werden konnten, wobei nicht ausgeschlossen werden kann, dass eine Kapsidvariante mit noch besseren Transduktionseigenschaften gefunden worden wäre, wenn die Möglichkeit der vollständigen Peptidabdeckung der NNK<sub>7</sub>-Peptidbibliothek bestanden hätte.

### **5.7 Vergleich der Plasmidbibliotheken AAV2-NNK<sub>7</sub> und AAV2-Trimer<sub>6</sub>-(-C)**

Mit Hilfe der NGS-Daten der verwendeten NNK-codierten Plasmidbibliothek konnten zwei Gründe für die Anreicherung der Codons für Glycin, Arginin und Glutaminsäure identifiziert bzw. verifiziert werden. Einer dieser Gründe ist bereits in der Wahl des Codierungsschemas und durch den degenerierten genetischen Code begründet. In einem NNK-Codierungsschema können die ersten beiden Positionen jedes der vier möglichen Nukleotide Adenin (A), Guanin (G), Cytosin (C) und Thymin (T) codiert sein, während die letzte Position entweder G oder C ist. Im Gegensatz zu einem NNN-Codierungsschema, bei dem 64 verschiedene Codons 20 Aminosäuren und drei Stoppcodons codieren, verteilen sich bei einem NNK-Schema 32 mögliche Codons auf 20 Aminosäuren und ein Stoppcodon. Folglich sind die NNK-codierten Codons besser aber nicht gleichverteilt. Drei Aminosäuren werden durch drei Codons und fünf durch zwei Codons codiert, während nur 12 Aminosäuren durch ein Codon codiert werden. Des Weiteren verringert das Vorhandensein von Stoppcodons die Diversität der Peptidbibliothek. Für die NNK<sub>7</sub>-Plasmidbibliothek konnte eine Anreicherung der Aminosäuren Arginin, Serin, Alanin, Glycin und Valin festgestellt werden. Die zugehörigen relativen Häufigkeiten dieser Aminosäuren wichen aber so stark vom Erwartungswert ab, dass das NNK-Codierungsschema nicht die einzige Ursache sein konnte.

Der zweite Grund für die Anhäufung bestimmter Aminosäuren in der Plasmidbibliothek lag in einer Ungleichverteilung der Codons aufgrund bevorzugter Nukleotide. Die Ursache für solche Präferenzen können unterschiedliche Bindungseffizienzen von Nukleosidanaloga während der Oligonukleotidsynthese oder die Klonierung der Plasmidbibliothek in *E.coli* sein<sup>169,170</sup>. Palfrey et al. haben Oligonukleotide mit einem randomisierten Nukleotid in Plasmide ligiert und anschließend in *E. coli* kloniert. Hierbei konnte gezeigt werden, dass die Verteilungen der Sequenzen unterschiedlicher Bibliotheken nach der Klonierung im Vergleich zu vorher und im Vergleich zueinander verschieden waren<sup>170</sup>. Aufgrund unterschiedlicher Verteilungen nach der Klonierung wurde ein systematischer Selektionsdruck als Grund ausgeschlossen. Für die in dieser Arbeit verwendeten NNK-codierte Plasmidbibliothek wurde eine Ungleichverteilung der Codons aufgrund unterschiedlich häufiger Nukleotide in der Reihenfolge T<C<A<G festgestellt. Zur Herstellung einer equimolaren Verteilung der Nukleotide sollten die Nukleotide unter Berücksichtigung unterschiedlicher Bindungseffizienzen in verschiedenen Verhältnissen gemischt werden (1,0 x T; 1,15 x G; 1,25 x C; 1,5 x A)<sup>169</sup>. Bei einem 1:1 Verhältnis hingegen würde man eine Reihenfolge von A<C<G<T erwarten. Diese entspricht aber nur z.T. der in der vorliegenden Arbeit beobachteten Reihenfolge von T<C<A<G. Folglich ist die Ungleichverteilung der Nukleotide der hier verwendeten NNK-Plasmidbibliothek vermutlich das Produkt aus unterschiedlichen Nukleotid-Bindungseffizienzen und Änderungen der Verteilungen durch die Klonierung in *E. coli*. Eine produktionsbedingte nicht-equimolare Verteilung der Nukleotide bei der Oligonukleotidsynthese kann dennoch nicht ausgeschlossen werden.

Eine Möglichkeit den genannten Ungleichverteilungen einer Peptidbibliothek entgegenzuwirken ist die Verwendung eines anderen Codierungsschemas. Für unsere neue AAV2-Trimer<sub>6</sub>-(-C)-Peptidbibliothek wurden anstatt von 32 NNK-codierten Codons 19 definierte Trimere verwendet. Eine Verschiebung der Häufigkeiten von Aminosäuren aufgrund des Codierungsschemas ist somit ausgeschlossen. Zusätzlich ist aufgrund der Synthese von Trimeren anstatt von einzelnen Nukleotiden und der Berücksichtigung von Trimer spezifischen Reaktionsfaktoren eine Ungleichverteilung von bestimmten Nukleotiden aufgrund von unterschiedlichen Bindungseffizienzen minimiert<sup>171</sup>. Eine Selektion unter Verwendung einer Trimer-codierten Plasmidbibliothek schafft somit für jede Aminosäure eine ähnliche Wahrscheinlichkeit, codiert zu werden.

Weitere Vorteile dieser Bibliothek sind die Möglichkeit der Rücksichtnahme auf die *codon usage* der Produktionszelle und die höhere Wahrscheinlichkeit eventuelle Sequenzierungsfehler zu erkennen. Zusätzlich können, wie es in der vorliegenden Arbeit getan wurde, bestimmte Codons wie Stoppcodons oder bestimmte Aminosäuren wie Cysteine von vornherein ausgeschlossen werden.

Ca. 9 % der Sequenzen der hier verwendeten NNK-codierten Plasmidbibliothek trugen Stoppcodons und ca. 7 % enthielten Cysteine. Übereinstimmend mit bisherigen Arbeiten in denen die Aminosäure Cystein in den untersuchten Virusbibliotheken unterrepräsentiert war<sup>125,129,162,172</sup> zeigte auch der hiesige Vergleich der Plasmidbibliothek mit den Virusbibliotheken eine klare Abnahme von Cysteinen in den Virusbibliotheken (**Abbildung 20c**). Folglich können bis zu 16 % der Sequenzen der NNK-codierten Plasmidbibliothek für den Zusammenbau eines Kapsids unbrauchbar sein, könnten durch funktionelle Peptide ersetzt werden und somit die Diversität funktioneller Peptide erhöhen.

### **5.8 Vergleich der nicht-selektierten Virusbibliotheken AAV2-NNK<sub>7</sub> und AAV2-Trimer<sub>6</sub>-(-C)**

Entsprechend der besseren Gleichverteilung der 2T<sub>6</sub>-Plasmidbibliothek waren auch die Aminosäuren der 2T<sub>6</sub>-Virusbibliothek gleichmäßiger verteilt als die der NNK<sub>7</sub>-Peptidbibliothek. Dennoch fand bei beiden Bibliotheken während der Produktion der Virusbibliothek eine Peptidanreicherung statt. Beim Vergleich der Aminosäure-Häufigkeiten der Plasmidbibliotheken mit den jeweiligen Virusbibliotheken wurde für beide Bibliotheken eine Abnahme von positiv geladenen (K, R, H), aromatischen (W, F, H, Y) Aminosäuren und Cystein (nur bei NNK<sub>7</sub>-Bibliothek) und eine Zunahme von negativ geladenen (D,E) und kleineren Aminosäuren (G, A, V, P, S) und den Aminosäuren Asparagin und Glutamin festgestellt. Perabo et al.<sup>172</sup> untersuchten eine AAV2-X<sub>7</sub>-Peptidbibliothek mit Insertion an der Position N587 durch die Sequenzierung von 80 zufällig ausgewählten Klonen. Im Fall der nicht selektierten Virusbibliothek fanden sie die Aminosäuren Glycin, Alanin und Prolin häufiger als erwartet, während die Aminosäuren Cystein, Phenylalanin, Tryptophan, Tyrosin und Leucin unterrepräsentiert waren<sup>172</sup>. Naumer et al.<sup>129</sup> untersuchten AAV2-NNK<sub>7</sub>-Peptidbibliotheken mit unterschiedlich langen Peptidinsertionen (5,12, 19 und 26 Aminosäuren) in der Position R588 (VP1). Auch sie fanden, dass die Aminosäuren Cystein, Phenylalanin und Tryptophan in den Virusbibliotheken leicht unterrepräsentiert waren<sup>129</sup>. Die Ergebnisse der vorliegenden Arbeit bestätigen und erweitern somit die Ergebnisse von Perabo et al.<sup>172</sup> und Naumer et al.<sup>129</sup>.

Aufgrund des NNK-Codierungsschemas war die Aminosäure Arginin in der NNK<sub>7</sub>-Plasmidbibliothek angereichert; in der NNK<sub>7</sub>-Virusbibliothek hingegen reicherte sich die Aminosäure mit allen drei Codons ab. Bezüglich der 2T<sub>6</sub>-Plasmidbibliothek war das Arginin leicht und in der 2T<sub>6</sub>-Virusbibliothek noch etwas stärker unterrepräsentiert. Die Aminosäure Asparagin war in beiden Plasmidbibliotheken unterrepräsentiert, hat sich in beiden Virusbibliotheken im Vergleich zu den Plasmidbibliotheken etwas angereichert, blieb aber insgesamt in den Virusbibliotheken unterrepräsentiert. Dennoch reicherten sich an den Positionen eins und vier der Peptidinsertionen der NNK<sub>7</sub>- und der 2T<sub>6</sub>-Virusbibliothek jeweils das Asparagin und das Arginin besonders unter den häufigsten Sequenzen

stark an. Die Anreicherung der Aminosäuren war somit nicht auf das gesamte Peptid verteilt, sondern im Wesentlichen auf bestimmte Positionen beschränkt. Folglich muss die Anreicherung dieser Aminosäuren entweder einen strukturellen Vorteil für das virale Kapsid bedeuten, oder aufgrund einer Vorselektion in der Produktionszelllinie HEK 293T zustande gekommen sein. Letzteres sollte aber aufgrund der Produktion von Transfer-Shuttles mit Mosaik-Kapsiden vor der finalen Virusproduktion verhindert werden<sup>122</sup>. Bei der Produktion dieser Transfer-Shuttles mittels Tripel-Transfektion werden die viralen Genome inklusive der randomisierten Peptidinsertion in *cis* und die genetische Information für das Codon-optimierte Wildtyp Kapsidprotein VP3 in *trans* zur Verfügung gestellt<sup>122,125</sup>. Die hierbei generierten Transfer-Shuttles zeigten eine Heparinbindung, die ca. 50 % der Bindung des Wildtypen entsprach, während Virusbibliotheken, die ohne Wildtyp-Helferplasmid produziert wurden eine Heparinbindung von ca. 12 % zeigten<sup>125</sup>. Hieraus wurde gefolgert, dass die auf diese Weise produzierten Transfer-Shuttles Mosaik-Kapside besitzen, die zum einen aus Proteinen der Bibliothek und zum anderen aus Wildtypprotein bestehen. Da HEK 293T Zellen permissiv für den AAV2 Wildtypen sind, können die Kapside der Transfer-Shuttles die Zellen über ihren Wildtypanteil binden und internalisiert werden. Die Wahrscheinlichkeit der Vorselektion in diesen Zellen sollte somit gering sein.

### 5.9 Die Sequenzmotive „NXXRXX<sup>E</sup>/<sub>X</sub>“ und „XXDXT<sup>R</sup>/<sub>X</sub>X“

Mit Hilfe der Bewertungskriterien konnte im Rahmen dieser Arbeit gezeigt werden, dass es sich bei dem Motiv „NXXRXX<sup>E</sup>/<sub>X</sub>“ im untersuchten Kontext um ein unspezifisches Motiv handelt. Sequenzen mit diesem Motiv wurden zwar unter den häufigsten Sequenzen aller untersuchten Selektionen gefunden, eine besondere Anreicherung oder gar Spezifität zeigten sie aber nicht. Gleichzeitig konnte gezeigt werden, dass das Motiv seinen Ursprung in der nicht-selektierten NNK<sub>7</sub>-Virusbibliothek hatte und das um eine Aminosäure verkürzte Motiv „NXXRXX“ auch in der 2T<sub>6</sub>-Virusbibliothek, trotz des Fehlens einer initialen Bevorteilung durch ein entsprechendes Codierungsschema, angereichert war.

Peptidsequenzen mit Sequenzmotiven wie „NXVXXXX“<sup>126</sup> oder „NXVRXXX“<sup>122,125</sup> und somit mit einem Asparagin an erster Position wurden zusätzlich zu den Selektionen dieser Arbeit auch *in vivo* in den ersten beiden Selektionsrunden in der Lunge gefunden<sup>147</sup> und in weiteren Selektionen *in vitro* selektiert. Sequenzen mit diesen Motiven wurden auch unter den zehn häufigsten Sequenzen aller drei NNK<sub>7</sub>-Virusbibliotheken gefunden und waren somit initial angereichert. Während das Motiv „NXVXXXX“ in der Selektion in der Lunge nicht komplett aber weitestgehend mit steigender Anzahl an Selektionsrunden verdrängt wurde, wurden Peptide mit den beschriebenen Motiven *in vitro* z.T. mit jeder Runde mehr angereichert und zeigten schließlich gegenüber Kontrollklonen oder dem Wildtypkapsid eine stärkere Transduktion bezüglich der Zielzellen<sup>122,125,126</sup>.

Die hohe Präsenz dieser Motive und auch des Motivs „NXXRXX<sup>E</sup>/<sub>X</sub>“ in den Selektionsrunden und besonders in den nicht selektierten Virusbibliotheken deutet auf einen unspezifischen generellen funktionellen Vorteil hin. Perabo et al.<sup>172</sup> trennten nicht-selektierte AAV2-Peptidbibliotheken mit sieben Aminosäure langen Peptidinsertionen an der Stelle N587 (VP1) in eine HSPG-bindende und eine nicht HSPG-bindende Fraktion auf und untersuchten sie mittels der Sequenzierung von jeweils 80 Einzelklonen. Hierbei fanden sie heraus, dass die Peptidsequenzen der HSPG-bindenden Fraktion, und somit der Kapsidvarianten die ähnliche Bindungseigenschaften wie das Wildtypkapsid aufwiesen, besonders häufig die Aminosäure Arginin trugen. Auch die Motive „NXXRXX<sup>E</sup>/<sub>X</sub>“ „NXVRXXX“ besitzen ein Arginin. Der generelle Vorteil dieses Motivs könnte somit in einer gewissen Ähnlichkeit mit dem AAV2-Wildtypkapsid begründet sein.

Da bestimmte Kombinationen von Aminosäuren in Zusammenhang mit diesen Motiven dennoch spezifisch sein könnten, sollten Peptide mit diesen Motiven nicht von vornherein von der Evaluation ausgeschlossen werden, wenn sie bezüglich der beschriebenen Bewertungskriterien Anzeichen von Spezifität aufweisen.

Im Gegensatz zu dem Motiv „NXXRXX<sup>E</sup>/<sub>X</sub>“ war das Motiv „XXDXT<sup>R</sup>/<sub>X</sub>X“ besonders im Pankreas und in den Langerhans-Inseln angereichert und einzelne das Motiv tragende Klone zeigten laut der *Scores* Spezifität. Alle diese Sequenzen haben sich durchschnittlich während der Selektion angereichert und alle, bis auf zwei, haben durchschnittlich im Zielorgan eine höhere relative Häufigkeit als in den Kontrollorganen. Dennoch waren gemäß der Berechnung des *C-Scores* auch hier nicht alle motivtragenden Sequenzen auch potenziell spezifisch.

Sechs der 30 häufigsten Sequenzen der Selektion in der Niere trugen ebenfalls das Motiv „XXDXT<sup>R</sup>/<sub>X</sub>X“. Diese reicherten sich z.T. sogar durchschnittlich an, waren aber im *C-Score* bezüglich der Niere allesamt unspezifisch. Diese Beobachtungen lassen vermuten, dass auch das Motiv „XXDXT<sup>R</sup>/<sub>X</sub>X“, wie höchstwahrscheinlich alle Sequenzmotive, für eine gewisse Stabilität für das Kapsid sorgt und manche Aminosäure-Kombinationen mit diesem Motiv unspezifisch sind, aber dennoch einen Vorteil haben und sich unspezifisch, besonders in der Niere (aus beschriebenen Gründen, **5.2.1**), anreichern können.

Das Beispiel zeigt des Weiteren, dass die unspezifische Amplifikation von unspezifischen Sequenzen z.B. während der Isolierung der viralen DNS aus den Zielgeweben mittels PCR oder auch während der Klonierung der sekundären Plasmidbibliotheken in *E. coli* und bei der Produktion der Virusbibliothek nicht unterschätzt werden sollten. Stabile Kapsidvarianten, die sich, wenn auch aufgrund unspezifischer Amplifikation, einmal angereichert haben, sind schwierig wieder aus der Bibliothek zu eliminieren und könnten andere Kapsidvarianten verdrängen.

Dennoch besteht zwischen den Sequenzmotiven „NXXRXX<sup>E</sup>/<sub>X</sub>“ und „XXDXT<sup>R</sup>/<sub>X</sub>X“ mindestens ein entscheidender Unterschied. Während das Motiv „NXXRXX<sup>E</sup>/<sub>X</sub>“ eindeutig in den nicht selektierten Virusbibliotheken angereichert war und somit einen initialen quantitativen Vorteil hatte, konnte, wie für das im Gehirn selektierte Motiv „XDGXXWX“, keine besondere Anreicherung des im Pankreas selektierten Motivs „XXDXT<sup>R</sup>/<sub>X</sub>X“ in diesen Bibliotheken festgestellt werden. Zusätzlich zeigte das Motiv „XXDXT<sup>R</sup>/<sub>X</sub>X“ wie auch das Motiv „XDGXXWX“ eine starke Anreicherung im Zielorgan und wurde somit tatsächlich selektiert. Außerdem konnte anhand der Evaluation des selektierten Klons rAAV2-GDDATRD eine zum Zielorgan Pankreas verschobene Vektorverteilung gezeigt werden.

### 5.10 Praktischer Vergleich der Peptidbibliotheken AAV2-NNK<sub>7</sub> und AAV2-Trimer<sub>6</sub>-(-C) mittels *in vitro*-Selektion

Bevor eine neue Peptidbibliothek *in vivo* verwendet werden kann, sollte sie *in vitro* im Hinblick darauf getestet werden, ob selektierte Kapsidvarianten die Zielzellen besser transduzieren als zufällig aus der nicht selektierten Bibliothek gewählte Klone. Um zusätzlich herauszufinden, ob die neue Peptidbibliothek gegenüber der NNK<sub>7</sub>-Peptidbibliothek einen praktischen Vorteil hat, wurden beide Bibliotheken *in vitro* auf den gleichen vier Zelllinien über fünf Selektionsrunden selektiert. Bei einer *in vitro*-Selektion findet zwar keine Negativselektion durch andere Gewebe statt, dennoch gibt es zusätzlich zur Selektion aufgrund der Rezeptorbindung einen weiteren Selektionsdruck. Durch die Überinfektion mit replikationskompetenten Adenoviren, einem Helfervirus des AAV, können auch die AAV der Bibliothek replizieren. Entsprechend werden die Viren im Gegensatz zur *in vivo*-Selektion amplifiziert und durchlaufen den gesamten Prozess der Transduktion und Replikation.

Die jeweils häufigsten Klone, bzw. die die sich am stärksten anreicherten, wurden anschließend als rekombinante Vektoren mit Luciferase als Reportergen auf den entsprechenden Zelllinien hinsichtlich der Luciferaseexpression und somit ihrer Transduktion getestet. Hierbei zeigte sich, dass sich mit der 2T<sub>6</sub>-Bibliothek wie mit der NNK<sub>7</sub>-Bibliothek Klone selektieren ließen, die die Zielzellen wesentlich besser als zufällig gewählte Klone transduzierten. Des Weiteren zeigte einmal ein Klon aus der NNK<sub>7</sub>-Bibliothek und einmal ein Klon aus der 2T<sub>6</sub>-Bibliothek die stärkste Transduktion, einmal war der Unterschied zwischen Klonen der beiden Bibliotheken nicht signifikant und einmal zeigte das Wildtypkapsid die stärkste Luciferaseexpression. Dementsprechend sind weder die Klone der einen noch der anderen Bibliothek besser oder schlechter. Wie Naumer et al. aber bereits bei dem Test von Bibliotheken mit verschiedenen langen Peptidinsertionen beschrieben haben, kann es aber sein, dass Peptide unterschiedlicher Bibliotheken verschiedene Zellrezeptoren des gleichen Zelltyps binden<sup>129</sup>. Des Weiteren können Vektoren der neuen 2T<sub>6</sub>-Bibliothek evtl. Rezeptoren von Zellen binden, die

bisher nicht transduziert werden konnten und somit das Zielspektrum der bisherigen Bibliotheken erweitern.

Zusätzlich wurde beobachtet, dass die das Motiv „NXXRXX“ tragende Sequenz „NPQRAA“ zwar unter den zehn häufigsten Sequenzen aller vier Zelllinien gefunden wurde, sich aber nur bei einer Selektion mit großem Abstand durchsetzte und beim Test der Transduktion die stärkste Reporterexpression zeigte. Entsprechend können die das Motiv „NXXRXX“ tragenden Sequenzen zumindest bei *in vitro*-Anwendungen von Vorteil sein.

### **5.11 Ausblick**

Im Rahmen dieser Arbeit wurden AAV2-Peptidbibliotheken *in vivo* in Mäusen nach systemischer Injektion selektiert und die hierbei generierten sekundären Peptidbibliotheken, sowie die primären nicht-selektierten Plasmid- und Virusbibliotheken im Hinblick auf die Anreicherung, Diversität und Spezifität von Sequenzmotiven und Peptidsequenzen mittels NGS untersucht. Zusammenfassend wurde hierbei festgestellt, dass bei der Selektion von AAV-Peptidbibliotheken nicht nur die Selektion für ein spezifisches Peptid, sondern auch die Gegenselektion gegen verschiedene Ungleichverteilungen eine große Rolle spielen. So muss nicht nur gegen die Ungleichverteilung einiger Aminosäuren aufgrund des NNK-Codierungsschemas, sondern auch gegen eine weitere Ungleichverteilung der Nukleotide auf Plasmid-Ebene selektiert werden. Während diese Ungleichverteilungen durch die Verwendung einer Trimer codierten Plasmidbibliothek umgangen werden können, ist die wichtigste Gegenselektion aber die Selektion gegen bereits in der nicht-selektierten Bibliothek angereicherte Sequenzen und Sequenzmotive. Dennoch war es in der Vergangenheit möglich, gewebespezifisch transduzierende AAV-Vektoren *in vivo* zu selektieren. Den Selektionen in der Lunge und dem Gehirn waren hierbei eine starke Peptidanreicherung und somit eine starke Gegenselektion gemein. Aber auch mit einer weniger starken Peptidanreicherung ist es in der vorliegenden Arbeit gelungen, einen AAV-Klon zu identifizieren, der in der untersuchten Selektion die höchste Wahrscheinlichkeit auf Spezifität besaß und sich auch praktisch im Zielorgan Pankreas der Maus nach systemischer Injektion anreicherte. Diese Kapsidvariante hat somit neben der Voraussetzung eines stabilen Kapsids eine weitere Voraussetzung eines spezifischen AAV-Vektors erfüllt und könnte mittels gewebespezifischer Promotoren auch bezüglich der Genexpression spezifisch gemacht und als solcher Vektor eingesetzt werden.

Durch die Produktion einer AAV-Peptidbibliothek mittels Transfer-Shuttles und anschließender Virusproduktion durch Ad5 Überinfektion lassen sich Virusbibliotheken mit hohen Titern von  $1 \times 10^{12}$  vG produzieren. Eine Möglichkeit, Virusbibliotheken herzustellen ohne den Zwischenschritt der Produktion von Transfer-Shuttles durchzuführen, ist die Produktion einer Virusbibliothek mittels Transfektion. Um die Genom-Kapsid Korrelation zu wahren, werden hierbei sehr geringe Mengen des



Bibliotheksplasmid eingesetzt (ein Plasmid pro Zelle). Auf diese Weise lassen sich allerdings nur Virusbibliotheken mit relativ geringem Titer (ca.  $1 \times 10^9$  vG) produzieren. Da nicht gänzlich ausgeschlossen werden kann, dass die starke Anreicherung des Sequenzmotivs „NXXRXX(<sup>E</sup>/x)“ in den nicht-selektierten Virusbibliotheken aufgrund einer gewissen Vorselektion während der Produktion der Virusbibliothek in der Produktionszelllinie HEK 293T zustande gekommen ist, sollen demnächst Peptidbibliotheken ohne den Zwischenschritt der Transfer-Shuttles nur mittels Transfektion produziert und mittels NGS auf das Vorhandensein des Sequenzmotivs untersucht werden.

Zusätzlich wurde mit der 2T<sub>6</sub>-Peptidbibliothek eine neue Bibliothek vorgestellt und im Zuge dessen die Vorteile einer Trimer codierten Peptidbibliothek verdeutlicht. Diese Bibliothek wurde *in vitro* getestet und Tropismus modifizierte Kapsidvarianten konnten selektiert werden. Entsprechend konnte die 2T<sub>6</sub>-Peptidbibliothek nun für *in vivo*-Selektionen eingesetzt werden. Aufgrund der erfolgreichen Selektion des vielversprechenden gehirnspezifischen AAV2-Vektors in der Maus, wird zurzeit die Übertragbarkeit der AAV-Peptidselektionsmethode in einem dem Menschen ähnlicheren Organismus, dem Weißbüscheläffchen (*Callithrix jacchus*), untersucht. Da die 2T<sub>6</sub>-Peptidbibliothek eine Ergänzung zur bereits eingehender untersuchten NNK<sub>7</sub>-Peptidbibliothek darzustellen scheint, werden hierbei beide Bibliotheken co-selektiert. Die in der vorliegenden Arbeit gewonnenen Erkenntnisse bezüglich NGS-Probenvorbereitung und Auswertung der Daten finden hier bereits Anwendung.

## 6 Literaturverzeichnis

1. Xu, Q., Wang, C.-H. & Pack, D. W. Polymeric Carriers for Gene Delivery: Chitosan and Poly(amidoamine) Dendrimers. *Curr Pharm Des* **16**, 2350–2368 (2010).
2. Chen, W., Li, H., Liu, Z. & Yuan, W. Lipopolyplex for Therapeutic Gene Delivery and Its Application for the Treatment of Parkinson's Disease. *Front Aging Neurosci* **8**, 68 (2016).
3. Miller, A. D. Cationic liposome systems in gene therapy. *IDrugs* **1**, 574–583 (1998).
4. Durzyńska, J. *et al.* Viral and other cell-penetrating peptides as vectors of therapeutic agents in medicine. *J. Pharmacol. Exp. Ther.* **354**, 32–42 (2015).
5. Fernandes, A. R. & Chari, D. M. Part I: Minicircle vector technology limits DNA size restrictions on ex vivo gene delivery using nanoparticle vectors: Overcoming a translational barrier in neural stem cell therapy. *J Control Release* (2016). doi:10.1016/j.jconrel.2016.06.024
6. Limberis, M. P. Phoenix rising: gene therapy makes a comeback. *Acta Biochim. Biophys. Sin. (Shanghai)* **44**, 632–640 (2012).
7. Gene Therapy Clinical Trials Worldwide. (2016).
8. Blaese, R. M. *et al.* T lymphocyte-directed gene therapy for ADA- SCID: initial trial results after 4 years. *Science* **270**, 475–480 (1995).
9. Flotte, T. R. & Carter, B. J. Adeno-associated virus vectors for gene therapy of cystic fibrosis. *Meth. Enzymol.* **292**, 717–732 (1998).
10. Kay, M. A. *et al.* Evidence for gene transfer and expression of factor IX in haemophilia B patients treated with an AAV vector. *Nat. Genet.* **24**, 257–261 (2000).
11. Mandel, R. J., Rendahl, K. G., Snyder, R. O. & Leff, S. E. Progress in direct striatal delivery of L-dopa via gene therapy for treatment of Parkinson's disease using recombinant adeno-associated viral vectors. *Exp. Neurol.* **159**, 47–64 (1999).
12. Michelfelder, S. & Trepel, M. in *Advances in Genetics* (ed. Renata Pasqualini and Wadih Arap) **Volume 67**, 29–60 (Academic Press, 2009).
13. Atchison, R. W., Casto, B. C. & Hammon, W. M. ADENOVIRUS-ASSOCIATED DEFECTIVE VIRUS PARTICLES. *Science* **149**, 754–756 (1965).
14. Hoggan, M. D. Presence of small virus-like particles in various adenovirus type 2, 5, 7 and 12 preparations. *Federation Proc.* **24**, (1965).

15. 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. U.S.A.* **55**, 1467–1474 (1966).
16. Berns, K. I. Parvovirus replication. *Microbiol. Rev.* **54**, 316–329 (1990).
17. Gonçalves, M. A. F. V. Adeno-associated virus: from defective virus to effective vector. *Virology* **2**, 43 (2005).
18. Gao, G. *et al.* Clades of Adeno-associated viruses are widely disseminated in human tissues. *J. Virol.* **78**, 6381–6388 (2004).
19. Mori, S., Wang, L., Takeuchi, T. & Kanda, T. Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsid protein. *Virology* **330**, 375–383 (2004).
20. Wu, Z., Asokan, A. & Samulski, R. J. Adeno-associated virus serotypes: vector toolkit for human gene therapy. *Mol. Ther.* **14**, 316–327 (2006).
21. 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).
22. Salganik, M., Hirsch, M. L. & Samulski, R. J. Adeno-associated Virus as a Mammalian DNA Vector. *Microbiol Spectr* **3**, (2015).
23. Lisowski, L., Tay, S. S. & Alexander, I. E. Adeno-associated virus serotypes for gene therapeutics. *Curr Opin Pharmacol* **24**, 59–67 (2015).
24. Gao, G.-P. *et al.* Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc Natl Acad Sci U S A* **99**, 11854–11859 (2002).
25. Asokan, A., Schaffer, D. V. & Samulski, R. J. The AAV vector toolkit: poised at the clinical crossroads. *Mol. Ther.* **20**, 699–708 (2012).
26. Mayor, H. D. & Melnick, J. L. Small deoxyribonucleic acid-containing viruses (picodnavirus group). *Nature* **210**, 331–332 (1966).
27. Rose, J. A., Berns, K. I., Hoggan, M. D. & Koczot, F. J. Evidence for a single-stranded adenovirus-associated virus genome: formation of a DNA density hybrid on release of viral DNA. *Proc. Natl. Acad. Sci. U.S.A.* **64**, 863–869 (1969).
28. Srivastava, A., Lusby, E. W. & Berns, K. I. Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J. Virol.* **45**, 555–564 (1983).

29. Sonntag, F., Schmidt, K. & Kleinschmidt, J. A. A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc. Natl. Acad. Sci. U.S.A.* **107**, 10220–10225 (2010).
30. Lusby, E., Fife, K. H. & Berns, K. I. Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. *J. Virol.* **34**, 402–409 (1980).
31. 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. U.S.A.* **70**, 215–219 (1973).
32. Laughlin, C. A., Westphal, H. & Carter, B. J. Spliced adenovirus-associated virus RNA. *Proc Natl Acad Sci U S A* **76**, 5567–5571 (1979).
33. Marcus, C. J., Laughlin, C. A. & Carter, B. J. Adeno-associated virus RNA transcription in vivo. *Eur. J. Biochem.* **121**, 147–154 (1981).
34. 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).
35. Im, D. S. & Muzyczka, N. The AAV origin binding protein Rep68 is an ATP-dependent site-specific endonuclease with DNA helicase activity. *Cell* **61**, 447–457 (1990).
36. Im, D. S. & Muzyczka, N. Partial purification of adeno-associated virus Rep78, Rep52, and Rep40 and their biochemical characterization. *J. Virol.* **66**, 1119–1128 (1992).
37. Smith, R. H. & Kotin, R. M. The Rep52 gene product of adeno-associated virus is a DNA helicase with 3'-to-5' polarity. *J. Virol.* **72**, 4874–4881 (1998).
38. Collaco, R. F., Kalman-Maltese, V., Smith, A. D., Dignam, J. D. & Trempe, J. P. A biochemical characterization of the adeno-associated virus Rep40 helicase. *J. Biol. Chem.* **278**, 34011–34017 (2003).
39. Chejanovsky, N. & Carter, B. J. Mutagenesis of an AUG codon in the adeno-associated virus rep gene: effects on viral DNA replication. *Virology* **173**, 120–128 (1989).
40. 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).
41. 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).

42. Vogel, R., Seyffert, M., Pereira, B. de A. & Fraefel, C. Viral and Cellular Components of AAV2 Replication Compartments. *Open Virol J* **7**, 98–120 (2013).
43. Rose, J. A., Maizel, J. V., Inman, J. K. & Shatkin, A. J. Structural proteins of adenovirus-associated viruses. *J. Virol.* **8**, 766–770 (1971).
44. 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).
45. B&uuml;ning, H., Braun-Falco, M. & Hallek, M. Progress in the Use of Adeno-Associated Viral Vectors for Gene Therapy. *Cells Tissues Organs* **177**, 139–150 (2004).
46. Johnson, F. B., Ozer, H. L. & Hoggan, M. D. Structural proteins of adenovirus-associated virus type 3. *J. Virol.* **8**, 860–863 (1971).
47. 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).
48. Xie, Q. *et al.* The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *PNAS* **99**, 10405–10410 (2002).
49. Kaludov, N. *et al.* Production, purification and preliminary X-ray crystallographic studies of adeno-associated virus serotype 4. *Virology* **306**, 1–6 (2003).
50. 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).
51. 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).
52. Nam, H.-J. *et al.* Structure of Adeno-Associated Virus Serotype 8, a Gene Therapy Vector. *J Virol* **81**, 12260–12271 (2007).
53. 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).
54. Walters, R. W. *et al.* Structure of Adeno-Associated Virus Serotype 5. *J Virol* **78**, 3361–3371 (2004).

55. 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).
56. Kern, A. *et al.* Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. *J. Virol.* **77**, 11072–11081 (2003).
57. Opie, S. R., Warrington, Jr., K. H., Agbandje-McKenna, M., Zolotukhin, S. & 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).
58. Girod, A. *et al.* Genetic capsid modifications allow efficient re-targeting of adeno-associated virus type 2. *Nat. Med.* **5**, 1052–1056 (1999).
59. Grifman, M. *et al.* Incorporation of tumor-targeting peptides into recombinant adeno-associated virus capsids. *Mol. Ther.* **3**, 964–975 (2001).
60. Nicklin, S. A. & Baker, A. H. Tropism-modified adenoviral and adeno-associated viral vectors for gene therapy. *Curr Gene Ther* **2**, 273–293 (2002).
61. 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).
62. 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).
63. 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).
64. Kashiwakura, Y. *et al.* Hepatocyte growth factor receptor is a coreceptor for adeno-associated virus type 2 infection. *J. Virol.* **79**, 609–614 (2005).
65. 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).
66. 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).
67. Nonnenmacher, M. & Weber, T. Intracellular transport of recombinant adeno-associated virus vectors. *Gene Ther.* **19**, 649–658 (2012).
68. Weinberg, M. S. *et al.* Recombinant adeno-associated virus utilizes cell-specific infectious entry mechanisms. *J. Virol.* **88**, 12472–12484 (2014).

69. Di Pasquale, G. & Chiorini, J. A. AAV transcytosis through barrier epithelia and endothelium. *Mol. Ther.* **13**, 506–516 (2006).
70. Sanlioglu, S. *et al.* Endocytosis and Nuclear Trafficking of Adeno-Associated Virus Type 2 Are Controlled by Rac1 and Phosphatidylinositol-3 Kinase Activation. *J Virol* **74**, 9184–9196 (2000).
71. 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).
72. 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).
73. 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).
74. 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).
75. 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).
76. Akache, B. *et al.* A two-hybrid screen identifies cathepsins B and L as uncoating factors for adeno-associated virus 2 and 8. *Mol. Ther.* **15**, 330–339 (2007).
77. 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).
78. Xiao, W., Warrington, K. H., Hearing, P., Hughes, J. & Muzyczka, N. Adenovirus-facilitated nuclear translocation of adeno-associated virus type 2. *J. Virol.* **76**, 11505–11517 (2002).
79. 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).

80. Casto, B. C., Atchison, R. W. & Hammon, W. M. Studies on the relationship between adeno-associated virus type I (AAV-1) and adenoviruses. I. Replication of AAV-1 in certain cell cultures and its effect on helper adenovirus. *Virology* **32**, 52–59 (1967).
81. Buller, R. M., Janik, J. E., Sebring, E. D. & Rose, J. A. Herpes simplex virus types 1 and 2 completely help adenovirus-associated virus replication. *J. Virol.* **40**, 241–247 (1981).
82. Samulski, R. J. *et al.* Targeted integration of adeno-associated virus (AAV) into human chromosome 19. *EMBO J* **10**, 3941–3950 (1991).
83. Kotin, R. M., Linden, R. M. & Berns, K. I. Characterization of a preferred site on human chromosome 19q for integration of adeno-associated virus DNA by non-homologous recombination. *EMBO J* **11**, 5071–5078 (1992).
84. Schnepf, B. C., Jensen, R. L., Chen, C.-L., Johnson, P. R. & Clark, K. R. Characterization of Adeno-Associated Virus Genomes Isolated from Human Tissues. *J Virol* **79**, 14793–14803 (2005).
85. Hamilton, H., Gomos, J., Berns, K. I. & Falck-Pedersen, E. Adeno-Associated Virus Site-Specific Integration and AAVS1 Disruption. *J Virol* **78**, 7874–7882 (2004).
86. Janik, J. E., Huston, M. M. & Rose, J. A. Locations of adenovirus genes required for the replication of adenovirus-associated virus. *Proc Natl Acad Sci U S A* **78**, 1925–1929 (1981).
87. Richardson, W. D. & Westphal, H. A cascade of adenovirus early functions is required for expression of adeno-associated virus. *Cell* **27**, 133–141 (1981).
88. Samulski, R. J. & Shenk, T. Adenovirus E1B 55-Mr polypeptide facilitates timely cytoplasmic accumulation of adeno-associated virus mRNAs. *J. Virol.* **62**, 206–210 (1988).
89. Ni, T.-H. *et al.* Cellular Proteins Required for Adeno-Associated Virus DNA Replication in the Absence of Adenovirus Coinfection. *J. Virol.* **72**, 2777–2787 (1998).
90. Yalkinoglu, A. O., Heilbronn, R., Bürkle, A., Schlehofer, J. R. & zur Hausen, H. DNA amplification of adeno-associated virus as a response to cellular genotoxic stress. *Cancer Res.* **48**, 3123–3129 (1988).
91. Tattersall, P. & Ward, D. C. Rolling hairpin model for replication of parvovirus and linear chromosomal DNA. , *Published online: 09 September 1976; | doi:10.1038/263106a0* **263**, 106–109 (1976).



92. Nash, K., Chen, W. & Muzyczka, N. Complete In Vitro Reconstitution of Adeno-Associated Virus DNA Replication Requires the Minichromosome Maintenance Complex Proteins. *J Virol* **82**, 1458–1464 (2008).
93. Ryan, J. H., Zolotukhin, S. & Muzyczka, N. Sequence requirements for binding of Rep68 to the adeno-associated virus terminal repeats. *J. Virol.* **70**, 1542–1553 (1996).
94. Hunter, L. A. & Samulski, R. J. Colocalization of adeno-associated virus Rep and capsid proteins in the nuclei of infected cells. *J Virol* **66**, 317–324 (1992).
95. 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).
96. Bleker, S., Pawlita, M. & Kleinschmidt, J. A. Impact of Capsid Conformation and Rep-Capsid Interactions on Adeno-Associated Virus Type 2 Genome Packaging. *J Virol* **80**, 810–820 (2006).
97. Berns, K. I. & Adler, S. Separation of two types of adeno-associated virus particles containing complementary polynucleotide chains. *J. Virol.* **9**, 394–396 (1972).
98. Tratschin, J. D., West, M. H., Sandbank, T. & Carter, B. J. A human parvovirus, adeno-associated virus, as a eucaryotic vector: transient expression and encapsidation of the procaryotic gene for chloramphenicol acetyltransferase. *Mol. Cell. Biol.* **4**, 2072–2081 (1984).
99. Hermonat, P. L. & Muzyczka, N. Use of adeno-associated virus as a mammalian DNA cloning vector: transduction of neomycin resistance into mammalian tissue culture cells. *Proc. Natl. Acad. Sci. U.S.A.* **81**, 6466–6470 (1984).
100. Tratschin, J. D., Miller, I. L., Smith, M. G. & Carter, B. J. Adeno-associated virus vector for high-frequency integration, expression, and rescue of genes in mammalian cells. *Mol. Cell. Biol.* **5**, 3251–3260 (1985).
101. Flotte, T. *et al.* A phase I study of an adeno-associated virus-CFTR gene vector in adult CF patients with mild lung disease. *Hum. Gene Ther.* **7**, 1145–1159 (1996).
102. Carter, B. J. Adeno-associated virus and the development of adeno-associated virus vectors: a historical perspective. *Mol. Ther.* **10**, 981–989 (2004).
103. Peng, Z. Current status of gendicine in China: recombinant human Ad-p53 agent for treatment of cancers. *Hum. Gene Ther.* **16**, 1016–1027 (2005).

104. van der Loo, J. C. M. & Wright, J. F. Progress and challenges in viral vector manufacturing. *Hum. Mol. Genet.* **25**, R42-52 (2016).
105. Burnett, J. R. & Hooper, A. J. Alipogene tiparovec, an adeno-associated virus encoding the Ser(447)X variant of the human lipoprotein lipase gene for the treatment of patients with lipoprotein lipase deficiency. *Curr. Opin. Mol. Ther.* **11**, 681–691 (2009).
106. 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**, 3096–3101 (1987).
107. Zolotukhin, S., Potter, M., Hauswirth, W. W., Guy, J. & Muzyczka, N. A ‘humanized’ green fluorescent protein cDNA adapted for high-level expression in mammalian cells. *J Virol* **70**, 4646–4654 (1996).
108. Grimm, D., Kern, A., Rittner, K. & Kleinschmidt, J. A. Novel tools for production and purification of recombinant adenoassociated virus vectors. *Hum. Gene Ther.* **9**, 2745–2760 (1998).
109. Xiao, X., Li, J. & Samulski, R. J. Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. *J Virol* **72**, 2224–2232 (1998).
110. Smith-Arica, J. R. *et al.* Infection efficiency of human and mouse embryonic stem cells using adenoviral and adeno-associated viral vectors. *Cloning Stem Cells* **5**, 51–62 (2003).
111. Sen, S. *et al.* Gene delivery to the vasculature mediated by low-titre adeno-associated virus serotypes 1 and 5. *J Gene Med* **10**, 143–151 (2008).
112. Müller, O. J. *et al.* Improved cardiac gene transfer by transcriptional and transductional targeting of adeno-associated viral vectors. *Cardiovasc Res* **70**, 70–78 (2006).
113. Zincarelli, C., Soltys, S., Rengo, G. & Rabinowitz, J. E. Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol. Ther.* **16**, 1073–1080 (2008).
114. Kwon, I. & Schaffer, D. V. Designer gene delivery vectors: molecular engineering and evolution of adeno-associated viral vectors for enhanced gene transfer. *Pharm. Res.* **25**, 489–499 (2008).
115. Bartlett, J. S., Kleinschmidt, J., Boucher, R. C. & Samulski, R. J. Targeted adeno-associated virus vector transduction of nonpermissive cells mediated by a bispecific F(ab’gamma)2 antibody. *Nat. Biotechnol.* **17**, 181–186 (1999).

116. Ponnazhagan, S., Mahendra, G., Kumar, S., Thompson, J. A. & Castillas, M. J. Conjugate-based targeting of recombinant adeno-associated virus type 2 vectors by using avidin-linked ligands. *J. Virol.* **76**, 12900–12907 (2002).
117. Huttner, N. A. *et al.* Genetic modifications of the adeno-associated virus type 2 capsid reduce the affinity and the neutralizing effects of human serum antibodies. *Gene Therapy* **10**, 2139–2147 (2003).
118. Wu, P. *et al.* Mutational Analysis of the Adeno-Associated Virus Type 2 (AAV2) Capsid Gene and Construction of AAV2 Vectors with Altered Tropism. *J Virol* **74**, 8635–8647 (2000).
119. Trepel, M., Arap, W. & Pasqualini, R. Exploring vascular heterogeneity for gene therapy targeting. *Gene Ther.* **7**, 2059–2060 (2000).
120. Trepel, M., Arap, W. & Pasqualini, R. In vivo phage display and vascular heterogeneity: implications for targeted medicine. *Curr Opin Chem Biol* **6**, 399–404 (2002).
121. Arap, W. *et al.* Steps toward mapping the human vasculature by phage display. *Nat. Med.* **8**, 121–127 (2002).
122. 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).
123. Bupp, K. & Roth, M. J. Targeting a Retroviral Vector in the Absence of a Known Cell-Targeting Ligand. *Human Gene Therapy* **14**, 1557–1564 (2003).
124. Perabo, L. *et al.* In vitro selection of viral vectors with modified tropism: the adeno-associated virus display. *Mol. Ther.* **8**, 151–157 (2003).
125. Waterkamp, D. A., Müller, 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).
126. Michelfelder, S. *et al.* Vectors selected from adeno-associated viral display peptide libraries for leukemia cell-targeted cytotoxic gene therapy. *Exp. Hematol.* **35**, 1766–1776 (2007).
127. Sellner, L. *et al.* Generation of efficient human blood progenitor-targeted recombinant adeno-associated viral vectors (AAV) by applying an AAV random peptide library on primary human hematopoietic progenitor cells. *Exp. Hematol.* **36**, 957–964 (2008).
128. Stiefelhagen, M. *et al.* Application of a haematopoietic progenitor cell-targeted adeno-associated viral (AAV) vector established by selection of an AAV random peptide library on a leukaemia cell line. *Genet Vaccines Ther* **6**, 12 (2008).

129. Naumer, M. *et al.* Development and validation of novel AAV2 random libraries displaying peptides of diverse lengths and at diverse capsid positions. *Hum. Gene Ther.* **23**, 492–507 (2012).
130. Durr, E. *et al.* Direct proteomic mapping of the lung microvascular endothelial cell surface in vivo and in cell culture. *Nat. Biotechnol.* **22**, 985–992 (2004).
131. Smith, T. L. *et al.* AAVP displaying octreotide for ligand-directed therapeutic transgene delivery in neuroendocrine tumors of the pancreas. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 2466–2471 (2016).
132. 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**, 980–990 (2010).
133. 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).
134. Körbelin, J. *et al.* A brain microvasculature endothelial cell-specific viral vector with the potential to treat neurovascular and neurological diseases. *EMBO Mol Med* **8**, 609–625 (2016).
135. Körbelin, J. *et al.* Pulmonary Targeting of Adeno-associated Viral Vectors by Next-generation Sequencing-guided Screening of Random Capsid Displayed Peptide Libraries. *Molecular Therapy* (2016). doi:10.1038/mt.2016.62
136. Chen, S. *et al.* Gene delivery in renal tubular epithelial cells using recombinant adeno-associated viral vectors. *J. Am. Soc. Nephrol.* **14**, 947–958 (2003).
137. Favre, D., Ferry, N. & Moullier, P. Critical aspects of viral vectors for gene transfer into the kidney. *J. Am. Soc. Nephrol.* **11 Suppl 16**, S149-153 (2000).
138. Kapturczak, M. H., Chen, S. & Agarwal, A. Adeno-associated virus vector-mediated gene delivery to the vasculature and kidney. *Acta Biochim. Pol.* **52**, 293–299 (2005).
139. Lipkowitz, M. S. *et al.* Transduction of renal cells in vitro and in vivo by adeno-associated virus gene therapy vectors. *J. Am. Soc. Nephrol.* **10**, 1908–1915 (1999).
140. Akbulut, T. & Park, F. Gene therapy to the kidney using viral vectors. *Paidiatrike* **71**, 177–185 (2008).
141. Flotte, T. *et al.* Efficient ex vivo transduction of pancreatic islet cells with recombinant adeno-associated virus vectors. *Diabetes* **50**, 515–520 (2001).

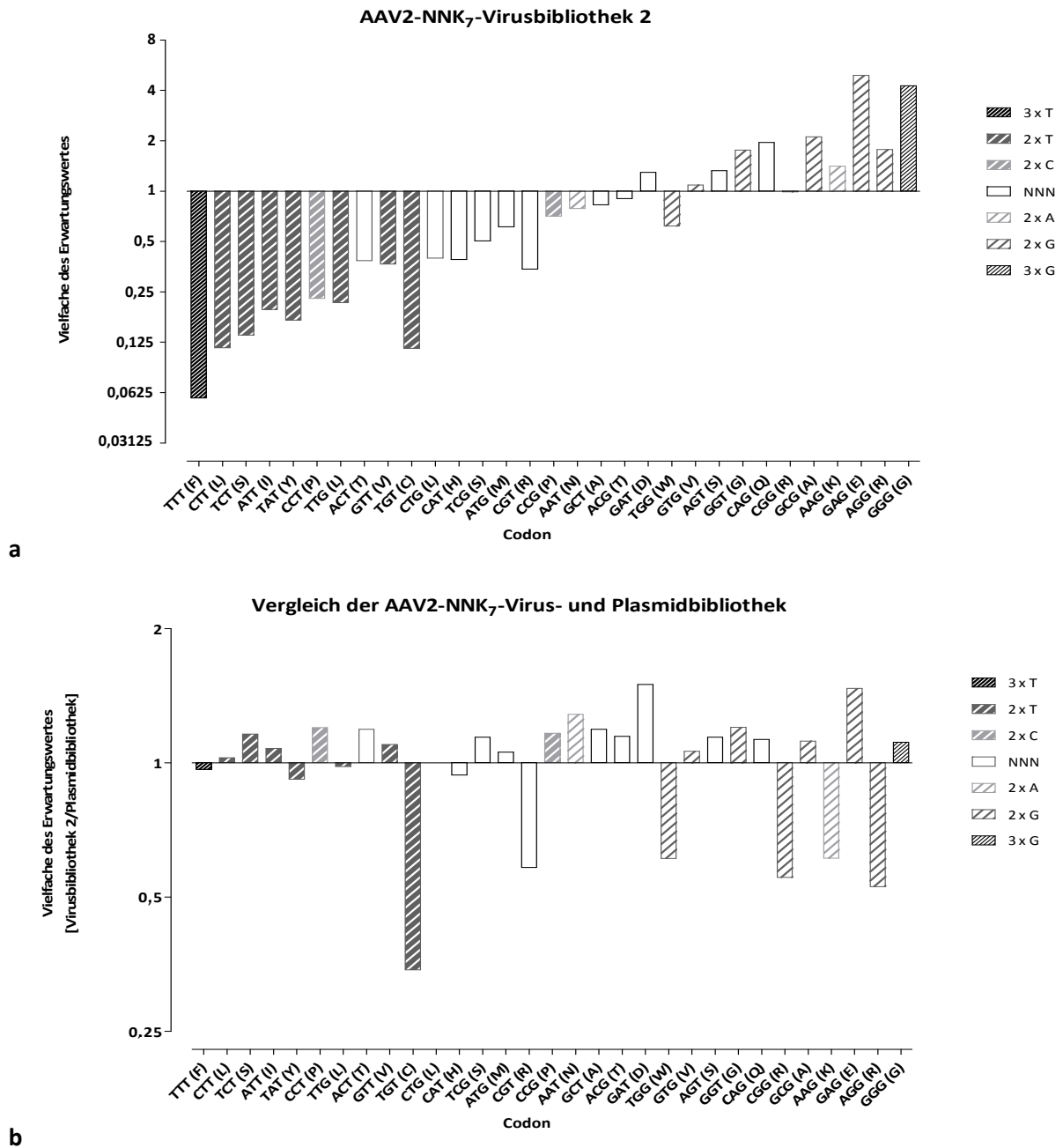
142. Rehman, K. K. *et al.* Efficient gene delivery to human and rodent islets with double-stranded (ds) AAV-based vectors. *Gene Ther.* **12**, 1313–1323 (2005).
143. Wang, Z. *et al.* Widespread and stable pancreatic gene transfer by adeno-associated virus vectors via different routes. *Diabetes* **55**, 875–884 (2006).
144. Maxam, A. M. & Gilbert, W. A new method for sequencing DNA. *Proc. Natl. Acad. Sci. U.S.A.* **74**, 560–564 (1977).
145. Sanger, F. & Coulson, A. R. A rapid method for determining sequences in DNA by primed synthesis with DNA polymerase. *J. Mol. Biol.* **94**, 441–448 (1975).
146. van Dijk, E. L., Auger, H., Jaszczyszyn, Y. & Thermes, C. Ten years of next-generation sequencing technology. *Trends Genet.* **30**, 418–426 (2014).
147. Körbelin, J. Generation of tissue-specific vectors by in vivo selection of random peptide libraries displayed on the surface of adeno associated virus type 2. (Dissertation, Universität Hamburg, Fachbereich Biologie). (2013).
148. 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).
149. Hanahan, D. Studies on transformation of *Escherichia coli* with plasmids. *Journal of Molecular Biology* **166**, 557–580 (1983).
150. Bullock, W.O., Fernandez, J. M. an Short, J. M. XI1-Blue: A High Efficiency Plasmid Transforming *recA* *Escherichia coli* Strain With Beta-Galactosidase Selection. **5**, 376–378 (1987).
151. Graham, F. L., Smiley, J., Russell, W. C. & Nairn, R. Characteristics of a human cell line transformed by DNA from human adenovirus type 5. *J. Gen. Virol.* **36**, 59–74 (1977).
152. 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).
153. Aslakson, C. J. & Miller, F. R. Selective events in the metastatic process defined by analysis of the sequential dissemination of subpopulations of a mouse mammary tumor. *Cancer Res.* **52**, 1399–1405 (1992).
154. Fogh, J., Wright, W. C. & Loveless, J. D. Absence of HeLa cell contamination in 169 cell lines derived from human tumors. *J. Natl. Cancer Inst.* **58**, 209–214 (1977).

155. Soule, H. D., Vazquez, 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).
156. Kalinina, T. *et al.* Establishment and characterization of a new human pancreatic adenocarcinoma cell line with high metastatic potential to the lung. *BMC Cancer* **10**, 295 (2010).
157. Rohr, U.-P. *et al.* Fast and reliable titration of recombinant adeno-associated virus type-2 using quantitative real-time PCR. *J. Virol. Methods* **106**, 81–88 (2002).
158. Taketo, M. *et al.* FVB/N: an inbred mouse strain preferable for transgenic analyses. *Proc. Natl. Acad. Sci. U.S.A.* **88**, 2065–2069 (1991).
159. Xiao, X., Li, J. & Samulski, R. J. Production of High-Titer Recombinant Adeno-Associated Virus Vectors in the Absence of Helper Adenovirus. *J Virol* **72**, 2224–2232 (1998).
160. Lacy, P. E. & Kostianovsky, M. Method for the isolation of intact islets of Langerhans from the rat pancreas. *Diabetes* **16**, 35–39 (1967).
161. Hermens, W. T. *et al.* Purification of recombinant adeno-associated virus by iodixanol gradient ultracentrifugation allows rapid and reproducible preparation of vector stocks for gene transfer in the nervous system. *Hum. Gene Ther.* **10**, 1885–1891 (1999).
162. 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).
163. Takeda, S. *et al.* Successful gene transfer using adeno-associated virus vectors into the kidney: comparison among adeno-associated virus serotype 1-5 vectors in vitro and in vivo. *Nephron Exp. Nephrol.* **96**, e119-126 (2004).
164. Ito, K. *et al.* Adeno-associated viral vector transduction of green fluorescent protein in kidney: effect of unilateral ureteric obstruction. *BJU Int.* **101**, 376–381 (2008).
165. Lüllmann-Rauch, R. *Taschenlehrbuch Histologie.* (Thieme, 2006).
166. Tandle, A. *et al.* Tumor vasculature-targeted delivery of tumor necrosis factor-alpha. *Cancer* **115**, 128–139 (2009).
167. Dolenšek, J., Rupnik, M. S. & Stožer, A. Structural similarities and differences between the human and the mouse pancreas. *Islets* **7**, (2015).
168. Sieber, T., Hare, E., Hofmann, H. & Trepel, M. Biomathematical description of synthetic peptide libraries. *PLoS ONE* **10**, e0129200 (2015).

169. Ho, S. P. *et al.* Potent antisense oligonucleotides to the human multidrug resistance-1 mRNA are rationally selected by mapping RNA-accessible sites with oligonucleotide libraries. *Nucleic Acids Res.* **24**, 1901–1907 (1996).
170. Palfrey, D., Picardo, M. & Hine, A. V. A new randomization assay reveals unexpected elements of sequence bias in model ‘randomized’ gene libraries: implications for biopanning. *Gene* **251**, 91–99 (2000).
171. Kayushin, A. L. *et al.* A convenient approach to the synthesis of trinucleotide phosphoramidites--synthons for the generation of oligonucleotide/peptide libraries. *Nucleic Acids Research* **24**, 3748 (1996).
172. Perabo, L. *et al.* Heparan sulfate proteoglycan binding properties of adeno-associated virus retargeting mutants and consequences for their in vivo tropism. *J. Virol.* **80**, 7265–7269 (2006).

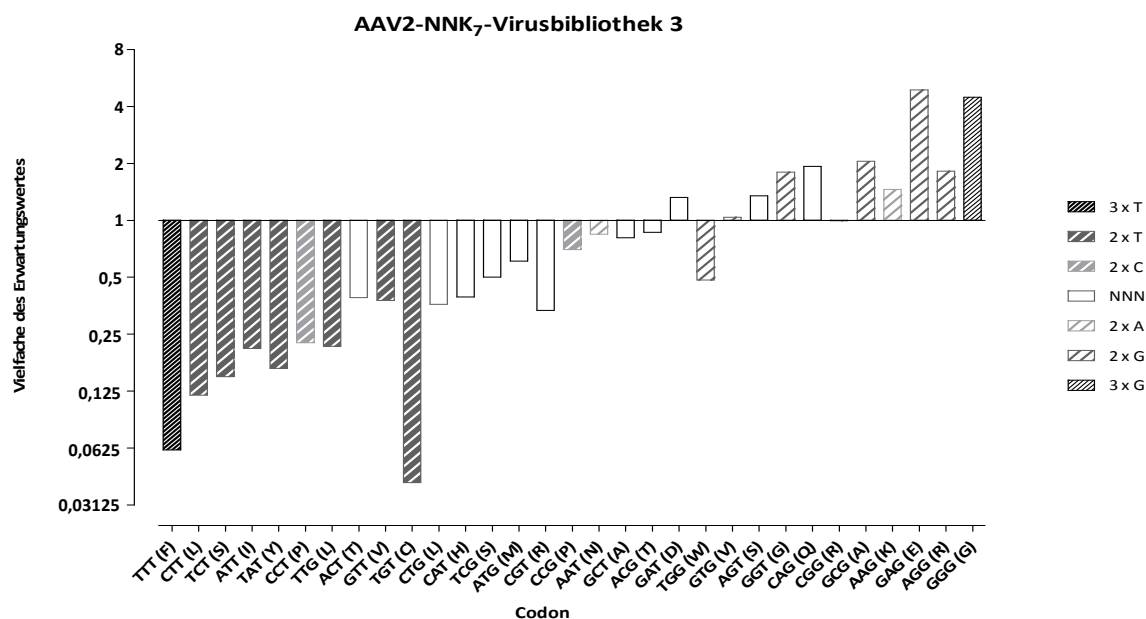
## 7 Anhang

### 7.1 Anhang zu den Ergebnissen zu 4.3.4

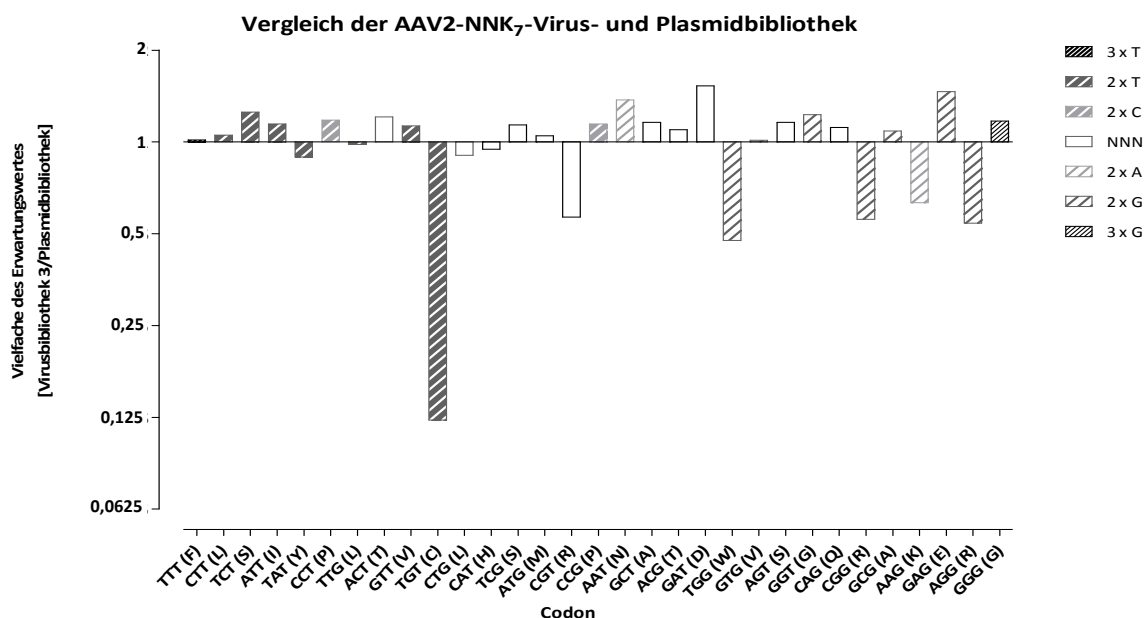


**Abbildung 29: Codonhäufigkeit und Abweichungen vom Erwartungswert (AAV2-NNK<sub>7</sub>-Virusbibliothek 2).** **a** AAV2-NNK<sub>7</sub>-Virusbibliothek 2. Die Vielfachen des Erwartungswertes wurden für jedes Codon bestimmt und von der höchsten negativen Abweichung zur höchsten positiven Abweichung sortiert. **b** Vergleich der AAV2-NNK<sub>7</sub>-Plasmidbibliothek mit der AAV2-NNK<sub>7</sub>-Virusbibliothek 2. Durch die Berechnung der Verhältnisse der Vielfachen des Erwartungswertes von Virusbibliothek und Plasmidbibliothek wird die Veränderung durch die Produktion der Virusbibliothek deutlich.





a



b








**Abbildung 30: Codonhäufigkeit und Abweichungen vom Erwartungswert (AAV2-NNK<sub>7</sub>-Virusbibliothek 3).** **a)** AAV2-NNK<sub>7</sub>-Virusbibliothek 3. Die Vielfachen des Erwartungswertes wurden für jedes Codon bestimmt und von der höchsten negativen Abweichung zur höchsten positiven Abweichung sortiert. **b)** Vergleich der AAV2-NNK<sub>7</sub>-Plasmidbibliothek mit der AAV2-NNK<sub>7</sub>-Virusbibliothek 3. Durch die Berechnung der Verhältnisse der Vielfachen des Erwartungswertes von Virusbibliothek und Plasmidbibliothek wird die Veränderung durch die Produktion der Virusbibliothek deutlich.

## 7.2 Auflistung der verwendeten Gefahrenstoffe nach GHS (global harmonisiertes System)

Die in dieser Arbeit verwendeten Gefahrenstoffe nach GHS sind in **Tabelle 29** aufgelistet.

**Tabelle 29:** Auflistung der verwendeten Gefahrenstoffe nach GHS.

Chemikalie	CAS-RN	Stoffname (IUPAC)	Piktogramm	H-Sätze	P-Sätze
<b>Ampicillin</b>	69-52-3	(2 <i>S</i> ,5 <i>R</i> ,6 <i>R</i> )-6-(2-Amino-2-phenyl-acetyl)amino-3,3-dimethyl-7-oxo-4-thia-1-azabicyclo[3.2.0]heptan-2-carbonsäure		H315-H317-H319-H334-H335	P261-P280-P305+P351+P338-P342+P311
<b>Calciumchlorid</b>	10043-52-4	Calciumchlorid		H319	P305+P351+P338
<b>EDTA</b>	60-00-4	2-[2-[bis(carboxymethyl)amino]ethyl-(carboxymethyl)amino]acetic acid		H319	P305+P351+P338
<b>Ethanol</b>	64-17-5			H225-H319	P210-P240-P305+P351+P338-P403+P233
<b>Ethidiumbromid</b>	1239-45-8	5-ethyl-6-phenylphenanthridin-5-ium-3,8-diamine;bromide		H330-H341-H302-H315-H319-H335	P281-P302+P352-P305+P351+P338-P304+P340-P309-P310
<b>Hydrogenchlorid</b>	7647-01-0	Hydrogenchlorid		H314-H290-H335-	P234-P260-P305+P351+P338-P303+P361+P353-P304+P340-P309+P311-P501
<b>Isofluran</b>	26675-46-7	2-chloro-2-(difluoromethoxy)-1,1,1-trifluoroethane		H336-H373	P304+P340+P312
<b>Isopropanol</b>	67-63-0	Propan-2-ol		H225-H319-H336	P210-P233-P305+P351+P338
<b>Methanol</b>	67-56-1	Methanol		H225-H301-H311-H331-H370	P210-P233-P280-P302+P352
<b>Natriumhydroxid</b>	1310-73-2	Natriumhydroxid		H290-H314	P280-P301+P330+P331-P305+P351+P338-P308+P310
<b>Paraformaldehyd</b>	30525-89-4	Paraformaldehyd		H228-H302-H332-H351-H335-H315-H319-H317	P302+P352-P305+P351+P338-P308+P313-P304+P340
<b>Penicillin G, Natriumsalz</b>	69-57-8	Natrium[2 <i>S</i> -(2 $\alpha$ ,5 $\alpha$ ,6 $\beta$ )]-3,3-dimethyl-7-oxo-6-(phenylacetamido)-4-thia-1-azabicyclo[3.2.0]heptan-2-carboxylat		H315-H320-H317-H334-H335	P302+P352-P333+P313-P304+P340-P342+P311-P305+P351+P338-P337+P313-P312-

					P261-P285-P280-P362-P264-P271
<b>PEI</b>	9002-98-6	Polyethylimin	 	H301	P301+P310
<b>Streptomycin Sulfate</b>	3810-74-0	1-[(1S,2R,3R,4S,5R,6R)-2-[(2R,3R,4R,5S)-3-[(2S,3S,4S,5R,6S)-4,5-dihydroxy-6-(hydroxymethyl)-3-(methylamino)tetrahydropyran-2-yl]oxy-4-formyl-4-hydroxy-5-methyl-tetrahydrofuran-2-yl]oxy-5-guanidino-3,4,6-trihydroxycyclohexyl]guanidine; Schwefelsäure		H315-H320-H317-H334-H335	P302+P352-P333+P313-P304+P340-P342+P311-P305+P351+P338-P337+P313-P312-P261-P285-P280-P362-P264-P271
<b>TRIS</b>	77-86-1	2-Amino-2-(hydroxymethyl)-propan-1,3-diol		H315- H319-H335	P261-P305+P351+P338
<b>Triton X-100</b>	9002-93-1	2-[4-(2,4,4-trimethylpentan-2-yl)phenoxy]ethanol	  	H302-H318-H411	P273-P280-P305+P351+P338

In **Tabelle 30** sind die Erklärungen zu den aufgelisteten GHS-Piktogrammen aufgeführt.

**Tabelle 30: Legende zu den aufgelisteten Piktogrammen.**

<b>Piktogramm</b>	<b>Beschreibung</b>	<b>Kodierung</b>	<b>Gefahrenklasse</b>
	Flamme	GHS02	Entzündbar, selbsterhitzungsfähig, selbstzersetzend, pyrophor, Organische Peroxide
	Ätzwirkung	GHS05	Auf Metalle korrosiv wirkend, hautätzend, schwere Augenschädigung
	Totenkopf mit gekreuzten Knochen	GHS06	Akute Toxizität
	Dickes Ausrufezeichensymbol	GHS07	Gefährliche Substanzen
	Gesundheitsgefahr	GHS08	Diverse Gesundheitsgefahren
	Umwelt	GHS09	Gewässergefährdend

## Erklärungen zu den H- und P-Sätzen

### H-Sätze:

H225	Flüssigkeit und Dampf leicht entzündbar.
H228	Entzündbarer Feststoff.
H290	Kann gegenüber Metallen korrosiv sein.
H301	Giftig bei Verschlucken.
H302	Gesundheitsschädlich bei Verschlucken.
H311	Giftig bei Hautkontakt.
H314	Verursacht schwere Verätzungen der Haut und schwere Augenschäden.
H315	Verursacht Hautreizungen.
H317	Kann allergische Hautreaktionen verursachen.
H318	Verursacht schwere Augenschäden.
H319	Verursacht schwere Augenreizung.
H320	Verursacht Augenreizung
H330	Lebensgefahr bei Einatmen.
H331	Giftig bei Einatmen.
H332	Gesundheitsschädlich bei Einatmen.
H334	Kann bei Einatmen Allergie, asthmaartige Symptome oder Atembeschwerden verursachen.
H335	Kann die Atemwege reizen.
H336	Kann Schläfrigkeit und Benommenheit verursachen.
H341	Kann vermutlich genetische Defekte verursachen.
H351	Kann vermutlich Krebs erzeugen
H370	Schädigt die Organe.
H373	Kann die Organe schädigen bei längerer oder wiederholter Exposition.
H411	Giftig für Wasserorganismen, mit langfristiger Wirkung.

### P-Sätze:

P210	Von Hitze, heißen Oberflächen, Funken, offenen Flammen sowie anderen Zündquellenarten fernhalten. Nicht rauchen.
P233	Behälter dicht verschlossen halten.
P234	Nur im Originalbehälter aufbewahren.
P240	Behälter und zu befüllende Anlage erden.
P260	Staub / Rauch / Gas / Nebel / Dampf / Aerosol nicht einatmen.
P261	Einatmen von Staub / Rauch / Gas / Nebel / Dampf / Aerosol vermeiden.
P264	Nach Handhabung Hände gründlich waschen.
P271	Nur im Freien oder in gut belüfteten Räumen verwenden.
P273	Freisetzung in die Umwelt vermeiden.
P280	Schutzhandschuhe / Schutzkleidung / Augenschutz / Gesichtsschutz tragen.
P285	Bei unzureichender Belüftung Atemschutz tragen.
P310	Sofort Giftinformationszentrum, Arzt oder ... anrufen.
P312	Bei Unwohlsein Giftinformationszentrum, Arzt oder ... anrufen.
P362	Kontaminierte Kleidung ausziehen und vor erneutem Tragen waschen.
P301 + P330 + P331	Bei Verschlucken: Mund ausspülen. Kein Erbrechen herbeiführen.
P301 + P310	Bei Verschlucken: Sofort Giftinformationszentrum oder Arzt anrufen.
P302 + P352	Bei Kontakt mit der Haut: Mit viel Wasser und Seife waschen.
P303 + P361 + P353	Bei Kontakt mit der Haut (oder dem Haar): Alle beschmutzten, getränkten Kleidungsstücke sofort ausziehen. Haut mit Wasser abwaschen/duschen.

P304 + P340 + P312	Bei Einatmen: An die frische Luft bringen und in einer Position ruhigstellen, die das Atmen erleichtert. Bei Unwohlsein Giftinformationszentrum, Arzt oder ... anrufen. (Keine offizielle Kombination laut GHS, aber von Herstellern verwendet)
P305 + P351 + P338	Bei Kontakt mit den Augen: Einige Minuten lang behutsam mit Wasser spülen. Vorhandene Kontaktlinsen nach Möglichkeit entfernen. Weiter spülen.
P308+P310	Bei Exposition oder falls betroffen: Sofort Giftinformationszentrum, Arzt oder ... anrufen. (Keine offizielle Kombination laut GHS, aber von Herstellern verwendet)
P308+P313	Bei Exposition oder falls betroffen: Ärztlichen Rat einholen / ärztliche Hilfe hinzuziehen.
P333 + P313	Bei Hautreizung oder -ausschlag: Ärztlichen Rat einholen/ärztliche Hilfe hinzuziehen.
P304 + P340	Bei Einatmen: An die frische Luft bringen und in einer Position ruhigstellen, die das Atmen erleichtert.
P337 + P313	Bei anhaltender Augenreizung: Ärztlichen Rat einholen/ärztliche Hilfe hinzuziehen.
P342 + P311	Bei Symptomen der Atemwege: Giftinformationszentrum oder Arzt anrufen.
P403 + P233	Behälter dicht verschlossen an einem gut belüfteten Ort aufbewahren.
P501	Inhalt / Behälter ... zuführen. (Die vom Gesetzgeber offen gelassene Einfügung ist vom Inverkehrbringer zu ergänzen)

### 7.3 Publikationen

Bisherige Publikationen:

**Lampe M**, Steinhoff A., Depping R, Jelkmann W. *“Importin  $\alpha$  Mutants - Characterization of Binding Properties to Oxygen-Sensing Proteins“*, International EPO-Conference, Lübeck, 2009, Poster

Körbelin J, Dogbevia G, Michelfelder S, Ridder DA, Hunger A, Wenzel J, Seismann H, **Lampe M**, Bannach J, Pasparakis M, Kleinschmidt JA, Schwaninger M, Trepel M. *A brain microvasculature endothelial cell-specific viral vector with the potential to treat neurovascular and neurological diseases*. EMBO Mol Med.(2016).

## Danksagung

Ich möchte mich ganz herzlich bei Prof. Dr. Martin Trepel, Chefarzt der II. Medizinischen Klinik und Direktor des Interdisziplinären Cancer Centers am Klinikum Augsburg, für die Betreuung und Begutachtung meiner Arbeit und seine Unterstützung bedanken. Des Weiteren danke ich Prof. Dr. Elke Oetjen, Fachärztin für Pharmakologie und Toxikologie und stellvertretende Institutsdirektorin Pharmakologie für Pharmazeuten, für die Begutachtung meiner Arbeit und die Übernahme des Vorsitzes meiner Disputation.

Ein großer Dank geht an Dr. Henning Seismann und Dr. Stefan Michelfelder für die anfängliche Betreuung meiner Arbeit und ihre tollen Ideen. Ein weiterer besonderer Dank gilt Dr. Jakob Körbelin für seine Unterstützung, unsere wissenschaftlichen Diskussionen, seine Hilfsbereitschaft, seine lustige Art und die interessanten Mittagessen. Außerdem danke ich Dr. Timo Sieber für seine Ideen, seine projektinitiierende Kreativität, seine Bereitschaft für Kooperationen und sein „Vermächtnis“.

Ich möchte mich ganz herzlich bei der gesamten AG Trepel, besonders bei Dr. Elmar Spies, Agnes Hunger und Barbara Gösch, aber auch bei denen die im Laufe der Jahre die Arbeitsgruppe verlassen haben, für ihre Hilfe bei Versuchen und der Organisation des Labors bedanken.

Danken möchte ich auch Malik Alawi für die bioinformatische Auswertung der NGS-Daten und seine Geduld, sowie der Arbeitsgruppe von Prof. Dr. Adam Grundhoff, besonders Dr. Daniela Indenbirken für die Durchführung der NGS-Experimente.

Vielen Dank auch an Prof. Dr. Mascha Binder für die zwischenzeitliche Betreuung und vielen Dank an ihre Arbeitsgruppe, insbesondere Friederike Braig und Christopher Ford, für die NGS-Kooperation. Außerdem danke ich Dr. Axel Leingärtner von der Core Facility „*In vivo optical imaging*“ für die Bereitstellung der IVIS.

Des Weiteren danke ich Roholla Babaeikelishomi aus der Arbeitsgruppe von Prof. Dr. Oetjen für die praktische Hilfe bei der Isolierung der Langerhans-Inseln.

Schließlich möchte ich mich bei meiner ganzen Familie, insbesondere meiner Mutter Martina Lampe und meinem Mann Stephan Melchert für ihre Unterstützung und ihre Durchhalteparolen bedanken! Außerdem danke ich allen meinen Freunden, besonders Dr. Manuela Gorgel, Dr. Sonia Singh, Janna Jacubczik, Louise Thieme, Judith Thunemann und Frauke Degenhardt, für ihre immer offenen Ohren und ihre tolle freundschaftliche Unterstützung!

## **Eidesstattliche Versicherung:**

*Hiermit versichere ich an Eides statt, die vorliegende Dissertation selbst verfasst und keine andere als die angegebenen Hilfsmittel benutzt zu haben.*

*Die eingereichte Fassung entspricht der auf dem elektronischen Speichermedium. Ich versichere, dass diese Dissertation nicht in einem früheren Promotionsverfahren eingereicht wurde.*

*Hamburg, 02.08.2016*

*Melanie Lampe*

## **Änderungsvermerk**

In der deutschen und der englischen Zusammenfassung wurde das Wort Cytosin bzw. cytosine durch Thymin bzw. thymine ersetzt.