UNDERSTANDING AND MANIPULATING AAV-GLYCAN INTERACTIONS

Shen Shen

A dissertation submitted to the faculty of the University of North Carolina at Chapel Hill in partial fulfillment of the requirements for the degree of Doctor of Philosophy in the Department of Biochemistry and Biophysics.

Chapel Hill
2013

Approved by:

Aravind Asokan, Ph.D.

Ronald Swanstrom, Ph.D.

Hengming Ke, Ph.D.

Jian Liu, Ph.D.

Nikolay Dokholyan, Ph.D.

ABSTRACT

SHEN SHEN: Understanding and Manipulating AAV-Glycan Interactions
(Under the direction of Dr. Aravind Asokan)

Viral infections generally begin with virus-glycoconjugate interactions on the host cell surface. Adeno-associated Virus (AAV) serotypes have been shown to bind to sialylated glycans and heparan sulfates for cellular attachment. However, the underlying molecular mechanism and its impact on AAV tropism in hosts have not been well understood. In this dissertation, terminal *N*-linked galactose was identified as the primary receptor for AAV serotype 9 (AAV9). Moreover, recombinant sialidase was explored as an adjuvant to dramatically enhance AAV9 gene transfer both *in vitro* and *in vivo*. More in-depth analysis of the effects of sialidase on the transduction, tropism, and pharmacokinetics profiles of AAV9 indicates that higher binding avidity between AAV9 and galactose alters its tropism, shifting it from a systemic towards a hepatic phenotype. Using AAV9 as a model system, the functions of glycan receptors in determining AAV transduction efficiency *in vitro* and its tropism *in vivo* were elucidated. In addition, it was also showed that the blood circulation kinetics of AAV vectors are tightly regulated by their binding avidities with cognate glycan receptors on tissues and blood components. In parallel, key amino acids in the galactose binding footprint on the AAV9 capsid were identified by screening of a random mutagenesis library of AAV9. Furthermore, grafting of the galactose binding footprint from AAV9 onto other AAV strains yields a series of novel AAV variants with enhanced transduction efficiency, as well as a rapid on-set of

transgene. This work also demonstrates the orthogonality and modularity of glycan binding motifs from several naturally occurring AAV serotypes, as well as the adaptability of the AAV capsid for the incorporation of new functionalities. By establishing the link between the first step of AAV infection and its tissue targeting profiles, a novel strategy was proposed to design and construct AAV vectors with organ-targeting or detargeting phenotypes. In conclusion, this dissertation sheds light on the crucial roles of glycan receptors in AAV infectious pathways, and paves a new avenue for vector development for gene therapy applications.

To my aunt, Xiaojian Shen,

and

To my parents, Hongqi Fa and Xuan Shen,

with love…

ACKNOWLEDGEMENTS

Being Dr. Aravind Asokan's first graduate student is an honor and an adventure. His dedication and enthusiasm in science and the nurture of young scientists made me who I am today, and will influence my entire career in the forthcoming years. He introduced me to virology, glycobiology, and gene therapy, and gave me the intellectual and scientific freedom to explore what really intrigues me. I am extremely grateful for his constant trust in me for the past four years. I would never forget those days when we were working on the discovery of the glycan receptor for AAV9 and how we managed to publish it in time. What Aravind taught me over the years also goes above and beyond science. I would like to thank him for shaping my personalities and guiding me to be a better person. Thank you very much!

I am extremely fortunate to have a fabulous thesis committee, including top-notch virologist, computational biologist, structure biologist, and glycobiologist. I would like to thank Drs. Ronald Swanstrom, Nikolay Dokholyan, Hengming Ke, and Jian Liu for their inspiration, encouragement, and setting role models for me in various aspects. Their insightful mentorship, generous support, and visionary guidance are instrumental for my entire graduate study and my future endeavors in science.

I would also like to express my sincere appreciation to all the co-authors on my publications, especially Andrew Troupes, Sarah Brown, and Kelli Bryant. I would like to thank these three extremely talented and lovely technicians for their gracious support,

caring encouragement, and endless laughter. Furthermore, I would like to wholeheartedly thank Drs. Nagesh Pulicherla and Eric Horowitz for their day-to-day guidance and inspiration during my graduate training. I very much enjoy the productive discussions with both of you, and thank you very much for teaching me persistence, hard-working, and creativity. In addition, I benefit tremendously from the collaborative efforts in UNC Gene Therapy Center, constituted by the Samulski lab, Kafri Lab, Monahan Lab, Asokan lab, and UNC Vector Core. Specifically, I would like to thank Dr. Tal Kafri, Dr. Chengwen Li, Dr. Junjiang Sun, Erin Borchardt, Giridhar Murlidharan, Garrett Berry, Swati Yadav, Lavanya Bachaboina, Dr. Eun-Young Choi, Ping Zhang, and Sophia Shin.

Most importantly, I am extremely indebted to my aunt, Xiaojian Shen, and my parents, Hongqi Fa and Xuan Shen, for their unconditional love. They are my first mentors in life, supporting me to pursuit happiness on the other side of the globe. In addition, I own a great deal of gratitude to Weizhong Zhang, Stephanie Zhang, Dr. Yan Xia, and Zhi Zhou for their tremendous caring and love over the years. Last but not least, I would like to thank my best friends Drs. Yu Lei, Gaoyang Liang, Wenqi Pan, Han Dai, Ningqi Hou, Jun Zhang, Yun Jian, Ran Jin, Liheng Cai, Yuying Xie, Yangfan Liu, Dongfen Yu, Shaoying Wang, Fan Liu, and Zhou Li for our enduring friendship and all the joy you brought to me. I am truly blessed.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATIONS

µM	Micromolar
3'NST	α2,3-*N*-sialyltransferase
3'OST	α2,3-*O*-sialyltransferase
3'-S-Di-LN	α2,3-sialylated di-LN
3'-SLN	α2,3-sialylated LN
6'NST	α2,6-N-sialyltransferase
AAP	Assembly-activating Protein
AAV	Adeno-associated virus
ALS	Amyotrophic Lateral Sclerosis
Arg	Arginine
BAAV	Bovine Adeno-associated Virus
B_{max}	Number of Binding Sites
BoV	Bocavirus
BSA	Bovine Serum Albumin
$CaCl_2$	Calcium chloride
CBA	Chicken Beta Actin
CHO	Chinese Hamster Ovary
CKII	Casein Kinase II
CMV	Cytomegalovirus
CNS	Central Nervous System
ConA	Concanavalin A
CPV	Canine parvovirus

CsCl	Cesium Chloride
DMEM	Dulbecco's Modified Eagles Medium
ECL	*Erythrina Cristagalli* Lectin
EDTA	Ethylenediaminetretraacetic acid
EGFR	Epidermal Growth Factor Receptor
EGFR-PTK	Epidermal Growth Factor Receptor Protein Tyrosine Kinase
EM	Electron Microscopy
FGFR	Fibroblast Growth Factor Receptor
FITC	Fluorescein
FPV	Feline Parvovirus
Fuc	Fucose
GAG	Glycosaminoglycan
Gal	Galactose
GFP	Green Fluorescent Protein
Glc	Glucose
GPV	Goose Parvovirus
GSL	Glycosphingolipid
HA	Hemagglutinin
HAE	Human Airway Epithelium
HBV	Hepatitis B Virus
HCV	Hepatitis C Virus
HGFR	Hepatocyte Growth Factor Receptor
HPI	Hours Post Infection/Injection

HS	Heparan Sulfate
IFWB	Immunofluorescence Wash Buffer
IGF-1	Insulin-Like Growth Factor 1
IM	Intramuscular
IP	Intraperitoneal
ITR	Inverted Terminal Repeat
IV	Intravascular
k_{12}	Blood-To-Organ Distribution Rate Constant
k_{21}	Organ-To-Blood Distribution Rate Constant
K_d	Binding Affinity
K_d'	Observed Binding Affinity
k_e	Elimination From Blood
LacNAc	N-acetyllactosamine
LN	LacNAc
Luc	Luciferase
Lys	Lysine
MAL-I	*Maackia Amurensis* Lectin
$MgCl_2$	Magnesium Chloride
mLam	Mouse Lamin Gene
mM	Millimolar
MOI	Multiplicity of Infection
MVM	Minute Virus of Mice
MVMi	Immunosuppressive Minute Virus of Mice

MVMp	Prototype Minute Virus of Mice
NAb	Neutralizing Antibody
Neu5Ac	N-acetylneuraminic Acid
NLS	Nuclear Localization Sequence
NPC	Nonparenchymal Cells
ORF	Open Reading Frames
PBS	Phosphate Buffered Saline
PC	Parenchymal Cells
PDB	Protein Data Bank
PDGFR	Platelet Derived Growth Factor Receptor
PEG8000	Polyethylene Glycol 8000
PEI	Polyetherimide
PI	Post Injection/Infection
PLA_2	Phospholipase A_2
PPV	Porcine Parvovirus
qPCR	Quantitative Polymerase Chain Reaction
RBC	Red Blood Cell
RBE	Rep Binding Element
RES	Reticuloendothelial System
RLU	Relative Light Unit
RMS	Rostral Migratory Stream
RPE	Retinal Pigmented Epithelium
RT	Room Temperature

S.E.M.	Standard Error Mean
SA	Sialic Acid
SC	Self-complementary
SDS	Sodium Dodecyl Sulfate
Ser	Serine
SNA	*Sambucus Nigra* Lectin
$t_{1/2,dis}$	Distribution Half-Life
$t_{1/2,e}$	Elimination Half-Life
TBSV	Tomato Bushy Stunt Virus
TCV	Turnip Crinkle Virus
tdTom	tdTomato
TEM	Transmission Electron Microscopy
Thr	Threonine
Tyr	Tyrosine
V. Cholera	*Vibrio Cholera*
VEGF	Vascular Endothelial Growth Factor
VG	Vector Genomes
VP	Viral Protein
VP1u	VP1 Unique Region
WGA	Wheat Germ Agglutinin

CHAPTER 1

Introduction

1.1 Glycans in Viral Infection

The glycocalyx on cell surface is composed of complex glycoconjugates that are critical in physiological and pathological conditions. It plays important roles in maintaining the integrity of cellular morphology, intracellular signaling, cell adhesion and migration, regeneration and differentiation, as well as mediating immune responses and tumor metastasis. While glycocalyx forms a shield against microbial invasion, pathogens also exploits these glycans on the host cell surface as anchoring points to initiate infection.

Three major classes of glycoconjugates on the cell surface are utilized by viruses as primary receptors [1]. The first class is *glycosphingolipid* (also called *glycolipid*), a polysaccharide chain covalently linked to the terminal hydroxyl group of the ceramide lipid anchored in the cellular membrane. The second type is *glycoprotein*, formed by oligosaccharides covalently linking to the extracellular domains of polypeptide backbones. Glycoproteins can be classified as *N*-glycans and *O*-glycans based on the amino acid residues the sugar chains attach to. An *N*-glycan forms when an asparagine residue of an Asn-X-Ser/Thr motif on the polypeptide chain is attached by an oligosaccharide molecule. An *O*-glycan is linked to the polypeptide backbone via the hydroxyl group of a serine or threonine residue. The third class of glycoconjugates is *proteoglycan*, which is formed by covalently linking glycosaminoglycans (GAGs) to serine residues on a peptide backbone via xylose

moieties. Glycolipids and glycoproteins contain linear or branched oligosaccharide chains that are usually capped with sialic acids. These sugar chains often consist of repeating disaccharide units Galβ1-4GlcNAcβ1- or GalNAcβ1-4GlcNAcβ1-, and are modified with sulfate groups, fucose (Fuc), or β-GalNAc. In contrast, GAG are linear polysaccharide chains composed of repeating disaccharide units, including GalNAcβ1-4GlcA (in chondroitin), GalNAcβ1-4IdoA (in dermatan sulfate), GlcNAcα1-4GlcA (in heparin), GlcNAcα1-4IdoA (in heparan sulfate), or GalNAcβ1-4GlcNAc (in keratin sulfate).

What forces drive the virus-glycan interactions? Since the sulfate groups on GAGs and the carboxyl groups of the terminal sialic acid moieties are negatively charged, these glycans usually bind to basic patches on the exterior surface of viral structural proteins. Especially in the case of heparan sulfate-recognizing proteins, common motifs with continuous positive charges XBBXBX or XBBBXXBX (B: lysine or arginine; X: any other amino acid) were proposed [2]. Therefore, electrostatic forces are the key forces involved in virus-glycan interactions. In addition, direct hydrogen bonding as well as H_2O mediated hydrogen bonding between hydroxyl groups on glycan residues and hydrophobic or aromatic amino acids contribute to the specificity of glycan-virus interactions.

Although the interactions between viruses and their cognate glycan receptors are highly specific, it is noteworthy to mention that the binding affinity (K_d) between a single viral particle and a single glycan receptor is very low, sometimes within the micromolar (μM) to millimolar (mM) range [3]. Therefore, *avidity*, the cooperative synergy of multiple virus-glycan interactions, increases the affinity and specificity of this binding event [4]. This multivalency of virus-glycan binding is attributed to the icosahedral symmetry of identical viral protein (VP) subunits, and the multiantennary glycan structures. The spatial symmetry

of icosahedral viruses allows for the presence of 12, 20, 30, or even more identical receptor footprints on one macromolecule assembly.

The first step of viral infection is the virion-receptor interactions on cell surface, usually mediated by carbohydrates. Once attached by cell surface glycans, viruses are concentrated on cell membrane. At the same time, this virus-glycan interaction primes the viral particles for protein receptor recognition by inducing global conformational changes on viral proteins. Then the surface bound viruses are guided into cytoplasm and hijack the cellular replication machinery. Because these virus-glycan interactions are very specific, the distribution and presence of the corresponding glycans plays a determining role in the permissibility of viruses in different cell types, tissue types, and species.

An example of how glycan binding ability impacts the species specificity of viral infection and the antigenic drift among species is influenza viruses. Influenza A viruses, the major cause of annual flu epidemics (www.cdc.gov/flu), infect several animal species, including birds, swine, and humans [5-7]. Although all influenza A strains recognize terminal sialic acids (SA) as primary receptors, it was not until 1983 that Rogers and Paulson discovered that the influenza hemagglutinin (HA) possesses specificity towards different linkage forms between sialic acid and the penultimate galactose [8]. Tumpey et al. further demonstrated that by mutating two amino acids, A/South Carolina/1/18 (SC18), specific to α2-6 SA, can switch to a α2-3 SA preference displayed by AV18 (avianized 1918 influenza virus). This altered glycan recognition pattern of SC18 completely abrogates its transmission between ferrets, a model for human respiratory system [9]. X-ray crystallography and high-throughput glycan array screening revealed that α2-6 SA adopts an *umbrella*-like topology,

allowing for human-adapted HA binding. The *cone*-topology, displayed by α2-3 SA and short α2-6 SA chains, is more amenable towards avian or swine-adapted HAs [10,11].

Other examples of viruses binding to sialylated glycans for infection include Adenovirus type 37 [12], Coronavirus OC43 [13,14], Measle virus [15], Sendai virus [16], Newcastle disease virus [17], Enterovirus 70 [18] and Rhinovirus 87 [19,20] in the Picornavirus family, Reovirus 3 [21-23], as well as Rotavirus [24-26], JC [27] and BK viruses [28] in Polyomaviridae (**Table 1**). The distribution and availability of their cognate sialylated glycan receptors on tissues partially determine the tropism and transmissibility of these viruses in hosts. For examples, Norovirus, the causative agent for acute nonbacterial gastroenteritis, binds to ABO(H) histo-blood group antigens on red blood cells (RBCs) and intestinal mucosa to infect hosts [29]. A key enzyme in the biosynthesis pathways of histo-blood group antigens in the intestines is FUT2, which transfers fucose to galactose through an α1-2 covalent bond [30]. Due to a FUT2 deficiency in ~20% of Caucasians and Africans, these populations lack the glycan receptor Fucα1-2Gal for most Noroviruses in gut. Therefore, the same populations also possess a natural resistance to most Noroviruses transmission [30].

Glycoaminoglycans (GAGs) is another group of carbohydrate receptors for viral infections as shown in **Table 1**. The differentiated degree of sulfation on heparan sulfates in major organs partially explains the hepatic tropism of hepatitis A, B, D, E, and hepatitis C virus (HCV). According to a study led by Vongchan et al., heparan sulfate in human liver has an elevated level of sulfation compared to that in other human organs, porcine liver, or HepG2 cell culture [31]. *In vitro* GAG inhibition assays demonstrate that HCV binds preferentially with highly sulfated heparan sulfates or heparin [32]. In addition, only the heparan sulfates isolated from liver, not kidney, can efficiently blocks HCV binding. It

prompted us to speculate that the preferential binding of HCV on highly sulfated heparin sulfate prevents HCV from being absorbed in other organs and ultimately leads to its hepatic accumulation. Although HCV belongs to a different viral family different from HAV and HBV, they all requires heparan sulfates for efficient binding and cellular uptake of viruses [33-36]. Therefore, analogy is likely to be drawn among them with regards to the tropism-determining role of heparan sulfates.

1.2 Glycan Receptors for Parvoviruses

Parvoviruses are small, non-enveloped, single-stranded DNA viruses [37]. Its members include canine parvovirus (CPV), feline parvovirus (FPV), goose parvovirus (GPV), porcine parvovirus (PPV), bocavirus (BoV), B19, AAV, and minute virus of mice (MVM), etc. Similar to all other pathogens, interactions with glycan receptors is a prerequisite for efficient infection by parvoviruses.

Two parvoviruses, bovine AAV (BAAV) and B19 are the only strains known to bind to glycosphingolipid (GSL) to engage productive infection [38,39]. Schmidt and Chiorini showed that the transduction and binding of BAAV is SA-dependent as sialidase treatments abolish its infection and binding *in vitro* [38]. On the other hand, incorporation of GSL in GSL-depleted CHO cells or C6 cells restored BAAV infection and binding, while disruption of glycoprotein using trypsin and dispase does not alter its infection profiles. Cryo-EM structure of B19-globoside complex published in 1996 suggests that globosides bind to the depression around the three-fold axes [39]. Mutations around the globosides binding footprints yield mutant B19 that can escape neutralizing antibodies.

Parvoviruses that bind to heparan sulfates for infection include AAV2 [40], AAV3B [41,42], and AAV6 [43]. In addition, AAV1 [44], AAV5 [45], and AAV6 [44] use *N*-linked sialic acids to attach to cell surface. It is noteworthy to mention that AAV6 is the only naturally occurring strain of parvovirus that is known to bind alternatively to two types of glycan receptors, heparan sulfate and sialic acid. With a difference of six amino acids (AAV6 VP1 numbering: F129, D418, K531, L584, V598, H642) out of more than 700 residues on AAV6 VP3, AAV1 shows suppressed heparan sulfate binding profiles and the subsequential inability to transduce liver parenchymal cells. Interestingly, a single lysine-to-glutamate (K531E) mutation on AAV6 attenuates its heparan sulfate binding efficiency, while the converse mutation (E531K) on AAV1 imparts its heparan sulfate binding ability. Furthermore, this E531K also bestows AAV1 the ability to transduce liver parenchymal cells. This hepatotropism induced by AAV-heparan sulfate interactions echoes the aforementioned phenotype displayed by HCV.

Another AAV serotype that engages with sialoglycans is AAV4, a cardiopulmonary tropic isolate. AAV4 utilizes *O*-linked sialylated glycoprotein, mucin, to initiate infection [46,47]. **The molecular mechanism of how sialylated glycans in heart and lung contribute to its cardiopulmonary phenotype is discussed in detail in Appendix I of this dissertation.** Briefly, intravenous injection of sialidase disrupts the transduction of AAV4 in murine heart and lung, which prerents abundant *O*-linked sialylated glycans on cell surface. AAV4 mutants that are deficient in SA binding also demonstrate loss-of-transduction in the cardiopulmonary system. Blood circulation assay of the parental and mutant AAV4 reveals that attenuated SA binding ability leads to reduced binding on erythrocytes and other blood

cells. It is the formation of AAV-erythrocytes in circulation and high affinity towards *O*-lined SA in heart and lung that eventually result in the cardiopulmonary tropism of AAV4.

Previous studies on the prototype MVM (MVMp) and immunosuppressive MVM (MVMi) also imply the determining roles of SA in viral tropism. A battery of biophysical characterizations revealed that high binding affinity between MVMi and sialylated glycans contributes to the neurotropism of MVMi [48,49]. Specifically, mutations on residues 399, 460, 553, and 558 of MVMp confer a fibroblast-tropism similar to that of MVMi. X-ray crystallography structure suggests that the "shoulder" region of the spikes surrounding the three-fold axes of MVM and the "dimples" along the two-fold axes harbors the residues involved in SA binding and fibrotropism [50].

Recent studies from our group and others discovered that terminal *N*-linked galactose is the primary receptor for AAV9, and the binding avidity between AAV9 and galactose drives the systemic tropism of AAV9 in animal [51-53]. **Details are discussed in Chapters 2 and 3 of this dissertation.** Enzymatic desialylation on a spectrum of cell lines demonstrates elevated AAV9 binding and transduction profiles. This observation was corroborated by transduction and cell surface binding assays on CHO-Pro5 and CHO-Lec2 cells. In conjunction with lectin inhibition and biophysical assays, this led to the conclusion that the penultimate galactose moiety is the glycan receptor for AAV9. In a murine model, enhanced availability of galactose in major organs shifts the systemic tropism of AAV9 towards liver, while both its blood distribution and blood retention time are shortened due to absorption into tissues. The scarcity of free terminal galactose residues in physiological conditions prevents AAV9 particles from being depleted by a single major organ like AAV2. As a result, AAV9 tends to circulate longer in blood stream, allowing for non-specific low-affinity interactions

with sialylated glycoconjugates or secondary receptors. Ultimately, AAV9 are evenly distributed into all major organs, and displays its unique systemic transduction profiles in animals. Molecular simulation, together with alanine-scanning on AAV9 VP3, indicates that N470, D271, N272, Y446, and W503 form a three-dimensional binding pocket for galactose recognition. Grafting of this pocket and flanking residues enables multiple AAV serotypes to bind to galactosyl-glycans, suggesting the modularity of this glycan receptor footprint and the adaptability of AAV capsid for new functionalities. **The engineering of G mutants (G represents galactose) is presented in Chapter 4 of this dissertation.**

1.3 Genomic Composition of AAV

AAV genome is a single-stranded DNA, 4,679 bases in length [54]. At each end of the AAV genome, there is a 145-base long inverted terminal repeat (ITR) flanking the rest of the DNA sequence. Due to its high GC content, ITRs form secondary hairpin structures, discernible using tungsten-shadowing EM [55]. These ITRs are the minimal requirement for AAV genome packaging [56], and also play critical roles in genome integration and second-strand DNA synthesis [57].

As shown in **Figure 1**, there are two Open Reading Frames (ORFs) between the two ITRs in AAV genome. They encode replication components and viral protein of AAV, respectively. The *Rep* region on AAV genome encodes four Rep proteins that are instrumental in wild-type AAV life cycle. Alternative splicing of the mRNA transcribed from p5 promoter yields Rep78 and Rep68. P19 drives the transcription of two smaller Rep proteins, Rep52 and Rep40. While Rep78 and Rep68 facilitate genome replication and integration into chromosome 19 of the human genome [58], Rep52 and Rep40 are required for

efficient encapsidation of AAV genome [59]. Both p5 and p19 promoter regions contain Rep Binding Elements (RBEs), which bind to Rep proteins in order to self-regulate transcription [60].

The *Cap* ORF encodes two mRNAs, one of which translates into VP1. The second mRNA can be translated into VP2 from an alternate start codon (ACG), or VP3 from a conventional start codon (ATG). Due to the leakiness of the ACG start codon, the ratio between VP2 and VP3 is 1:10. VP1, VP2, and VP3 share the same C-terminus, while VP1 unique region (VP1u) harbors a phospholipase A_2 (PLA_2) domain and nuclear localization sequence (NLS), necessary for endosomal escape and nuclear entry during AAV infectious pathway(s) [61-63]. The alternate start codon also encodes for a 23 kD assembly-activating protein (AAP), which targets newly synthesized VPs from cytoplasm into nucleus and facilitates the assembly of AAV particles [64-66].

1.4 Icosahedral Structures of Viral Capsids

Viruses are small bionanoparticles. To generate a protective protein shell for nucleic acids with minimal structural information, only a few types of protein capsid morphologies are constructed by viruses, including helical tubes and icosahedral shells [67]. Crick and Watson first proposed the hypothesis that construction of identical subunits in an ordered form is the most efficient way to use the limited amount of structural information provided by viral genome [68,69]. This hypothesis has been verified by X-ray crystallography and cryo-EM structures of many viruses. It was in 1957 that the first icosahedral viral structure, tomato bushy stunt virus (TBSV), was solved by X-ray crystallography [70], followed by another icosahedral capsid of turnip yellow mosaic virus [71]. From then, a series of nonevenloped

RNA viruses were claimed to adopt the same icosahedron shape of capsid assembly, including turnip crinkle virus (TCV), picornavirus, rhinovirus, and reovirus [3,72,73]. The collection of icosahedral viruses further expands to DNA viruses, for example, SV40 and murine polyomaviruses. At the same time, it was also shown by geometrical theorems that icosahedral symmetry is the most efficient form of packing with the lowest energy [74].

1.5 Capsid Structures of AAV

Parvoviruses have a common icosahedral protein shell with a triangulation number T=1, and AAV is no exception. A 5:5:50 ratio of VP1 (87kD), VP2 (83kD), and VP3 (61kD) constitutes this icosahedral capsid with a size of 25 nm in diameter. There are 12 representative members of AAV serotype, AAV1-12. Due to high amino acid sequence homology (60-90%), VPs of these 12 AAV serotypes adopt similar 3D structures. With technological advancements over the past decades, atomic structures of most representative AAV serotypes are available in RCSB Database as shown in **Table 2**. The core structure of AAV VP is an eight-stranded jelly-roll β-barrel, observed in many icosahedral viral capsids. As depicted in **Figure 2A and 2B**, βBIDG and βCHEF constitute two antiparallel β-sheets. A short β-strand A (βA) is located at the *N*-terminal region of VP3, in antiparallel interactions with βB. There is also a short α-helix facing the interior of viral capsid. All eight β strands and the α helix are conserved among all AAV serotypes. In contrast, the loops connecting neighboring β strands are highly dynamic in both amino acid sequence and structural morphology. These loops are named by the two β strands they connect. For instance, the loop that links βH and βI is called the HI loop. The shortest loop is the FG loop, only 2-3 amino

acids in length. The longest loop is the GH loop (~230 amino acids), where most sequence variation resides.

When 60 subunits assemble into an intact AAV capsid, these dynamic loops demonstrate their pivotal roles in intra-subunit interactions and the formation of exterior surface topology. First of all, the five-fold axes of symmetry are formed by DE loops and HI loops from five adjacent monomers. The most prominent feature of the pentamers is their iris-like five-fold pores in the center. Extensive mutagenesis of the amino acid constituting the five-fold pore reveals that it is not only instrumental in packaging single-stranded DNA into AAV capsid, but also critical for extrusion of VP1u during post-entry steps in infection [75-77]. Bleker S. et al. further demonstrated that the interactions between the five-fold pores and Rep proteins promote the encapsidation of AAV genomes in preformed capsid shells [78]. Deletion, insertion, and substitution mutations on the HI loop by DiPrimio et al. shed light on the functional roles of the canyon structures surrounding the five-fold pore [79]. F661 on the HI loop and P373 on the EF loop of AAV2 are critical residues involved in the incorporation and activity of the VP1 unique domain. Substitution of the HI loop with a hexa-histidine stretch yields AAV2 particles that are amenable to nanogold labeling and purification through nickel column [80].

According to structural analysis of subunit interfaces, the two-fold axes of symmetry of AAV hold the least energy among three types of interfaces [81,82]. This area with dimple structures was also proposed to contain heparan sulfate binding motif as well as high tolerance of peptide insertion [77,83]. Site-directed mutagenesis of tyrosine-to-phenylalanine changes of all seven Tyr residues clustered near two-fold axes of AAV2 (Y252F, Y272F, Y444F, Y500F, Y700F, Y704F, Y730F) showed enhanced transduction profiles both *in vivo*

and *in vitro*. The underlying mechanism is that these Tyr mutants are able to avoid the phosphorylation by epidermal growth factor receptor protein tyrosine kinase (EGFR-PTK) [84]. Similar results were observed with two tyrosine-to-phenylalanine mutations on AAV6 (Y445F, Y731F), indicating a ubiquitous function of tyrosine residues in AAV life cycle [85]. Chemically crosslinking Y704 from two adjacent monomers on AAV2 results in dityrosine adducts with reduced flexibility of both the two-fold depression and five-fold interface [86]. The reduction of transduction and nuclear entry induced by this artificial constrain further supports the notion that the two-fold axes on AAV capsid play a role in VP1u extrusion. In addition, amino acids along the dimple structure of the two-fold axes of symmetry on AAV2 are also implicated in the neutralizing antibody (NAb) recognition as well as affinity towards skeleton muscles. An AAV1-AAV2 hybrid variant, AAV-221-IV, first drew the attention towards the aa213-423 region on AAV1. Swapping this domain from AAV1 onto AAV2 promotes higher transduction of AAV1 in skeleton muscles. Later on, Bowles D.E. et al. constructed AAV2.5, an AAV2 mutant with five amino acids swapped from AAV1 [87]. Although no changes were observed in *in vitro* transduction profiles, AAV2.5 demonstrates elevated transduction efficiency in skeleton muscles in mice and attenuated cross-reactivity with NAbs. More in-depth mutagenesis focused on two specific amino acids (Q263A, and 265T insertion) in this area [88]. It implies a functional role of the 265 region in stimulating humoral immune response. Given its unique antibody-neutralizing profile, AAV2.5 can serve as a candidate for repetitive administration of AAV vectors.

The three-fold axes of symmetry on AAV capsids are the hot spot of structural variance, which further dictates AAV functionalities. As mentioned earlier, the GH loops forming the three-fold axes are the longest loop on AAV VP. On the trimeric assembly, GH

loops constitute three mounds of spike protrusions, the inner loops, middle loops, and outer loops. Glycan and protein receptor binding footprints disseminate among these surface exposed loops.

Extensive mutagenesis studies discovered the heparan sulfate binding site on AAV2, constituted with R484, R487, K532, R585, and R588 at the walls and valleys of the three-fold spikes [77,89-91]. The three-dimensional footprint has a positive electrostatic potential and concaved structure, readily accommodating highly negatively charged heparan sulfate oligosaccharide chains. The repetitive appearance of this footprint on the icosahedral AAV capsid allows the flexible heparan sulfate chains to wrap around the virions with enhanced binding avidity. Further cryo-EM study on the AAV2-heparan sulfate complexes visualizes conformational shifts at the three-fold and five-fold area of AAV2 [92]. Our knowledge of the glycan receptor footprint on AAV surface is expanded with co-crystal structure of AAV3b and heparan sulfate [93]. In addition to conserved residues R475, R485, R488, K533 from AAV2, R447 and the AAV3b-unique R594 are also in contact with heparan sulfate on AAV3b capsids. The N500 also contributes to the positive electrostatic potential in close proximity of three-fold spikes. A similar continuous basic patch is also conceivable on the three-fold spikes of AAV6 [43]. K531 on AAV6 three-fold axes, which is absent on AAV1, connects residues R487, K528, K533, and K567 to compose a positively-charged surface exposed region. K531E mutation on AAV6 not only reduces its retention time on heparin columns, but also diminishes its transduction efficiency in hepatocytes.

In case of SA binding AAV serotypes, their glycan receptor footprints are also located in close proximity to the three-fold axes of symmetry. Our recent studies on AAV4 identified K492, K503, M523, G581, Q583, N585 as key residues in contact with sialylated

glycans on CV-1 cell culture and murine erythrocytes. Three-dimensional structural alignment of AAV4 and AAV2 trimers showed that the heparan sulfate and SA binding pockets occupy similar regions on AAV capsids. In addition, site-saturation mutagenesis and alanine-scanning confirmed the notion that M569, Y583, and L585 on the inner loop of AAV5 three-fold area are indispensable for productive infection of AAV5 on cell culture (Shen Shen and Aravind Asokan, unpublished data).

On the exterior surface of AAV9, galactosylated glycans are indicated to bind to amino acids in close proximity to the three-fold axes. With a combination of molecular simulation, site-directed mutagenesis, and cell-surface binding assays, N470, D271, N272, Y446, and W503 were identified as the galactose binding site on AAV9 [94]. In parallel, screening of a random mutagenesis library on AAV9 GH loop reveals that single mutations of W503 or N498 are sufficient to abrogate its galactose binding ability [95]. Further investigation of transduction and pharmacokinetics profiles of these individual AAV9 mutants confirms that mutations on W503 and N498 are also detrimental to the systemic tropism of AAV9 [51,95].

Interestingly, AAV2 and AAV9 also engage with their cognate protein receptors via domains on the trimer protrusions. Integrins αVβ5 and α5β1 recognize AAV2 through specific interactions with its conserved binding domain NGR/RGD (residues 511-513 on AAV2 VP1) [96]. Located on the outer loops, this domain is conserved across several AAV serotypes (except AAV4, AAV5, and AAV11) as shown by sequence alignment. Akache et al. discovered that 37/67 kD laminin receptors mediates transduction of AAV8, AAV2, AAV3, and AAV9 [97]. Coincidentally, the structure on AAV capsids that interact with LamR

were mapped onto the GH loops---stretches 491-547 and 593-623 (AAV8 VP1 amino acid numbering).

The protrusions surrounding the three-fold axes of AAV serotypes not only contain receptor binding sites, but can also be attached by monoclonal NAbs. Cryo-EM and 3D reconstruction of AAV8 virion, together with docking simulation of a generic Fab model, provide the structural hint that residues 586-591 on AAV8 (VP1 numbering) interact directly with the neutralizing monoclonal antibody ADK8 [98]. In-depth biochemical analysis reveals that binding of ADK8 prohibits AAV8 intracellular trafficking. With the recent discovery of ADK9 [66], whether residues 586-591 on the inner loops of other AAV serotypes are in direct contact with their corresponding NAbs are yet to be determined.

1.6 Infectious Pathways of AAV

Upon attachment to its cognate glycan receptors on the cell surface, AAV undergoes conformational changes to prime for interactions with its secondary receptors [92]. A spectrum of growth factor receptors, integrins, and LamR are claimed to be the co-receptors for AAV serotypes (listed in **Table 2**). Specifically, AAV2 utilizes αVβ5 and α5β1 Integrins, hepatocyte growth factor receptor (HGFR), fibroblast growth factor receptor (FGFR), LamR to engange infection [96,97,99,100]. Similarly, AAV3b also exploits HGFR and LamR for infection [97,101]. Bioinformatics technologies and cell-based assays discovered that AAV5 and AAV6 interact with platelet-derived growth factor receptor (PDGFR) and epidermal growth factor receptor (EGFR) as co-receptors for transduction, respectively [102,103]. As described earlier, AAV8 and AAV9 also uses LamR [97]. Recent study in our lab indicates that αVβ3 integrins are indispensable for efficient gene transfer by AAV9 both *in vitro* and *in vivo*

(Shen Shen and Aravind Asokan, unpublished data). This work also shed light on the roles of co-receptors in AAV life cycles in animals.

Receptor binding of AAV serotypes triggers signaling pathways in the cytoplasm to embark efficient uptake of virions, which is a rate-limiting step for most viral infection [104]. Most AAV serotypes enter cells through dynamin-dependent budding of clathrin-coated vesicles [105-107]. A recent study in Thomas Weber's group incorporates the CLIC/GEEC pathway in the picture of AAV2 internalization [108]. This cellular entry step is both serotype- and cell-type dependent. For instance, in well-differentiated human airway epithelial cells (HAE), AAV tends to infect the basolateral side 200-fold more efficiently than from the apical side of polarized cells [109]. AAV5 is an exception. It employs both clathrin- and caveoli-dependent routes of cellular uptake to enter primary human embryo fibroblasts [110]. Another study from John Engelhardt's group demonstrated distinct transduction pathways of AAV1 and AAV6 on polarized HAE, supporting the notion that receptor recognition of individual AAV serotype also plays an role in post-entry steps of AAV lifecycle [111].

Following cellular entry is the cytoplastic processing of AAV particles. Colocalization of AAV virions with endosomal compartments highlighted by fluorescent markers reveal that AAV travels through early, late, and recycling endosomes before nuclear entry [112-114]. In HeLa cells, AAV2 traffics through the late endosome marked by Rab7 at low multiplicity of infection (MOI). When excessive AAV virions are present, Rab11-positive routes serve as the express highways for AAV2 transduction [112]. During this endosomal trafficking, AAV capsids also undergo structure alterations and chemical modifications. As described earlier, VP1u with a PLA_2 domain and an NLS externalize from the five-fold axes of symmetry prior to endosomal escape. This conformational rearrangement in post-entry

steps is paramount for effective AAV transduction as microinjections of intact AAV particles do not guarantee infection [115]. Later, Arun Srivastava's group discovered that epidermal growth factor receptor protein tyrosine kinase (EGFR-PTK) phosphorylates some tyrosine residues on AAV2 capsids by casein kinase II (CKII), which is further proven to be an prerequisite step for nuclear entry [84,116,117]. AAV2 mutants evading tyrosine phosphorylation have higher transduction efficiency *in vitro*. Other components in signaling pathways involved in AAV2 intra-cellular trafficking include, but not limited to, rac1 and phosphatidylinositol-3 kinase activation, phosphorylate FKBP52 protein, and ubiquitin-proteasome pathways [114,118,119].

Cytoskeleton is the key player facilitating the translocation of AAV from cellular membrane to perinuclear region, as demonstrated in HeLa cells using quantitative 3D imaging techniques [108,119-122]. Direct visualization of the intact AAV particles in the nucleus indicates that uncoating and genome replication of AAV occur inside the nucleus [106,118,123].

1.7 Reengineering AAV Capsids

A prominent advantage of AAV vectors in gene therapy applications is its broad collection of tissue tropism, attained by its ever-growing collection of naturally occurring and artificially generated variants. In murine models, AAV2 and AAV8 are highly hepato-centric, while AAV1 transduces skeleton muscle ~100-fold more efficiently [124,125]. AAV4 is prevalent in transducing the cardiopulmonary system after systemic administration, while it targets astrocytes in the rostral migratory stream (RMS) if injected intraventrically into the brain. AAV5 transduces the hippocampus after CNS administration [126]. AAV6 demonstrates high transduction efficiency in both liver and skeleton muscle. AAV9 has a systemic

transduction profile after intravenous administration. Intriguingly, AAV9 is also able to cross the blood-brain barrier besides its systemic phenotype. Recently, AAV.rh8 and AAV.rh10, isolated from rhesus monkey tissues, show compelling brain transduction efficacy similar or greater than AAV9 in CNS [127-129]. Although the arsenal of naturally occurring AAV strains is expanding, pre-existing NAbs in human population, cross-species disparity of tropism, and dose-related toxicity still hurdles its transition from bench to bedside into clinical applications. To conquer these issues, the next generation of AAV vectors was introduced by viral protein engineering and directed evolution of AAV libraries.

In 2003, Muller and colleagues generated the first AAV capsid library by inserting a random peptide library on AAV2 capsid [130]. Panning and selection on human coronary artery endothelial cells gave rise to AAV mutants with cell line specific transduction profile. The next breakthrough in AAV library construction and screening took place in 2006. In this study led by N. Maheshri, random mutations were introduced to AAV2 VP1-3 by combining error-prone PCR and staggered extension process [131-133]. High-throughput screening of this library with $>10^6$ individual clones successfully isolate mutants with altered affinities towards heparan sulfates and acquired ability to evade antibody-neutralization compare to the parental AAV2. AAV libraries based on other serotypes sprout since then. Based on different selection strategies, AAV mutants with specific tissue-targeting properties were characterized. Directed evolution of a chimera library of AAV2 and AAV5 yields a novel AAV vector (AAV2.5T) efficient in transducing well-differentiated polarized human airway epithelium (HAE), projecting a promising candidate for cystic fibrosis treatment [134,135]. Mechanistic studies of AAV2.5T indicates that its increased infectivity on HAE is due to its higher affinity towards α2,3-*N*-linked SA on the respirotary airway than AAV5 and AAV2. Other

studies form David Schaffer's group developed AAV vectors targeting retina glia, Müller cells, pig airway epithelium, and neural stem cells for cell type specific applications [136-140].

Another strategy to construct AAV capsid library is DNA shuffling. Li W. et al. shuffled the VP sequences from AAV1-9 with exception of AAV7 to generate an AAV shuffling library [141]. Directed evolution of this library on the ciliated HAE resulted in two variants, HAE-1 and HAE-2, both of which possess higher transduction efficiency on HAE compared to the parental strains. Selection and screening of the same library further yielded AAV isolates capable of crossing seizure-compromised blood brain barrier [142].

During the endeavor to decipher AAV9 structure-function relation, Pulicherla et al. created an AAV9 capsid library with random mutations on the GH loop by error-prone PCR [95]. 10 out of 96 individual AAV9 clones were isolated and characterized in mice. Three variant, 9.24, 9.45, and 9.61, were identified as class II mutants displaying liver-detargeted phenotype. These vectors are well suitable for clinical applications to evade liver toxicity. In vitro analysis of the glycan binding property of each mutant also discovered that W503 and N498 are key residues involved in AAV9-galacotse interactions, as mentioned earlier.

Despite the prevalence of AAV libraries with random swaps/mutations/insertions, rational engineering based on receptor usage information has been an important strategy [143]. By substitute partial heparan sulfate binding footprint on AAV2 (585-RGNRQA-590) with corresponding residues from AAV8 (585-QQNTAP-590), AAV2i8 is unable to bind to heparan sulfate. The same variant also demonstrates muscle-tropic phenotype both in mice and non-human primates. Results from this study not only showed the versatility of AAV2 capsid, but also correlate AAV-receptor interactions with hepatotropic phenotype of AAV in general.

1.8 AAV in Gene Therapy

AAV2 was first discovered in an Adenovirus type 12 preparation as contaminant in 1965 [144]. AAV1, AAV3, and AAV4 were sequentially isolated from a simian Adenovirus 15 stock, an Adenovirus 7 stock, and African green monkey [145-147]. AAV5 and AAV9 were discovered from humans, while AAV7 and AAV8 were isolated from rhesus monkeys [124,148]. In addition, AAV10 and AAV11 were cloned from cynomolgus monkeys [149]. In 2004, a study led by Gao G. expand the collection of naturally occurring AAV to more than 100 members [124]. The pursuit of more naturally occurring AAV strains and reengineered variants is fueled by AAV's encouraging successes as gene therapy vectors.

The idea of using AAV to package transgene traces back to 1982 [150,151]. The era of gene therapy flourishes when therapeutic genes can be introduced to target cells or organs. Compared to other viral vectors, AAV has the advantage of no currently-known pathogenicity, prolonged transgene expression kinetics, and the ability to transduce both dividing and non-dividing cells. Transition of AAV vectors from bench-to-bedside was enabled by the extensive knowledge of the basic virology of the vectors described earlier.

Preclinical and clinical trials have been carried out with many AAV strains and emerging AAV variants [152]. For instance, AAV8 has been tested in nonhuman primates and dog models to express factor IX in liver, with the hope of alleviating hemophilia symptoms. AAV1 packaging the *SERCA2a* gene is in phase I/II clinical trial to treat cardiac failure. As described earlier, AAV5, HAE-1/2, and AAV2.5T have been investigate to develop cures for pulmonary diseases, such as cystic fibrosis. In the past few years, extremely exciting results of recombinant AAV vectors come from clinical trials of Leber's congenital amaurosis. AAV4 and AAV8 restoring retinal pigmented epithelium (RPE) functions showed success in

expression *RPE65* gene into RPE65-deficient dogs and nonhuman primate models. Patient treatments demonstrated over five orders of magnitude of progression in rod photoreceptor function. Using AAV as delivery machinery into the brain also saw encouraging results in treating Canavan's, Parkinson's, and Batten's diseases in long-term assessments.

Despite persistent successes in the clinical front, there are still concerns and challenges regarding to the pre-existing humoral immunity, cross-species difference in tropism, immune response to both AAV vector cassette and transgene products. Further improvement of existing vectors and development of new AAV vectors counts on a persistent accumulation of basic knowledge about AAV biology.

Table 1. Examples of Viruses and Their Cognate Glycan Receptors[1]

Virus Family	Virus Type	Glycan Receptor
Adenoviridae	Adeno 37 Adenovirus 2,5	α2-3 SA Heparan Sulfate
Arenaviridae	Lassa Virus	Dystroglycan Glycans
Caliciviridae Noroviruses	 Norwalk and others	 Histo-blood group
Coronaviridae	Coronavirus OC43	9-*O*-acetyl SA
Flaviviridae Hepaciviruses Flavivirus	 Hepatitis C Dengue Virus Japanese encephalitis Virus West Nile Virus	 Heparan Sulfate Heparan Sulfate Heparan Sulfate
Herpesviridae α-herpesviruses β-herpesviruses γ-herpesviruses	Herpes simplex virus types 1 & 2 Varicella-zoster virus Cytomegalovirus Human Herpesvirus Types 6 & 7 Human Herpesvirus type 8	Heparan Sulfate Heparan Sulfate Heparan Sulfate Heparan Sulfate
Ortomyxoviridae	Influenza A virus Influenza B virus Influenza C virus	α2-3 SA, α2-6 SA α2-3 SA, α2-6 SA 9-*O*-acetylsialic acid
Papillomaviridae Papillomavirus	Human papillomavirus types 11, 16, 33	Heparan sulfate
Paramyxoviridae Respirovurus Pneumovirus Metapneumov.	 Paramyxovirus 1-3 Respiratory syncytial virus Human metapneumovirus	 Sialic acid Heparan sulfate (chondroitin sulfate) Heparan sulfate
Parvoviridae Erythrovirus Dependovirus	(see section 1.2) B19 Adeno associated virus (AAV) types 4 & 5 AAV type 2	 Globoside/Histo-blood group P substance SA Heparan sulfate
Picornavirus Enterovirus Rhinovirus	 Enterovirus 70 Rhinovirus 87	 SA SA
Polyomaviridae polyomavirus	 JC and BK virus	 SA
Poxviridae Ortopoxvirus Reoviridae	 Vaccinia virus	 Heparan Sulfate (Chondroitin Sulfate)
Ortoreovirus	Reovirus 3	SA
Rotavirus	Rotavitus	SA
Retroviridae	HIV-1	Sulfatide; Galactosylceramide, HS

Table 2 Structure, Receptor, and Tropism of AAV1-9

AAV Serotype	VP Structure (PDB#)	Glycan Receptor	Protein Receptor	Tropism
1	3NG9	*N*-linked SA		Muscle, Lung
2	1LP3	Heparan Sulfate	αVβ5 & α5β1 Integrins, HGFR, FGFR, LamR	Liver, Neuron
3B	3KIE	Heparan Sulfate	LamR	Liver
4	2G8G	*O*-linked SA		Heart, Lung, Astrocytes
5	3NTT	*N*-linked SA	PDGFR	Muscle
6	3OAH	*N*-linked SA Heparan Sulfate	EGFR	Muscle
7				Muscle, Liver
8	2QA0			Liver
9	3UX1	*N*-linked Galactose	LamR	Systemic

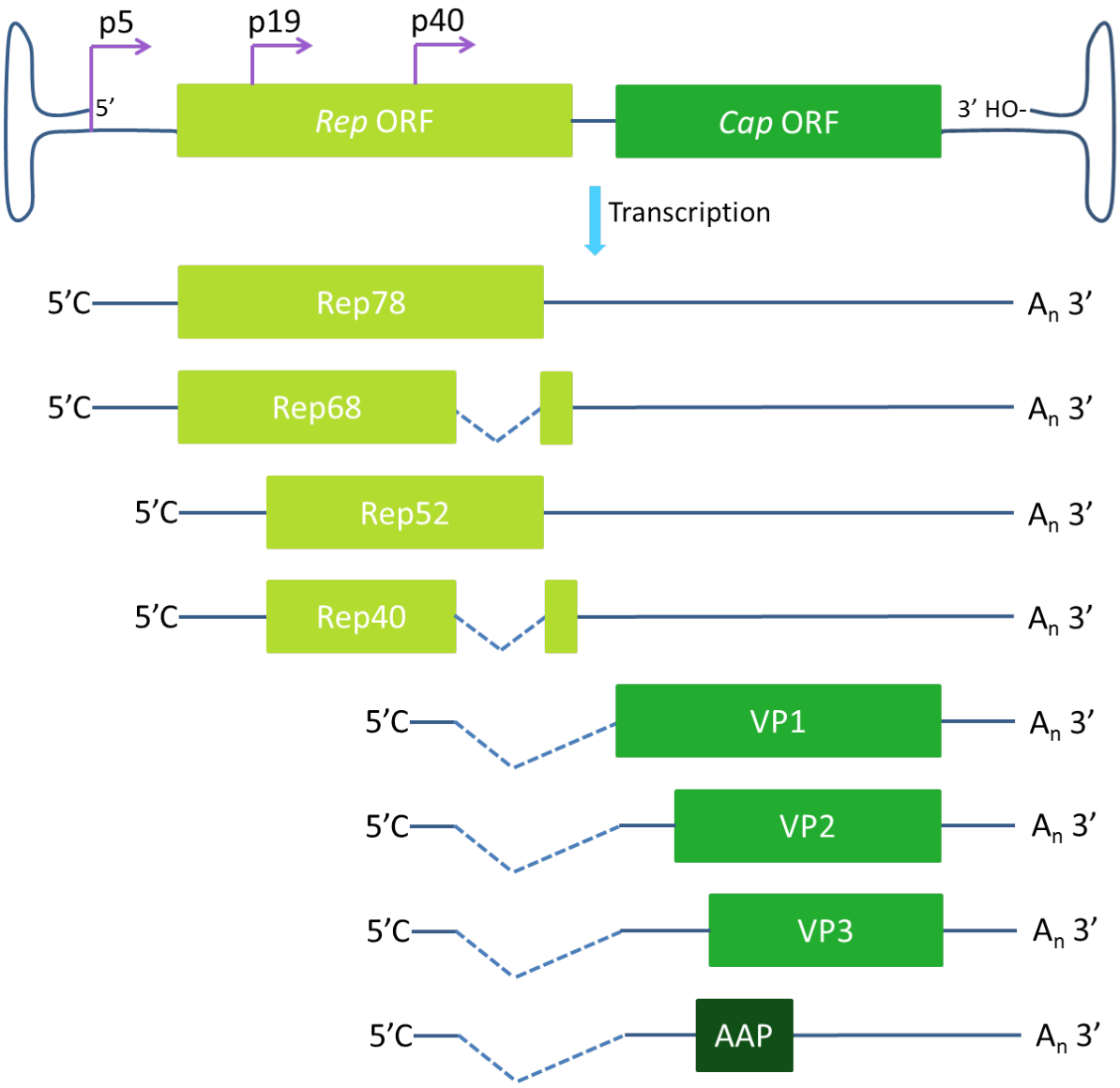

Figure 1. Genomic Map of AAV. There are two ORFs between ITRs, coding non-structural and structural proteins of AAV. The p5 promoter initiates transcription of Rep78 and Rep68, while p19 drives that of Rep52 and Rep48. Transcripts of non-structural protein, AAP, and structural proteins, VP1, 2, and 3, are driven by p40 promoter. Alternative splicing results in mRNAs coding for VP1 and VP2. Leaky scanning of the stop codon in the middle of VP2 yields VP3.

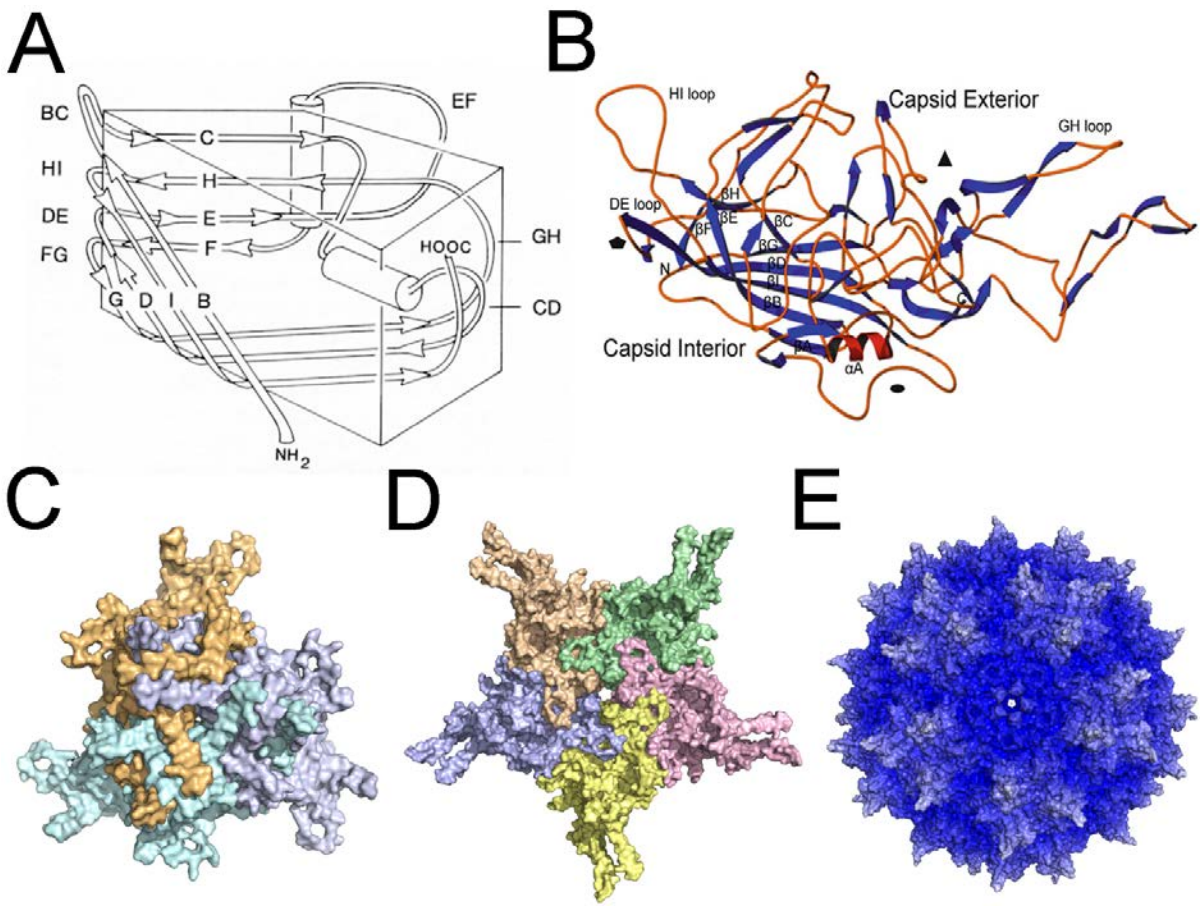

Figure 2. Structural Components of AAV Capsid. (A) Schematic diagram of the eight-strand jelly-roll β-barrel structure displayed by many icosahedral viral capsids. Adapted from reference 153. Eight β strands are sequentially numbered B-I. Loops connecting the neighboring two strands are named by numbers of the two connecting strands. **(B)** Ribbon rendering of AAV2 VP3 (PDB# 1LP3) showing β strands corresponding to (A) in blue. The conserved short α-helix is colored in red. DE loop and HI loop forming the five-fold axes of symmetry, and GH loop on the three-fold axes are labeled. Two-, three-, five-fold axes of symmetry are marked by oval, triangle, and pentagon, respectively to orient the view. **(C)** A trimer model of AAV2 VP3, with individual chain colored in palecyan, bluewhite, and lightorange in PyMOL. **(D)** A pentamer model of AAV2 VP3, showing the five-fold axes of symmetry. Each VP monomer is colored wheat, palegreen, lightblue, paleyellow, and light pink, respectively. **(E)** Visualization of a full capsid model of AAV2, looking into the five-fold axes of symmetry.

CHAPTER 2

Terminal *N*-linked Galactose is the Primary Receptor for AAV9[1]

2.1 Overview

A broad range of pathogens engage with cell surface sialoglycans to initiate infection, and parvoviruses are no exception. In the current study, we discovered that asialoglycans, instead of sialylated glycoconjugates, mediate cell surface binding and infectivity of AAV serotype 9, an isolate with systemic tropism in animals. Enzymatic removal of sialic acid, but not heparan sulfate or chondroitin sulfate, increased AAV9 transduction on multiple cell types. Viral binding and transduction assays on a series mutant Chinese Hamster Ovary (CHO) cell lines defective in glycan chain synthesis or translocation, revealed a pivotal role of core glycan residues beneath sialic acid moieties in AAV9 transduction. Treatment with chemical inhibitors of *N*- or *O*-glycosylation and competitive inhibition studies with a spectrum of lectins suggest that *N*-linked glycans with terminal galactosyl residues facilitate cell surface binding and transduction by AAV9. In addition, resialylation of galactosylated glycans on the sialic acid-deficient CHO Lec2 cell line with different sialyltransferases partially blocked AAV9 transduction. Quantitative binding curves of AAV9 on parental, sialidase-treated or sialic acid-deficient mutant CHO cells revealed a 3 to 10-fold increase in relative binding potential of AAV9 particles upon desialylation. Lastly, pretreatment of well-

[1]This chapter includes the original publication: Shen, S., Bryant, K.D., Brown, S.M., Randell, S.H., Asokan, A. *Terminal N-linked Galactose is the Primary Receptor for Adeno-associated Virus 9*, Journal of Biological Chemistry, 2011, 286(15):13532-40 (PMID: 21330365).

differentiated human airway epithelial cultures and intranasal instillation of recombinant sialidase in murine airways enhanced transduction efficiency of AAV9 by over an order of magnitude. In corollary, the studies described herein provide a molecular basis for low infectivity of AAV9 *in vitro* and a biochemical strategy to enhance gene transfer by AAV9 vectors in general.

2.2 Introduction

Cell surface glycans have been shown to play a critical role in the infectious pathways of viruses [1]. Detailed studies of virus-glycan interactions have yielded significant insight into mechanisms underlying emergence and transmission of viral pathogens in different hosts. Amongst various glycolipids, glycoproteins or proteoglycans anchored to the plasma membrane, sialylated glycans and heparan sulfate proteoglycans appear to serve as predominant substrates for viral attachment to the cell surface. For instance, heparan sulfate serves as a primary receptor for Herpesviridae [154] as well as certain adenoviruses [155] and parvoviruses [40,156]. Interactions between sialylated glycans and members of the Orthomyxoviridae [157], Reoviridae [158], Polyomaviridae [159] families and certain parvoviruses are also well known [45,48,160].

Adeno-associated viruses (AAV) are small, single-stranded DNA viruses that belong to the genus Dependovirus of the Parvoviridae family [144]. Recombinant AAV vectors, by virtue of their lack of pathogenicity and low immunogenicity, are currently being evaluated as lead candidates in clinical gene therapy trials [161]. The discovery of a large number of AAV isolates over the past decade has accelerated efforts to exploit tissue tropisms displayed by different strains for therapeutic gene transfer applications [162,163]. Successful translation from

bench to bedside will require a thorough understanding of molecular mechanisms underlying AAV infection. As with other viruses, attachment to cell surface glycans constitutes the first step in the AAV infectious pathway. For instance, several AAV serotypes have been shown to bind heparan sulfate proteoglycans (AAV2 [40]; AAV6 [44]), while others utilize sialic acid for cell surface binding and entry (AAV4 [46]; AAV5 [45]; AAV1/6 [44]; Bovine AAV [38]).

Sialylated glycans that serve as primary receptors for the latter AAV strains vary at the level of *N*-acetylneuraminic acid (Neu5Ac) linkage to underlying sugars, i.e., *N*-2,3 or *N*-2,6 linked to galactose residues [44,46]. Further receptor specificity has been demonstrated at the level of *N*-linked or *O*-linked glycans displayed on the cell surface [44,46]. Selective recognition of such linkages and underlying core glycan types [164] is likely enabled by differences in the capsid surface topology of AAV serotypes [82,165]. In general, dependence of AAV infectivity on sialic acid has been demonstrated using a battery of chemical, biochemical and genetic tools to desialylate cell surface glycans. The current study is focused on glycan interactions of a human AAV isolate, AAV serotype 9/Hu.14 (Clade F; [124]). Recombinant AAV9 vectors display widespread and robust transduction following systemic administration in animal models, but fail to infect cells in culture [124,125]. We demonstrate that efficient gene transfer by AAV9 vectors requires an atypical interaction with non-sialylated cell surface glycans. The results described herein provide a molecular basis for the low infectivity of AAV9 observed in cell culture.

2.3 Materials and Methods

Plasmids and viruses. All plasmids were obtained from the UNC vector core. The triple plasmid transfection protocol [166] utilized for production of AAV9 vectors, includes (i)

the AAV helper plasmid, pXR9, containing AAV2 *Rep* and AAV9 *Cap* genes, (ii) the Adenoviral helper plasmid, pXX6-80 and (iii) the vector genome cassette, pTR-CBA-Luc or pHpa-trs-SK, containing the firefly luciferase gene driven by the chicken beta-actin (CBA) promoter or self-complementary GFP cassette driven by the cytomegalovirus (CMV) promoter, respectively. The vector genome cassette is flanked by inverted terminal repeats (ITRs) required for packaging. The ITRs are the only elements within the vector genome cassette derived from the wild-type AAV genome, thereby eliminating 96% of viral elements. Recombinant AAV9 vectors generated thus allow quantitation of viral infectivity (or transduction efficiency) through luciferase transgene expression assays. HEK293 cells utilized for production of recombinant AAV9 vectors were obtained from the UNC vector core. Sonicated cell lysates and PEG8000 precipitates from supernatant were pooled and subjected to cesium chloride ultracentrifugation as described earlier [166]. Dialyzed peak fractions were subjected to quantitative PCR using a Roche Light Cycler instrument with luc transgene-specific primers to determine viral vector titers (forward 5'-AAA AGC ACT CTG ATT GAC AAA TAC-3'; reverse 5'-CCT TCG CTT CAA AAA ATG GAA C-3').

Cell lines. All cell lines were cultured in DMEM supplemented with 10% fetal bovine serum (FBS) and 1% penicillin, streptomycin, amphotericin B (Sigma) and maintained in 5% CO_2 at 37°C unless mentioned otherwise. COS-1 (monkey kidney), Neuro2a (mouse neuroblastoma), U87 (human glioma) and Huh7 (human hepatocarcinoma) were obtained from the UNC tissue culture facility and utilized in viral transduction assays. Chinese hamster ovary (CHO) Pro5 and mutant Lec1, Lec2 cell lines were a gift from Dr. Jude Samulski (UNC-Chapel Hill) and the CHO Lec8 cell line purchased from ATCC. All CHO cells, utilized for viral binding and transduction assays, were cultured in α-MEM

(GIBCO) supplemented with 10% FBS and penicillin, streptomycin, amphotericin B as outlined above. Well-differentiated human airway epithelial (HAE) cultures (4-6 weeks) grown on permeable membrane supports (Millipore, Corning, NY) at the air-liquid interface were provided by the Cell Culture Models Core and the UNC Cystic Fibrosis/Pulmonary Research Center.

Transduction assays. Different cell lines were seeded at 10^5 cells/well in 24-well plates and allowed to adhere overnight at 37°C. Plates were then pre-chilled at 4°C for 30 min and incubated with AAV9 vectors at a multiplicity of infection (MOI) of 1000 vector genomes per cell (vg/cell) to allow binding to the cell surface for 1.5 hrs at 4°C. Unbound virus was then removed by washing three times with ice cold 1x phosphate-buffered saline (1xPBS) and 0.5mL of DMEM added to each well. Luciferase transgene expression levels were quantitated after incubation for 24 hrs from cell lysates using a Victor 2 luminometer (Perkin Elmer). For studies with HAE, well-differentiated cultures were pretreated with 1.25 U/ml of sialidase A (Prozyme #GK80040) at 37°C for 3hrs followed by three washes with ice-cold 1xPBS. Cultures were then incubated with scAAV9-CMV-GFP vectors (MOI = 10^5 vg/cell) as outlined above. Fluorescence micrographs of green fluorescent protein (GFP) expression in HAE cultures at 2 weeks post-transduction were obtained using an Olympus epifluorescence microscope equipped with a 20x objective and a Hamamatsu camera.

Enzymatic desialylation assays. Different cell lines were treated with Heparinase I and III (from *Flavobacterium heparinum*; Sigma #H2519 and #H8891), Chondroitinase ABC (from *Proteus vulgaris* Sigma #C2905), Neuraminidase Type III (from *Vibrio cholerae* Sigma #N7885), endo-β-galactosidase (from *Pseudomonas* sp., Sigma #G6920) and α-fucosidase (from bovine kidney, Prozyme #GKX-5006) to determine the role of cell surface

glycans in AAV9 infection. Briefly, Cos1 cells were seeded at a density of $1x10^5$ cells/well in 24-well plates and pretreated with 50mU/ml neuraminidase, 3 U/ml heparinase I, 1.5 U/ml heparinase III, 1.5 U/ml chondroitinase ABC, 80 mU/ml endo-β-galactosidase, and 50 mU/ml α-fucosidase in DMEM at 37 °C for 2 hrs. Neuro2a, U87, HEK293 and Huh7 cells were treated with neuraminidase alone. Cells were then washed three times with 1xPBS and subjected to AAV9 infection at an MOI of 1000 vg/cell. Luciferase transgene expression assays were carried out as described earlier at 24 hrs post-infection.

Internalization assays. Human glioma U87 cells were untreated or pretreated with Neuraminidase Type III from *Vibrio Cholerae* followed by incubation with AAV1 or AAV9 at MOI = $1x10^5$ vg/cell. At 60 min post incubation on ice, unbound virions were removed with three 1xPBS washes. Cells were then incubated at 37°C for 30 min to allow uptake of AAV9 particles. Cells were then treated with 100µl/well of 0.05% Trypsin EDTA (Cellgro #25-052-CI) for 5min at 37°C to remove cell surface bound virions. Cells were then pelleted by centrifugation and washed thrice with 1xPBS. Cell lysates were further analyzed using Q-PCR to quantitate cell-associated (internalized virions) as described earlier. Internalized virions are expressed as vector genome copy numbers (vg) per cell. All experiments were carried out at n=5 and error bars represent standard error mean.

Chemical inhibition assays. CHO Lec2 cells were seeded at 10^5 cells/well in 24-well plates and pretreated for 24 hrs with small molecule inhibitors of glycosylation, Swainsonine (10µM; Sigma, #S8195) and α-benzyl-GalNAc (1µg/mL; Sigma, #B4894) to determine the role of *N*- and *O*-glycans in AAV9 infection. Cells pretreated with chemicals were subjected to AAV9 infection at an MOI of 1000 vg/cell and luciferase transgene expression assays carried out as described earlier.

Enzymatic resialylation assays. The sialic acid-deficient cell line, CHO Lec2 was treated with 50mU/mL each of α2,3-(N)-sialyltransferase (Calbiochem #566218), α2,6-(N)-sialyltransferase (Calbiochem #566222) or α2,3-(O)-sialyltransferase (Calbiochem, #566227) and 1mM CMP-sialic acid (Sigma) in serum-free media for 3 hrs at 37°C. Untreated CHO Pro5 cells with endogenous sialic acid were included as control. Cells were then rinsed three times with 1xPBS and subjected to AAV9 infection at an MOI of 1000 vg/cell and luciferase transgene expression assays carried out as described earlier.

Lectin inhibition and staining assays. Competitive inhibition of AAV9 infection was carried out using a panel of lectins, Concanavalin A (Con A), Wheat germ agglutinin (WGA), *Maackia amurensis* lectin (MAL I), *Sambucus nigra* lectin (SNA) and *Erythrina cristagalli* lectin (ECL) obtained from Vector Labs (Burlingame, CA). Briefly, prechilled CHO Pro5 or mutant Lec2 cells were incubated with 100mg/mL FITC-conjugated lectin (fluorescent labeling assay) in serum-free α-MEM media or along with AAV9 particles (transduction assay at MOI = 10,000 vg/cell) for a period 1.5 hrs at 4°C. After three washes with ice cold 1xPBS to remove unbound virus and/or lectins, cells were imaged using an Olympus epifluorescence microscope equipped with a Hamamatsu camera or incubated for 24 hrs at 37°C prior to luciferase transgene expression analysis.

Cell surface binding assays. CHO Pro5 cells were seeded at a density of 10^4 cells/well in 96-well plates prior to treatment with with 50mU/mL neuraminidase type III from *Vibrio cholerae* for 2 hrs at 37°C. Untreated Pro5 and Lec2 cells were included as negative and positive controls respectively. Cells were then pre-chilled at 4°C for 30 minutes, followed by incubation with AAV9 particles at MOI of $1x10^2$, $5x10^2$, $1x10^3$, $5x10^3$, $1x10^4$, $5x10^4$, $1x10^5$, $5x10^5$ vg/cell for 1.5 hrs at 4°C. Cells were then subjected to three washes with

ice cold 1xPBS to remove unbound virions. Cell surface-bound virions were collected along with cell lysates following three freeze-thaw cycles and vector genome copy numbers per cell determined using Q-PCR as outlined earlier. Binding curves were generated using GraphPad Prism 5 software by applying the single site binding model (**$Y = B_{max}*X/(K_d + X)$**), where **Y** represents number of bound virions/cell determined by Q-PCR; **X** represents MOI; **B_{max}** is the maximum binding capacity and **K_d'**, the observed disassociation constant.

Animal studies. All experiments were carried out with 6-8 week old female BALB/c mice (Jackson Labs, Bar Harbor, ME) maintained and treated in accordance with National Institutes of Health guidelines and as approved by IACUC at UNC-Chapel Hill. Mice were administered via intranasal instillation with either 100 µl PBS (50 ul/nostril) or 100 µl Neuraminidase Type III from *Vibrio cholerae* (200 mU, Sigma, St. Louis, MO). At 2 hrs post-treatment, a dose of $5x10^{10}$ AAV9 particles in 1xPBS (50 µl/nostril) was administered. Luciferase transgene expression in live animals was obtained using an Xenogen IVIS Lumina® imaging system (Caliper Lifesciences, CA) after intranasal instillation of luciferin substrate (120 mg/kg; Nanolight). Image analysis was carried out using the Living Image software® (Caliper Lifesciences) and luciferase expression reported in relative light units (photons/sec/cm^2/sr).

2.4 Results

Neuraminidase treatment selectively increases AAV9 transduction in different cell types. Enzymatic removal of different cell surface glycans was achieved by treating cells with different glycosidases. Removal of terminal sialic acid residues using neuraminidase (from *V. Cholerae*) enhanced AAV9 transduction by > 1 log unit when compared to virus

alone on COS1 cells **(Figure 3A)**. In contrast, hydrolysis of cell surface heparan sulfate or chondroitin sulfate proteoglycans using heparinase I/III or chondroitinase ABC, respectively had no effect on viral transduction compared to control. Neuraminidase treatment abrogated AAV1 transduction, while AAV2 remained unaffected under these conditions **(Figure 4A)**. In addition, treatment of cell lines derived from different tissues with neuraminidase (grey bars) prior to AAV9 infection enhanced transduction efficiency by nearly 2 log units in contrast to infection with virus alone (white bars) **(Figure 3B)**. These results were corroborated by increased binding and internalization of AAV9, but not AAV1 vectors upon sialidase treatment in these cell lines **(Figure 4B & 4C)**. Thus, sialic acid appears to mask cell surface glycans that selectively facilitate AAV9 infection *in vitro*. Further, enzymatic desialylation might serve as a facile biochemical strategy to enhance transduction efficiency of AAV9 vectors *in vitro* and might enable detailed analysis of intracellular trafficking pathways.

Mutant CHO Lec2 cells lacking terminal sialic acid are highly permissive to AAV9 infection. Analysis of cell surface binding and infectivity of AAV1 and AAV9 on parental (Pro5) and mutant CHO cell lines was carried out to further understand the role of core glycans under sialic acid in AAV9 infection. The CHO Lec2 cell line lacks terminal sialic acid due to a defect in CMP-sialic acid transport [167], while Lec8 and Lec1 cell lines are defective in translocation of UDP-galactose and *N*-acetylglucosaminyl transferase activity [168,169], respectively. Correspondingly, cell surface glycans on CHO Lec2 cells contain terminal galactosyl residues, while Lec8 and Lec1 cell lines predominantly display terminal *N*-acetylglucosamine and mannosylated glycans, respectively (**Figure 6**). As seen in **Figure 6B & C**, cell surface binding and transduction of AAV9 particles (grey bars) on Lec2 cells is

significantly increased (> 1 log unit) when compared to the parental Pro5 cell line. In contrast, AAV1 (white bars), which requires sialic acid for infection shows ~ 10-fold decrease in binding and transduction in all CHO Lec mutant cell lines. In addition, no major changes in binding and infectivity are observed in case of Lec8 and Lec1 cells for AAV9 particles. These results suggest that galactosylated glycans immediately underlying sialic acid can facilitate AAV9 cell surface binding and entry. The results also support the notion that AAV9 particles might exploit an inefficient and non-specific uptake mechanism in the parental Pro5, mutant Lec8 and Lec1 cell lines.

Sialylation of *N*-linked glycans blocks AAV9 infection. In order to further elucidate the nature of glycans required for AAV9 infection, we utilized small molecule inhibitors of glycosylation and sialyltransferases to modify terminal galactosyl residues on the sialic acid-deficient Lec2 cell surface. Swainsonine [170] and α-benzyl-*O*-GalNAc [171] are chemical inhibitors of *N*-linked and *O*-linked glycosylation, respectively. Treatment with these reagents results in a corresponding decrease in cell surface expression of *N*-linked glycans and *O*-linked glycans. As seen in **Figure 7A**, AAV9 infection is significantly blocked by swainsonine (~ 75%), while α-benzyl-*O*-GalNAc has a modest inhibitory effect (~25%). These results suggest that AAV9 prefers *N*-linked cell surface glycans for infection. In addition, resialylation of the sialic acid-deficient Lec2 cell surface was carried out to understand the nature of the sialic acid linkage that blocks AAV9 infection (**Figure 7B**). The abilities of different sialyltransferases to partially block AAV9 infection was observed to follow the order, α2,3-*N*-sialyltransferase (3'NST) > α2,6-*N*-sialyltransferase (6'NST) > α2,3-*O*sialyltransferase (3'OST). Alteration of cell surface glycans upon resialylation was confirmed by staining with lectins recognizing different glycan residues and linkages (**Figure**

7C). As expected Pro5 expressing sialylated glycans and Lec2 cells expressing asialoglycans demonstrate preferential staining by FITC-MAL I and FITC-ECL, respectively. Resialylation of Gal(β1,4)GlcNAc residues with 3'NST, but not 3'OST partially restores FITC-MAL-I staining concurrent with decrease in AAV9 transduction. The lack of MAL I staining in 6'NST-treated Lec2 cells is expected due to lack of recognition of α2,6-sialylated glycans by MAL I. Taken together, these results not only corroborate the important role played by core *N*-linked glycans in AAV9 infection, but also the potential for α2,3- and α2,6-sialic acid linkages to block infection by masking underlying glycoconjugates on the cell surface.

Terminal galactosyl residues are critical for AAV9 infection. To better understand the nature of AAV9-glycan interactions that mediate infection, we carried out competitive inhibition studies with lectins that recognize different glycan linkages on the cell surface (**Figure 8A & B**). Specifically, we utilized (i) Maackia amurensis lectin (MAL I), which recognizes native, *N*-2,3-sialylated or sulfated glycoconjugates having Gal-(β1,4)-*N*-GlcNAc structures [172,173]; (ii) Sambucus nigra lectin (SNA), which binds preferentially to α2,6-sialylated galactose residues [174]; (iii) Erythrina cristagalli lectin (ECL), which demonstrates specificity towards galactose residues, in particular, galactosyl-(β1,4)-*N*-acetylglucosamine (Gal-(β1,4)-*N*-GlcNAc) [175]; (iv) Wheat germ agglutinin (WGA), which binds *N*-acetylglucosamine (*N*-GlcNAc) and also tolerates glycoconjugates containing sialic acid with different linkages [176] and (v) Concanavalin A (Con A), which recognizes mannose residues [177]. The SNA lectin had no effect on AAV9 infection and can be explained by low levels of α2,6-sialylated glycans in both cell lines of rodent (hamster) origin [177]. On the other hand, the MAL I lectin demonstrated 5 to 10-fold inhibition of AAV9 infection in both Pro5 and Lec2 cell lines. More importantly, a striking difference in AAV9 transduction was

observed in case of ECL-treated cells with 5-fold inhibition in parental Pro5 cells and nearly 200-fold inhibitory activity in the Lec2 cell line demonstrating the importance of core Galβ1-linked residues. The ConA and WGA lectins demonstrated broad inhibitory activity in both parental Pro5 cells and the sialic acid-deficient Lec2 cell line suggesting that underlying *N*-GlcNAc and core mannose residues might contribute to AAV9-glycan interactions.

These results were corroborated by fluorescent staining of Pro5 and Lec2 cell lines with aforementioned lectins (**Figure 8C**). As expected, MAL I and WGA lectins, which bind to α2,3-sialylated glycans, preferentially label Pro5 cells. In contrast, preferential ECL staining is noted in Lec2 cells. The SNA lectin, which recognizes α2,6-sialic acid does not stain either Pro5 or Lec2 cells confirming the lack (or modest expression) of α2,6-sialylated glycans on these hamster-derived cell lines. Further evidence supporting the important role of galactose in mediating AAV9 infection was obtained by treatment of CHO Pro5 cells with endo-β-galactosidase (**Figure 7A & B**). Removal of galactose residues, but not fucose abrogated AAV9 transduction. Taken together, these results suggest that terminal galactosyl residues serve as the primary receptor for AAV9.

Desialylation increases the cell surface binding potential of AAV9 particles. Cell surface binding curves were generated to establish a quantitative biochemical rationale for the observed increase in AAV9 transduction following desialylation. Binding of AAV9 particles to the surface of parental CHO Pro5 cells, sialidase-treated Pro5 cells or sialic acid-deficient Lec2 cells was carried out over a range of multiplicities of infection (genome-containing viral particles (vg) per cell). As seen in **Figure 9**, enzymatic removal of sialic acid partially recapitulates the effect of genetic desialylation by increasing the number of cell surface-bound AAV9 particles. Non-linear regression analysis of binding data was carried

out using the single site binding model ($Y = B_{max}*X/(K_d' + X)$), where X and Y represent multiplicity of infection and number of bound AAV9 particles, respectively; B_{max} is the maximum binding capacity and K_d' is the relative (observed) binding affinity. Calculation of aforementioned parameters reveals B_{max} values averaging 375 vg/cell on wild type Pro5 cells, a ~3-fold increase to 1075 vg/cell on sialidase-treated Pro5 cells and a ~5-fold increase to 1971 vg/cell on the sialic acid-deficient Lec2 cell line. These results support the observation that AAV9 prefers asialo *N*-glycans with terminal Galβ1-linked residues. In addition, an apparent decrease in K_d' (~3-fold) is observed in Lec2 cells as well as a corresponding 10-fold increase in relative binding potential (B_{max}/K_d'; **Table 3**). These results support the notion that enhanced avidity plays a role in mediating AAV9 binding to asialo *N*-glycans.

Sialidase pretreatment increases AAV9 gene transfer efficiency in HAE and murine airways. To evaluate the potential of enzymatic desialylation as a strategy to enhance gene transfer by AAV9 vectors, we evaluated the effect of sialidase pretreatment on AAV9 transduction efficiency in well differentiated HAE cultures and murine airways. A marked increase in GFP positive HAE was observed upon treatment with neuraminidase from *Arthrobacter Ureafaciens* (recombinant Sialidase A) prior to infection with scAAV9/CMV-GFP vectors (**Figure 10A & B**). Further, quantitation of live animal bioluminescent images of Balb/c mice following intranasal instillation of sialidase showed ~ 10-fold increase in luciferase transgene expression in murine airways mediated by AAV9/CBA-Luc vectors (**Figure 10C & D**). These results further validate the role of asialo *N*-glycans as the primary receptor for AAV9 in mouse and human species as well as provide a novel adjuvant strategy to enhance AAV9 gene transfer.

2.5 Discussion

Glycomic profiles of major *N*-and *O*-glycans expressed in the parental CHO Pro5 and the CHO Lec2 mutant [178,179] were instrumental towards delineating the nature of glycans that play a role in cell surface binding and infection by AAV9 vectors. The *N*-glycan profile of CHO Pro5 cells has been shown to possess complex bi-, tri-, and tetra-antennary structures bearing multiple *N*-acetyllactosamine (LacNAc) extensions, capped with sialic acid (NeuAc) residues. In addition, the *O*-glycan profile contains Gal-(β1,3)-GalNAc core structures that are mono- or di-sialylated. In contrast, the most abundant glycans produced by the sialic acid-deficient Lec2 cell line are asialo *N*-glycans, possessing between 2 and 7 LacNAc units. The *O*-glycan profile is known to be similarly affected by altered sialylation [179]. These observations suggest that LacNAc units might serve as cell surface attachment factors for AAV9. Treatment of the Lec2 cell line with chemical inhibitors and sialyltransferase enzymes confirmed that Galβ1-linked *N*-glycans rather than *O*-glycans were the preferred substrate for attachment. These results were further corroborated by competitive inhibition studies with the LacNAc-specific Erythrina cristagalli lectin (ECL). In addition, ConA and WGA lectins revealed a potential contributing role for mannose and GlcNAc residues underlying terminal galactose in mediating AAV9 infection. Further evidence supporting the aforementioned results was also obtained from competitive inhibition studies of AAV9 particles co-incubated with different glycan residues. Although only a modest inhibitory effect was observed potentially due to low affinity interactions, monomeric LacNAc (LN) units, but not α2,3-sialylated LN (3'-SLN) or α2,3-sialylated di-LN (3'-S-Di-LN) glycans appear to selectively block AAV9 infection [52].

Cell surface binding studies constitute a direct approach to determine the relative binding potential (B_{max}/K_d') of AAV9 particles under physiological conditions. Upon desialylation, the observed increase in number of binding sites (B_{max}) is likely due to increased availability of asialo *N*-glycans on the cell surface. The increase in relative binding affinity (K_d') is potentially a consequence of the increased number of binding sites and can be explained by enhanced avidity arising from multivalent interactions. The aforementioned results suggest that the relatively low abundance of asialo *N*-glycans provides a molecular basis for the low transduction displayed by AAV9 in cell culture. In contrast to *in vitro* assays, studies in rodent, canine and primate models have demonstrated robust and widespread systemic gene transfer using AAV9 vectors following intravenous administration [125,180,181]. As such, the nature of AAV9-glycan interactions does not provide sufficient rationale for the observed lack of *in vitro-in vivo* correlation outlined above. As seen in case of other viruses, potential interactions with blood components [182] or co-receptors such as integrins [183] might play an important role in dictating the transduction profile observed in animal models. Nevertheless, it is tempting to speculate that the relative abundance of asialo *N*-glycans in various animal tissues likely contributes to the broad biodistribution pattern and tissue tropism displayed by AAV9 vectors in vivo. Studies with animal models lacking sialyltransferases [184,185] might provide further insight into the mechanism(s) underlying tissue tropism displayed by AAV9.

Carbohydrate receptors utilized by members of different AAV clades appear to fall under two classes, namely, heparan sulfate proteoglycans and sialylated glycans [186]. For instance, AAV2, a Clade B member utilizes heparan sulfate proteoglycan as a primary receptor [40]. The closely related strains, AAV1 and AAV6 of Clade A appear to equally prefer

for α2,3- and α2,6- *N*-linked sialic acid for infection [44]. Further, AAV6 has been shown to bind heparin implying a potential dual mechanism of interaction with cell surface glycans [43]. Serotypes AAV4 and AAV5, which have not been assigned to any clade thus far, require α2,3-*O*- and α2,3-*N*-linked sialic acid for cell surface binding, respectively [46]. The current study identifies a third class of glycan receptors with terminal galactose utilized by the AAV strain Hu.14/AAV9 for infection. The latter serotype has been classified under Clade F within the AAV phylogenetic tree [124]. The major capsid protein (VP3, viral protein subunit 3) of other AAV isolates within Clade F is largely similar to Hu.14/AAV9 (GenBank# AY530579.1). Specifically, isolate Hu.31 (GenBank# AY530596.1) differs from Hu.14/AAV9 by 2 amino acid residues (S386G, N716S), while the VP3 subunit of isolate Hu.32 (GenBank# AY530597.1) is identical to Hu.14/AAV9. These observations suggest that at least one other member of Clade F (Hu.32) might utilize non-sialylated glycans for cell surface binding and entry. Lastly, Akache et al. [97] have previously reported the role of laminin receptor (LamR) in mediating transduction by AAV9. Whether the 2.5-fold increase in AAV9 transduction mediated by LamR in the previous report within transfected NIH 3T3 cells is due to enhanced binding or internalization is unknown. Nevertheless, low levels of LamR expressed in CHO cells [187] might account for residual transduction of AAV9 seen in the current study.

Recombinant sialidase has been proposed as a therapeutic agent in several clinical modalities. For instance, a sialidase fusion protein has been proposed as an antiviral agent for protection from influenza viruses [188,189]. The latter approach hinges on the inhibition of cell surface binding by influenza viruses due to desialylation of cell surface glycans. In another recent study, intrathecal infusion of recombinant sialidase, which reverses inhibition of axon

outgrowth in neurons, has been proposed as a candidate therapy for spinal cord injury [190]. In the current study, enzymatic desialylation was found to markedly enhance binding and transduction of AAV9 across different cell types regardless of host or tissue type. In addition, intranasal instillation of sialidase markedly enhanced gene transfer efficiency of AAV9 in murine airways supporting the potential application of recombinant sialidase as an adjuvant in therapeutic gene transfer applications in the lung/airways. Specifically, co-administration or pretreatment of different tissue types such as the lung, CNS or eye with recombinant sialidase might serve as (i) a strategy to expose high avidity glycan receptors and consequently restrict AAV9 transduction to these specific tissue types; (ii) a facile biochemical strategy to increase gene transfer efficiency of AAV9 vectors and (iii) a method to evaluate AAV9 vectors in desialylated preclinical models eliminating cross-species variation in sialylation patterns.

Table 3. Binding Parameters for AAV9 Interactions with Cell Surface Glycans

	Pro5	Pro5 + Sialidase	Lec2
B_{max} (vg/cell)	$3.75x10^{2} \pm 8.60x10^{1}$	$1.03x10^{3} \pm 1.88x10^{2}$	$1.97x10^{3} \pm 2.52x10^{2}$
K_d' (vg/cell)	$3.58x10^{5} \pm 1.64x10^{5}$	$3.34x10^{5} \pm 1.26x10^{5}$	$1.22x10^{5} \pm 4.12x10^{4}$
B_{max}/K_d'	$1.05x10^{-3}$	$3.07x10^{-3}$	$1.62x10^{-2}$
R^2	0.97	0.97	0.96

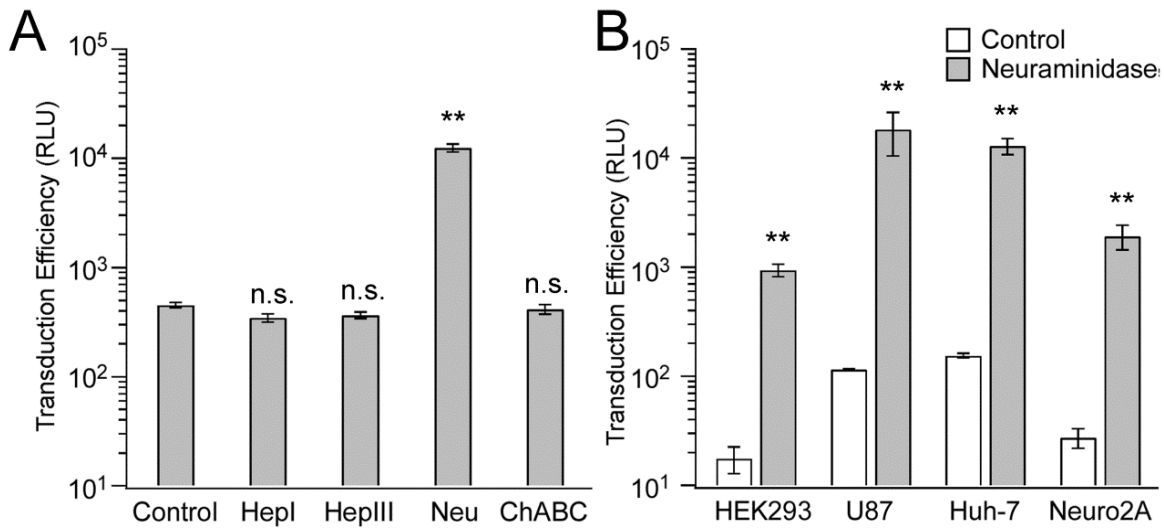

Figure 3. Enzymatic desialylation enhances AAV9 infectivity on multiple cell lines. (A) A series of glycosylases (heparinase I; heparinase III, neuraminidase from *Vibrio Cholerae* and chondroitinase ABC) were utilized to modify the glycoconjugates prominently expressed cell surface glycans on COS-1 cells. Transduction assays with AAV9 vectors (MOI = 1000 vg/cell) were carried out 2 hrs post-enzymatic treatment. **(B)** Different cell lines (HEK293, U87, Huh-7 and Neuro2a) were untreated (white bars) or pretreated with neuraminidase from *Vibrio Cholerae* (grey bars) followed by subsequent infection with AAV9 vectors (MOI = 1000 vg/cell). Luciferase transgene expression (Relative Light Units, RLU) was quantified for both studies at 24 hrs post-infection. All experiments were carried out in triplicate. Error bars represent standard error mean. Statistical significance was analyzed using the one-tailed Student's *t*-test (* $p < 0.05$; ** $p < 0.01$).

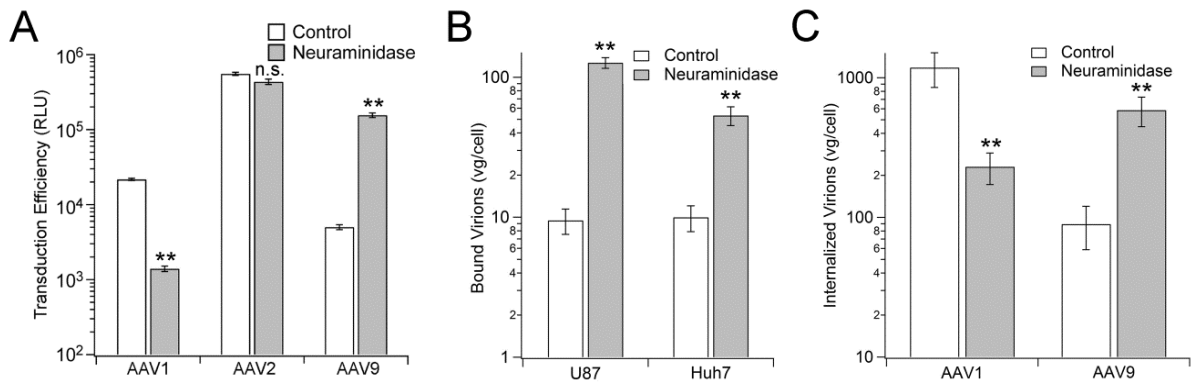

Figure 4. Effects of enzymatic desialylatoin on transduction, binding, and internalization of AAV9 on mammalian cell cultures. (A) COS-1 cells were pretreated with PBS or neuraminidase prior to infection with AAV1, AAV2 or AAV9 vectors (MOI = 1000 vg/cell). 24 hrs post infection, luciferase assays were carried out to demonstrate serotype-specific effects of enzymatic desialylation on transduction efficiency. **(B)** Effect of enzymatic desialylation on AAV9 binding to different human cell lines. Briefly, U87 and Huh-7 cells were seeded at a density of $1x10^5$ cells/well on a 24-well plate. Cells were untreated (white bars) or pretreated with 50 mU/ml of Neuraminidase Type III from *Vibrio Cholerae* (grey bars) prior to pre-chilling and incubation with AAV9-CBA-Luciferase vectors (MOI = 1000 vg/cell) at 4°C for 1.5 hrs. Unbound virions were removed by three washes with ice-cold 1xPBS. The amount of cell surface-bound AAV9 virions was quantified using Q-PCR. The amount of bound virions was expressed as vector genome copy numbers (vg) per cell. All experiments were carried out in quadruplicate. Error bars represent standard error mean. **(C)** Effect of desialylation on the internalization of AAV1 and AAV9 in U87 cells. U87 cells were untreated (white bars) or pretreated with Neuraminidase Type III from *Vibrio Cholerae* (grey bars) followed by incubation with AAV1 or AAV9 at MOI = 10^5 vg/cell. At 60 min post incubation on ice, unbound virions were removed with three 1xPBS washes. Cells were then incubated at 37°C for 30 min to allow uptake of AAV9 particles. Incubation with 100 µl/well of 0.05% Trypsin EDTA (Cellgro #25-052-CI) for 5 min at 37°C to remove cell surface bound virions. Cells were then pelleted by centrifugation and washed thrice with 1xPBS. Cell lysates were then analyzed using Q-PCR to quantitate cell-associated (internalized virions). Internalized virions are expressed as vector genome copy numbers (vg) per cell. Data are shown as mean ± s.e.m. (n=5). Statistical significance was analyzed using the one-tailed Student's *t*-test (* $p < 0.05$; ** $p < 0.01$).

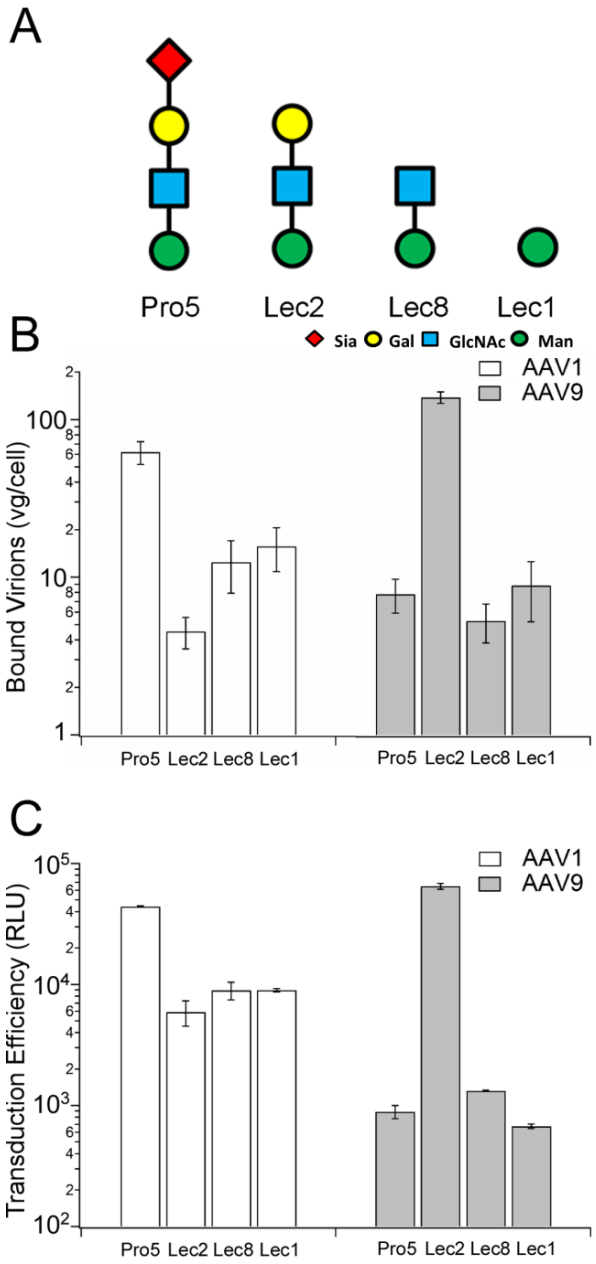

Figure 5. Glycan chain composition on cell surface affects binding and transduction of AAV9. (A) Schematic representation of *N*-glycan compositions of the parental CHO Pro5 cell line and mutants Lec2, Lec8 and Lec1 using nomenclature proposed by the Consortium for Functional Glycomics nomenclature committee. **(B)** Binding of AAV1 (white bars) and AAV9 (grey bars) particles to wild type and parental CHO cell lines was analyzed using Q-PCR as described in Materials and Methods. Number of bound virions is expressed as vector genome copy number (vg) per cell. **(C)** Transduction efficiency (Relative Light Units, RLU) of AAV1 (white bars) and AAV9 (grey bars) particles (MOI = 1000 vg/cell) on parental and mutant CHO cell lines were analyzed by quantifying luciferase transgene expression at 24 hrs post-infection. All binding experiments were carried out in quadruplicate and infectivity assays in triplicate. Error bars represent standard error mean.

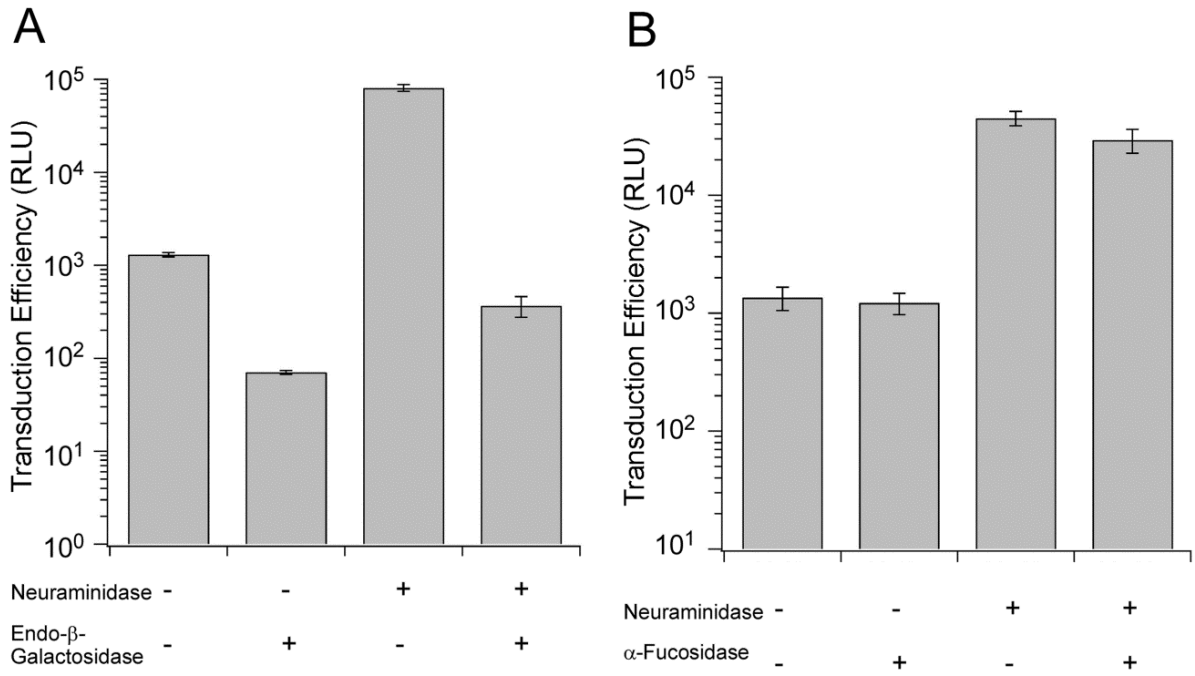

Figure 6. Endo-β-galactosidase and α-fucosidase decreases AAV9 transduction efficiency on cell culture. Cos-1 cells were untreated (-) or pretreated (+) with Neuraminidase Type III from *Vibrio Cholerae* followed by no treatment (-) or treatment (+) with **(A)** endo-β-galactosidase (80 mU/ml, from *Pseudomonas* sp., Sigma #G6920) and **(B)** α-fucosidase (50 mU/ml, from bovine kidney, Prozyme #GKX-5006). Transduction of AAV9 were carried out with MOI = 1000 vg/cell. Transduction efficiency (RLU) was quantified at 24 hrs post-infection. All experiments were carried out in triplicate. Error bars represent standard error mean.

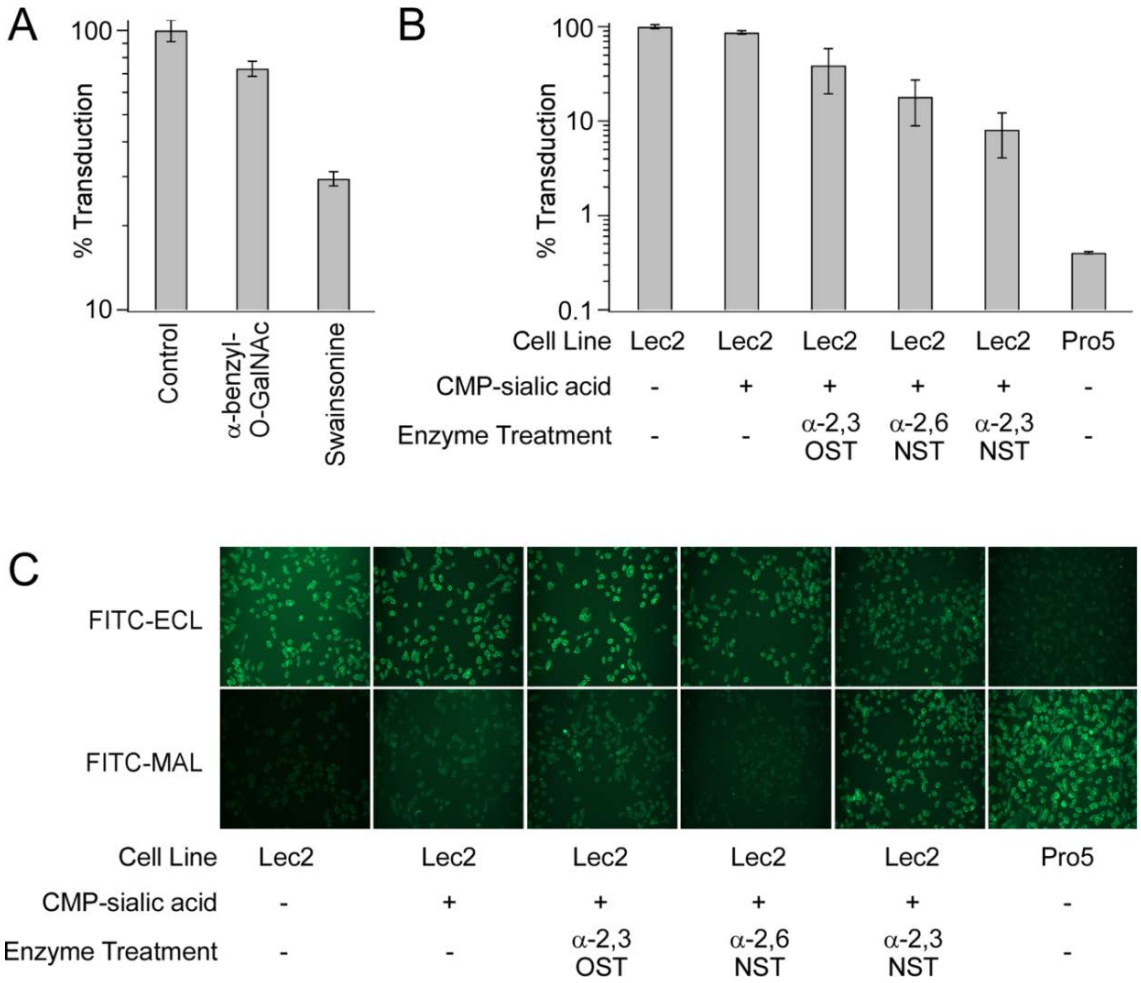

Figure 7. *N*-linked galactosylated glycans are prerequisite for AAV9 transduction. (A) Effect of glycosylation inhibitors on AAV9 transduction. Prior to infection with AAV9 vectors (MOI = 1000 vg/cell), CHO Lec2 cells were cultured in media containing α-Benzyl-GalNAc (1 μg/mL) or Swainsonine (10 μM), inhibitors of *O*- and *N*-glycosylation, respectively. **(B)** Effect of enzymatic resialylation on AAV9 transduction. CHO Lec2 cells were treated with 50 mU/mL each of α2,3-(N)-sialyltransferase (α2,3NST), α2,6-(N)-sialyltransferase (α2,6NST) or α2,3-(O)-sialyltransferase (α2,3OST) and 1 mM CMP-sialic acid (Sigma) to resialylate cell surface asialo glycans. Luciferase transgene expression was quantified at 24 hrs post-infection and expressed as % infectivity with respect to untreated or wild type (CHO Pro5) control. All experiments were carried out in triplicate. Error bars indicate standard error mean. **(C)** Lectin staining of CHO cell lines subjected to enzymatic resialylation. CHO Lec2 cells treated with CMP-sialic acid alone or with CMP-sialic acid and different sialyltransferases were subjected to lectin staining using FITC-labeled ECL, which exclusively recognizes Gal(β1,4)GlcNAc or FITC-labeled MAL, which recognizes α2,3-Sialylated Gal(β1,4)GlcNAc. Untreated wild type CHO Pro5 cells, which show high levels of FITC-MAL I staining and untreated Lec2 cells, which show high levels of FITC-ECL staining were included as control.

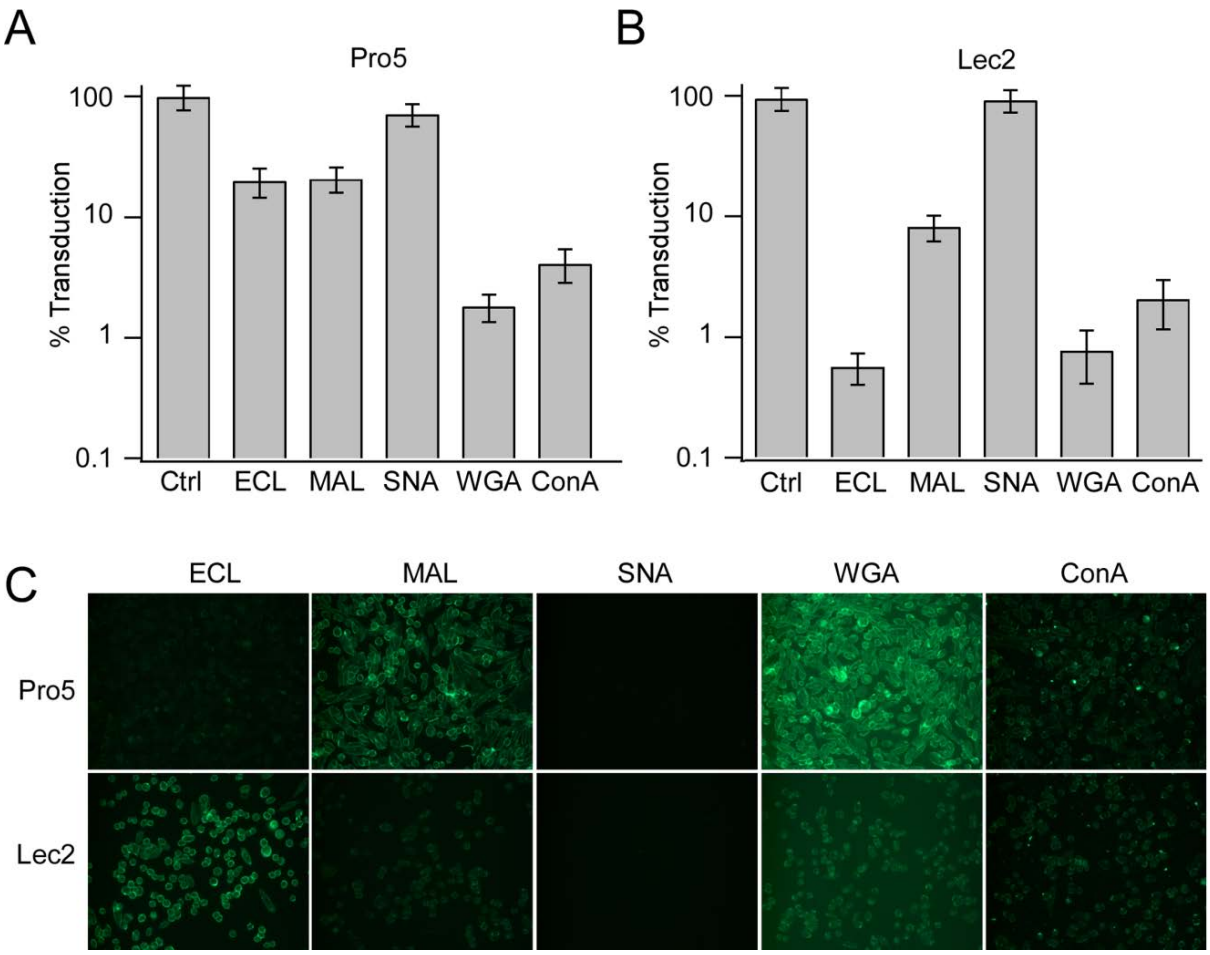

Figure 8. Galactose-specific lectin competitively inhibits AAV9 transduction on Lec2. Wild type CHO Pro5 **(A)** and sialic acid-deficient Lec2 **(B)** cell lines were co-incubated with lectins recognizing different glycan linkages (100 mg/mL) and AAV9 vectors (MOI = 10,000 vg/cell). Briefly, ConA lectin recognizes mannose residues, while WGA lectin recognizes GlcNAc residues as well as sialic acid residues. MAL and SNA lectins recognize α2,3- and α2,6-linked sialic acid residues, respectively, while ECL exclusively recognizes Gal(β1,4)GlcNAc residues. Luciferase transgene expression (% Infectivity) was quantified at 24 hrs post-infection. All experiments were carried out in triplicate. Error bars represent standard error mean (n=4). **(C)** Lectin staining of Pro5 and Lec2 cell lines were carried out using FITC-labeled lectins and fluorescent images obtained as described in Materials and Methods.

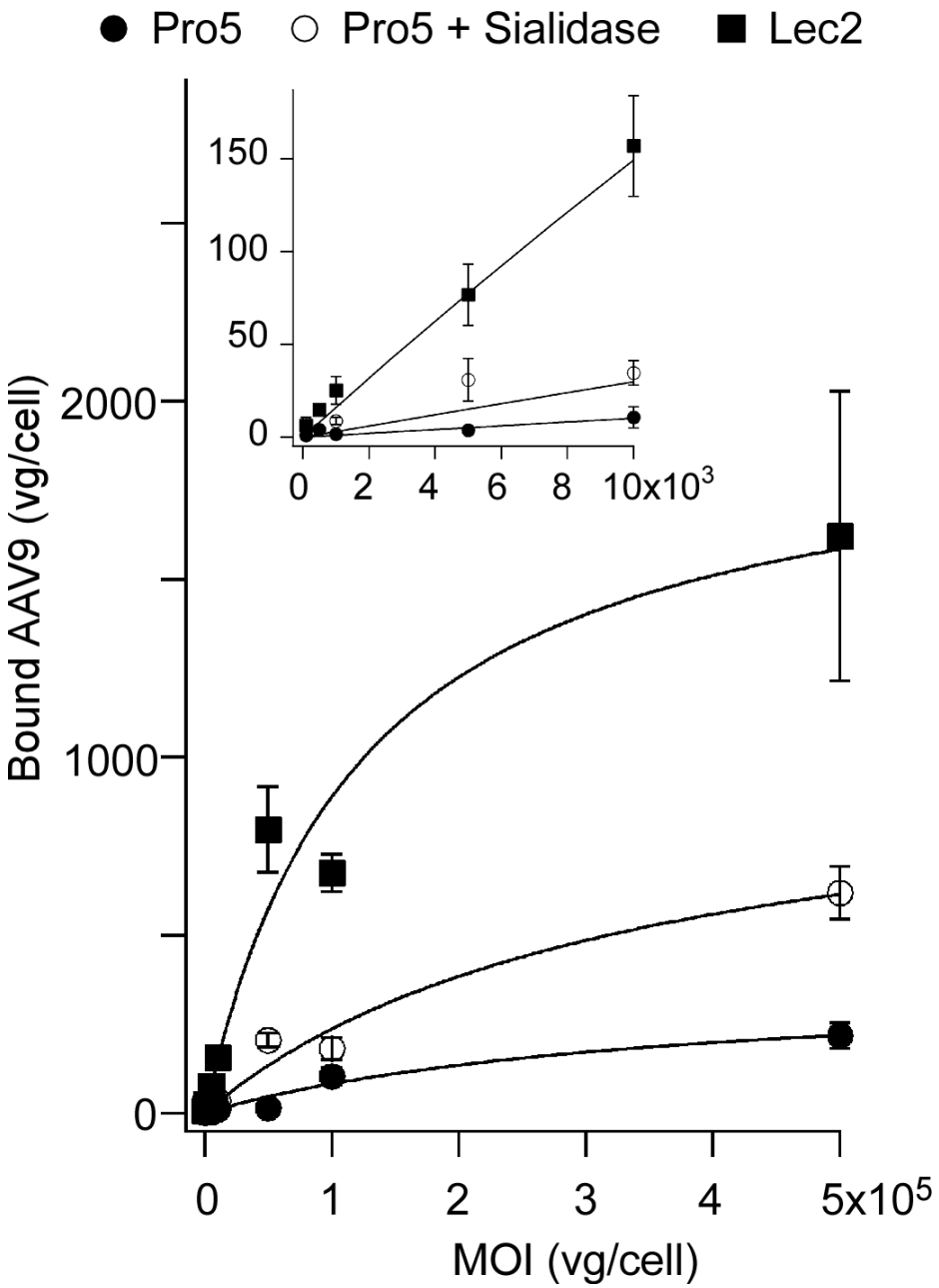

Figure 9. Effect of desialylation on cell surface binding of AAV9 particles. Sialic acid-deficient Lec2 cells (■); untreated wild type Pro5 cells (●) and sialidase-treated Pro5 cells (○) were prechilled and incubated with AAV9 particles at different MOI ranging from 100 to 500,000 (across 4.5 orders of magnitude) at 4˚C to allow binding, without cellular uptake. Quantitative analysis of dose-dependent AAV9 binding to cell surface asialo *N*-glycans was carried out by generating binding curves using a single-site binding model as described in Materials and Methods. Inset shows linear range of the binding curve from X-axis values ranging from 100 to 10,000 vg/cell. Calculated binding parameters are listed in Table 1. All experiments were carried out in triplicate. Error bars represent standard error mean.

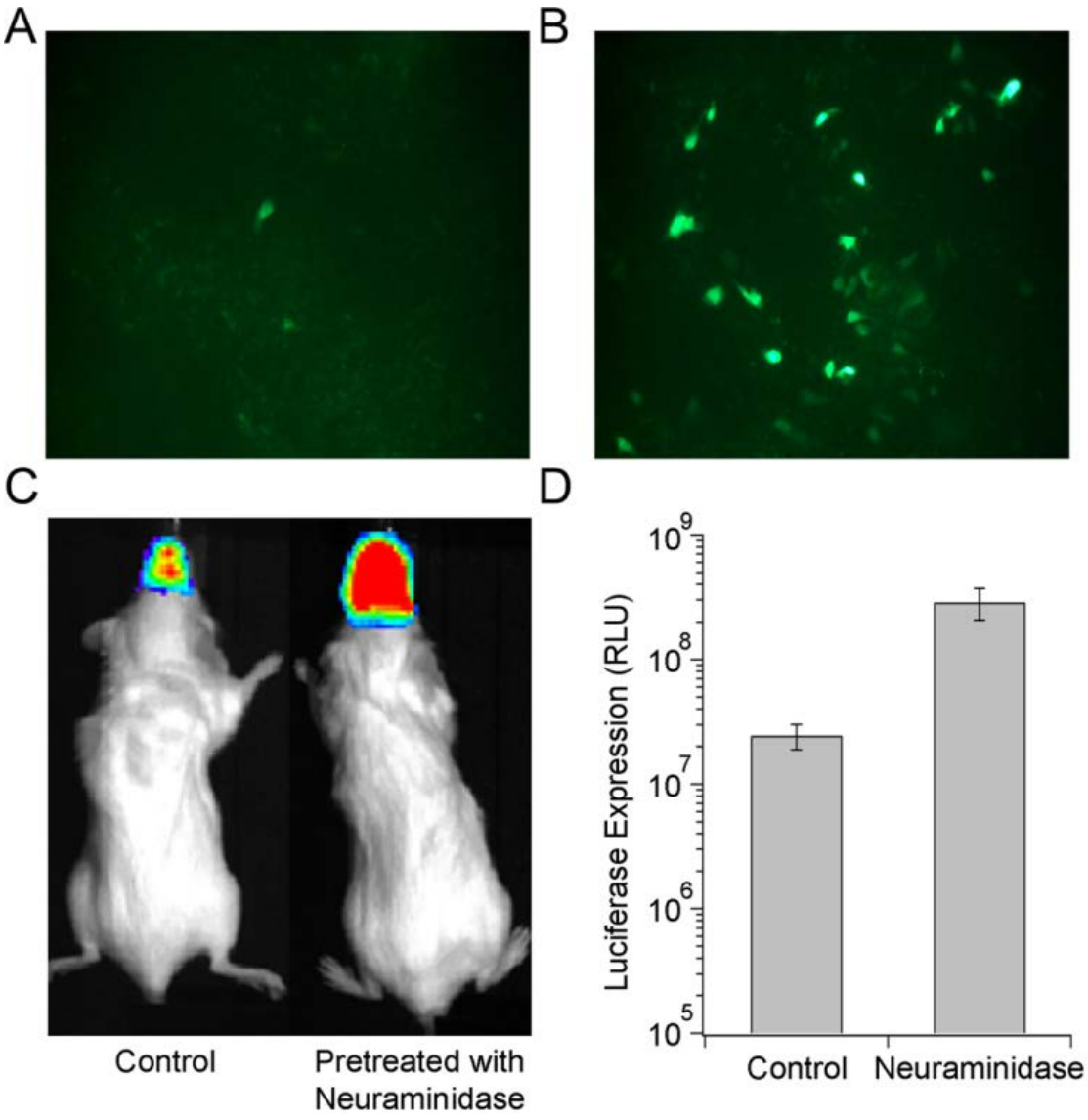

Figure 10. Enzymatic desialylation enhances AAV9 transduction efficiency in HAE cultures *in vitro* and murine airways *in vivo*. Well differentiated human airway epithelial (HAE) cells were untreated **(A)** or pretreated with recombinant sialidase A **(B)** followed by subsequent infection with self-complementary (sc) AAV9-CMV-GFP vectors (MOI = 10^5 vg/cell). Representative fluorescence micrographs of GFP transgene expression were obtained using an Olympus microscope equipped with a 20x objective and a Hamamatsu camera. **(C)** Representative live animal bioluminescent images of luciferase expression in Balb/c mice pretreated with intranasally administered PBS or neuraminidase (200 mU/50µl/nostril). Intranasal instillation of AAV9-CBA-Luciferase vectors ($5x10^{10}$ vg/50µl/nostril) was carried out 2 hrs post-sialidase treatment and bioluminescent images obtained at 4 weeks post-administration. **(D)** Bioluminescence intensity was quantified using Living Image ® software and expressed as relative light units (RLU). Error bars represent standard error mean.

CHAPTER 3

Glycan Binding Avidity Determines the Systemic Fate of AAV9[2]

3.1 Overview

Glycans are key determinants of host range and transmissibility in several pathogens. In the case of adeno-associated viruses (AAV), different carbohydrates serve as cellular receptors *in vitro*; however, their contributions in vivo are less clear. A particularly interesting example is adeno-associated virus serotype 9 (AAV9), which displays systemic tropism in mice, despite low endogenous levels of its primary receptor (galactose) in murine tissues. To understand this further, we studied the effect of modulating glycan binding avidity on the systemic fate of AAV9 in mice. Intravenous administration of recombinant sialidase increased tissue levels of terminally galactosylated glycans in several murine tissues. These conditions altered the systemic tropism of AAV9 into a hepatotropic phenotype, characterized by markedly increased sequestration within the liver sinusoidal endothelium and Kupffer cells. In contrast, an AAV9 mutant with decreased glycan binding avidity displayed a liver-detargeted phenotype. Altering glycan binding avidity also profoundly affected AAV9 persistence in blood circulation. Our results support the notion that high glycan receptor binding avidity appears to impart increased liver tropism, while decreased avidity favors systemic spread of AAV vectors. These findings may not only help predict

[2]This chapter includes the original publication: Shen, S., Bryant, K.D., Sun, J., Brown, S.M., Troupes, A., Pulicherla, N., Asokan, A. *Glycan Binding Avidity Determines the Systemic Fate of Adeno-associated Virus 9*, Journal of Virology, 2012 (PMID: 22787229).

species-specific differences in transduction profile for AAV9 on the basis of tissue glycosylation profile, but also provide a general approach to tailor AAV vectors for systemic or hepatic gene transfer by reengineering capsid-glycan interactions.

3.2 Introduction

The cell surface glycocalyx contains many glycans terminated with sialic acid moieties. Variability in the extent, type and linkage of sialylation between different hosts is exploited by several pathogens during infection [191,192]. Sialo/asialoglycans have also been implicated as cell surface receptors for different strains of adeno-associated virus (AAV) *in vitro* [193,194]. The latter are non-pathogenic, helper-dependent parvoviruses, currently being evaluated in phase I-III clinical trials as vectors for gene therapy [195]. Over 150 different strains have been isolated from human and primate tissues till date and these AAV isolates display striking differences in capsid structure, receptor usage and tissue tropism within different hosts [152,193,196]. A better understanding of the mechanisms underlying AAV tropism in different hosts could help guide the selection of preclinical animal models for evaluating AAV vectors and possibly predict outcomes in human gene therapy clinical trials.

Glycan receptors for different AAV serotypes include heparan sulfate for AAV2 [40], *N*-linked α2,3-/α2,6-sialic acid for AAV1/AAV6 [44], *N*-linked α2,3-sialic acid for AAV5 [45], *O*-linked α2,3-sialic acid for AAV4 [46] and β1,4-galactose for AAV9 [52,53]. In addition, various co-receptors such as integrins and growth factor receptors have been implicated in AAV infection *in vitro* [193,194]. Structural analysis of AAV capsids in conjunction with mutagenesis studies has provided a better understanding of the molecular determinants of AAV capsid-glycan interactions. For instance, crystallographic and/or cryoelectron microscopy-based

analysis have enabled mapping the role of critical amino acid residues involved in heparan sulfate recognition by AAV2, AAV-DJ, AAV3B and AAV6 capsids [83,92,93,165,197,198]. Recently the binding site of galactose on AAV9 capsids was mapped using a combination of mutagenesis and computational modeling tools [94]. The structural/molecular basis for recognition of sialo/asialoglycans by the corresponding AAV serotypes is currently unknown.

Despite the accruing knowledge of AAV-glycan interactions *in vitro*, the underlying molecular basis for differential tissue tropisms displayed by AAV serotypes in vivo is not well understood. For instance, AAV2, which utilizes heparan sulfate proteoglycans for cell surface attachment is a hepatotropic strain, while AAV1, which binds *N*-linked α2,3-/α2,6-sialic acid displays cardiac/muscle tropism [152]. In contrast, AAV9, which displays systemic, multi-organ tropism within different hosts [125,180,199] utilizes terminal galactose as its primary receptor [52,53]. The current study seeks to understand exactly how variations in host glycosylation patterns can affect AAV tissue tropism in mice, using AAV9-galactose interactions as a model. The study hinges on a two-pronged approach, the first of which exploits our preliminary observation that tissue sialo/asialoglycan levels can be readily modulated by enzymatic treatment in vivo. Briefly, intravenously administered sialidase afforded a dramatic increase in the density of asialoglycans (terminal galactose residues) across multiple tissues, which markedly altered the systemic behavior of AAV9 in mice. In the second part of this study, we carried out the systemic evaluation of a mutant AAV9 displaying low glycan binding avidity in normal mice. Our results demonstrate that altered tissue glycosylation patterns and/or glycan binding affinity can result in sharply contrasting liver uptake and blood circulation profiles of AAV in vivo.

3.3 Materials and Methods

Biochemical reagents. Recombinant sialidase (neuraminidase type III from *V. cholerae*) was obtained from Sigma-Aldrich (St. Louis, MO). Erythrina Cristagalli (ECL) and Maackia Amurensis I (MAL-I) lectins were purchased from Vector Labs (Burlingame, CA). Monoclonal antibodies (MAb) against endothelium-specific marker FcγRII/III/IV (rat 2.4G2 anti-CD16/CD32) [200] was obtained from BD Biosciences (San Jose, CA), and Kupffer cell-specific marker (rat anti-F4/80) [200] was obtained from Abcam (Cambridge, MA). Fluorophore-conjugated secondary antibodies (goat anti-rat or anti-mouse) were obtained from Invitrogen (Carlsbad, CA). The ADK9 mouse anti-capsid MAb recognizing intact AAV9 capsids was a kind gift from Dr. Jürgen Kleinschmidt (German Cancer Research Centre, DKFZ, Heidelberg, Germany).

AAV production. HEK293 cells were maintained at 37˚C in 5% CO2 in Dulbecco's modified Eagle's medium supplemented with 10% fetal bovine serum and penicillin-streptomycin-amphotericin B. The following plasmids were generated in the lab as described earlier [95] or obtained from the UNC vector core facility: (1) the pXR9 plasmid encoding rep2 and cap9; (2) the plasmid pXX6-80 containing adenoviral helper genes; and (3) the vector cassettes, pTR-CBA-Luc or pTR-CBA-tdTomato, encoding the Chicken beta-actin (CBA) promoter and luciferase or tdTomato reporter transgene, respectively. AAV9 vectors packaging TR-CBA-Luc or TR-CBA-tdTomato were produced by the triple transfection method [166] followed by cesium chloride density gradient ultracentrifugation and dialysis against 1x phosphate buffered saline (PBS) overnight with two changes. Viral titers were determined by quantitative PCR using primers (IDT Technologies, Ames, IA) specific for the CBA promoter (forward 5′-CGT CAA TGG GTG GAG TAT TT-3′; reverse 5′-GCG ATG

ACT AAT ACG TAG ATG-3′) or Luc transgene region (forward 5′-AAA AGC ACT CTG ATT GAC AAA TAC-3′; reverse 5′-CCT TCG CTT CAA AAA ATG GAA C-3′). Viral titers for different vector preps typically ranged from 2-5 x 10^{12} vector genome-containing particles per mL (vg/ml).

Cell surface binding assay. CHO Lec2 cells deficient in terminal sialic acid moieties were seeded at a density of 10^4 cells/well in 96-well plates and maintained in an incubator at 37 ˚C and 5% CO2 overnight. Prior to adding virus particles, cells were pre-chilled to 4°C for 30 min, followed by incubation with AAV9 or AAV9-W503R particles at MOI of $1x10^2$, $5x10^2$, $1x10^3$, $5x10^3$, $1x10^4$, $5x10^4$, $1x10^5$, $5x10^5$ vg/cell for 1.5 hrs at 4°C. Cells were then subjected to three washes with ice-cold 1x PBS to remove unbound virions. Cell surface-bound virions were collected along with cell lysates following three freeze-thaw cycles and vg copy numbers/cell determined using quantitative PCR as outlined earlier. Binding curves were generated using GraphPad Prism 5 software by applying the single site binding model ($Y=B_{max}*X/(K_d'+X)$), where Y represents the number of bound virions/cell determined by quantitative PCR; X represents MOI; B_{max} is the maximum binding capacity, and K_d', the observed disassociation constant. Binding potential is defined as B_{max}/K_d' and serves as a measure of glycan binding avidity.

Live animal studies. All in vivo experiments were carried out using 6–8-week old female BALB/c mice (Jackson Laboratories, Bar Harbor, ME) maintained and treated in accordance with National Institutes of Health guidelines and as approved by IACUC at UNC-Chapel Hill. Sialidase (neuraminidase type III from *V. cholerae*, 200 milliunits) or PBS was administrated through intravenous injection 3 hrs prior to AAV9-CBA-Luc ($5x10^{10}$ vg) through the same route. Luciferase transgene expression in live animals was monitored at

indicated time points for each experiment using an Xenogen IVIS Lumina® imaging system (Perkin Elmer/Caliper Life Sciences, Waltham, MA) following intraperitoneal injection of D-luciferin substrate (120 mg/kg; Nanolight, Pinetop, AZ). Bioluminescent image analysis was conducted using the Living Image software® (Perkin Elmer/Caliper Life Sciences, Waltham, MA) and luciferase expression reported in relative light units (photons/s per cm^2 per steradian).

Quantitation of luciferase expression. The same group of mice was sacrificed at indicated time points and the heart and liver harvested. Approximately 50 mg of each tissue was homogenized in 150 µl of passive lysis buffer (Promega, Madison, WI) using a Tissue Lyser II system (Qiagen, Valencia, CA). Tissue lysates were centrifuged at 8,000 rpm for 2 min to pellet debris and 50 µl of the supernatant transferred to 96-well plates for luminometric analysis using a Victor2® luminometer (Perkin Elmer, Waltham, MA). Total protein concentration in tissue lysates was determined using the Bradford assay (BioRad, Hercules, CA).

Quantitation of viral genome biodistribution. Approximately 100 µl of supernatant from tissue lysates obtained as mentioned above was processed using a DNeasy kit (Qiagen, Valencia, CA) to extract host and vector genomic DNA. Vector genome (Luc) and mouse lamin gene (housekeeping gene serving as internal standard) copy numbers were determined from 100 ng of total extracted DNA using quantitative PCR as described earlier.

Immunohistology. For lectin staining studies, sialidase (neuraminidase type III from *V. cholerae*, 400 milliunits) in 150 µl of PBS or PBS alone was first administrated via tail vein injection. At 3 hrs post-administration, mice were anesthetized with intraperitoneal injection of avertin (0.2 ml/10 g of a 1.25% solution) and perfused transcardially with 20 ml

of PBS, then 20 ml of freshly prepared 4% paraformaldehyde (Sigma-Aldrich, St. Louis, MO) in PBS. Cardiac, liver, brain, and musculartissues were fixed at 4°C overnight, and then sectioned into 40 μm slices using a Leica vibrating blade microtome (Leica Microsystems, Buffalo Grove, IL). The sections were blocked in 10% goat serum in 1xPBS containing 0.1% Triton X-100 (PBS-T) for 1 hr at room temperature prior to incubation with primary antibodies (10 μg/ml CD16/CD32 or F4/80 MAb and 1:10 dilution of ADK9 in hybridoma media) for 24 hrs at 4°C. Antibody binding was localized using fluorophore-conjugated secondary antibodies diluted 1:1000 in 3% goat serum in PBS-T for 1 hr at 4°C. Specifically, Alexa Fluor 647 goat anti-rat IgG (H+L) (Invitrogen A21247, 2 mg/ml, 1:500 dilution) was used to recognize rat anti-mouse-CD16/CD32 and rat anti-mouse-F4/80, and Alexa Fluor 594 goat anti-mouse IgG (H+L) (Invitrogen A11005, 2 mg/ml, 1:500 dilution) was utilized to detect ADK9 MAb, which binds to intact AAV9 capsids. When indicated, the sections were further stained in 30 μg/ml of FITC-labeled ECL or MAL-I lectins in 3% blocking solution. After three washes with PBS to get rid of unbound antibodies and lectins, the sections were mounted under cover slips in Prolong® gold anti-fade reagent with DAPI (Invitrogen P36935). Tissues harvested from untreated mice were utilized as controls. Four color fluorescence images were acquired using a Zeiss 710 confocal laser scanning microscope equipped with a spectral detection system for finer separation of fluorochromes. Image processing was carried out using LSM viewer and Image J® softwares.

To compare the transduction profile of AAV9 and AAV9-W503R in live animals, PBS (mock treatment), AAV9-CBA-tdTomato ($5x10^{10}$ vg), and AAV9-CBA-tdTomato ($5x10^{10}$ vg) were injected to BALB/c mice through tail vein. 2 weeks post infection, animals were anesthetized using intaperitoneal avertin (0.2 ml/10 g of a 1.25% working solution),

followed by transcardial perfusion with 20 ml of PBS and 20 ml of 4% paraformaldehyde in PBS. Tissues were fixed in 4% paraformaldehyde in PBS for overnight at 4˚C, and sectioned into 40 μm slices using a Leica vibrating blade microtome. Tissue sections were mounted under cover slips using Prolong® gold anti-fade reagent with DAPI. Red fluorescent images were taken using an Olympus epifluorescence microscope equipped with a 10x objective and a Hamamatsu digital camera.

Pharmacokinetic analysis. Sialidase (neuraminidase type III from *V. cholera*, 200 milliunits) or 100 μl of PBS was injected through intravenous injections via the mouse tail vein. AAV9-CBA-Luc ($5x10^{10}$ vg) was injected through the same route 2 hrs later. For pharmacokinetic analysis, 10 μl of blood was drawn from the tail vein through nicking at the different time intervals into heparinized glass capillary tubes (Fisherbrand Hematocrit®). Viral DNA was extracted from the blood samples using a DNeasy kit (Qiagen, Valencia, CA) and further quantified by qPCR as described earlier. Blood clearance of viral genomes was plotted as a function of viral genome copy number vs. time and the data fit into a bi-exponential model using least-square regression analysis in GraphPad Prism® software. Distribution and elimination half-life as well as different rate constants based on a two compartment pharmacokinetic model were calculated using equations outlined in **Table 4** as described earlier [201]. Pharmacokinetics parameters, standard errors, and goodness of fit from curve-fitting are listed in **Table 5 and Table 6**.

Statistics. All data are expressed as mean ± s.e.m. unless indicated otherwise. The two-tailed unpaired t test was utilized for all statistical analysis and P values < 0.05 were considered significant unless indicated otherwise. Statistical evaluation of pharmacokinetic parameters and goodness of fit are provided in **Table 5**.

3.4 Results

Intravenous sialidase alters tissue surface glycosylation patterns in mice. Using lectin staining, we analyzed the endogenous constitution of sialylated glycans in major mouse organs as well as the effect of administering intravenous sialidase on tissue surface glycosylation patterns. First, BALB/c mice were intravenously injected with PBS or recombinant sialidase from *V. cholerae* 3 hrs prior to tissue harvesting/fixation. Sections of major organs were then stained with FITC-labeled lectins, specifically, ECL, which recognizes terminal galactose moieties on Gal(β1,4)*N*-GlcNAc or MALI, which binds both α2,3-sialylated and sulfated forms of Gal(β1,4)*N*-GlcNAc. As shown in **Figure 11A**, FITC-ECL staining is barely discernible, while substantial FITC-MALI staining is observed in murine heart, liver, brain and muscle. These observations suggest that the endogenous expression of asialoglycans containing terminal galactose is low when compared to that of α2,3-sialylated glycans in normal mouse tissues. In contrast, intravenous injection of recombinant sialidase markedly enhanced FITC-ECL staining intensity in heart, liver and muscle tissue, while MALI staining showed a modest decrease (**Figure 11B**). Interestingly, murine brain displayed ECL staining potentially localized to blood vessels alone, while MALI staining remained more or less unaffected upon sialidase treatment (**Figure 11B**). These results correlate well with earlier *in vitro* studies [52] and suggest that intravenous sialidase can effectively desialylate glycoproteins and glycolipids revealing the underlying β1,4-galactose moiety in major organs *in vivo*. We then exploited these observations to further evaluate the impact of increasing cognate glycan receptor density on AAV9 tropism *in vivo*.

High glycan binding avidity redirects systemic AAV9 to the liver. We analyzed the effect of enzymatic desialylation on the transduction profile of AAV9 vectors packaging a chicken β-actin (CBA) promoter-driven firefly luciferase transgene *in vivo*. Briefly, AAV9 vectors were injected into the tail vein at 3 hrs following intravenous administration of recombinant sialidase. To track the kinetics of luciferase transgene expression, live animal bioluminescence images were taken at different time intervals post-injection (**Figure 12A**). As seen in **Figure 12A**, AAV9 displays a systemic transduction profile in the PBS treated mouse group (left panel), consistent with previous reports [95,125,202]. In the sialidase treated group (**Figure 12A**, right panel), a predominantly liver-centric bioluminescent signal is observed, accompanied by decreased signal in other systemic regions. Luciferase transgene expression and viral genome copy number in different tissue lysates were quantified at 2 weeks and 3 days post-injection, respectively. As shown in **Figure 12B**, the transduction efficiency of AAV9 in the liver is ~5-fold higher upon sialidase treatment. Further, a substantial decrease in transduction efficiency within the heart (~10-fold) and skeletal muscle (~14-fold) is observed. Transgene expression levels in the murine brain are relatively low in comparison with other organs and show no significant change upon sialidase treatment. Consistent with these observations, the biodistribution of viral genomes in the liver is increased by ~7-fold in the liver (**Figure 12C**). Correspondingly, viral genome copy numbers are decreased in the brain, heart and skeletal muscle are decreased (ranging from 2 to 5-fold) upon sialidase treatment. These results suggest that surface glycan density in host tissues strongly influences AAV tissue tropism. In particular, increased availability of the galactose receptor in host tissue appears to redirect AAV9 capsids from major systemic organs towards the liver.

High glycan binding avidity results in AAV9 sequestration within liver endothelial and Kupffer cells. To understand the mechanism(s) underlying increased liver accumulation of AAV9, we analyzed liver sequestration of viral capsids through immunohistology. Briefly, mice were injected with AAV9 vectors 3 hrs post-sialidase or PBS treatment. At 15 minutes post-vector administration, liver and heart were harvested and processed as described in materials and methods. Following PBS pre-treatment, we observed sporadic immunostaining for AAV9 capsids partially overlapping with endothelial cells defined by FcγRII/III/IV staining (anti-CD16/CD32) or other cell types within the liver (**Figure 13A**). In contrast, sialidase pre-treatment resulted in increased liver uptake of AAV9 capsids consistent with increased galactose density (**Figure 13A**). Further, increased immunofluorescence associated with AAV9 capsids almost completely co-localized with liver sinusoidal endothelial cells upon sialidase treatment (**Figure 13A**). We also observed substantially increased co-localization of AAV9 capsids with Kupffer cells (anti-F4/80) in addition to sinusoidal endothelial cells within the liver (**Figure 13B**). These results confirm the earlier observation that high glycan density in host tissues redirects AAV9 from systemic organs to the liver. The underlying mechanism can be attributed to markedly increased sequestration of AAV9 capsids within the liver sinusoidal endothelium and Kupffer cells.

Decreased glycan binding avidity detargets AAV9 from the liver. We previously engineered a class of AAV9 mutants, which have selectively lost their ability to transduce the liver [95]. As shown in the previous study, one such mutant, AAV9-W503R displays markedly lower transduction efficiency (> 1 log unit) in liver when compared to wild type AAV9. In the current study, we evaluated the ability of this mutant to bind surface-exposed galactose residues on Chinese hamster ovary (CHO) Lec2 cells, which are permissive to AAV9

infection *in vitro* [52,53]. As seen in **Figure 14A**, the AAV9-W503R mutant is deficient in cell surface binding, with a binding potential (B_{max}/K_d') that is ~4.5-fold lower than that observed for wild type AAV9. Correspondingly, this mutant displays transduction efficiency that is several orders of magnitude lower when compared to wild type AAV9 in CHO Lec2 cells at different multiplicities of infection (MOI) (**Figure 14B**). We then investigated the impact of such decreased glycan binding avidity on liver sequestration of AAV9 and the W503R mutant following intravenous administration. As seen in **Figure 14C**, immunohistological analysis reveals markedly lower immunostaining for AAV9-W503R in comparison with AAV9 capsids. Moreover, while AAV9 capsids co-localized with FcγRII/III/IV (anti-CD16/CD32 staining), which is specific for endothelial cells as well as liver parenchyma, the AAV9-W503R mutant weakly co-localized with endothelial cells alone (**Figure 14C**). In addition, consistent with earlier studies in our lab [95], we observed a lack of tdTomato reporter gene expression in hepatocytes for AAV9-W503R in contrast to wild type AAV9 (**Figure 14D**). Taken together, these results suggest that decreased glycan binding is the underlying basis for the liver-detargeted phenotype displayed by the AAV9-W503R mutant.

Vascular endothelium constitutes a limiting barrier for cardiac uptake of AAV9. We also evaluated the impact of altering glycan binding avidity on cardiac tropism of AAV9 in mice. As seen in Figure 5, immunostaining for AAV9 capsids (MAb ADK9, red) as well as the endothelial marker FcγRII/III/IV (anti-CD16/CD32 staining) is markedly lower in comparison to that observed in liver sections (**Figure 13**). Moreover, unlike liver, where viral particles are dispersed amongst parenchymal and non-parenchymal cell types, AAV9 capsids predominantly co-localize with the vascular endothelium in the murine heart. This

observation supports the notion that transvascular transport is a likely prerequisite for viral entry into cardiomyocytes as proposed earlier [203]. Interestingly, despite the dramatic increase in galactose receptor density as demonstrated by FITC-ECL staining throughout cardiac tissue following sialidase pre-treatment, no changes in the pattern or extent of cardiac uptake of AAV9 capsids were observed when compared to PBS-treated control. Instead, AAV9 capsids co-localized with cardiac endothelium in a manner similar to PBS-treated control animals. In addition, immunostaining for the low binding affinity mutant (AAV9-W503R) was significantly lower at this early time interval potentially due to decreased endothelial surface interactions.

Altered glycan binding avidity affects the blood circulation profile of AAV9. Several reports have demonstrated that AAV9 displays prolonged blood circulation in comparison with other AAV strains in mice [125,203]. We hypothesized that the latter phenomenon could arise from low endogenous levels of terminal galactose residues in murine tissue. To test this hypothesis, we first treated mice with intravenous sialidase prior to AAV9 administration in order to increase galactose receptor density in murine tissues as described earlier. Viral genome copies in blood were then collected at different time intervals, quantitated by Q-PCR analysis and the data fit to a bi-exponential, two compartment model (**Table 4**) to obtain different pharmacokinetic parameters (**Table 5 and Table 6**). As seen in **Figure 16A**, high glycan binding avidity results in rapid clearance of AAV9 from the blood (solid circles), while the low binding avidity mutant AAV9-W503R appears to persist longer in blood circulation in comparison to AAV9 circulating under normal conditions. Specifically, increasing tissue glycan density by sialidase treatment results in ~4 times faster elimination half-life ($t_{1/2,e}$). In contrast, the AAV9-W503R mutant, which is deficient in galactose

binding, is eliminated twice as slowly (~10 hrs) from the blood circulation when compared to AAV9 alone (~4.7 hrs). Further, while the half-life of viral dissemination from blood to tissue ($t_{1/2,dis}$) is reduced by ~2-fold upon sialidase treatment, this parameter is extended from ~0.4 hrs to ~1 hr (2.5-fold increase) for the W503R mutant. Consistent with these observations, we note that the blood-to-organ distribution rate constant (k_{12}) and organ-to-blood distribution rate constant (k_{21}) are comparable in mice injected with AAV9 alone. In contrast, k_{12} for AAV9 is ~7-fold higher than k_{21} in mice pretreated with sialidase supporting rapid tissue dissemination of injected virus particles. Further, for the AAV9-W503R mutant, both the aforementioned rate parameters are lower in comparison with AAV9, consistent with prolonged blood residence time. Taken together, these results confirm that the capsid-glycan interactions profoundly affect the pharmacokinetics of AAV9.

3.5 Discussion

The liver is a prominent systemic organ featuring hepatocytes as well as a large proportion of cells belonging to the reticuloendothelial system. Due to the large volume of circulating blood that filters through the liver, the sinusoidal endothelial and Kupffer cells in this organ play a major scavenging role in the uptake of various macromolecules and pathogens [204,205]. A particularly well-studied example in this regard is the adenoviral sequestration within the liver [200,206-208]. Most, if not all naturally occurring isolates of AAV studied to date, display a moderate-to-high transduction efficiency within the liver [125]. However, whether liver sequestration mechanisms similar to those observed for adenoviruses prevail in the case of AAV is not known.

In the current study, we investigated the potential role of glycan receptor binding avidity on AAV liver sequestration and systemic spread using AAV9-galactose interactions as a model. Notably, the transduction profile and *in vivo* biodistribution of AAV9 were altered from systemic to hepatotropic under high glycan binding avidity conditions. These observations are corroborated by immunostaining studies, which revealed that increasing glycan receptor binding avidity potentiates AAV9 sequestration within the liver sinusoidal endothelium and Kupffer cells. In corollary, the AAV9-W503R mutant displays low galactose binding affinity and displayed decreased liver uptake in mice. Further, our results suggest that high affinity receptor sites for AAV9 in mice are constitutively low. This observation is not surprising, since it is well known that glycans containing β1,4-linked galactose are extended by sialylation or sulfation, although formation of poly-*N*-acetyllactosamine (Galβ1,4GlcNAcβ1-) units has also been demonstrated earlier [209]. The resulting modest glycan binding avidity could explain the underlying basis for widespread, systemic tropism and prolonged blood circulation displayed by AAV9 in comparison to other AAV serotypes *in vivo*. Strikingly, high glycan binding avidity appears to adversely affect AAV9 uptake and transduction in peripheral tissues such as cardiac and skeletal muscle. This effect can likely be attributed to the rapid blood clearance of AAV9 under such conditions and highlights the limiting role of vascular endothelium in viral dissemination. Our conclusions are further supported by earlier studies by Nakai and others [203], which demonstrate that trans-endothelial transport of AAV9 is a capacity-limited process that constitutes a significant barrier towards viral uptake by cardiomyocytes, but not hepatocytes.

Why is the AAV9-W503R mutant selectively detargeted from the liver, while transduction in peripheral tissues such as the heart and skeletal muscle remain unaffected [95]?

In the current study, we determined that this mutant (as well as other liver-detargeted AAV9 mutants, Shen and Asokan, unpublished) displays low galactose binding avidity *in vitro* and reduced endothelial binding *in vivo*. These results are consistent with W503R being recently identified as one of the key residues involved in galactose recognition by AAV9 capsids [94]. A potential explanation for this phenomenon is that low glycan binding avidity results in the doubling of distribution half-life ($t_{1/2,dis}$ ~1 hr) as well as elimination half-life ($t_{1/2,e}$ ~9 hrs) of the mutant when compared to wild type AAV9 (**Table 3**). Thus, it is likely that the AAV9-W503R mutant significantly benefits from repeated circulation through cardiac tissue, where traversing the vascular endothelium appears to be a capacity-limited process. Such transendothelial transport of viral particles might be redundant for transduction in liver, which contains open fenestrae that allow free exchange of material between blood and the space of Disse [204]. These results are further corroborated by earlier studies with AAV serotype 2, wherein, abrogation of capsid binding to heparan sulfate yields liver-detargeted AAV2 mutants [89]. One such mutant, AAV2i8, which persists longer in blood than other AAV strains, displays robust cardiac and skeletal muscle tropism with ability to traverse the vascular endothelium [143]. Based on a collective interpretation of results, we propose mathematical models that will help predict the impact of capsid-glycan interactions on AAV blood clearance as well as transport between the central (blood) compartment and major peripheral (liver) compartment (**Figure 16B and Figure 16C**). High glycan receptor binding avidity will likely increase the rate of viral dissemination from blood to liver (k_{12}) as well as rapid viral elimination from blood (k_e) by excretory mechanisms (**Figure 16B**). In contrast, when glycan binding avidity is low, both dissemination and elimination rate constants are decreased resulting in significantly longer blood circulation times (**Figure 16C**). Although

virus-glycan interactions alone are unlikely to provide a complete understanding of the biology of AAV, it is tempting to speculate that the latter models could potentially help predict liver sequestration of AAV in different species on the basis of their respective glycosylation patterns.

In summary, our studies offer new insight into the role of glycans in determining AAV tropism *in vivo*. In addition to resolving the mechanistic basis of liver-detargeted AAV vectors, we provide a potential general strategy for engineering any AAV strain by modulating capsid-glycan interactions. Further, the results described herein underscore the importance of understanding species-specific variation in glycosylation patterns. For instance, striking differences in sialylation patterns between different tissues in mice, rats, chimpanzees and humans have been reported [210]. Such information could guide rational selection of preclinical animal models for evaluating AAV vectors and might help better predict clinical gene therapy outcomes.

Table 4. Pharmacokinetic Equations for a Two Compartment Model [a]

Bi-exponential equation describing AAV9 pharmacokinetics:	$C_p = A * e^{-\alpha x} + B * e^{-\beta x}$
Tissue-to-blood compartment distribution rate constant:	$k_{21} = \frac{A*\beta+B*\alpha}{A+B}$
Blood elimination rate constant:	$k_e = \frac{\alpha*\beta}{k_{21}}$
Blood-to-tissue compartment distribution rate constant:	$k_{12} = \alpha + \beta - k_{21} - k_e$
Distribution half-life:	$t_{\frac{1}{2},distribution} = \frac{ln2}{k_{12}}$
Elimination half-life:	$t_{\frac{1}{2},elimination} = \frac{ln2}{k_e}$

a C_p is the viral genome copy number in plasma, x is the time interval

Table 5. Pharmacokinetic Parameters and Goodness of Fit from Curve-fitting Blood Circulation Data to a Bi-exponential Two-compartment Model

Treatment	PBS	Sialidase	PBS
Virus	**AAV9**	**AAV9**	**AAV9-W503R**
Y_0	$6.86 \times 10^{10} \pm 9.81 \times 10^{9}$	$9.89 \times 10^{10} \pm 1.65 \times 10^{10}$	$8.62 \times 10^{10} \pm 2.46 \times 10^{10}$
A	$3.75 \times 10^{12} \pm 5.55 \times 10^{10}$	$8.79 \times 10^{12} \pm 2.49 \times 10^{10}$	$3.94 \times 10^{12} \pm 8.83 \times 10^{11}$
B	$3.11 \times 10^{12} \pm 5.55 \times 10^{10}$	$1.10 \times 10^{12} \pm 2.49 \times 10^{10}$	$4.68 \times 10^{12} \pm 8.83 \times 10^{11}$
α	3.509 ± 1.12	4.602 ± 0.7836	1.752 ± 3.566
β	$6.826 \times 10^{-2} \pm 4.874 \times 10^{-3}$	$7.055 \times 10^{-2} \pm 7.696 \times 10^{-3}$	$3.884 \times 10^{-2} \pm 6.423 \times 10^{-2}$
R^2	0.9995	0.9999	0.7250

Table 6. Parameters Describing the Pharmacokinetics of AAV9 Under Different Glycan Binding Avidity Conditions *In Vivo*

Treatment	AAV9	AAV9+Sialidase	AAV9-W503R
Glycan binding avidity	++	+++	+
k_{21} (h^{-1})	1.63	0.57	0.97
k_e (h^{-1})	0.15	0.57	0.07
k_{12} (h^{-1})	1.80	3.53	0.75
$t_{1/2,dis}$ (h)	0.38	0.20	0.92
$t_{1/2,e}$ (h)	4.72	1.22	9.87

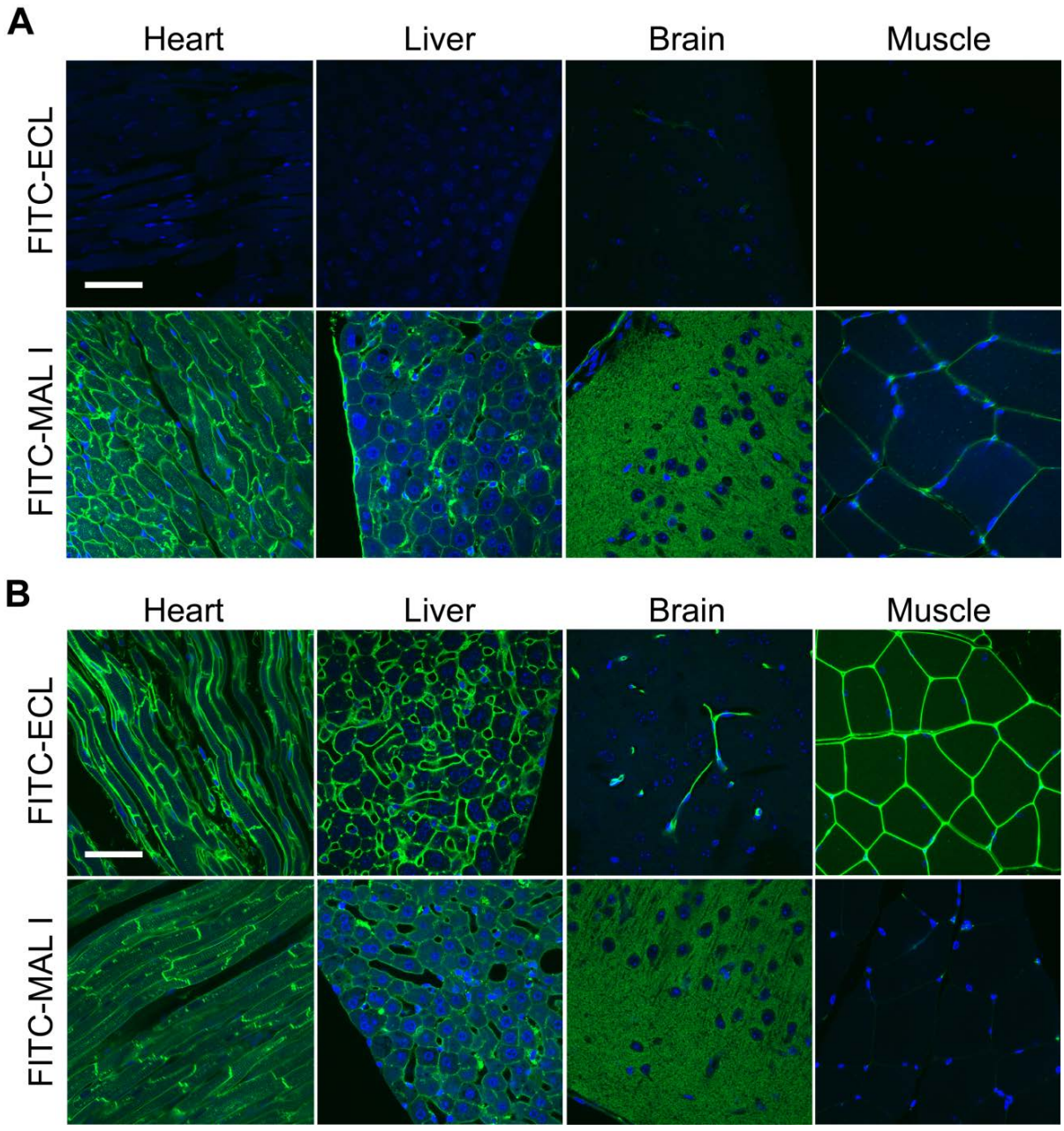

Figure 11. Intravenous sialidase alters endogenous tissue glycosylation patterns *in vivo*. **(A)** Fluorescent lectin staining of heart, liver, brain and skeletal muscle tissue harvested from Balb/C mice. FITC-ECL was used to detect tissue glycans containing terminal β1,4-Galactose and FITC-MALI detects α2,3-sialylated glycans. Confocal micrographs were obtained using a Zeiss 710 Confocal Laser Scanning Microscope with a 40x objective at zoom 0.6. Scale bar = 50 μm. **(B)** Fluorescent lectin staining of heart, liver, brain and skeletal muscle harvested from Balb/C mice administered intravenous sialidase to enzymatically remove α2,3-sialic acid residues *in vivo*.

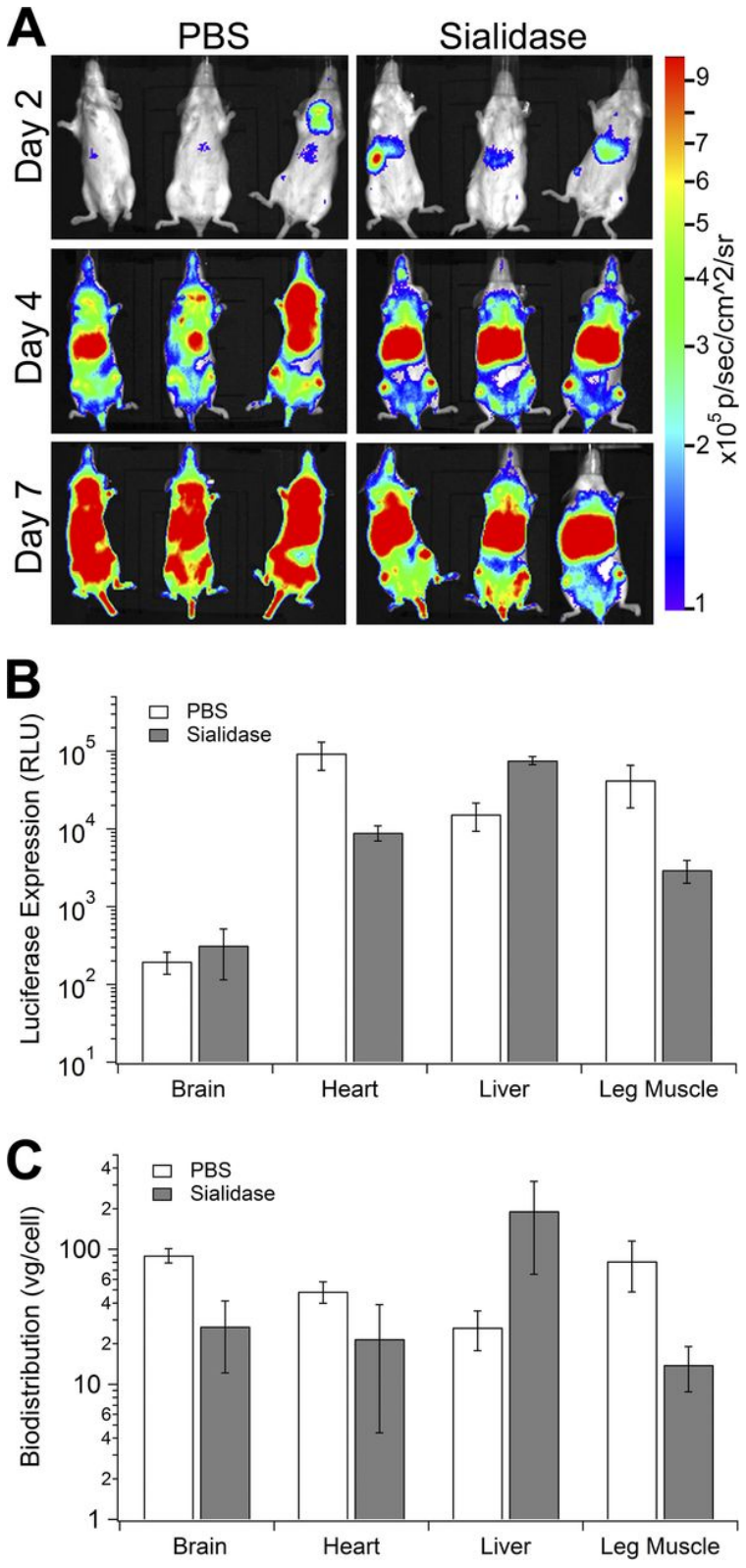

Figure 12. High glycan binding avidity redirects systemic AAV9 to the liver. (A) Live animal bioluminescence images of AAV9-mediated luciferase transgene expression in

Balb/C mice injected prior with PBS (control, upper panel) or recombinant sialidase (lower panel) at 2 weeks post-administration (n=3). Images were obtained using a Xenogen IVIS Lumina system. Rainbow scale represents relative light units (RLU) expressed in photons per sec per cm^2 per steradian. Kinetics of luciferase transgene expression on days 2, 4 and 7 was also recorded through live animal bioluminescence imaging (See Supplementary Figure S1). **(B)** Luciferase transgene expression in brain, heart, liver, skeletal muscle tissue lysates harvested from mice injected with PBS (white bars) or sialidase (grey bars) prior to AAV9 at 2 weeks post-administration. Luciferase transgene expression (RLU) is normalized to total cellular protein in tissue lysate (n=3). Data are represented as mean ± s.e.m.. **(C)** AAV9 vector genome biodistribution in brain, heart, liver and skeletal muscle tissues harvested from Balb/c mice pre-treated with PBS (white bars) or sialidase (grey bars) at 3 days post-administration (n=3). Viral genome copy numbers in different tissues/animals were obtained using quantitative PCR and normalized to total copy number of the constitutive mLamin gene in host genomic DNA. Primer sequences utilized for QPCR are listed under experimental procedures. Data are represented as mean ± s.e.m..

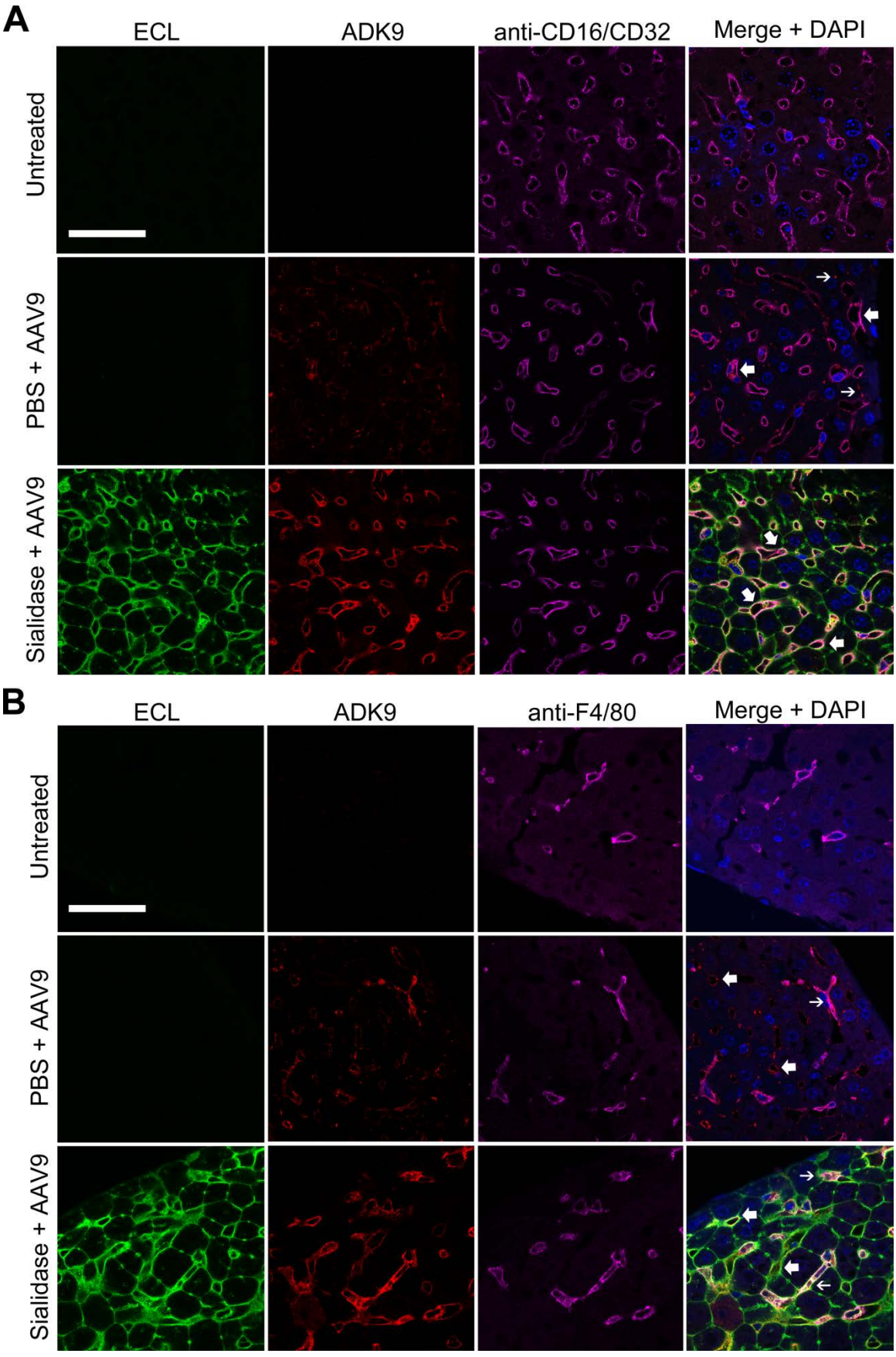
A
ECL
ADK9
anti-CD16/CD32
Merge + DAPI
Untreated
PBS + AAV9
Sialidase + AAV9
B
ECL
ADK9
anti-F4/80
Merge + DAPI
Untreated
PBS + AAV9
Sialidase + AAV9

Figure 13. High glycan avidity potentiates AAV9 sequestration by liver sinusoidal endothelium and Kupffer cells. (A) Liver sections were obtained 15min post-injection of AAV9 in Balb/C mice pre-treated with intravenous PBS (middle row) or sialidase (bottom row). Tissue sections were stained with fluorescent lectin to detect β1,4-galactose, FITC-ECL (green); anti-AAV9 capsid antibody, ADK9 (red) and anti-CD16/CD32 endothelial cell marker antibody (magenta). Color images were merged along with nuclear DAPI staining (blue) to generate the overlay panel. Untreated mouse liver is shown as control (top row). Large white arrows indicate liver sinusoidal endothelial cells and small white arrows indicate non-endothelial cells. Fluorescent micrographs were obtained using a Zeiss 710 Confocal Laser Scanning Microscope with a 63x objective. Scale bar = 50 μm. **(B)** Liver sections obtained from Balb/C mice pre-treated with intravenous PBS or sialidase and immunostained with FITC-ECL (green), ADK9 (red), anti-F4/80 Kupffer cell marker antibody (magenta) and nuclear DAPI staining (blue). Large white arrows indicate liver sinusoidal endothelial cells and small white arrows indicate Kupffer cells.

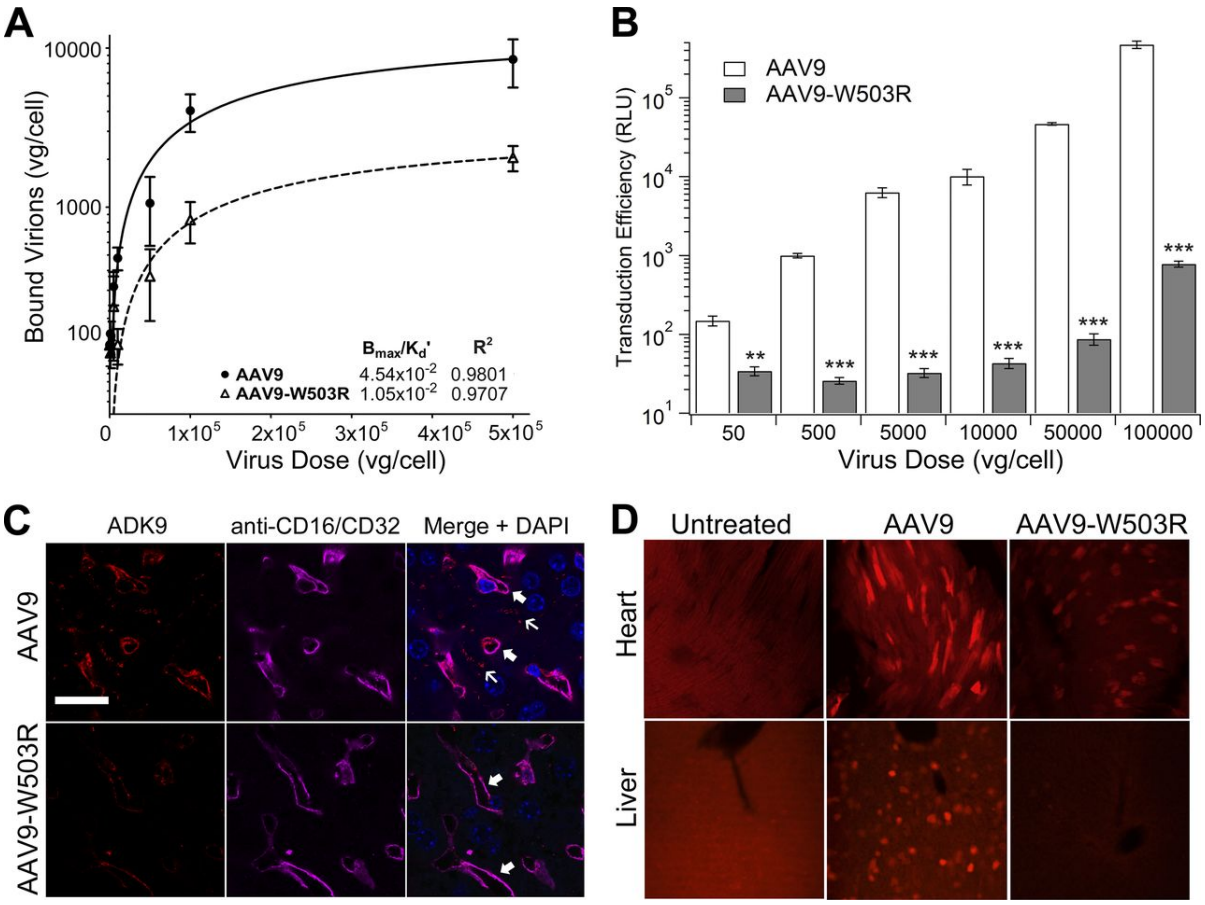

Figure 14. Low glycan binding avidity detargets AAV9-W503R from the liver. (A) The AAV9-W503R mutant displays decreased binding affinity to terminal galactose residues on the surface of CHO Lec2 cells lacking sialic acid (n=4). Curve fitting based on a single-binding site model was utilized to obtain binding potential values (B_{max}/K_d') for the mutant and parental AAV9. Data are represented as mean ± s.e.m.. **(B)** Transduction efficiency of AAV9 and AAV9-W503R on CHO Pro5 (constitutive glycan composition) and Lec2 cells (sialic acid-deficient) at different multiplicities of infection (n=4). Luciferase transgene expression in cell lysates was analyzed 24 hrs post-infection **p < 0.01; ***p < 0.005. Data are represented as mean ± s.e.m.. **(C)** AAV9 and AAV9-W503R accumulation in the liver of Balb/C mice at 15 min post-intravenous administration. Immunostaining was carried out using anti-AAV9 capsid antibody, ADK9 (red); anti-CD16/CD32 endothelial cell marker (magenta) and nuclear DAPI staining (blue). Fluorescent images were collected as described earlier. Large white arrows indicate liver sinusoidal endothelial cells and small white arrows indicate non-endothelial cells. Scale bar = 20 μm. **(D)** Transgene expression (tdTomato reporter) mediated by AAV9 and AAV9-W503R in liver and heart in BALB/c mice (n = 3). Tissues were harvested and imaged 2 weeks post-infection with a 10× objective.

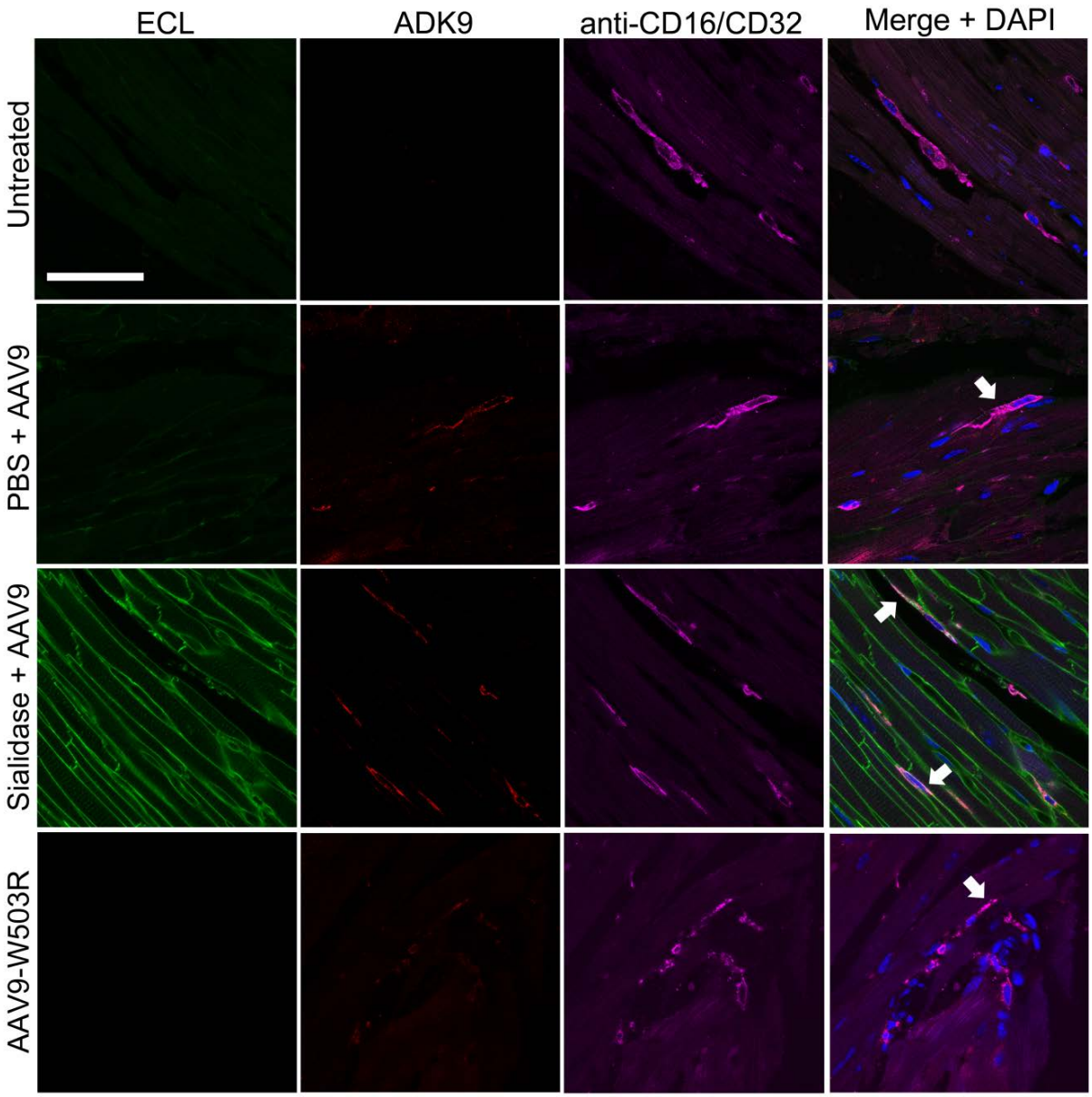

Figure 15. Vascular endothelial cells limit cardiac uptake of AAV9 capsids. Confocal micrographs of cardiac tissue sections harvested from Balb/C pre-treated with PBS (second row) or sialidase (third row) obtained at 15min post-intravenous AAV9 injection. Immunostaining was carried out using FITC-ECL (green), ADK9 (red), anti-CD16/CD32 (magenta) and nuclear DAPI staining (blue). Co-localization of mutant AAV9-W503R with vascular endothelium in the heart (bottom row) and cardiac tissue sections from untreated Balb/C mice (top row) are also shown. Large white arrows indicate cardiac endothelium. Scale bar represents 50 μm.

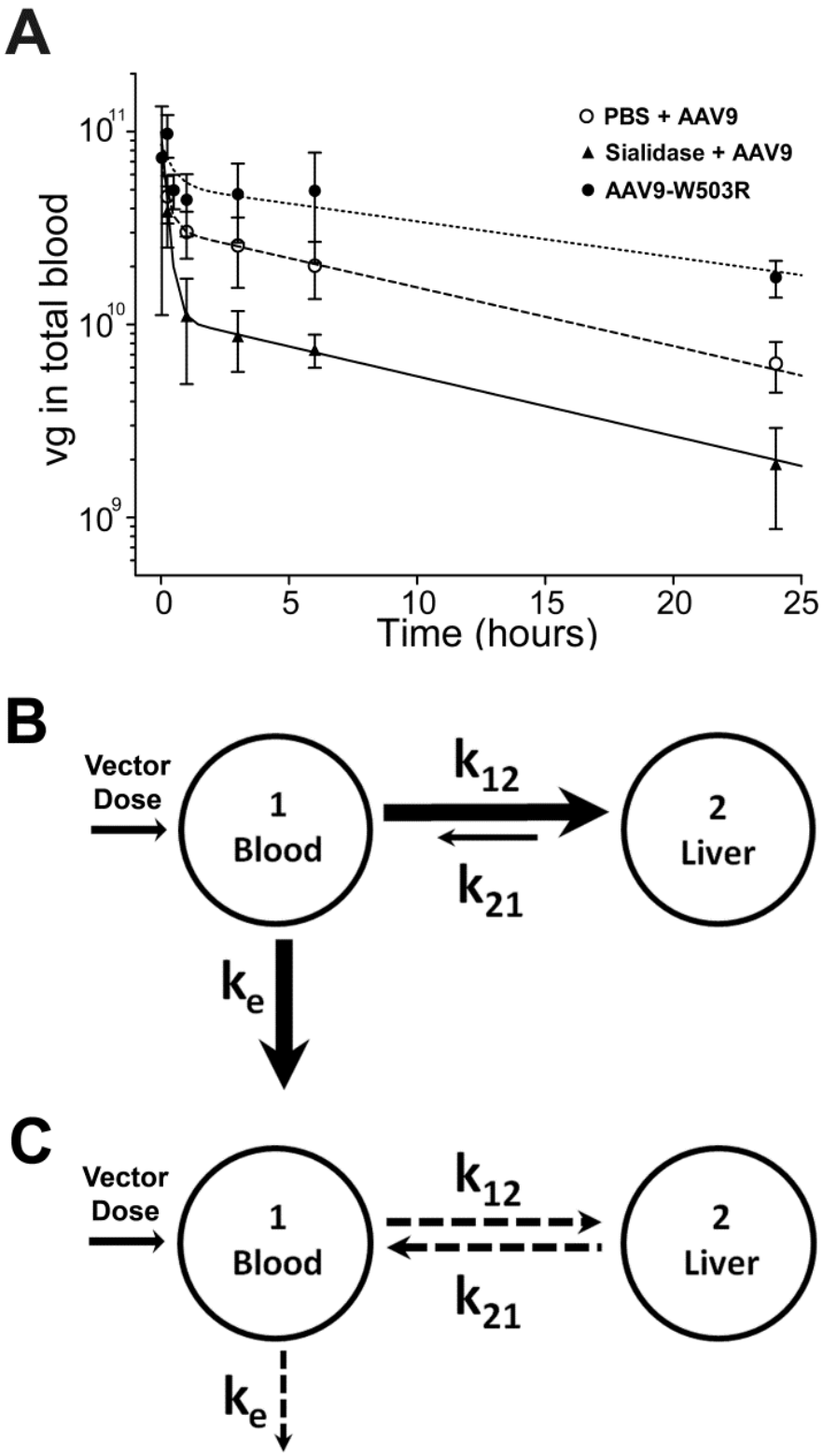

Figure 16. Glycan binding avidity affects the blood circulation profile of AAV9. (A) Blood circulation kinetics of the AAV9-W503R mutant (●) or AAV9 in Balb/c mice pre-treated with PBS (o) or sialidase (▲). Viral genome copy numbers in blood obtained at

15min, 30min, 1 hr, 3 hrs, 6 hrs and 24 hrs post-administration were obtained by quantitative PCR (n=3). Curve fitting based on a bi-exponential, two-compartment model was utilized to obtain pharmacokinetic parameters listed in Table 4. Data are represented as mean ± s.e.m.. Equations, goodness of fit and statistical analysis are listed in Tables 5 and 6. **(B)** Proposed mathematical model for predicting the impact of high glycan binding avidity on viral blood clearance and liver sequestration. Rates of viral elimination from blood (k_e), viral dissemination from blood-to-liver (k_{12}) and liver-to-blood (k_{21}) are shown. Bold arrows indicate faster rates of dissemination and clearance from blood. This model specifically explains the pharmacokinetics of wild type AAV9 following enzymatic desialylation in mice. **(C)** Proposed mathematical model for predicting the impact of low glycan binding avidity on viral blood clearance and liver sequestration. Different rates (k values) are shown as above. Dotted arrows indicate slower rates of dissemination and clearance from blood. This model specifically explains the pharmacokinetics of mutant AAV9-W503R in mice.

CHAPTER 4

Engineering Novel Dual Glycan Binding AAVs for Enhanced Gene Transfer[3]

4.1 Overview

New viral strains can be evolved to recognize different host glycans through mutagenesis and experimental adaptation. However, reassortants generally harbor mutations that affect viral binding to a single class of carbohydrate receptors. We describe the rational design and synthesis of a novel class of adeno-associated viruses (AAV) that exploit two distinct glycan receptors for cell entry. A prototypical dual glycan binding AAV strain was engineered by "grafting" an orthogonal galactose (Gal) binding footprint from AAV serotype 9 onto the heparan sulfate (HS) binding AAV serotype 2. The resulting chimera, AAV2G9, interchangeably exploits Gal and HS as evidenced by competitive inhibition assays with lectins and glycans. Further, AAV2G9 mediates rapid onset and sustained, higher transgene expression *in vivo* compared to parental AAV serotypes. In addition to demonstrating the modularity of glycan receptor footprints on viruses, our approach provides design parameters to upgrade the current AAV vector toolkit for clinical gene therapy.

[3]This chapter includes the original manuscript in preparation: Shen, S., Horowitz, E.D., Troupes, A., Brown, S. M., Asokan, A. *Engineering of a Dual-glycan Binding Aden-associated Virus Variant Using Molecular Grafting.*

4.2 Introduction

Virus-glycan interactions are critical determinants of host cell invasion. Cell surface carbohydrates such as sialic acids, gangliosides or heparan sulfate are exploited by a vast number of viruses such as influenza, herpesvirus, SV40, polyomavirus, papillomavirus and other pathogens [1,192]. In most cases, a single class of glycans primarily serves as the cell surface attachment factor for viruses, leading to sequential or parallel engagement of other receptors/co-receptors for cell entry. Adeno-associated viruses (AAV) are helper-dependent parvoviruses that exploit heparan sulfate (HS), galactose (Gal) or sialic acids (SA) as primary receptors for cell surface binding [152,193]. For instance, AAV serotypes 2 and 3b utilize HS; AAV1, 4 and 5 bind SA with different linkage specificities; while AAV9 exploits Gal for host cell attachment. Different AAV strains also require subsequent interaction with co-receptors such as integrin αVβ5 or α5β1, fibroblast growth factor receptor (FGFR), platelet-derived growth factor receptor (PDGFR), epidermal growth factor receptor (EGFR), hepatocyte growth factor receptor (HGFR) or the laminin receptor for cellular uptake [152,193].

A notable exception to the monogamous relationship between a specific AAV strain and a single class of carbohydrates is AAV serotype 6, which recognizes both SA and HS [211]. However, only SA has been shown essential for viral transduction. Structural studies have now established that the K531 residue in conjunction with R488, K528 and K533 in the VP3 subunit of the AAV6 capsid form a continuous basic patch for electrostatic recognition of HS glycosaminoglycans [43,165,198]. Similarly, the structural basis for HS recognition by AAV2 and AAV3b is well known and attributed to similar clusters of basic amino acid residues located at the three-fold axis of symmetry [83,90,92,93]. The SA binding footprints for AAV1, AAV4, AAV5 and AAV6 remain to be determined. More recently, key amino acid residues involved

in Gal recognition by AAV9 capsids were identified by using a combination of molecular docking and site-directed mutagenesis [94]. This abundance of structural information and knowledge pertaining to host receptor usage by AAV serotypes provided the rationale for designing new dual glycan binding AAV strains.

In the current study, a new, prototypical dual glycan binding strain, AAV2G9 was generated by incorporating the Gal binding footprint from AAV9 onto the AAV2 VP3 backbone using structural alignment and site-directed mutagenesis. *In vitro* binding and transduction assays confirmed the exploitation of both HS and Gal receptors by AAV2G9 for cell entry. Subsequent *in vivo* characterization of the kinetics of transgene expression and viral genome biodistribution profiles indicate fast, sustained and enhanced transgene expression by this novel rationally engineered chimera.

4.3 Materials and Methods

Structural modeling. Coordinates for the AAV2 and AAV9 viral protein (VP) crystal structures were obtained from RCSB Protein Databank (PDB accession# 1LP3 and 3UX1, respectively) [81,212]. Using the SWISS-MODEL protein structure modeling server (http://swissmodel.expasy.org/) [213], homology models of the 2G9 VP3 monomer were generated with crystal structures of AAV2 VP3 as template. A three-dimensional icosahedral model of an intact 2G9 capsid was created using the Oligomer Generator utility in VIPERdb-Virus Particle ExploreR2 (http://viperdb.scripps.edu/) [214]. Similarly, illustration of the AAV2 VP3 trimer, 2G9 trimer, and AAV9 trimer were obtained using the Oligomer Generator utility. All structural models were visualized using PyMOL with residues forming the galactose binding site (AAV9 VP3 numbering: D271, N272, Y446, N470, A472, V473,

W503) [94] and heparan sulfate binding site (AAV2 VP3 numbering:R487, K527, K532, R585, R588) [83,89,90,92] highlighted in orange and purple, respectively. Different VP3 monomers were colored in pale green, light blue and light pink.

Generation of dual glycan binding AAV strains. Helper plasmids pXR1, 2, 6, 8 and 9 were obtained from UNC vector core. The prototypical pXR2G9 chimera plasmid construct was generated by substituting amino acid residues directly involved or flanking the Gal recognition site on the AAV9 VP3 capsid protein subunit onto corresponding residues on the VP3 subunit of AAV2 (AAV2 VP3 numbering: A266S, Q464V, A467P, D469N, I470M, R471A, D472V, S474G, Y500F, S501A). Substitutions were generated using the QuikChange® Lightning site-directed mutagenesis kit (Agilent) using the following primers (IDT): 5'- GGAACCACCACGCAGTCAAGGCTTCAGTTTTCTGTGGCCGGACC CAGTAACATGGCTGTCCAGGGAAGGAACTGGCTTCCTGGACCCTGTTACCGC-3' and 5'- GACATCTGCGGATAACAACAACAGTGAATTTGCTTGGACTGGAGCTACC AAGTACCACCT-3'. Recombinant AAV vectors packaging the CBA-Luc transgene cassettes were generated as described previously [166]. Viral titers were obtained by quantitative PCR.

***In vitro* binding, transduction and competitive inhibition assays.** CHO Lec2 cells were cultured in αMEM (Thermo Scientific) supplemented with 10% fetal bovine serum (FBS), 100 U/ml of penicillin (Cellgro), 100 µg/ml of streptomycin (Cellgro), and 2.5 µg/ml of amphotericin B (Sigma). Cells were seeded at a density of $1x10^5$ cells/well in 24 well plates. For competitive inhibition assays, cells were pre-chilled at 4ºC for 30 minutes and incubated with 100 µg/ml of FITC-labeled *Erythrina Cristagalli* Lectin (FITC-ECL, Vector Labs) in αMEM at 4ºC for 1 hr. Alternatively, different viral capsids were incubated with

100µg/ml of soluble heparin (Sigma) or 1xPBS (control) at room temperature for 1 hr. Mock-treated or FITC-ECL treated cells were then infected with HS-bound or mock-treated AAV2, AAV2G9, or AAV9 capsids packaging a CBA-Luc transgene cassette at an MOI of 1000 vector genome (vg) copies/cell. Following incubation in the cold room for 1 hour, unbound virions were removed by three washes with ice cold 1xPBS. For cell surface binding assays, the number of bound virions was measured by quantifying vector genome copy numbers/cell in each well using quantitative PCR. For transduction assays, infected Lec2 cells were moved to 37ºC and incubated for 24 hrs prior to quantitation of luciferase transgene expression from cell lysates.

For competitive inhibition with parental AAV2 or AAV9 capsids, vectors packaging CBA promoter-driven tdTomato transgene cassette were utilized. Briefly, Lec2 cells were seeded in 24 well plates overnight at a density of $1x10^5$ cells/well. After being pre-chilled at 4ºC for 30 minutes, Lec2 cells were pre-incubated with either AAV2-tdTomato or AAV9-tdTomato vectors at multiplicities of infection (MOI) ranging from 500 to 100,000 vg/cell at 4ºC for another 30 minutes. Cells were then super-infected with AAV2G9-CBA-Luc at an MOI of 1000 vg/cell for 45 minutes at 4ºC, followed by removal of unbound virions using ice cold PBS. Infected cells were then incubated at 37ºC for 24 hrs prior to luciferase expression analysis. Controls included AAV2-CBA-Luc or AAV9-CBA-Luc vectors.

Kinetics of transgene expression *in vivo*. Female BALB/c mice (6-8 weeks old) were purchased from Jackson Laboratories and handled in accordance with NIH guidelines using IACUC approved protocols at UNC Chapel Hill. Different AAV vectors packaging the CBA-Luc cassette were injected intravenously into the tail vein at a dose of $1x10^{11}$ vg/mouse. At indicated time intervals post-administration (3, 7, and 18 days), mice were

intraperitoneally injected with luciferin (120 mg/kg; Nanolight) and bioluminescent images obtained using an Xenogen IVIS® Lumina system (Caliper Lifesciences). Quantitation of light output from liver and whole animal images was carried out using Wavemetrics® software. Further quantitation of luciferase transgene expression and vector genome biodistribution in different tissues was carried out in two different groups of mice that were sacrificed at days 3 and 18 post-vector administration. Luciferase transgene expression was monitored in different tissue lysates as described earlier. Vector genome biodistribution was determined by first extracting genomic DNA from tissue lysates using a DNeasy® Kit (Qiagen). Luciferase transgene copy number was determined using qPCR and normalized to the number of copies of the mouse lamin gene to determine vg/cell in each tissue. Specific primer sets were 5'-AGG GCA CCT CCA TCT CGG AAA C-3' / 5'-GGA CCC AAG GAC TAC CTC AAG GG-3' (for mouse lamin) and 5'-AAA AGC ACT CTG ATT GAC AAA TAC-3'/ 5'-CCT TCG CTT CAA AAA ATG GAA C-3'(for CBA-Luc) , respectively.

Statistical Analysis. All data is expressed as mean ± standard error mean and the number of replicates for each experiment is provided in the corresponding figure legends. Statistical significance was determined using the unpaired one-tail student's *t*-test and p-values less than 0.05 considered statistically significant for different experiments unless indicated otherwise.

4.4 Results

Construction of the dual glycan binding prototype strain

To explore the feasibility of "grafting" the Gal footprint of AAV9 onto several AAV strains, we first compared the three-dimensional structures of VP3 subunit trimers of AAV

serotypes 1, 2, 6 and 8 in alignment with that of AAV9 (**Figure 18**). Amino acid residues on the template capsids that overlapped with corresponding AAV9 VP3 residues directly involved in binding or immediately flanking the Gal receptor footprint were modified by multiple rounds of site-directed mutagenesis. All of the chimeric AAV strains generated were prepared as recombinant vectors packaging a chicken beta-actin promoter driven firefly luciferase (CBA-Luc) reporter transgene cassette using previously established protocols [166]. Amino acid residues involved in Gal recognition and other flanking residues from AAV9 were remarkably well tolerated on different AAV serotype capsids as the packaging efficiencies of these AAV chimeras are comparable with parental strains. Multiple AAV chimeras based on AAV serotypes 1, 2, 6, 8 and the previously engineered AAV2i8 mutant [143] were obtained (at titers ranging from 5×10^{11} to 5×10^{12} viral genome copies/mL) and observed to exploit Gal as a novel primary receptor in transducing CHO Lec2 cells *in vitro* (**Figure 20**). We then carried out a detailed characterization of a prototypical dual glycan binding AAV chimera, dubbed AAV2G9 (where G stands for the Gal footprint and the numbers identify the recipient and donor capsid serotypes, respectively).

Three-dimensional models of synthetically engineered AAV2G9 (full capsid in **Figure 17A** and VP3 trimer in **Figure 17D**) with the putative dual glycan receptor binding sites (HS and Gal) highlighted in purple and orange were generated by homology modeling using Swiss Model®. The molecular model of AAV2G9 full capsids demonstrates the geometrical distribution and orthogonality of HS and Gal binding sites located around the three-fold symmetry axis on the icosahedral capsid. Close-up views of HS and Gal receptor footprints from the three-fold axes further support the observation that grafting orthogonal Gal binding sites on the backbone of AAV2 capsid can be tolerated with regard to capsid

assembly. Three-dimensional structures of the AAV2 VP3 subunit trimer with side chains of positively charged residues involved in HS recognition (highlighted in purple, **Figure 17A**) as well as the side chains of amino acid residues comprising the Gal recognition site on the AAV9 VP3 subunit trimer (highlighted in orange, **Figure 17C**) are also shown.

AAV2G9 exploits HS and Gal receptors interchangeably *in vitro*

The first line of evidence supporting the usage of dual glycan receptors by AAV2G9 was obtained from competitive inhibition assays of virus binding on cell surface involving soluble heparin and *Erythrina Cristagalli* lectin (ECL), which selectively binds terminally galactosylated glycans. As seen in **Figure 19A & Figure 19B**, HS, but not ECL significantly inhibits AAV2 transduction in CHO Lec2 cells (dark grey bars), while ECL selectively blocks AAV9 transduction by nearly two log units (white bars). These results are consistent with the expected transduction profiles for AAV2 and AAV9 [40,52,53]. In contrast, AAV2G9 can only be effectively neutralized by pre-treatment with a combination of both ECL and HS (light grey bars, **Figure 19C**). A small, yet significant inhibitory effect is observed for ECL.

Transduction profiles for AAV2 and AAV9 were further corroborated by inhibition of cell surface binding by each strain using ECL or HS (**Figure 19D & Figure 19E**). The unique cell surface attachment of the chimeric AAV strain is further supported by competitive inhibition of cell surface attachment of AAV2G9 exclusively by a combination of ECL and HS, but neither reagent alone (**Figure 19F**). In addition, confocal immunofluorescence micrographs (**Figure 22**) obtained using monoclonal antibodies against different AAV capsids suggest that AAV2G9 binds more robustly to the surface of CHO Lec2 cells than AAV2 or AAV9. Such a scenario can be expected based on the apparent ability of AAV2G9 to bind two different glycans interchangeably.

In order to further interrogate the exploitation of alternate transduction pathways by AAV2G9, we conducted competition assays with the parental serotypes, AAV2 and AAV9. As shown in **Figure 21A & Figure 21B**, pre-incubation with AAV2-CBA-tdTom or AAV9-CBA-tdTom competing vectors at MOIs ranging from 500 to 100,000 vg/cell efficiently blocks transduction by AAV2-CBA-Luc or AAV9-CBA-Luc, respectively as measured by luciferase transgene expression. However, both AAV2 and AAV9 are unable to effectively block AAV2G9 transduction at 10-fold excess multiplicities of infection (MOI). At higher MOI (100-fold excess), AAV2 appears to compete less effectively than AAV9 in neutralizing AAV2G9 transduction. Taken together, these results support the notion that AAV2G9 is indeed a novel, dual glycan binding strain with the unique ability to exploit both HS and Gal as primary receptors for transduction.

AAV2G9 mediates rapid onset of transgene expression

We then investigated whether dual glycan binding confers specific advantages to viral transduction *in vitro* and *in vivo*. Monitoring the time course of luciferase reporter expression in CHO Lec2 cells revealed that AAV2G9 mediates rapid onset and improved gene transfer *in vitro* (**Figure 23**). Live animal imaging studies were then carried out to monitor luciferase transgene expression following systemic administration of different AAV strains in BALB/c mice (**Figure 24**). Bioluminescent images and quantitative assessment of light output within the liver and the whole animal obtained at days 3, 7 and 18 post-injection correlate with *in vitro* data and support the notion that AAV2G9 can mediate rapid onset and enhanced gene expression (**Figure 24B andFigure 24C**). Interestingly, the kinetic profile displayed by AAV2G9 mirrors that of AAV9 but not AAV2. In contrast, the transduction profile/tissue tropism of AAV2G9 appears to be primarily hepatotropic, similar to AAV2 and unlike the

systemic tropism displayed by AAV9 as established previously [125,152,215,216]. Thus, dual glycan receptor engagement appears to improve the transduction efficiency of AAV strains, but does not alter tissue tropism.

Transduction and vector biodistribution profile of AAV2G9 *in vivo*

To further evaluate the *in vivo* transduction and biodistribution profiles of AAV2G9, quantitative analysis of tissue lysates from BALB/c mice were carried out at days 3 and 18 post-administration. Specifically, AAV2G9 displays markedly higher luciferase transgene expression in liver compared to AAV2 (nearly two log units) and AAV9 (~1 log unit) at 3 days post-administration (**Figure 25A**). While AAV9 displays more than 10-fold higher transduction efficiency in heart than AAV2G9, a modest increase in cardiac transduction by AAV2G9 compared to AAV2 is also observed. At day 18, cardiac and liver tissues harvested from mice treated with AAV2G9 continue to demonstrate higher transgene expression, although AAV9 emerges as the most efficient strain at this stage. Specifically, transduction efficiencies in cardiac tissue by AAV2, AAV2G9, and AAV9 maintain a similar trend as observed 3 days post-administration. In the liver, the differences between luciferase transgene expression by AAV2, AAV2G9, and AAV9 diminish upon progressing to 18 days post-injection. Specifically, AAV9 demonstrates between 5 to 10-fold higher transgene expression when compared to AAV2G9 and AAV2, respectively.

Quantitative analysis of vector genome copy numbers in liver and heart by AAV2G9 and the parental AAV strains at 3 days post-administration (**Figure 25B**) is consistent with the trends observed for transduction efficiencies shown in Figure 5A. Specifically, AAV2G9 accumulated to a higher extent in cardiac tissue compared to AAV2, but was still ~2 log units lower than AAV9. In liver, AAV2G9 copy number is comparable to that of AAV9, but over

one log unit higher than AAV2. At day 18, copy numbers for all serotypes were decreased presumably due to continuous cell turnover and degradation of single-stranded AAV genomes as reported previously [217,218].

4.5 Discussion

The current study demonstrates that grafting conformational receptor footprints from one AAV strain onto others is feasible and could result in novel chimeric AAV capsids with enhanced transduction profiles *in vitro* and *in vivo*. The dual glycan binding chimera AAV2G9 was synthesized by incorporating residues from within or around the Gal binding pocket on AAV9 onto the AAV2 capsids. Remarkably, a minor change in the primary amino acid sequence of the AAV2 VP3 subunit (~1.4%) was sufficient to impart Gal recognizing ability without affecting HS binding. These results are particularly unique from a structural perspective. For instance, in the case of influenza A, mutations within a single sialic acid binding site have been associated with altered recognition of glycan linkage specificities by new reassortant strains[219,220]. However, it is noteworthy to mention that the Gal and HS binding residues on AAV2G9 create two distinct footprints on the viral capsid surface. Further, AAV2G9 appears to utilize HS and Gal simultaneously and interchangeably. Based on similar rationale, our studies might provide design parameters for synthesizing a comprehensive panel of dual/multi-glycan binding AAV strains harboring various permutations and combinations of distinct glycan recognition patterns.

Our results also provide insights into the biology of AAV receptor usage *in vivo*. First, the predominant determinants of AAV2G9 tropism *in vivo* appear to be dictated by one of the parental AAV serotypes (i.e., in this case, AAV2) which composes a majority of the capsid

structure. This would suggest that the Gal footprint on AAV2G9 is compatible with utilization of fibroblast growth factor receptor (FGFR), integrins and/or other co-receptors thought to play a role in AAV2 cell entry [152,193]. An intriguing question that follows is whether interplay between primary glycan receptors and different AAV co-receptors are critical determinants of tissue tropism. Whether incorporation of a SA binding pocket onto AAV2 or HS binding sites on other SA/Gal binding AAV serotypes will yield viable chimeric strains remains to be determined. Further, as co-receptors for AAV serotypes continue to be identified, synthetic dual glycan binding strains might be extremely useful in understanding this specific aspect of AAV biology.

Alternatively, the predominant hepatotropism displayed by AAV2G9 can possibly be attributed to a higher affinity for HS than Gal moieties. Indeed, we observed that incorporation of the Gal footprint into a HS-binding defective mutant (AAV2i8) yielded a liver-detargeted mutant displaying higher transduction efficiency in heart and skeletal muscle (mutant AAV2i8G9, **Figure 26**). In addition, while AAV2G9 did not exhibit systemic tropism similar to AAV9, the transduction efficiency of AAV2G9 in liver appears to be as high as AAV9. This explanation is further corroborated by earlier studies demonstrating the critical role played by HS interactions in liver tropism of AAV serotype 6 [43]. Thus, it is tempting to speculate that while HS binding by AAV2G9 primarily determines liver accumulation, recognition of a second glycan receptor (i.e., Gal) appears to promote increased tissue uptake and rapid onset of transgene expression.

For the gene therapy community, the new technology reported herein provides a platform to develop improved clinical reagents. Dual glycan binding AAV vectors can initiate earlier gene expression that continues to increase at a faster rate and can sustain

higher transduction levels in comparison to naturally occurring parental strains. As a result, such lab-derived strains might help decrease the effective vector dose required for achieving therapeutically relevant transgene expression levels and help alleviate potential clinical concerns associated with vector dose-dependent toxicity [221]. Whether such vector engineering approaches can help achieve higher transduction efficiencies upon synergistic integration with other major advancements in AAV technology, e.g., self-complementary vectors [222] or Tyr-to-Phe capsid mutations [84], which enable bypassing second strand DNA synthesis and proteasomal degradation, respectively remains to be seen. In addition to such studies, validation of synthetic dual glycan binding AAV strains in large animal models is forthcoming and likely to shed light on cross-species variation in AAV tissue tropism.

In summary, rational engineering of capsid-glycan receptor interactions is a promising approach towards continued improvement of the preclinical/clinical portfolio of AAV vectors for gene therapy. Additional structural information pertaining to receptor binding footprints on different AAV capsids or other closely related parvoviruses could further expand the matrix of synthetic, dual glycan binding AAV vectors.

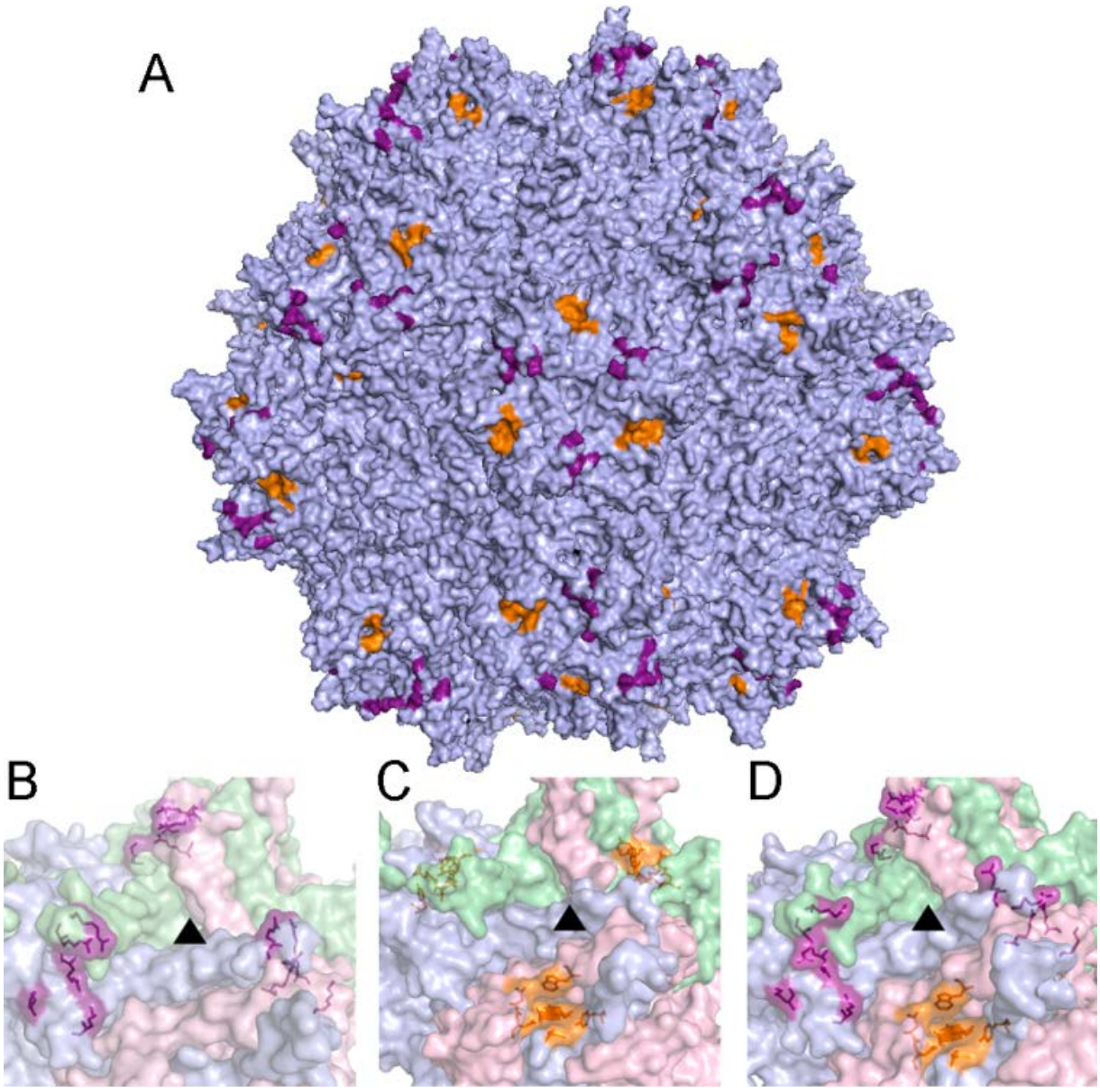

Figure 17. Three-dimensional models of the dual glycan binding AAV2G9 chimera and its parental strains AAV2 and AAV9. (A) Three-dimensional structural model of an intact AAV2G9 capsid with existing HS and "grafted" Gal binding sites colored in purple and orange, respectively. **(B-D)** Illustrations of the three-dimensional surface model of VP3 trimers at the three-fold symmetry axes of AAV2 **(B)**, AAV9 **(C)**, and AAV2G9 **(D)** capsids. Residues involved in HS binding (AAV2 VP3 numbering: R487, K527, K532, R585, R588) and Gal binding (AAV9 VP3 numbering: D271, N272, Y446, N470, A472, V473, W503) are highlighted as in **(A)**. Black triangles indicate the three-fold symmetry axes.

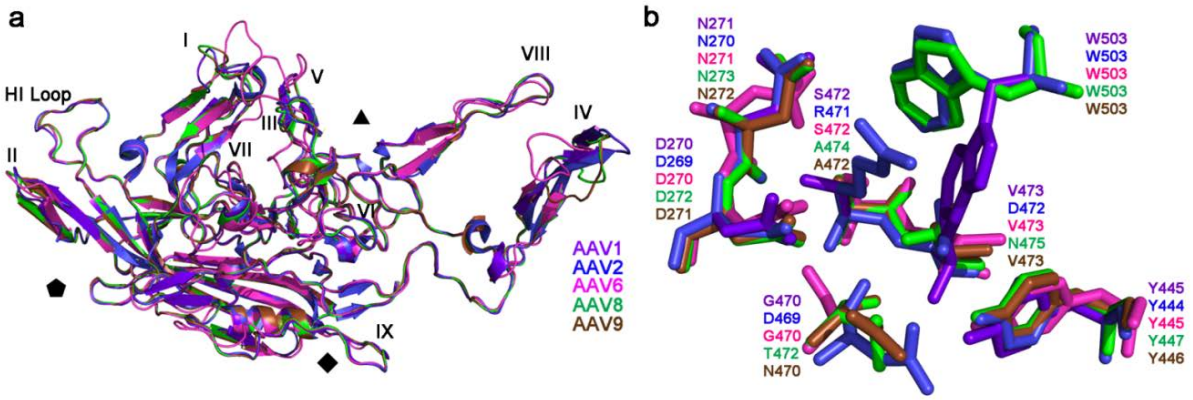

Figure 18. Structural alignment of AAV1, AAV2, AAV6, AAV8, and AAV9 VP3 monomers. (A) Superposition of the VP3 monomers of AAV1 (purple blue), AAV2 (deep blue), AAV6 (light magenta), AAV8 (green), and AAV9 (brown) with loops I-IX labeled and axes of symmetry indicated. **(B)** Close-up views of overlay of the galactose binding site on AAV9 and equivalent residues on AAV1, AAV2, AAV6, and AAV8. Amino acid residues are marked by the color code in **(A)**. Coordinates were obtained from X-ray crystallography structure of VP monomers (PDB accession#: AAV1-3NG9, AAV2-1LP3, AAV6-3OAH, AAV8-2QA0, AAV9-3UX1). Structure alignment was performed and visualized using PyMOL.

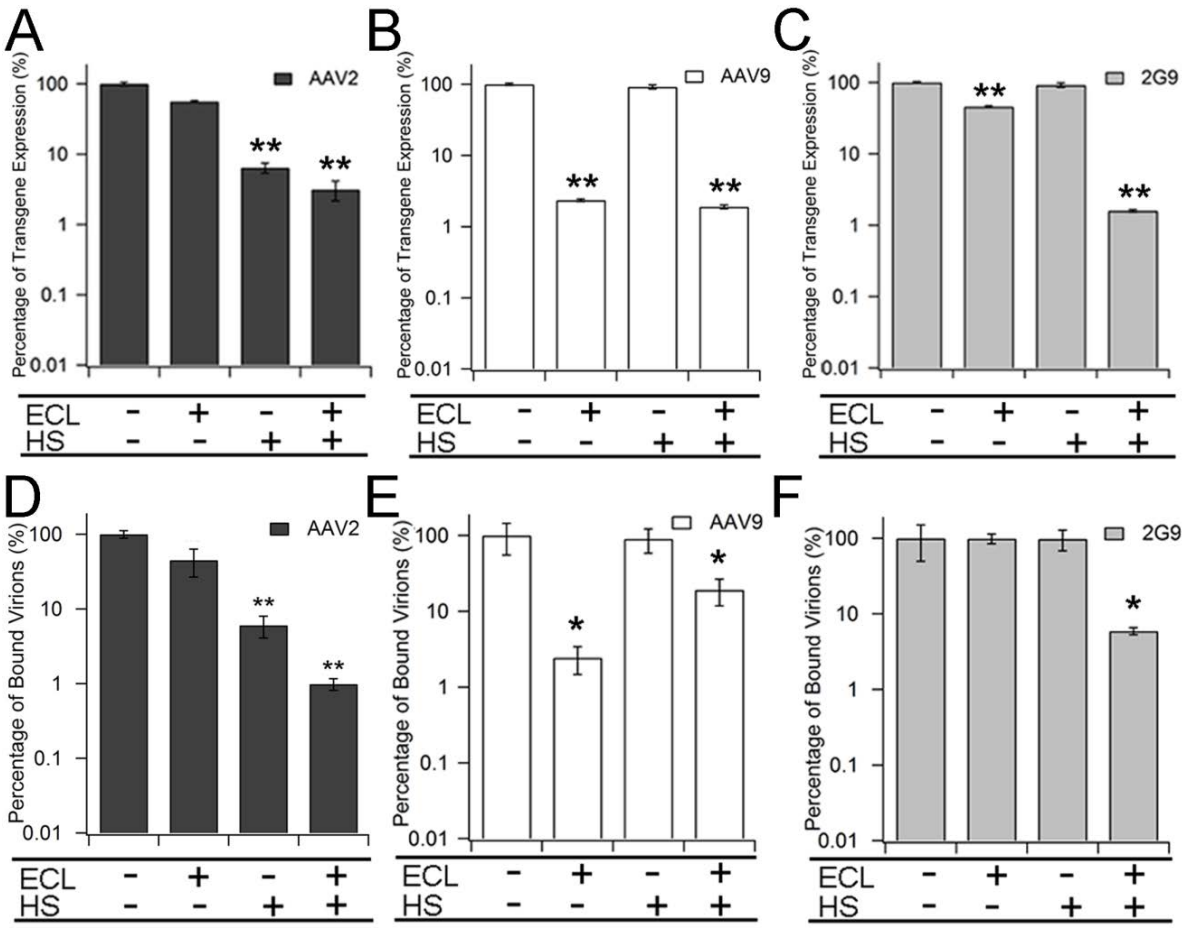

Figure 19. *In vitro* characterization of the dual glycan binding AAV2G9 chimera. (A-C) Inhibition of AAV2 **(A)**, AAV9 **(B)**, and AAV2G9 **(C)** transduction on CHO Lec2 cells with FITC-ECL and soluble heparin. CHO Lec2 cells were pre-chilled at 4ºC and incubated with FITC-ECL, soluble heparin or both prior to infection with AAV2, AAV9 or AAV2G9 packaging a CBA-luciferase reporter transgene cassette. Transduction efficiency was measured 24 hrs post infection as luciferase activity in relative light units (RLU). Percentage of transgene expression was calculated by normalizing transduction efficiency to RLU from controls. Results are presented as mean ± s.e.m. (n=4). **(D-F)** Competitive inhibition of cell surface binding of AAV2 **(D)**, AAV9 **(E)**, and AAV2G9 **(F)** on CHO Lec2 cells with FITC-ECL and soluble heparin. Different AAV particles were bound to cells pre-chilled at 4ºC and unbound virions removed by washing with cold PBS. Bound virions were quantified using qPCR after viral genome extraction. Percentage of bound virions was determined by normalizing number of bound virions to that of corresponding controls. Results are presented as mean ± s.e.m. (n=5). Statistical significance was analyzed using the one-tailed Student's *t*-test (* $p < 0.05$; ** $p < 0.01$).

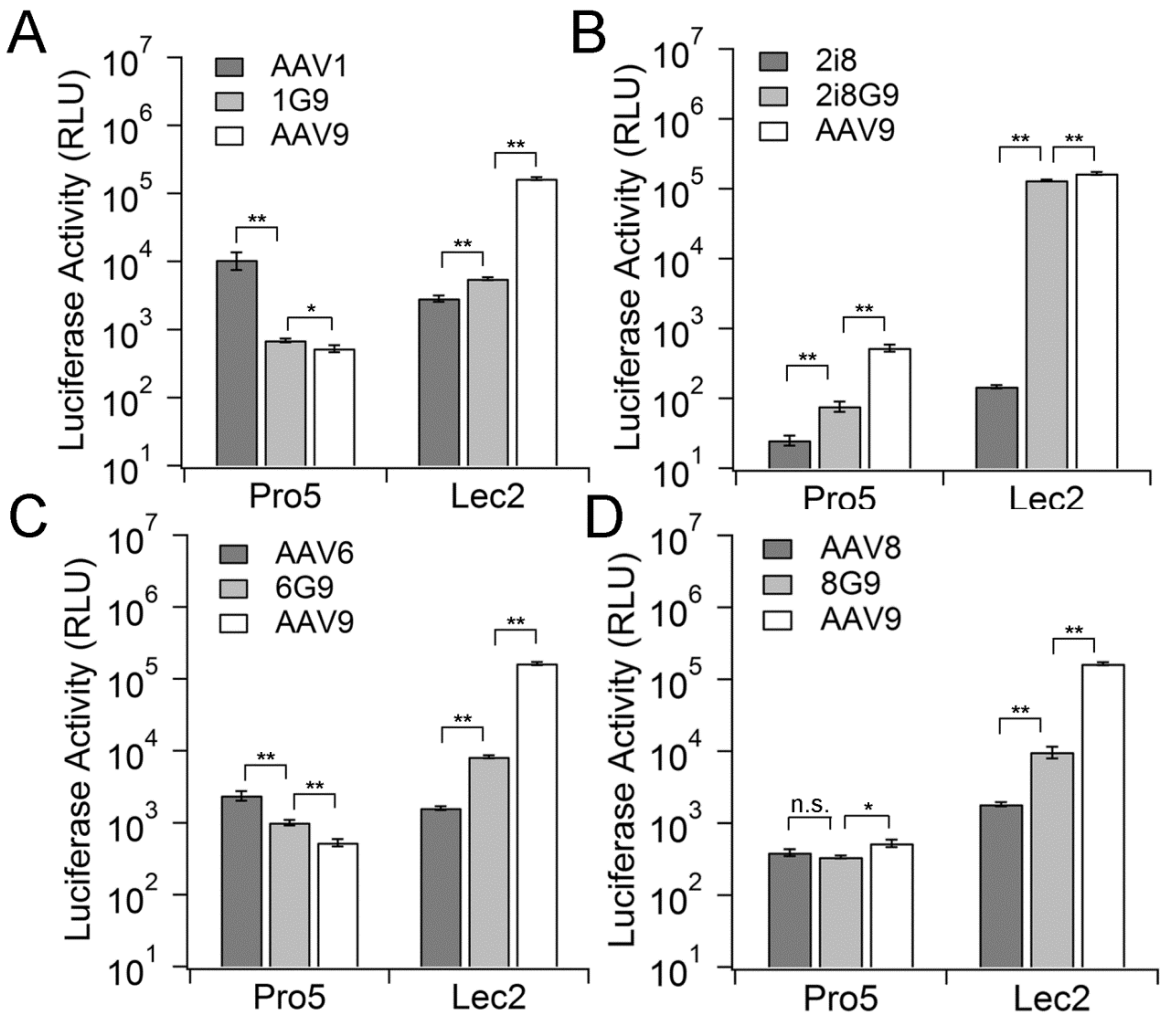

Figure 20. G-mutants utilize Gal as a novel glycan receptor to transduce cells *in vitro*. **(A)** Transduction efficiency of AAV1, 1G9, and AAV9 on Chinese Hamster Ovary (CHO) cell lines. Pro5 and Lec2 cells were pre-chilled to 4ºC for 30 minutes prior to AAV-CBA-Luciferase infection at an MOI of 1000 vg/cell at 4ºC for 60 minutes. After removing unbound virions by three washes with ice-cold PBS, infected cells were cultured in 37ºC incubator for 24 hrs. Luminometric analysis was performed to quantify the luciferase transgene expression efficiencies from cell lysates. **(B)** Transduction efficiency of AAV2i8, 2i8G9, and AAV9 on Pro5 and Lec2 cells. **(C)** Transduction efficiency of AAV6, 6G9, and AAV9 on Pro5 and Lec2 cells. **(D)** Transduction efficiency of AAV8, 8G9, and AAV9 on Pro5 and Lec2 cells. Results are presented as mean ± s.e.m. (n=4). Statistical significance was assessed using the one-tailed Student's *t*-test (n.s., not significant; * $p < 0.05$; ** $p < 0.01$).

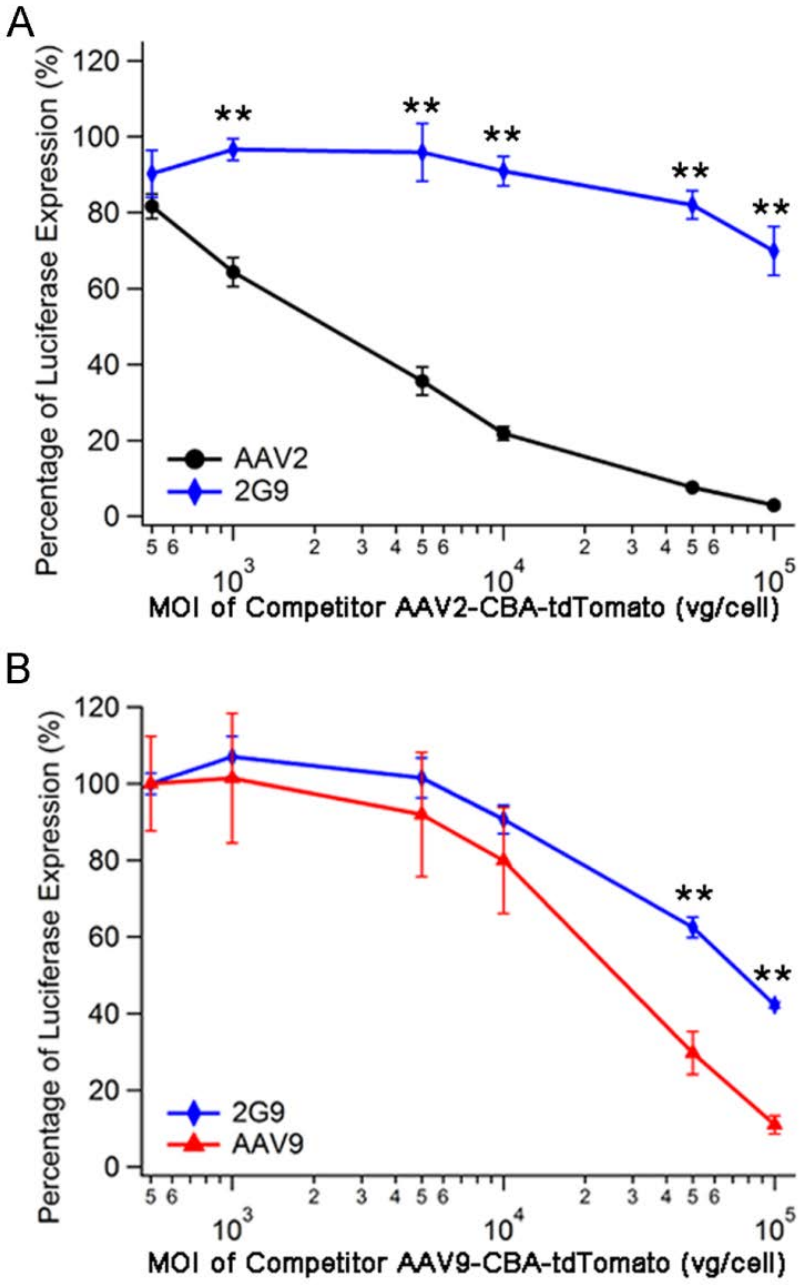

Figure 21. Competitive inhibition of AAV2G9 transduction on Lec2 cells by AAV2 capsids or AAV9 capsids. Lec2 cells were preincubated with **(*A*)** AAV2 or **(*B*)** AAV9-CBA-tdTomato at multiplicity of infection (MOI) ranging from 500 to 100,000 vg/cell for 2 hrs prior to infection with AAV2G9-CBA-Luc particles (MOI 1000 vg/cell). Percentage inhibition of AAV2G9 transduction was calculated by normalizing luciferase transgene expression levels to that of untreated control. Results are presented as mean ± s.e.m. (n=4). Statistical significance was analyzed using the one-tailed Student's *t*-test (* $p < 0.05$; ** $p < 0.01$).

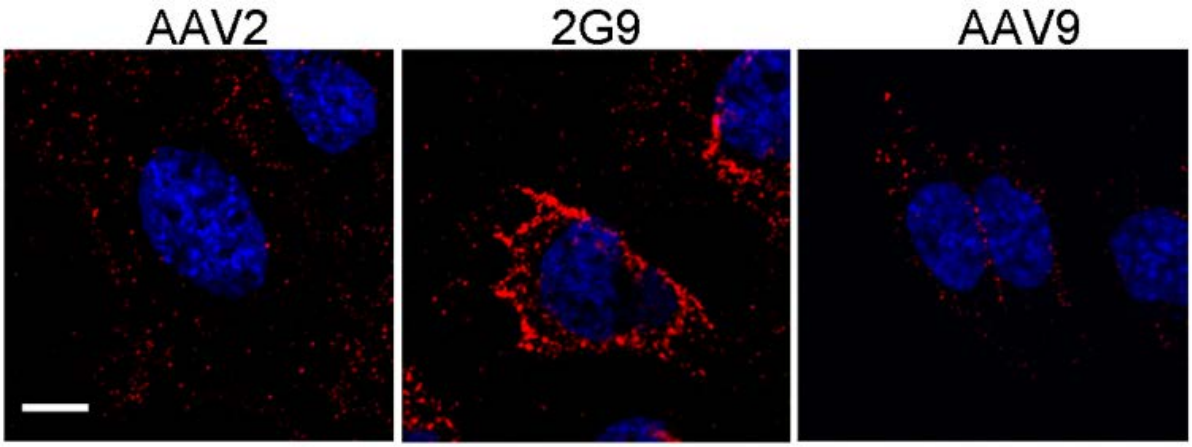

Figure 22. Immunofluorescence of bound virions on Lec2 cell surface at 0 hrs post-infection (hpi). CHO Lec2 cells were pre-chilled at 4ºC for 30 minutes, followed by infection with AAV2, AAV 2G9, and AAV9 at an MOI of 1000 vg/cell at 4ºC for 30 minutes. After removal of unbound virions, cells were fixed with 2% paraformaldehyde in 1xPBS. Intact virions were detected using the monoclonal antibodies (A20 for AAV2/AAV2G9 and ADK9 for AAV9). Alexa Fluor 594® goat anti-mouse IgG was utilized at a dilution of 1:1000 in IFWB as the secondary antibody for immunofluorescent detection. Fluorescent micrographs were acquired using a Zeiss® 710 confocal laser scanning microscope equipped with a 63x oil immersion objective and a spectral detection system. Image processing was carried out using LSM® viewer and Image J® software. The red color represents intact viral particles, and the blue color from DAPI marks nucleus. The white scale bar indicates 10 µm.

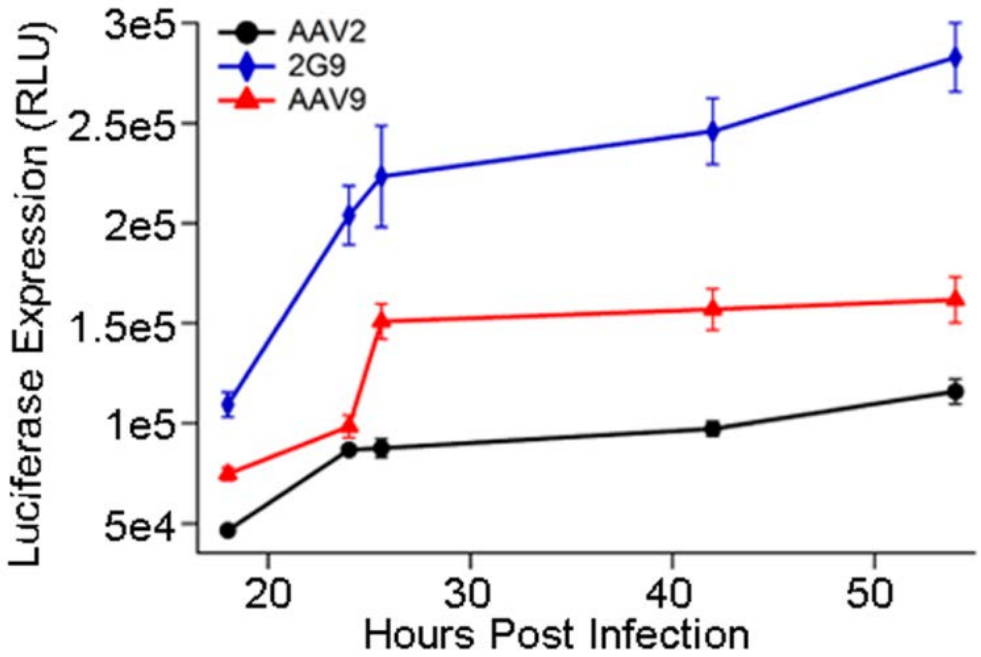

Figure 23. Kinetics of transduction efficiency profiles of AAV 2G9 compared to parental AAV2 and AAV9 on Lec2 cells at indicated time points post infection. Pre-chilled Lec2 cells were infected with AAV2, AAV2G9, or AAV9-CBA-luciferase vectors at an MOI of 1000 vg/cell as described earlier. At indicated time points (18, 24, 28, 42 and 54 hrs) post-infection, cells were lysed prior to luminometric analysis. Luciferase transgene expression was measure by luciferase activities of cell lysates in relative light units (RLU) (n=5). Statistical significance was assessed using the one-tailed Student's t-test (* $p < 0.05$; ** $p < 0.01$).

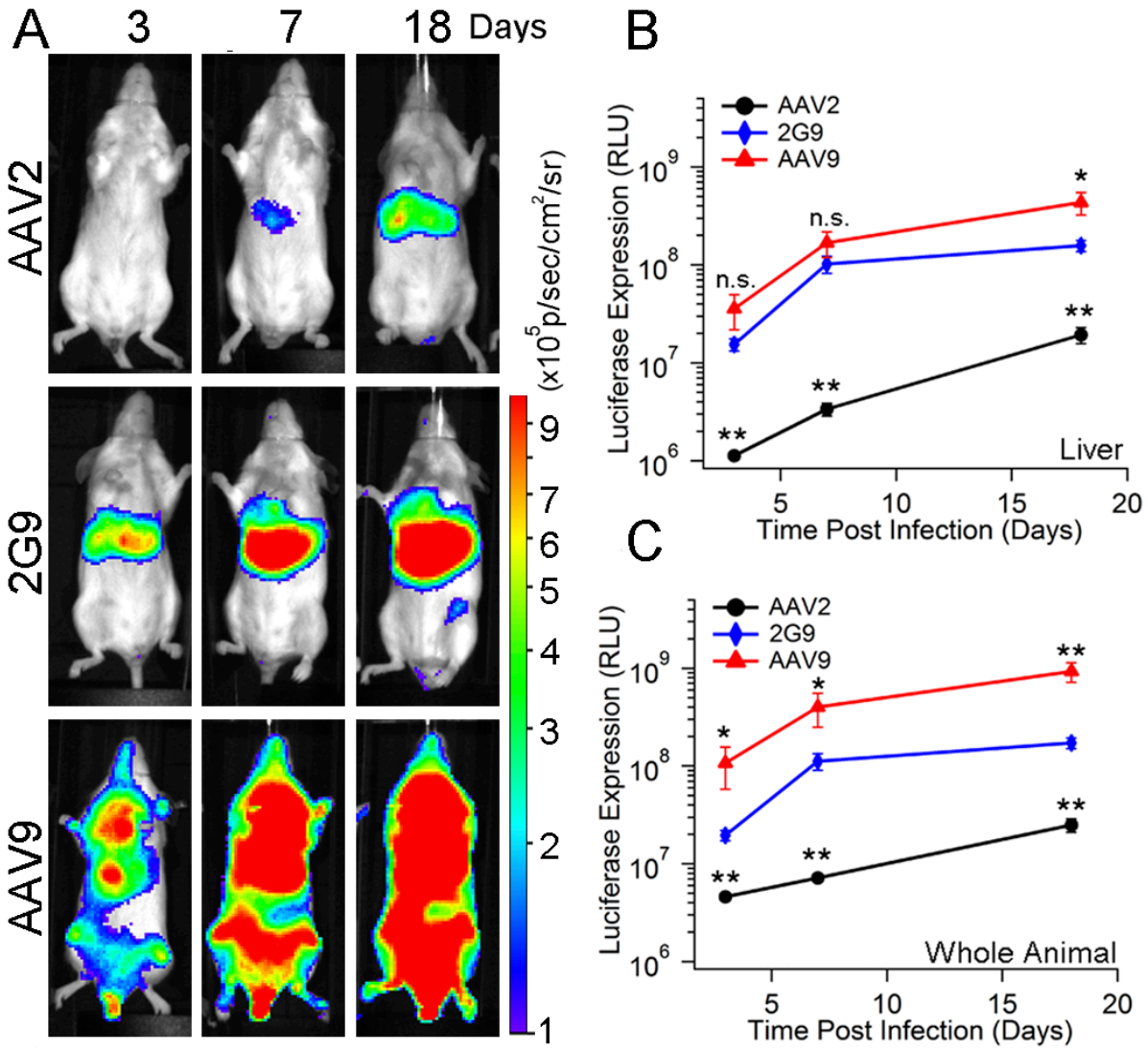

Figure 24. AAV2G9 mediates rapid onset and enhanced transgene expression *in vivo*. (A) *In vivo* transgene expression kinetics of AAV2, AAV 2G9, and AAV9 vectors packaging CBA-luciferase transgene cassette. BALB/c mice (n=4) were administered AAV vectors at a dose of $1x10^{11}$ vg/animal through the tail vein and bioluminescent images collected at 3, 7, and 18 days post-injection using an Xenogen® Lumina imaging system. Representative live animal images are shown with bioluminescence on a rainbow colored scale ($1x10^5$-$1x10^6$ photons/second/cm^2/steradian). AAV2G9 maintains the hepatic tropism of AAV2, but demonstrates a more rapid and robust luciferase signal than both parental AAV strains. **(B and C)** Quantitation of the kinetics of light signal output (expressed as photons/second/cm^2/steradian) was performed by marking regions of interest (ROIs) around images of the **(B)** liver region and **(C)** entire animals obtained at different time intervals (n=4). Statistical significance was assessed using the one-tailed Student's *t*-test (n.s., not significant; * $p < 0.05$; ** $p < 0.01$).

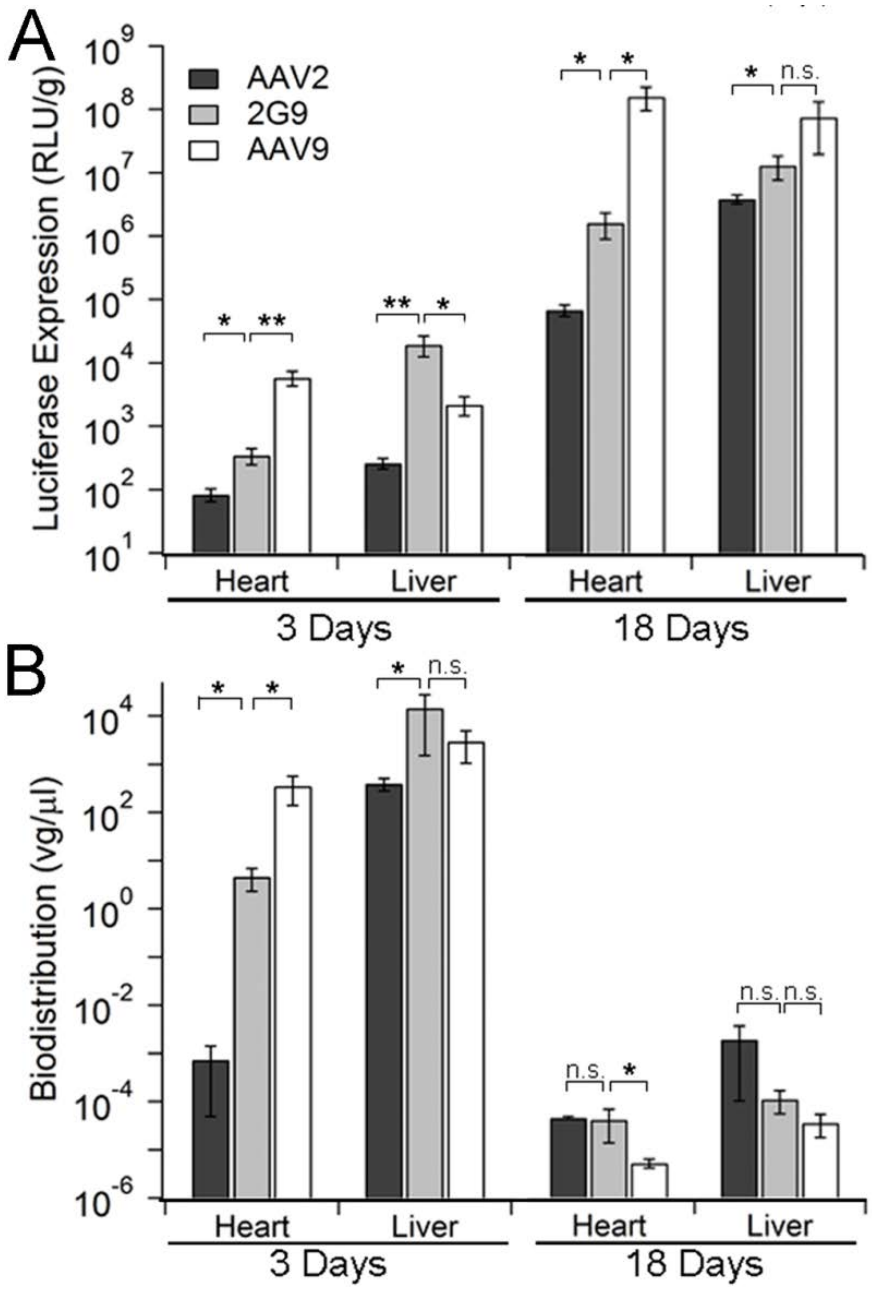

Figure 25. Quantification of transgene expression and biodistribution profiles of AAV2G9 in mice. (A) Quantitation of luciferase transgene expression from heart and liver tissue lysates of AAV2 (black), AAV2G9 (grey), or AAV9 (white) treated animals at days 3 and 18 (n=4). **(B)** Biodistribution of vector genomes in liver and heart lysates obtained from BALB/c mice administered with AAV2 (black), AAV2G9 (grey), or AAV9 (white) at days 3 and 18 (n=4). At indicated time points, host genomic DNA and viral genomes were isolated from tissue lysates and quantified using qPCR with primer sets specific to mouse lamin gene and luciferase transgene. Results are presented as mean ± s.e.m. (n=4). Statistical significance was assessed using the one-tailed Student's *t*-test (n.s., not significant; * $p < 0.05$; ** $p < 0.01$).

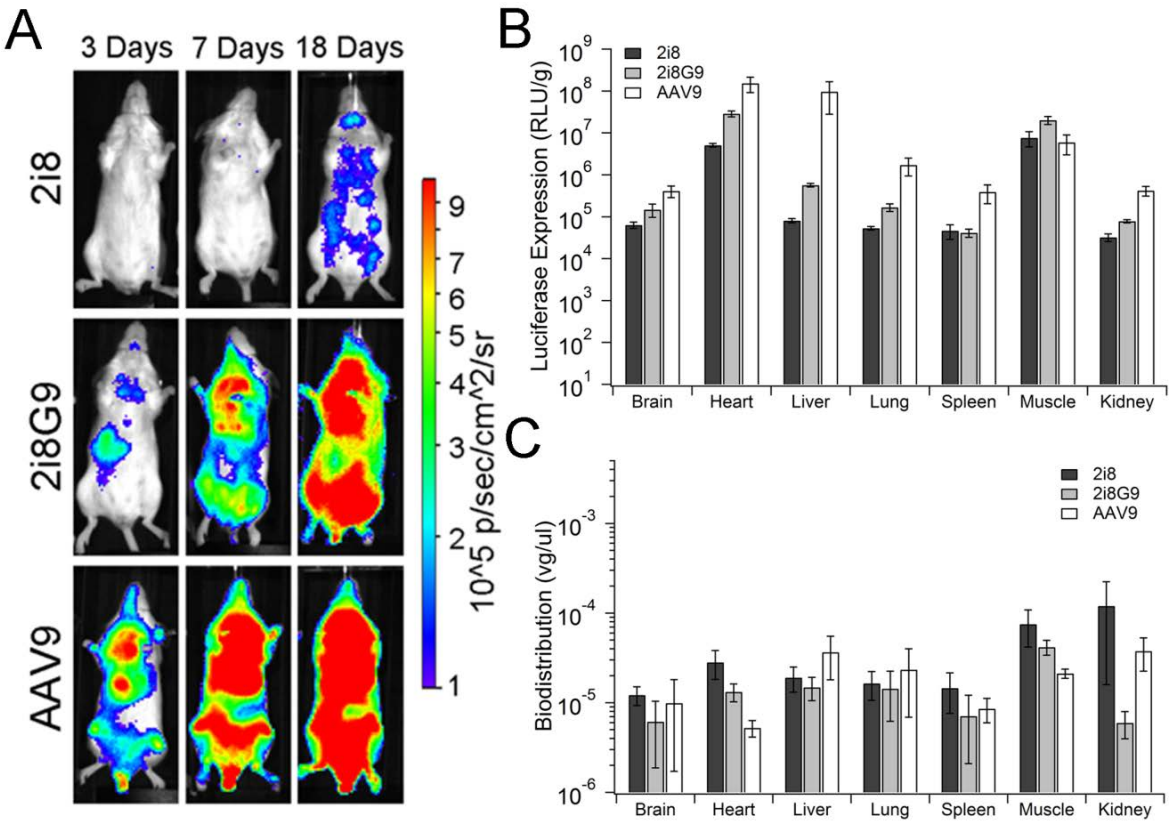

Figure 26. *In vivo* transgene expression kinetics of AAV2i8, 2i8G9, and AAV9 vectors packaging CBA-luciferase transgene cassette. (A) BALB/c mice (n=4) were administered AAV vectors at a dose of $1x10^{11}$ vg/animal through the tail vein and bioluminescent images collected at 3, 7, and 18 days post-injection using an Xenogen® Lumina imaging system. Representative live animal images are shown with bioluminescence expressed on a rainbow colored scale (10^5-10^6 photons/second/cm^2/steradian). **(B)** Luciferase transgene expression levels in tissue lysates obtained from **(A)** were quantified at 18 days pi. Data represents mean ± s.e.m. (n=4). **(C)** The biodistribution of vector genome copies of AAV2i8, 2i8G9, and AAV9 at 18 days pi were analyzed using qPCR. DNA contents were extracted from tissue lysates in (A). Luciferase transgene copy numbers in each tissue were normalized to the corresponding copy numbers of mLam housekeeping gene. Data represents mean ± s.e.m. (n=4).

CHAPTER 5

Synopsis and Future Directions

5.1 Summary

The past thirty years witnessed the accelerating exploration of AAV as gene therapy vectors. Encouraging results from preclinical and clinical studies call for a comprehensive understanding of AAV biology as well as expedited generation of novel AAV-based vectors. In this dissertation, we sought to unravel the mysterious systemic phenotype of AAV9 by gaining insights into the first step of AAV9 infection---glycan receptor interactions. The identification of *N*-linked terminal galactose as the primary receptor of AAV9, in conjunction with deciphering the structure-function relationship of a random AAV9 library, provides important avenue for understanding and application of AAV9 in the clinical frontiers [52,95]. The *in vivo* studies of how glycan-virus interactions affect the systemic and hepatic tropism of AAV9 expanded the interface between glycobiology and virology. In corollary with previous studies on influenza viruses, HCV, and many others, our results highlight the functional role of glycan receptors during viral infections in animal models. We also built the foundation to fine-tune and rationally design viral vectors with desired phenotypes by attributing virus tropism and blood circulation profiles partially to the binding avidity between viruses and their cognate glycans receptors. In addition, identification of key residues involved in galactose recognition enabled the molecular grafting of galactose binding footprint onto a series of well-studied AAV serotypes. The modularity of glycan

binding motif and flexibility of AAV backbone would further inspire the engineering of designer viral toolkits for future research and therapeutic advancement.

Nevertheless, AAV vector development sees obstacles to attenuate safety concerns and meet higher expectations of transduction efficiency and specificity. Multiple aspects of AAV biology are yet to be elucidated.

5.2 Glycan Receptors of AAV Serotypes

In Chapter 2 of this dissertation, the glycan receptor for AAV9 was identified using a battery of glycobiology, biochemistry, biophysical assays. Enzymatic digestion of cell surface glycoconjugates enhances AAV9 transduction efficiencies on multiple cell lines by exposing the penultimate Gal moiety, the glycan receptor for AAV9. This result was counter-intuitive. In case of HS- or SA-using viruses, heparinase or sialidase pre-treatment of cells prior to viral infection abrogates their infectivity due to the unavailability of attachment factors. However, as the first virus discovered to bind to terminal galactose, AAV9 transduces host 10-1000-fold more efficiently. This enhancement of infectivity was also seen with well-differentiated ciliated human airway epithelia cells and the murine nasal airways. A temporary, moderate increase of the number of receptors available allows virus to engage productive infection.

To test whether sialidase could increase AAV9 transduction in situations similar to cell culture, we investigated the effects of combined injection of sialidase and AAV9 vectors with localized injections. The first trial was carried out in the context of the eye. The eye is spatially restricted, compared to other major organs where lymphatic and circulation systems are more readily available. Nevertheless, eye-related diseases have become a focus for gene

therapy research in the past few years, including Leber's congenital amaurosis, color blindness, and macular degeneration [152]. As shown in **Figure 27**, 1 ul of the mixture of sialidase and AAV9-CBA-Luciferase were intravitreously injected into the left eyes of BALB/c mice, while the right eyes were injected with PBS and AAV9. At 5 days pi, significant difference of gene expression levels was observed (**Figure 27A**). To our surprise, this disparate luciferase transgene expression maintained for 5 months pi at our last time point (**Figure 27B**). Dissection of these eyes revealed that the scleras were not transduced efficiently by AAV9 regardless of sialidase treatment, while sialidase co-injection enhances AAV9 transduction in retina by ~50-fold (**Figure 27C**).

Another site of localized injection we tested was the joint space. Briefly, a mixture of sialidase or PBS (mock) and 1e9 vg of AAV9-CBA-Luciferase was injected into the joint space of the left leg on mouse. Two weeks pi, mice were sacrificed to harvest tissue samples from liver, leg muscle and the joint. According to luciferase assays on tissue lysates, the transgene expression in the joint pretreated with sialidase is ~500-fold higher than the one treated with PBS. Correspondingly, qPCR analysis also demonstrates an accumulation of AAV9 in sialidase treated joints, 40-fold higher than that in PBS treated ones. These results are corroborated with live-animal images shown in **Figure 27D**. In conclusion, results from murine nasal airway, ciliated HAE, murine retina, and joint support the notion that recombinant sialidase can be used as an adjuvant to enhance AAV9 transduction.

Despite our focus on sialylated glycans, heparan sulfates, and galactose, the other structures and compositions of glycoconjugates are extremely complicated. In this dissertation, glycan specific lectins were applied to discern the glycan pattern alterations. More accurate methods include mass spectrometry, glycan microarray, as well as radioactive

labeling. With limited quantification and verification technologies available, manipulating and monitoring cell surface glycans has been very challenging. Given the intertwining networks of glycan biosynthesis in all three domains of life, the artificial biosynthesis of any given glycan structures is still under development [223]. All the previous studies of the glycan receptors for pathogens are mainly focused on the terminal moieties of the long and complex glycoconjugates chains. However, the side chains, sulfate groups, the repeating disaccharide sequences might all play a role in glycan-protein interactions.

From a series of mutant CHO-K1 cell lines, we started our endeavor to investigate how sulfation on glycoproteins and glycolipid affect AAV transduction. PgsA-745 is a CHO-K1 mutant defective in GAG synthesis due to mutations in UDP-D-xylose:serine-1,3-D-xylosyltransferase. PgsD-677 is only unable to synthesis heparan sulfate because of the deficiency of both *N*-acetylglucosaminyltransferase and glucuronyltransferase. Another CHO-K1 mutant is pgsE-606 that produces under-sulfated forms of heparan sulfate due to the deficiency in *N*-sulfotransferase [224-227]. Furthermore, the Esko group in UCSD developed two pgsA-745 derivatives pgsA-745-t26 and pgsA-745-t489, which are highly sulfated at the same sites as sialic acids [173]. The underlying mechanism of this over-sulfated pgsA-745 is that the enzyme that transfer phosphoadenosine-5'-phosphosulfate to Galβ1-4(3)GlcNAcβ-*O*-naphthalenemethanol is increased. In addition, a moderate concentration of sodium chlorate in culturing media could inhibit the incorporation of sulfate groups in glycoconjugates chains in many cell lines. Other than sulfation, fucosylation might also contribute to glycan-protein interactions as most polysaccharides are decorated with fucose groups. Pamela Stanley and coleagues developed CHO-LEC11 and LEC12 that does not express α-1,3-fucosyltransferase. Lec13, the hyper fucosylated CHO variant, was generated by transfecting plasmids

containing α-1,3-fucosyltransferase into CHO cells [228]. Using an arsenal of cell lines with mutations or deficiencies in glycan synthesis, we are able to interrogate the impact of sulfation on parvovirus infection. Although no significant phenomena were observed with AAV1-9 on these cell lines compared to their parental cells (Shen Shen and Aravind Asokan, unpublished data), this collection of CHOs laid a roadmap for studies on emerging naturally occurring AAV and other parvoviruses.

Our previous experience with pharmacokinetics analysis and hemagglutinin assay prompt the speculation that glycans on the erythrocytes might play a role in the transportation of AAV in their target tissues. The most dominant glycosylation forms on erythrocyte surfaces are the ABH(O) blood group antigens and the Lewis antigens Le^a and Le^b. ABH(O) blood group antigens present *N*-acetyllactosamines on *O*-linked glycans or glycolipid on the surface of erythorcytes and many other tissues. For example, they are expressed in the vascular endothelium and the epithelial of some salivary and exocrine glands. They can also be excreted into blood circulation. A hallmark of ABH(O) blood group antigens is the fucose groups transferred by α-1,2-fucosyltransferases. Lewis antigens Le^a and Le^b carry α-1,3- and α-1,4-fucose moieties. Unlike ABH(O) antigens, they are not synthesized by erythrocytes. Instead, Le^a and Le^b antigens are expressed on some epithelia and release to body fluids, where they are passively incorporated into erythrocytes. With a high sialylation rate and an enormous population on erythrocytes, these blood type antigens might contribute to the circulation of AAV vectors *in vivo* and serve as vehicles for tissue targeting.

One intriguing characteristics of glycoconjugates in mammalian systems is its species-specificity. In 1962, Emile Zuckerkandl and Linus Pauling first postulated a hypothesis called “Molecular Clock Hypothesis” [229]. In this hypothesis, the occurrence and

disappearance of molecules is viewed as part of evolution. The adaptation of the hemagglutinnin subunit of Influenza A from avian-like α2,3-SA to α2,6-SA specifically expressed in the small airway epithelia of human lung was attributed to the antigen drift of Influenza A. Therefore, this process is also considered an evolution of the HA subunit. Another species-specific glycol-moiety is *N*-Glycolyl-SA (Neu5Gc), which only exist in small quantity in human beings [209]. Does any parvovirus have the specificity towards a species-specific glycol-moiety? How does this species-specific glycosylation pattern affect the translation of preclinical results generated from mice, dogs, nonhuman primates to be applied in curing human diseases? Many questions are yet to be answered.

5.3 Understanding the Roles of Glycan in Animal Models

A cornerstone of this dissertation is the analysis of glycosylation pattern on murine tissues using a series of lectins. Viral tropism is partially attributed to the differential distribution of glycan structures. Sialidase and heparinase (Shen Shen and Aravind Asokan, unpublished data) were injected systemically or locally to hydrolyze the terminal glycan moieties in animal. However, the effects of these enzymatic removals of glycan termini only persist for a short span of time. A great number of sialyltransferases, sialyltransporters, sulfotransferases, as well as polysaccharides obtained externally (i.e. through diets) reconstitute the lost glycan moieties back to the end of glycoconjugates chains. Therefore, mouse models with knockdowns of the key enzymes in glycan biosynthesis pathways would significantly elucidate the role of each glycan in a long-term observation.

Given that protein and lipid glycosylation takes place in the Golgi compartments, it is compelling to speculate that the glycol-synthesis intermediates in cytoplasm could impact the

trafficking of AAV viral particles. This effect can be a double-edged sword, as it can affect the uptake and cytoplastic processing of AAV virions into the nuclei, as well as the packaging and release of newly synthesized AAV particles out of the nuclei. One technical hurdle to answering this question is the precise marking and analysis of glycan structures in each organelle and subcellular fraction. Quantitative imaging technology would be instrumental in studying this aspect of AAV biology, but the accurate labeling of individual glycans still remains challenging for now.

5.4 Beyond the Glycans

A central theme of this dissertation is to emphasize the significant function of glycan receptors in determining viral tropism. Results from Chapters 2-4 and Appendix I supports this notion in the context of AAV biology. However, the roles of the secondary receptor in viral transduction and tropism can never be underestimated. Over the past ten years, the list of secondary receptors for AAV serotypes has expanded, but the function of these receptors *in vivo* are yet to be unraveled. How does the distribution of PDGFR impact on the respiratory tropism of AAV5? Why do six amino acid changes allow AAV6 to bind EGFR, but not AAV1? How does the affinity towards EGFR of AAV6 contribute to its different tropism compared to AAV1? Although more secondary receptors for emerging AAV strains are to be discovered, the roles of these known receptors in AAV life cycles remain elusive.

In attempt to dissect the systemic tropism of AAV9, a conserved integrin binding motif on AAV9 capsid was brought to our attention by sequence alignment. As shown in **Figure 29A**, the NGR motif common in all integrin α5β1 ligands are present on AAV1, AAV2, AAV3a, AAV3b, and AAV6-10. Homology simulation shows that this domain

resides on the concaved region immediately outside of the three-fold axes of symmetry (**Figure 29B and Figure 29C**). Site-directed mutagenesis on the arginine residue (R514A) yields mutant 9NGA, that is 10-100-fold less efficient in transducing Lec2 cells compared to the parental AAV9 (**Figure 29D**). Confocal laser scanning microscopy also suggests that this 9NGA mutant is deficient in getting internalized into cells with high efficacy. At 0 hr pi, similar numbers of virions can be visualized on AAV9 and 9NGA infected Lec2 cells. In contrary, at 1 hr pi, a good amount of internalized viral particles can only be observed with AAV9-infected Lec2, not 9NGA-infected ones (**Figure 29E**). These results and quantification of internalized virions on Lec2 by qPCR (Shen Shen and Aravind Asokan, unpublished data) strongly suggest that 9NGA is an internalization deficient mutant. How does this loss-of-function affect AAV9 transduction in mice? Intravenous injection of AAV9 and 9NGA into BALB/c mice was followed by live-animal bioluminescence imaging to monitor their transgene expression kinetics (**Figure 29E)**. The systemic tropism of AAV9 was absent in case of 9NGA, which demonstrates a decreased transduction efficiency in all major organs. As a corollary, this observation was substantiated by quantifying the luciferase transgene expression profiles in tissue lysates and the biodistribution of viral particles using qPCR (Shen Shen and Aravind Asokan, unpublished data). Taken together, our results show that the secondary receptor for AAV9 infection also plays a pivotal role in determining AAV9 tropism *in vivo*. How secondary receptors are further involved in intracellular trafficking and uncoating of AAV virions remains elusive. Whether this theory applies to other members of the AAV family would be of great importance for future studies.

5.5 Rational Engineering of AAV Vectors

Although the collection of naturally occurring AAV strains are expanding, discovery of engineered AAV variants with higher transduction efficiency and improved immunogenicity is moving forward rapidly under the limelight. The first rational engineering of AAV capsid in pursuit of variants with novel properties was achieved by swapping a hexamer-peptide from the hepatotropic AAV8 to the partial heparan sulfate binding motif on AAV2 [143]. Chapter 4 of this dissertation focused on the synthesis of a dual-glycan binding AAV variant by grafting the Gal binding footprint and flanking residues from AAV9 to the corresponding region on AAV2 capsid. Our data first confirmed that the resulting 2G9 chimera utilizes Gal and HS interchangeably by lectin and HS inhibition assay. *In vitro* and *in vivo* studies further described the transduction profiles and biodistribution patterns of this chimera. The phenotype displayed by 2G9 is the combination of the hepatotropism of AAV2 and the transduction efficiency of AAV9. More interestingly, 2G9 also demonstrates an early onset of transgene expression. Furthermore, incorporation of the G9 footprint onto several other AAV serotypes (AAV1, AAV3, AAV6, AAV8, and AAV2i8) yielded similar results (Shen Shen and Aravind Asokan, unpublished data). This work not only demonstrated the modularity of Gal binding domain, but also laid the foundation of a novel strategy for AAV vector design.

Our previous success with the G9 series suggests that a comprehensive understanding of structure-function relation of AAV strains is the prerequisite for rational reengineering of AAV vectors. During our endeavors to dissect the GH loops, three AAV9 mutants were constructed by swapping parts of the inner loop from AAV1 to the AAV9 backbone (**Figure 30A-C**). The first region swapped from AAV1 (the iA stretch: 568-ATNPVATER

FGTVAVNF-584) starts from the N-terminal end of the inner loop. The resulting mutant was dubbed 9i1A (i stands for “inner loop”). The second stretch of amino acids swapped includes the iA stretch and 15 more residues to its C-terminal (the iB stretch: 568-ATNPVATERFGT VAVNFQSSSTDPA-592). The third variant, 9i1C has the entire inner loop of AAV1 swapped onto AAV9 (the iC stretch: 568-ATNPVATERFGTVAVNFQSSSTDPATGDVHA MGA-601). **Figure 30A-C** depicts the altered domains in the context of AAV9 trimeric assemblies. *In vitro* transduction assay on Cos-1 cells (**Figure 30D**) shows a 10-fold enhancement of transduction efficiency of 9i1A compared to AAV9 regardless of sialidase treatment. Binding assays corroborated with this observation by showing a 100-1000-fold increase of binding ability of 9i1A compared to AAV9 on Cos-1 (**Figure 30E**). To our surprise, more components from the inner loops reduces the transduction efficiency of 9i1A to the level lower than the parental AAV9, suggesting more severe perturbations to the global conformation on viral capsids. Animal studies also show the consistent enhancement of 9i1A transduction profiles. At 3 days pi, 9i1A surpasses both AAV1 and AAV9 in transducing major organs (brain, heart, lung, and muscle) by 10-100-fold (**Figure 30G**). The biodistribution analysis of AAV9 and 9i1A in mice gave similar results, except a 5-fold increase of viral genome accumulation in liver which is absent in transduction assay (**Figure 30H**). The molecular mechanism of the superior infectivity of 9i1A is still under investigation. However, this study provides us with an AAV9 variant with startling transduction efficiency in both *in vitro* and *in vivo* settings. In addition, these results also imply the structural functions and the inadaptability of the inner loops at the three-fold axes of symmetry on AAV capsids. For future studies, more comprehensive mutagenesis attempts and structural approaches could be employed to identify the specific amino acids key to

AAV9 life cycle. In conjunction with G mutants, these results further testified the feasibility of incorporating multifunctionality into AAV capsid for a wide range of applications.

In contrast with the rational vector design, directed evolution of AAV libraries starts from random mutagenesis or random mutations onto AAV capsid. As described in Chapter 1, directed evolution of AAV2 with random mutations or AAV shuffling libraries generated AAV variants with desired transduction specificity towards a given cell type or organ. We approach this directed evolution strategy from a different angle. Herein, we aim to evolve AAV9-based variants that bind to alternate glycan receptors. The synthetic dual-glycan binding 2G9 demonstrated the modularity of the Gal binding motif on AAV9 capsid. It is the minimum cluster of amino acid that is necessary to convert any AAV serotype binding to Gal. Instead of aiming to generate AAV9 variants targeting specific cell types or organs, we pursuit the idea of switching AAV9 to a HS binding isolate by introducing random mutations on the Gal binding motif.

As shown in **Figure 31A-C**, random mutations were introduced onto three regions harboring residues critical for Gal-AAV interactions by NNN-induced site-directed mutagenesis: (i) 262-STSGGS-268; (ii) 498-NSEFAWP-504; (iii) 590-QAQTGW-595. These stretches cover the tips of the inner loops, middle loops, and outer loops near the three-fold axes of symmetry on AAV9 capsids. Three AAV9 libraries, each with one stretch mutated, were combined to subject to elution on heparin columns. AAV2 serves as a control. The HS-binding AAV2 was eluted from heparin columns at a NaCl concentration of 450-500 mM (**Figure 31D**). Therefore, after flowing through the combined AAV9 triple library through a heparin column, only the eluate at 450 mM were collected and amplified on HEK293 cells with the presence of Ad5. Four days later, viruses purified from this set of

HEK293 cells went through another selection on a heparin column. Again, only the eluate with 450 mM NaCl concentration was amplified on the next set of HEK293 cells. After one more "evolution" on the heparin column, the eluate at 450 mM NaCl concentration was preserved. As shown in **Figure 31E**, the initial AAV9 triple library has a mixed population of AAV9 variants with a wide range of HS affinities. QPCR of vg in each eluate with a varying NaCl concentration shows similar quantity of viral particles. As the "evolution" on heparin column proceeds, the population of AAV9 variants left in the triple library concentrates in the 450 mM NaCl fraction **Figure 31F and 31G**. This result indicates a shift of Gal preference towards a HS preference of those AAV9 mutants. Isolation of individual AAV9 variant from the preserved collection after the 3rd heparin column is ongoing. Once individual variants are identified, a series of *in vitro* and *in vivo* assays will be carried out to understand the phenotypes of these HS-binding AAV9 mutants.

Considering the large size of this AAV9 triple library (20^6+20^7+20^6-3) in theory, a myriad of evolution or selection strategies can be readily applied in sarch for AAV9 variants with novel features. One possible strategy is to evolve this triple library back and forth on ferret airway epithelial cells and human airway epithelial cells. The aim of this strategy is to evolve AAV9 variants highly efficient in transducing airways for treating respiratory diseases, such as asthma and cystic fibrosis. Studies led by Yoshihiro Kawaoka and Ron Fouchier [219,220] generated H5N1 influenza mutants transmissible within ferret populations by acquired mutations during passages of airborne transmission. Successes in H5N1 influenza suggest a high likelihood of obtaining airway-tropicAAV9 mutants from this triple library.

5.6 Final Remarks

Interaction with cell surface glycan receptors is the initial step of infection for most pathogens. AAV serotypes utilize three major categories of glycans to attach to host cells, sialylated glycans, heparan sulfates, and galactoses. In this dissertation, the terminal N-linked galactose was identified as the primary receptor for AAV9. Enhanced galactose binding ability by sialidase treatment leads to increased cell surface binding and transduction. Modulating this binding avidity between galactose and AAV9 *in vivo* by sialidase administration or mutation alters the pharmacokinetics profiles and systemic tropism of AAV9. Herein, we established the link between the glycan binding avidity of AAV and their respective phenotypes in animal. These results contribute to our understanding of the cross-species differences of AAV tropism in general and aid the rational design and engineering of novel AAV variants with desired tissue-targeting/detargetting profiles. Further incorporation of Gal binding footprint demonstrates a new avenue to obtain vectors for enhance gene transfer. Although deciphering the exquisite infection pathways of AAV remains challenging, our study shed light on the AAV biology field and the utilization of AAV as gene therapy vectors to its full potential.

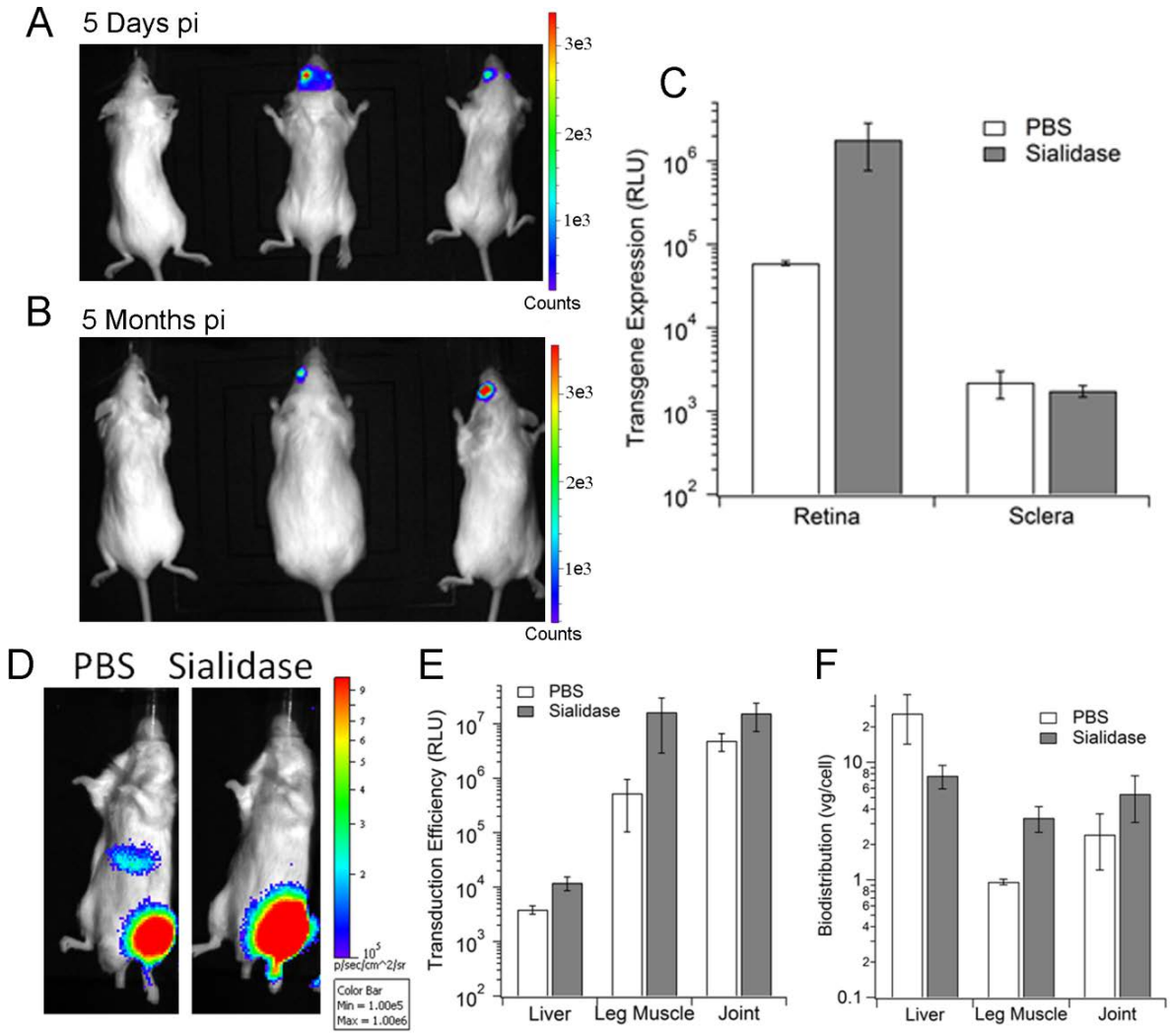

Figure 27. Sialidase can be used as adjuvant to enhance AAV9-mediated gene transfer. **(A-B)** PBS (left) or 1e9 vg of AAV9-CBA-Luciferase vectors were mixed with PBS (right eyes in the middle and right mice), sialidase (left eyes in the middle and right mice) were injected into female BALB/c mice intravitreally. To monitor luciferase transgene expression in animals, D-luciferin was administered through intraperitoneally at **(A)** 5 dpi and **(B)** 150 dpi. **(C)** Eyeballs harvested from mice at 150 dpi were dissected. Luciferase expression was measured from lysates of retina or sclera. Error bars represent s.e.m. (n=4). **(D)** 1e9 vg of AAV9-CBA-Luciferase vectors were injected into the joint space of the left leg in mice prior to PBS or sialidase through the same route. At 2 weeks post injection, luciferase expression profiles were recorded. Rainbow color represents quantification of RLU. Experiments were carried out with n=4. **(E)** Two weeks post AAV9 injection in the joint space, BALB/c mice were sacrificed to harvest tissues. Luciferase expression from liver, joint, and leg muscle were carried out. Error bars represent s.e.m. (n=4). **(F)** Viral genome copy numbers in liver, leg muscle, and joint from **(E)** were quantified using qPCR. Data are presented as mean±s.e.m. (n=4).

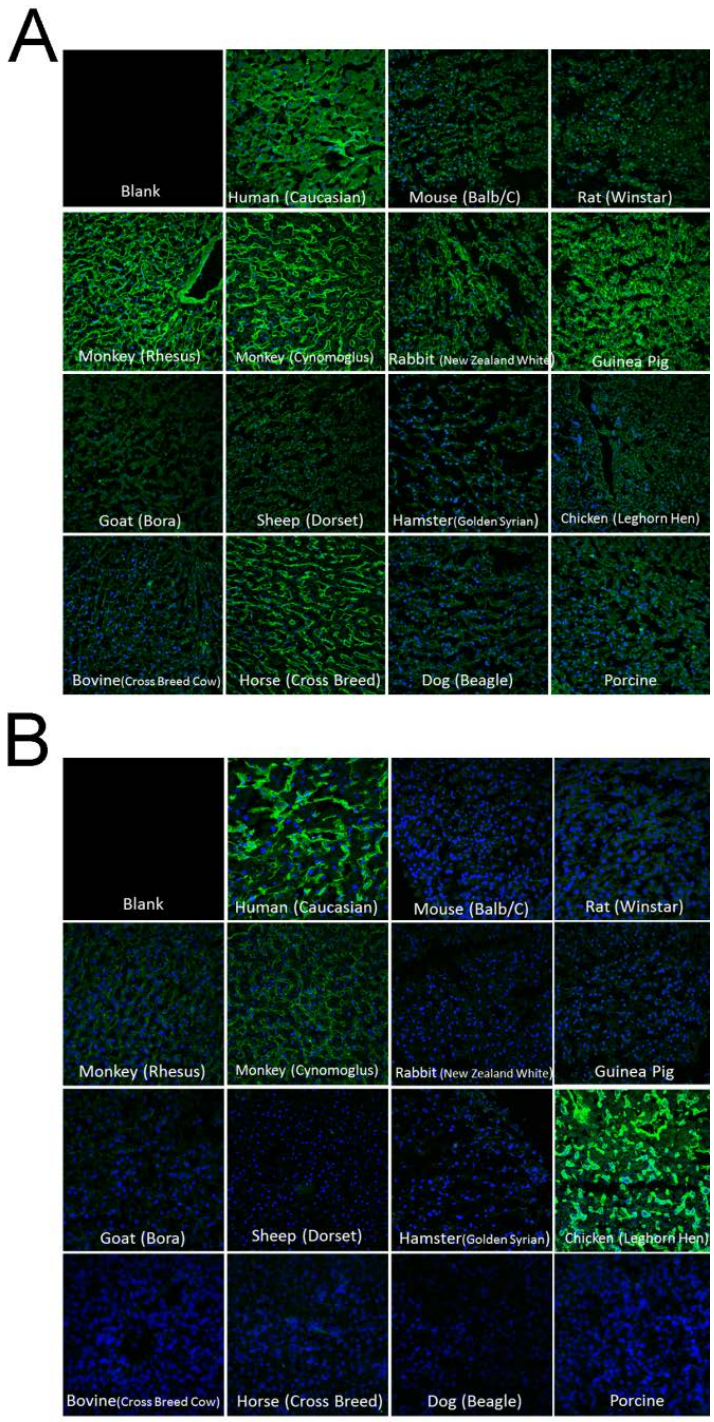

Figure 28. Lectin staining of a tissue array from multiple species. (A) FITC-MALI staining and **(B)** FITC-ECL staining of a tissue array of liver collected from 15 species (Amsbio #T6134149). 50 ul/ml FITC-MALI in 1xPBS were incubated with this slide of tissue array for 2 hours in rt prior to three washes with cold-PBS and covered with mounting media containing DAPI. Fluorescence images of individual tissue were obtained by imaging with a Zeiss 710 laser scanning confocal microscopy and an 40x oil objective at zoom 0.6. Representative images are shown here.

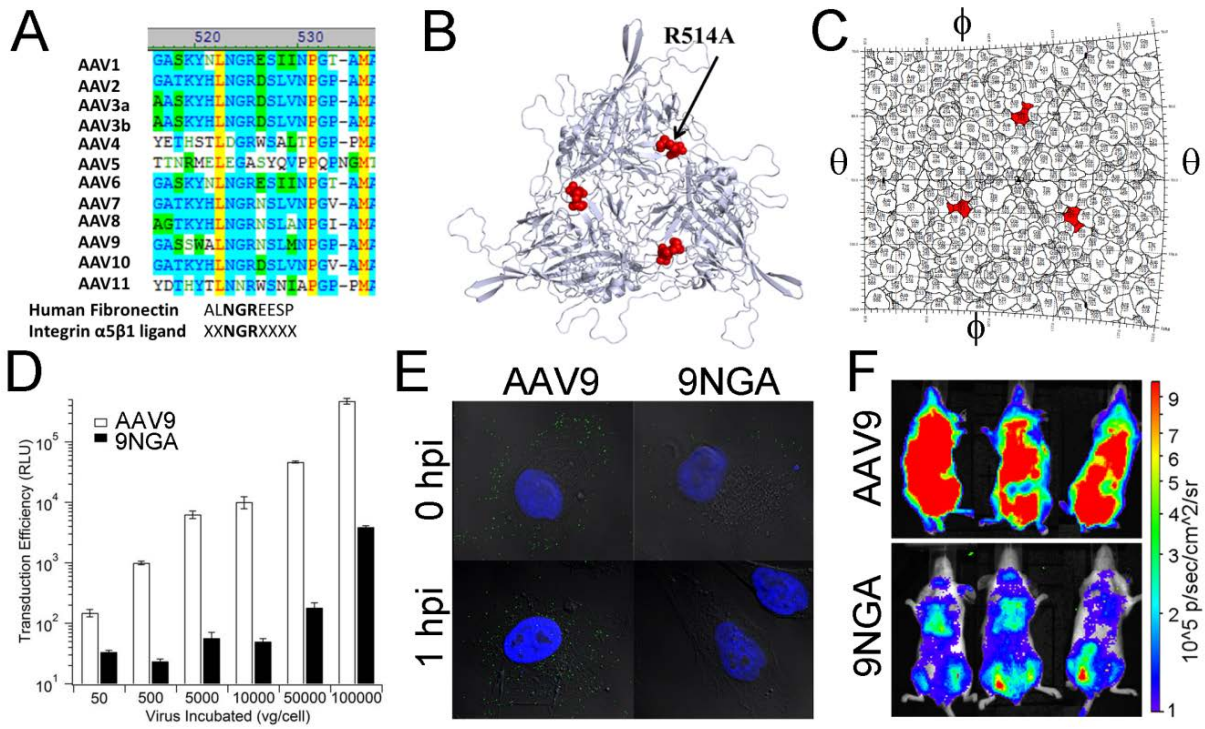

Figure 29. Integrin functions as co-receptors for AAV9. (A) Amino acid sequence alignment of AAV1-11 and human fibronectin reveals a conserved integrin binding domain close to the three-fold axes of sysmmetry. **(B)** Cartoon rendering of trimeric assembly of 9NGA variant, with the point mutation R514A highlighted in red. **(C)** RIVEM roadmap model of the three-fold axes region on AAV9 with R514 marked in red. **(D)** AAV9 and 9NGA packaging CBA-Luciferase cassette transduce Lec2 cells with MOIs ranging from 50 to 100,000 vg/cell. Luciferase activity of cell lysates were obtained 24 hrs pi. Error bars represent s.e.m. (n=4). **(E)** AAV9 and 9NGA-CBA-Luciferase were added to pre-chilled CHO-Lec2 cells culture. 1 hr post infection, unbound virions were removed by three washes with ice-cold PBS prior to incubation in 37ºC for 0 or 1 hr. Cells were then fixed with 2% paraformaldehyde and penetrated with tween-20. Intact virions were detected with monoclonal antibody ADK9, followed by secondary antibody Dylight488 (green). DAPI stains nuclei. **(F)** AAV9- or 9NGA-CBA-Luciferase were injected into BALB/c mice through tail vein. Bioluminescence images of live animals were obtained at 2 weeks pi (n=4).

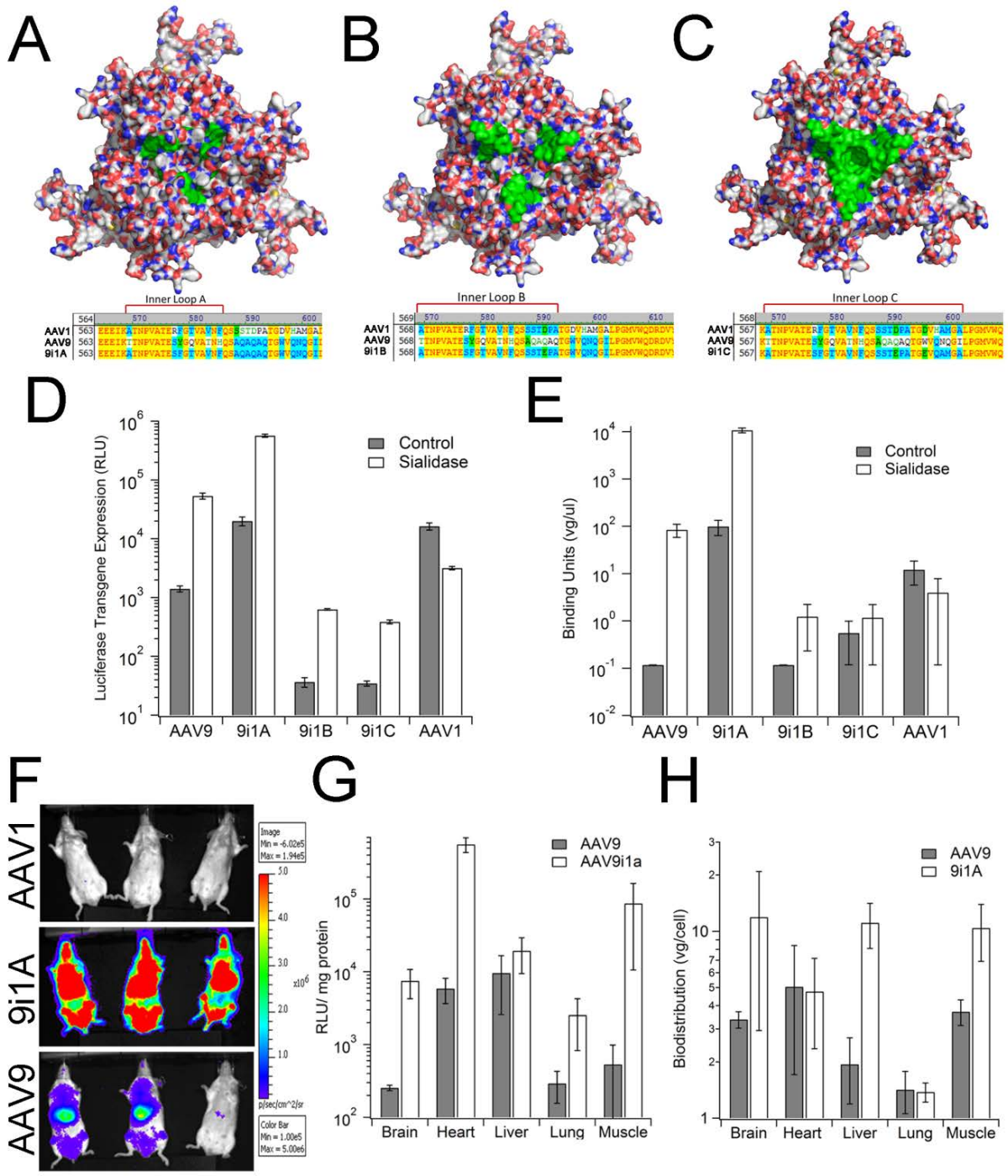

Figure 30. Swapping the inner loops from AAV1 onto AAV9 yields enhanced infectivity. **(A-C)** Surface models of the trimeric assemblies of **(A)** 9i1A, **(B)** 9i1B, and **(C)** 9i1C with AAV9 backbones colored as electrostatic potential (blue: positive charge, red: negative charge) and stretches from AAV1 highlighted in green (AAV9 VP1 numbering, 9i1A: 568-ATNPVATERFGTVAVNF-584; 9i1B: 568-ATNPVATERF GTVAVNFQSSS TDPA-592; 9i1C: 568-ATNPVATERFGTVAVNFQSSSTDPATGD VHAMGA-601). **(D)** Cos-1 cells

were pretreated with PBS or sialidase prior to infection with AAV9, 9i1A, 9i1B, 9i1C, and AAV1-CBA-luciferase at an MOI of 1000 vg/cell. 24 hrs pi, luciferase activities of cell lysates were obtained and quantified. Data represent mean ± s.e.m. (n=3). **(E)** Cos-1 cells pretreated with PBS or sialidase were infected with AAV9, AAV1 and chimeras at an MOI of 1000 vg/cell. After 1 hr incubation in 4 ºC, unbound virions were removed by three washes with ice-cold PBS. The amounts of virions bound to Cos-1 cell surface were quantified using qPCR. Data represent mean ± s.e.m. (n=5). **(F)** BALB/c mice were injected with AAV1, 9i1A, or AAV9 at a dose of 5e10 vg/mouse. 3 days pi, live-animal bioluminescence images were obtained using an Xenogen® Lumina imaging system. Live animal images are shown with bioluminescence expressed on a rainbow colored scale (10^5-10^6 photons/second/cm^2/steradian). **(G)** 1 month pi, mice from **(F)** were sacrificed to harvest tissue lysates from major organs. Luciferase transgene expression levels were measured and shown as mean ± s.e.m. (n=3). **(H)** Viral genome copy numbers from tissue lysates collected from mice in (F) at 1 month pi were quantified using qPCR. Biodistribution of viral particles were shown as number of viral genome copies normalized by housekeeping gene mLam. Data represent mean ± s.e.m. (n=3).

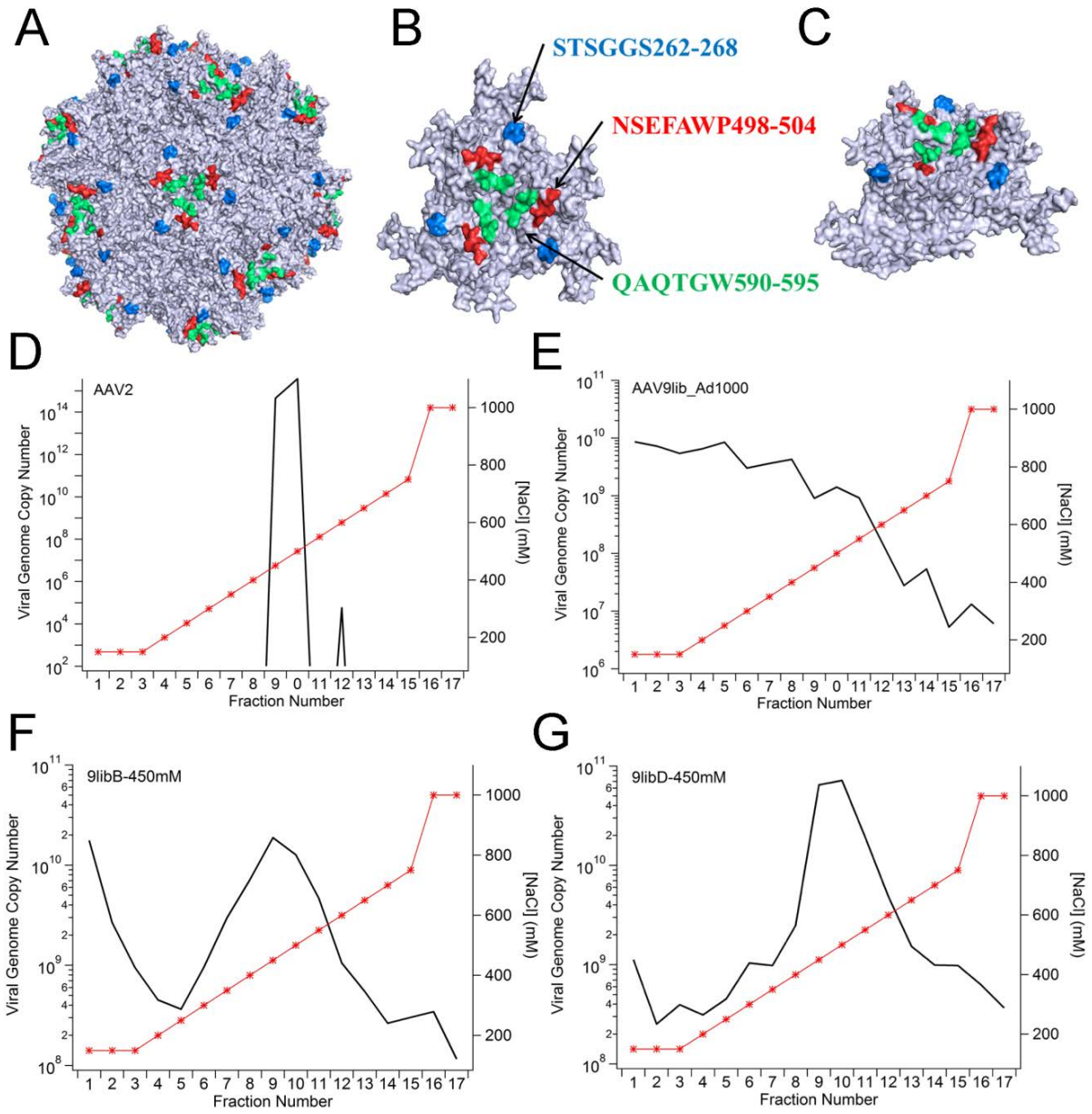

Figure 31. Construction and directed evolution of AAV9 libraries. (A) Surface model of an AAV9 full capsid with sites of random mutations (262-STSGGS-268, 498-NSEFAWP-504, 590-QAQTGW-595) colored in blue, red, and green, respectively. **(B and C)** Visualization of surface models of an AAV9 trimer from **(B)** above and **(C)** side of the three-fold axes of symmetry. Random mutations were introduced into three stretches highlighted in blue, red, and green separately to generate three AAV9 libraries. **(D)** Elution profiles of AAV2 from heparin columns. The red line with markers represents NaCl concentrations in elution buffer. The solid black line denotes quantity of AAV2 total vg in each eluate. **(E-G)** Elution profiles of the combined triple AAV9 library modeled in **(A-C)** on heparin columns. The combined triple AAV9 library was fractionized from a heparin column **(E)**. The eluate at 450 mM NaCl concentration was amplified on HEK293 with Adenovirus type 5, purified,

and eluted on a heparin column **(F)**. After another round of amplification, purification, and fractionation on heparin columns, majority of this combined triple AAV9 library demonstrates high heparin binding affinity **(G)**.

APPENDIX I

Sialic Acids Play a Multifunctional Role in Determining the Transduction Profile of Blood-Borne Adeno-Associated Virus Serotype 4[4]

AI.1 Overview

Adeno-associated virus serotype 4 (AAV4) is one of the most divergent strains amongst natural isolates of AAV characterized by its low sequence homology of viral proteins, distinct antigenicity, and selective cardiopulmonary tropism in mice after systemic administration. 2,3-linked sialylated *O*-glycans or mucins have been identified as primary receptors for cell surface attachment of AAV4 *in vitro*. However, little is known about the roles of sialic acids in determining AAV4 transduction *in vivo*. In the current study, Jacalin lectin staining reveals that *O*-linked sialic acids are expressed abundantly in cardiac and lung tissues. Enzymatic removal of sialic acids in BALB/c mice completely abrogates the cardiopulmonary tropism of AAV4 *in vivo* after systemic administration. These results were further corroborated using two loss-of-function AAV4 mutants obtained by screening a randomly mutated AAV4 library. Specifically, the two mutants, 4.18 (K492E, K503E, N585S) and 4.41 (M523I, G581D, Q583E), were defective in binding, uptake, and transduction in African green monkey kidney cells (CV-1) *in vitro*. As a result, these mutants demonstrated decreased transduction efficiency and vector genome copy numbers in murine heart and lung. *Ex vivo* virus distribution assay in blood and hemagglutination assay showed that both 4.18 and 4.41 bind to sialylated glycans on murine erythrocytes with lower

[4]This chapter includes the original manuscript in preparation: Shen, S., Troupes, N.A., Pulicherla, N., Asokan, A. *Sialic Acids Play a Multifunctional Role in Determining the Cardiopulmonary Tropism of Blood-borne Adeno-associated Virus Type 4*.

affinities. Therefore, both 4.18 and 4.41 were cleared faster from the blood circulation in comparison with the parental AAV4, similar to the pharmacokinetics profiles of AAV4 in sialidase-treated mice. In conclusion, these results indicate that sialylated glycans on erythrocytes and tissues play a multifunctional role in determining blood residence time, tissue uptake and transduction profile of AAV4 in mice.

AI.2 Introduction

Adeno-associated viruses (AAV) are small, non-enveloped, single-strand DNA viruses [37]. Several AAV serotypes have been developed into gene delivery vectors due to their diverse transduction profiles *in vivo* [152]. In addition to naturally occurring AAV strains, genetically modified AAV variants have been generated by DNA shuffling and directed evolution as well as rational and combinatorial mutagenesis techniques [152]. As the AAV toolkit continues to expand, a better understanding of AAV biology is essential to ensure vector safety and efficacy in the clinic.

Among more than one hundred isolates of AAV, AAV4 is particularly unique due to its divergent amino acid sequence of viral proteins (VPs), capsid topology [82,230], antigenicity, glycan receptor usage [46,47], and its systemic as well as CNS tropism in murine models [125,231]. The *Cap* gene product of AAV4 shares only 58% homology with that of AAV2 in comparison with higher sequence similarities between AAV2 and most other prominent AAV serotypes. The core structure of the AAV VP monomer is a seven-stranded beta sheet interconnected by highly dynamic loops [82]. The surface topology constituted by these highly dynamic loops on the exterior of AAV4 capsid is strikingly distinct when compared to other

AAV serotypes. This difference has been suggested as the underlying cause for the disparate properties of AAV4 in terms of receptor recognition, immunogenicity, and tissue tropism [212].

The first step in the AAV life cycle involves recognition of primary cell surface receptors by the viral capsid [37]. Currently known primary receptors for AAV serotypes include heparan sulfate for AAV2 [40], AAV3 [41,42], and AAV6 [43], galactose for AAV9 [52,53], and sialic acids for AAV1 [44], AAV4 [46,47], AAV5 [45], and AAV6 [44]. Amongst all currently known dependoviruses, AAV4 appears to be the only strain that binds to *O*-linked sialic acids/mucins. A particular interesting feature of AAV4 is the distinct transduction profiles after systemic administration or direct injection into central nervous system. AAV4 transduces heart and lung specifically after intravenous injection in mice [125]. If administered directly into the lateral ventricle, AAV4-mediated transgene expression accumulates in ependymal cells and astrocytes in the subventricular zone and rostral migratory stream (RMS) [231]. In addition, AAV4 also transduces specifically ependymal cells after intrastriatal injections [231]. This property of AAV4 rendered it an ideal candidate for gene delivery of insulin-like growth factor 1 (IGF-1) and vascular endothelial growth factor (VEGF) into mouse models of amyotrophic lateral sclerosis (ALS) [232].

Despite our growing knowledge of AAV4 in CNS, not much is understood pertaining to the biological mechanism behind its cardiopulmonary tropism. Meanwhile, reengineered strains of AAV have emerged for airway-specific transduction. For example, the pulmonary tropic AAV2.5T was discovered by directed evolution of combinatorial AAV libraries on well-differentiated Human Airway Epithelial Cells (HAE) [134]. In addition, DNA shuffling of AAV1, 6, and 9 in combination with directed evolution also resulted in two chimera AAV strains with higher transduction efficiency on *in vitro* HAE culture [141] Therefore, a thorough

understanding of AAV vectors in cardiopulmonary system also promotes the design and engineering of AAV vectors with enhanced transduction profiles.

In a recent study, we used AAV9 as model system to demonstrate that AAV-glycan receptor interactions are determining factors of hepatotropism and blood circulation profiles of AAV [51]. Specifically, higher binding avidity between AAV9 and its cognate glycan receptor drives its accumulation in liver and depletes viral particles from blood stream rapidly. On the contrary, lower binding avidity between AAV9 and galactoses detargets AAV9 from liver and leads to a longer blood retention time. In the current study, we hypothesized that *O*-linked sialylated glycans serve as a key determinant in AAV4 infectious pathway after cell surface attachment. To dissect the functional roles of sialylated glycans in AAV4 lifecycle, AAV4-sialic acid interactions were disrupted using two strategies: (1) enzymatically removal of sialic acids in mice, and (2) characterization of AAV4 mutants deficient in sialic acid binding *in vitro* and *in vivo*. Our data support the notion that sialylated glycans are indispensable for the cellular binding and uptake of AAV4 both *in vitro* and *in vivo*. These results also reveal that sialic acid-mediated interactions between AAV4 and erythrocytes are indispensable for the blood circulation and cardiopulmonary tropism of blood-borne AAV4 in mice.

AI.3 Materials and Methods

Random Mutagenesis. AAV4 helper plasmid (pXR4), containing Rep2 and Cap4, was obtained from UNC vector core. Random mutations were introduced to the GH loop (amino acids 428-617 on VP1, GeneBank no.: U89790.1) by error-prone PCR using the GeneMorph II EZ clone domain mutagenesis kit (Agilent Technologies, Santa Clara, CA).

Briefly, forward primer 5'-AACCCTCTCATCGACCAGTAC-3' and reverse primer 5'-TCCATCGGTATGAGGAATCTT-3' (IDT, Coralville, IA) were utilized in PCR cycles according to manufacturer's instruction to amplify regions on Cap4 encoding GH loop with random mutations. Individual pXR4 mutant clones were sequenced by the University of North Carolina Genome Analysis facility, and sequence alignments were carried out using VectorNTI software (Invitrogen, Carlsbad, CA).

Molecular Modeling. Surface models of AAV4 were displayed as a full icosahedral capsid or trimers using the crystal structure of AAV4-VP3 from RCSB Protein Data Bank (PDB-ID: 2G8G) in PyMOL. Homology models of 4.18 and 4.41 were obtained using the SWISS-MODEL online server (http://swissmodel.expasy.org/) [213] with the crystal structure of AAV4-VP3 (PDB-ID: 2G8G) as a template. Three-dimensional models of the three-fold axes of 4.18 and 4.41 were generated using the oligomer generator utility in VIPERdb (http://viperdb.scripps.edu/oligomer_multi.php) [214], and visualized in PyMOL Molecular Graphics System (Schrödinger, LLC., New York, NY).

Cell Culture. HEK293 cells were obtained from UNC Vector Core and cultured in Dulbecco's Modified Eagle's Medium supplemented with 10% fetal bovine serum (FBS), 100 U/ml of penicillin, 100 µg/ml of streptomycin, and 2.5 µg/ml of amphotericin B (Sigma-aldrich, St. Louis, MA). CV-1 cells (African green monkey fibroblast) were a gift from Dr. Jude Samulski (UNC-Chapel Hill) and propagated in Eagle's Minimum Essential Medium with 10% FBS, 100 U/ml of penicillin, 100 µg/ml of streptomycin, and 2.5 µg/ml of amphotericin B. All cell cultures were maintained in 5% CO_2 at 37°C.

Virus Production, Purification, and Quantification. Recombinant AAV4 and mutants were generated by triple-plasmid transfection method as described earlier [166]. All

three plasmids were obtained from UNC vector core unless mentioned otherwise. Briefly, a log-phase culture of HEK293 cells were transfected with the following three plasmids using polyethylene imine (Polysciences, Warrington, PA): (1) the AAV4 helper plasmids, pXR4, or mutant pXR4 generated from random mutagenesis described earlier, (2) pXX6-80 expressing Adenoviral helper genes, and (3) pTR-CBA-Luciferase, CBA-promoter-driven firefly luciferase transgene flanked by two Inverted-Terminal-Repeats (ITR) from AAV2. 60-72 hrs post transfection, HEK293 were pelleted by centrifugation and supernatant was collected. A stock solution of 40% [wt/vol] polyethylene glycol 8000 (PEG8000) and 2.5M NaCl was added in the supernatant to a final concentration of 8% PEG8000 and 0.5M NaCl. After overnight incubation at 4ºC, the precipitate was sedimented by centrifugation at 10k rpm for 30 min. This precipitate from supernatant and cell pellet were combined and subjected to cesium chloride untracentrifugation at 55k rpm in 17ºC overnight in a T-890 rotor (Thermo Scientific). Peak fractions with highest luciferase activity were collected and dialyzed against 1xPBS twice prior to titration of viral genome with quantitative PCR using a Roche Lightcycler 480 (Roche Applied Sciences, Pleasanton, CA). Quantitative PCR primers were designed to specifically recognize luciferase gene (forward 5'-AAA AGC ACT CTG ATT GAC AAA TAC-3'; reverse 5'-CCT TCG CTT CAA AAA ATG GAA C-3'). A virus stock with a previously known titer was included in qPCR as an internal control for viral genome titration.

***In Vitro* Transduction Analysis, Cell Surface Binding Assay, Internalization Assay.** CV-1 cells were seeded in 24-well plates at a density of 1e5 cells/well 18 hrs prior to experiment. The cells were first pre-chilled at 4ºC for 30 min, and then AAV4, 4.18, and 4.41 packing CBA-luciferase transgene were incubated on CV-1 cell surface at 4ºC for 60 min

followed by three washes with ice-cold 1x phosphate-buffered saline (1xPBS) to remove unbound virions. For transduction assays, 0.5 ml of DMEM with 10% FBS was replaced onto CV-1 cells before they were moved back to 37ºC incubator with 5% CO_2. 18 hrs later, cells were lysed to quantify the luciferase transgene expression in a Victor2® luminometer (Perkin Elmer, Waltham, MA).

For binding assays, after removal of unbound virions by washes with cold PBS as descried above, 200 µl of ddH2O was added to each well. Cells went through three freeze-thaw cycles followed by DNA extraction using DNeasy kit (Qiagen, Germantown, MD). The amount of transgene in each well was titrated by quantitative PCR as previously described.

In internalization assays, CV-1 cells were seeded at a density of 1e5 cells/well in 24-well plates the day before experiment to allow for attachment to the plates. During the internalization assay, cells were first chilled to 4ºC in cold room before adding viruses at an MOI of 1000 vg/cell. After incubation of viruses on cell surface at 4ºC for 60 min, unbound viruses were wash away by three changes of cold 1xPBS, and fresh DMEM was replaced for each well. Cells were brought into 37ºC incubator to synchronize virus internalization. 1 hour later, cells were treated with 150 µl/well of 0.25% trypsin at 37ºC for 5 min to release cell surface-associated AAV. Trypsinization was quenched by adding 150 µl/well of DMEM with 10% FBS into each well, followed by three washes of cell pellets with ice-cold PBS. DNA was extracted from cell pellets using DNeasy kit, and viral genome copy numbers were measured using qPCR as previously described in binding assays.

Hemagglutination Assay. All our animal experiments were performed on 6-8 weeks old female BALB/c mice purchased from Jackson Laboratories (Bar Harbor, ME). Animals were maintained and treated in accordance with National Institutes of Health guidelines and

as approved by Institutional Animal Care and Use Committee (IACUC) at UNC at Chapel Hill. 6-8 weeks old female BALB/c mice were anesthetized using 1.25% avertin through intraperitoneal injection before blood was drawn using cardiac puncture. Murine erythrocytes were collected by centrifugation at 5k rpm for 5min, washed three times in 1xPBS, and diluted by 10-fold with Alsever's solution (2.05% [wt/vol] glucose, 0.8% sodium citrate, 0.055% citric acid, 0.42% sodium chloride, pH 6.1). Hemagglutination assay was carried out on a V-bottomed 96-well clear plate in cold room. Briefly, 1e11 vg of AAV4, 4.18, and 4.41-CBA-luciferase were serially diluted in 2-fold steps with 1xPBS, respectively. 50 µl of 10% murine erythrocytes were added into each well, and incubated with virus for 1 hour in cold room. Hemagglutination titer was determined as the lowest fold-of-dilution of virus stocks that disrupt hemagglutination.

***Ex Vivo* Blood Cell Binding Assay.** Whole blood from BLAB/c mice were collected as described above. Whole blood were mixed with two volumes of Alserver's solution to prevent coagulation, and further diluted into 10% with 1xPBS and kept in 4ºC. 2e10 vg of AAV4, 4.18, and 4.41-CBA-luciferase were incubated with 100 µl of 10% whole blood on ice for 1 hour, followed by separation of plasma and blood cells by centrifugation. Blood cells were subjected to three washes with ice-cold 1xPBS prior to viral genome extraction from both blood cells and plasma using DNeasy kit. The amount of viral genome in each fraction was measured using qPCR with primers specific to luciferase transgene.

Transduction and Biodistribution Assays in Live Animals. To investigate the effect of sialic acid in AAV4 tropism *in vivo*, BALB/c mice were injected with either 400 mU/mouse of sialidase (Type III, from *V. Cholera*, Sigma) or 1xPBS through tail vein 3 hrs prior to administration of AAV4-CBA-luciferase (1e11 vg/mouse) via the same route. To

track luciferase transgene expression pattern in mice at indicated time points, aminals were anesthetized with 2% isoflourane prior to intraperitoneal injection of D-luciferin (120 mg/kg, Nanolight, Pinetop, AZ), and imaged using an Xenogen IVIS Lumina® imaging unit (Perkin Elmer/Caliper Life Sciences, Waltham, MA) at indicated parameters. To test transduction profiles of AAV4 mutants, 5e10 vg of each AAV4 variant was injected intravenous into BALB/c mice. Luciferase transgene expression in live animals was imaged as described above at indicated time points.

To quantify luciferase expression in individual organs, the same group of mice from imaging experiment was sacrificed to harvest about 50 mg tissues from indicated organs. 150 µl of 2x passive lysis buffer (Promega, Madison, WI) was added to each tissue before lysing them on a Tissue Lyser II 352 system (Qiagen, Valencia, CA). Tissue lysates were pelleted using centrifugation. 50 µl of supernatant from each lysate was subjected to luciferase assay on a Victor2® luminometer (Perkin Elmer, Waltham, MA). 2 µl of supernatant was used in Bradford assay (BioRad, Hercules, CA) to determine protein content in each sample.

For biodistribution analysis, a second group of BALB/c mice were injected with the same viruses and killed at 3 days post administration to harvest organs. 50 mg of tissue from each organ were pretreated with proteinase K prior to being homogenized in a Tissue Lyser. DNA was extracted from tissue lysates using DNeasy kit (Qiagen, Germantown, MD). Viral genome copies were measured using qPCR with primers binding to Luciferase transgene, while number of cells in each lysate sample was quantified using qPCR with primers specific to mLamin housekeeping gene. Data were shown as viral genome copies normalized by cell numbers.

Pharmacokinetics Assay. BALB/c mice were pretreated with intravenous sialidase or PBS as previously described. 3 hrs post administration, 1e11 vg copies of AAV4, 4.18, or 4.41-CBA-Luciferase were injected into animals through tail vein. 10 ul of blood was collected by tail nicking at the following time points post injection: 5 min, 15 min, 30 min, 1 hour, 3 hrs, 6 hrs, 24 hrs, 30 hrs, 48 hrs, and 72 hrs. DNA was extracted from each blood sample using a DNeasy kit, and viral genome copy numbers were measure using qPCR as described previously.

AI.4 Results

Intravenous sialidase abrogates the cardiopulmonary transduction profile and biodistribution of AAV4 in mice. To investigate how interactions with sialic acids impact on AAV4 tropism, a broad-spectrum sialidase from *Vibrio cholerae* was administered intravenously into BALB/c mice to remove cell-surface exposed sialic acid moieties on major organs. Previous study demonstrated that the density of terminal galactose on the surface of major organs increased significantly after systemic administration of this sialidase [51]. This increase of galactose-terminated glycans was accompanied by moderate decrease of terminal sialic acids in the same organs, as shown by Erythrina Cristagalli Lectin (ECL) and Maackia Amurensis Lectin I (MAL I) staining [51]. 3 hrs post sialidase treatment, AAV4 packaging chicken beta-actin promoter-driven firefly luciferase transgene (CBA-Luciferase) was injected into BALB/c mice intravenously. Bioluminescence images were obtained at 7 days post administration. As shown in **Figure 32A**, transgene expression of AAV4 is localized specifically in lung and heart regions in PBS-treated BALB/c mice, and this transduction profile is absent in mice pretreated with sialidase. This observation was

corroborated with luciferase expression assay and biodistribution analysis of tissue lysates from the same group of animals. In consistent with bioluminescence images, luciferase transgene expression (**Figure 32B**) in heart and lung shows stark decreases by ~300-fold and 600-fold respectively in sialidase-treated group compared to PBS mock-treated group. The decrease of luciferase expression was also observed in leg muscle (~15-fold), spleen (~2.5-fold), and kidney (~12-fold), but transgene expression in liver remained at the same level in sialidase-treated mice compared to PBS-treated group.

Quantitative analysis of viral genome copy numbers in major organs demonstrates the effect of sialidase on the biodistribution of AAV4 viral particles (**Figure 32C**). Normalized to the copy numbers of mouse genomic DNA, the biodistribution profiles of AAV4 reflects the trend displayed by the aforementioned transduction efficiency profiles. Specifically, there are > 250-fold and ~400-fold decreases of viral genome accumulation in heart and pulmonary system in sialidase-treated mice, respectively. Modest reductions of viral genome copies per cell were seen in leg muscle (~17-fold), spleen (~3-fold), and kidney (~12-fold) due to intravenous injection of sialidase, while no statistically significant changes observed in hepatic tissues. These results indicate that sialic acids play a critical role in mediating efficient infection of AAV4 in the cardiopulmonary system in mice.

***O*-linked sialic acids are highly expressed in the cardiopulmonary system in mice.** To interrogate the underlying mechanism of the differential transgene expression profiles and biodistribution patterns of AAV4, we hypothesized that glycan receptors plays a critical role in determining the tropism of AAV4. Using lectin staining, the endogenous distribution of the primary receptor for AAV4 – *O*-linked sialylated glycans – was analyzed in murine heart, lung, and liver. Briefly, sections of heart, lung and liver from BALB/c mice underwent

immunostaining by fluorescein labeled Jacalin lectin (FITC-Jacalin). Jacalin is a lectin isolated from jackfruit *Artocarpus integrifolia* [233]. It specifically recognizes glycopeptides that contain *O*-linked oligosaccharides with a disaccharide core of Galactose-(*N*-Acetyl-)Galactosamine (Gal-GalNAc) covalently linked to serine or threonine residues [234]. Therefore, Jacalin has been widely used as markers for *O*-glycosylation, such as mucins. As seen in the left panel of **Figure 33**, FITC-Jacalin shows markedly higher staining on the ciliated mouse airway epithelial lining and the alveoli. Similar intensity of FITC staining is evenly present in the cross-section of murine hearts (middle, **Figure 33**), indicating the outlines of cardiomyocytes. On the contrary, FITC-Jacalin does not stain liver sections with comparable intensity uniformly (right, **Figure 33**). Only the endothelial side of blood vessels in murine livers is stained by FITC-Jacalin, leaving both parenchymal and nonparenchymal cells unstained. Together with the aforementioned cardiopulmonary tropism of AAV4, these data suggest a pivotal role of *O*-glycosylation in the transduction profiles and biodistribution patterns of AAV4.

An AAV4 library was constructed by random mutagenesis on the GH loop near the three-fold axis of symmetry. The three fold axes of symmetry on AAV capsids have been claimed to be responsible for receptor recognition by AAV1, AAV2, AAV3b, AAV6, and AAV9 [43,83,90,92-94,165,198]. The protrusion and spike structures surrounding the three fold axes are composed of highly dynamic loops connecting seven strands of the core β-sheet structure. These loops share low amino acid sequence homology among AAV strains, and demonstrates high flexibility according to X-ray crystallography [81,82,165,198,212,235-237]. To identify key residues involved in sialic acid binding at three-fold axes on AAV4 capsids, a random library of AAV4 was generated by introducing mutations on the GH loop connecting

the G and H strand of the core beta sheet in VP (highlighted in blue in **Figure 34A**). 48 individual clones of plasmids encoding *Rep* from AAV2 and mutated *Cap* from AAV4 (pXR4) were selected arbitrarily from this library reservoir. By structural modeling, mutant isolates with the following features were excluded from further analysis: silent mutations, frame-shifts, deletions, mutations buried inside tertiary structure, or mutations only exposed on the interior side of capsid. As a result, 9 variants were selected out of 48 mutant AAV4 isolates for further characterization (**Table 7**). A collection of mutations on these 9 variants were clustered on the exterior of spikes forming the three-fold axes, as highlighted in blue in **Figure 34**.

4.18 and 4.41 are deficient in sialic acids binding and cellular internalization. In order to identify AAV4 mutants that are unable to bind to sialic acids, we carried out an *in vitro* screening of infectivity on CV-1 cells. CV-1 cells, kidney fibroblasts from African green monkey, are highly permissiveness to AAV4 infection. In this screening, CV-1 cells were infected with mutant AAV4 capsids packaging CBA-Luciferase reporter gene. Luciferase transgene expression was measured as the indicator of infectivity. As demonstrated in **Figure 35A**, the infectivities of both 4.18 and 4.41 are more than one log unit lower than that of the parental AAV4 in CV-1. To investigate the fundamental reasons for this loss of infectivity displayed by 4.18 and 4.41, cell surface binding and internalization assays were carried out on CV-1 cell cultures. As shown in **Figure 35B** (grey bars), both 4.18 and 4.41 displayed 2-4 –fold lower binding capacity on CV-1 cells surface. More interestingly, the differences between the internalization efficiencies of 4.18/4.41 and that of AAV4 (white bars, **Figure 35B**) are more significant. At the multiplicity infection (MOI) of 1000 vg/cell, there are 20 AAV4 bound onto CV-1 surface on average, 10 of which are

further internalized after 1 hour incubation in 37°C. In contrary, at the same MOI, only ~ 10 viral particles of 4.8/4.41 are bound to CV-1 surface, and ~1 virion is internalized later on average. This percentage of internalization depicted in **Figure 35C** strongly supports the notion that 4.18/4.41 are deficient in both cell surface binding and cellular uptake in an *in vitro* setting. Together with results from the infectivity screening, these data suggest that AAV4 mutants with low sialic acid binding affinity demonstrate less cellular attachment and uptake, which further leads to reduced infectivity *in vitro*.

4.18 and 4.41 are incapable in transducing the cardiopulmonary system *in vivo*. While the *in vitro* screening of AAV4 mutants were carried out, the same set of AAV4 variants were subjected to *in vivo* characterization by monitoring their transgene expression after intravenous injections. At 1 month post injection, bioluminescence images were collected, and two types of transduction pattern were observed: (1) Cardiopulmonary transduction similar as that of parental AAV4, such as 4.7, 4.15, 4.19, 4.29, 4.31, 4.38, and 4.49 (**Figure 36A**); (2) Lost of transduction efficiency *in vivo*, as displayed by 4.18 and 4.41 (**Figure 36A**). The classification of these two phenotypes was verified by luciferase assays on tissue lysates isolated from the same group of animals (**Figure 36B**). Compared to parental AAV4, only 4.18 and 4.41 demonstrate decreased luciferase activity across all major organs, including heart, lung, leg muscle, liver, and spleen (**Figure 36B**). Specifically, in case of 4.18, transduction efficiency in heart dropped by more than two orders of magnitude, while that in lung decreased by over a log; in case of 4.41, more dramatic lost in transduction efficiency was observed in lung tissue (>100-fold), while there was only a 10-fold decrease of transduction efficiency in heart.

These results are further corroborated by the biodistribution of viral genome of 4.18 and 4.41 as compared to AAV4 (**Figure 36C**). Consistent with the trend of transduction efficiency shown in **Figure 36B**, 4.18 and 4.41 displayed ~5-10 -fold decrease in viral genome accumulation in heart and lung compared to the parental AAV4. Similar loss of viral genome copies was also observed in muscle and kidney. The biodistribution of both 4.18 and 4.41 in hepatic tissues remain unaltered, while only 4.41 showed ~10-fold reduction in biodistribution in spleen. Taken together, these results suggested that 4.18 and 4.41 are unable to entry and transduce the cardiopulmonary system in mice due to surface exposed mutations on the three-fold axes. According to *in vitro* data on CV-1 cells, the loss of function of 4.18 and 4.41 results from their inability to bind and enter cells in target tissues marked by high *O*-glycosylation.

4.18 and 4.41 bind to erythrocytes with lower avidity. To correlate the *in vitro* binding and internalization data of 4.18 and 4.41 with their *in vivo* transduction profiles, an *ex vivo* blood binding experiment and hemagglutination assays were carried out with blood and erythrocytes from BALB/c mice, respectively. After incubating viruses with freshly prepared whole blood on ice, plasma and blood cells were separated using centrifugation, and the amount of viral genome distributed in each fraction were measured with qPCR. With the total amount of recovered viral genome as total virions, 75% of AAV4 was associated with blood cells, while only 25% of 4.18 and 40% of 4.41 were recovered from the fractions with blood cells (**Figure 37**). The majority of 4.18 and 4.41 remained in plasma. This observation reveals that 4.18 and 4.41 do not bind to blood cells as efficiently as AAV4.

These observations prompted us to carry out hemagglutination assays (**Table 7**) to quantitatively compare the avidities between 4.18/4.41 viruses and murine erythrocytes. As

negative control that does not bind to sialic acids, AAV2 has an HA titer of 2, the same as that of PBS. Consistent with previous publications [46], AAV4 has an HA titer of 128, indicating a high binding affinity towards sialic acids. Both 4.18 and 4.41 only showed HA titers of 4, suggesting that they do not bind to sialylated glycans as efficiently as the parental AAV4. This sialic acid binding deficiency leads to less association with blood cells (**Figure 37**), which are decorated with a spectrum of sialylated glycans. In corollary, 4.18 and 4.41 are taken up by tissues and organs with lower efficiencies *in vivo*, leaving most virions cleared from reticuloendothelial system (RES).

Lower erythrocyte binding avidity leads to faster blood clearance profiles of AAV4 in mice. In order to further understand the role of sialic acids in AAV4 infection *in vivo*, pharmacokinetics analysis of AAV4 in mice with and without sialidase treatment was performed (**Figure 38A**). At 24 hrs after virus administration, AAV4 viral genome in the blood of PBS-treated mice is more than 2-fold higher than that in the blood of desialylated mice. This difference became ~20-fold when it comes to 72 hrs post virus dosing. These results reveal that the availability of sialic acid receptors *in vivo* also alters the adsorption of AAV4 from blood to tissue, together with its clearance from blood circulation through unidentified routes. We then carried out pharmacokinetics analysis of 4.18/4.41 and AAV4 to elucidate the mechanism of their different biodistribution profiles (**Figure 38B**). Within 24 hrs post virus administration, the amount of viral genome in total blood for both 4.18 and 4.41 are notably lower than that of AAV4, indicating fast clearance rate from blood circulation. Therefore, lower erythrocytes binding avidity, resulted from either less sialic acid receptors or disruption of glycan receptor footprints on viral capsids, leads to faster pharmacokinetic profiles of AAV4 virions.

AI.5 Discussion

Sialylated glycans function as cellular attachment receptors for a broad spectrum of viruses, including influenza, rotaviruses, coxsackieviruses [1,192]. In the *Dependoviridae* genus of the *Parvoviridae* family, AAV4 is the only serotype that binds to *O*-linked 2,3-sialic acids, the polymeric form of which is called "mucin". Mucin, produced by goblet cells on the epithelia of tracheobronchial, gastrointestinal, and genitourinary tracts, constitutes a defensive mechanism against exogenous pathogens. Even though both AAV4 and AAV5 can hemagglutinate red blood cells by binding to cell surface sialic acids, only AAV4 can be efficiently inhibited from infecting well-differentiated human airway epithelial (HAE) cells from the apical side [47]. If infecting from the basolateral side, AAV4 transduces HAE efficiently through sialic acid-dependent pathway(s) [47]. Given the specific cardiopulmonary-tropism of blood-borne AAV4, the roles of sialic acids in AAV4 infectivity drew our attention.

The role of glycan receptors in virus lifecycles is to determine the host tropism and interspecies transmissibility of viruses. Pathogenic respiratory viruses, such as influenza viruses and rhinoviruses, evolved to recognize specialized forms of sialic acids on upper airway epithelial cells to initiate their infectious pathways in host [11,238]. Hepatitis B, C, and E viruses employ heparan sulfates, which is abundant on hepatocytes, to infect human liver [34-36]. A good example of how glycan binding profile of AAV translates into tropism difference is AAV1 and AAV6 [43]. By a difference of six out of >700 amino acids on VP1 sequence, AAV1 binds only to *N*-linked sialic acids, while AAV1 binds to both *N*-linked sialic acids and heparan sulfates. As a result, besides muscle and lung for AAV1, AAV6 also transduces hepatocytes with moderate efficiency. Our recent study indicates that the systemic

transduction profile of AAV9 is driven by the low glycan binding avidity and the low availability of cognate glycan receptors [51]. In current study, by disrupting sialic acid-AAV4 interactions, the endogenous cardiopulmonary transduction profile of AAV4 is completely lost, with pharmacokinetics profiles altered. Our data bridges the *in vitro* results of AAV4-erythrocytes interactions with its behaviors *in vivo*. After systemic administration, AAV4 enters the blood circulation, and has a preferential affinity towards organs expressing *O*-sialic acids abundantly. Upon intravenous injection of sialidase, the interactions between AAV4 and cell-surface sialic acids are interrupted. Therefore, AAV4 are unable to bind to blood cells or internalized into tissues, and eventually end up cleared from the blood circulation faster. The same reasoning also applies to AAV4 mutants with disrupted sialic acid binding footprint.

For the first time, it is suggested that sialylated glycans play a multi-facet role in the infectious pathway(s) of AAV4 *in vivo*. Reconstitute of sialic acid moieties on desialylated rhesus erythrocytes by α-2,3-*O*-sialyltransferases, but not α-2,3-*N*-sialyltransferases recovers the sialic acid binding avidity of AAV4 [46]. It suggests that polysaccharides linked to Ser/Thr residues on glycopeptides are critical for cell surface attachment of AAV4. However, the protein receptor(s) involved in the cellular uptake of AAV4 has yet to be determined. Of other sialic acid-binding AAV serotypes, platelet-derived growth factor receptor (PDGFR) and epidermal growth factor receptor (EGFR) were determined to be the secondary receptors for AAV5 and AAV6, respectively [102,103]. In addition, there is no integrin binding motif (RDG or NGR) on AAV4 capsid either. Based on the internalization assay of 4.18 and 4.41, it is tempting to speculate that sialylated glycans are also a determining factor in mediating the cellular internalization of AAV4.

Another critical finding in this study is the identification of key amino acids involved in glycan-AAV4 interactions. In case of heparan sulfate-binding AAV2, there is a basic strip (R484, R487, K532, R585, R588) on the protrusions of three-fold axes which is critical in interacting with glycans [83,89,90,92]. When it comes to AAV6, its heparan sulfate binding ability can be attributed to a single amino acid, K531 [43]. Sequence alignment of several heparin-binding proteins suggests a typical heparin-binding motif was proposed: XBBXBX or XBBBXXBX, where B is lysine or arginine, X is any other amino acids [2]. Unlike heparan sulfate binding motif, no conserved binding site for sialic acid has been discovered across several viruses, or different strains of AAV. Based on structure-function studies of canine parvoviruses (CPV) [239] and minute virus of mice (MVM) [240], amino acid residues at similar position as basic residues on AAV2/AAV6 are responsible for sialic acid binding [43,83,89,90,92]. From this current study, key residues on proposed sialic acid binding footprint is focused on K492, K503, M523, G581, Q583, N585 near the three-fold axes. By sequence alignment and structural comparison, G581-N585 geographically equivalent to the basic stretch on AAV2. K492 and K503 both constitute dominant basic patch on the exterior of protein shell at locations close to its counterpart on AAV2 capsid. Mutations at these positions, but not around 450-465 or 588-610, significantly affect the sialic acid binding profile of AAV4. Mutagenesis study with individual residues would provide more information on this region. Our attempts to dissect the functions of variable regions on AAV4 capsid by random mutagenesis provides substantial aids in understanding the structure-function relationship of AAV in general.

In conclusion, our study provides new insights into the functional roles of sialylated glycans in AAV4 biology. Sialylated glycans not only serve as attachment factors for AAV4,

but also induce its cellular uptake in both *in vitro* models and *in vivo* models. Thus, interactions between AAV4 and its cognate glycan receptors are indispensable for its infection in cardiopulmonary system. In addition, we also identified key residues (K492, K503, M523, G581, Q583, N585) on the three-fold axes of AAV4 capsid that compose part of sialic acid binding motif, and reconfirmed that trimer-forming GH loop is critical in receptor binding of AAV. With this random library of AAV4, more aspects of AAV4, such as immunogenicity determinants, can be studied to facilitate AAV4-based vector design. Furthermore, this study might also contribute to our understanding on the mechanism of how AAV4 transduces ependymal cells and astrocytes in CNS.

Table 7. List of AAV4 Mutants

Number	Point Mutations on AAV4-VP3
4.7	T454S
4.15	T496S, S515C, L564P, L588M, R593G, M603V, I610T
4.18	K492E, K503E, N585S
4.19	K479N, L501F, A569P
4.20	T446A, S584G
4.31	L588Q
4.38	T463A, W514L, Q537L, G581C
4.41	M523I, G581D, Q583E
4.49	R593D

Table 8. The Ability of AAV4 and Mutants to Hemagglutinate Murine Erythrocytes

Virus	Hemagglutination Titer
PBS	< 2
AAV2	< 2
AAV4	< 128
4.18	<4
4.41	< 4

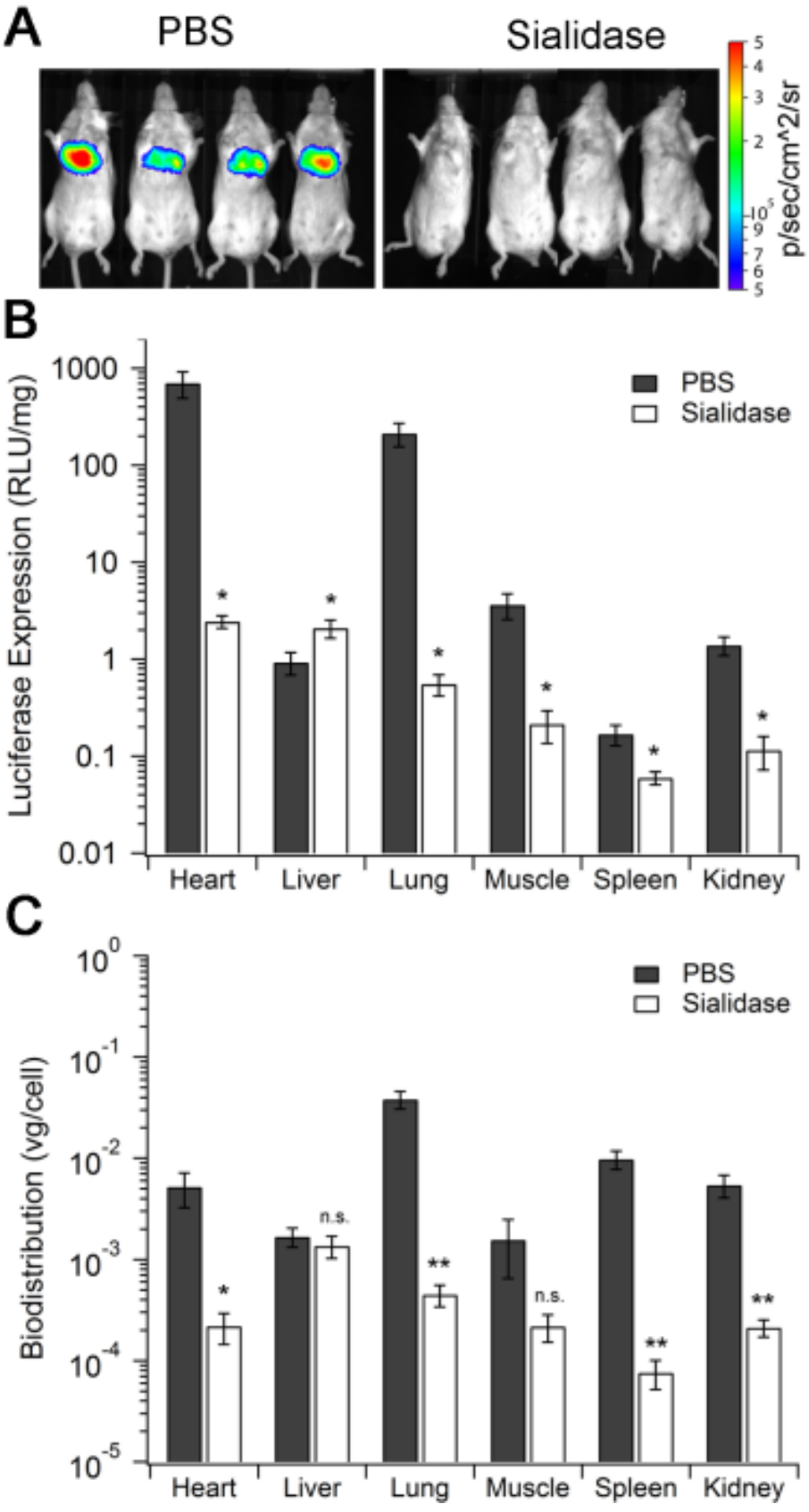
A
PBS
Sialidase
p/sec/cm^2/sr
B
Luciferase Expression (RLU/mg)
PBS
Sialidase
Heart
Liver
Lung
Muscle
Spleen
Kidney
C
Biodistribution (vg/cell)
PBS
Sialidase
n.s.
Heart
Liver
Lung
Muscle
Spleen
Kidney

Figure 32. Enzymatic removal of sialic acids *in vivo* disrupts AAV4 transduction. (A) Live animal images of AAV4-mediated luciferase transgene expression in BALB/c mice. Mice (n=4) were injected with PBS or sialidase from *Vibrio Cholera* (400 mU/mouse) through tail vein 3 hrs prior to AAV4-CBA-Luciferase (1e11 vg/mouse) injection via the same route. At 7 days post infection, live animal bioluminescence images were recorded with an Xenogen IVIS Lumina Imaging System. Scale stands for relative light units (RLU) as photons per sec per cm^2 per steradian. **(B)** Luciferase transgene expression in tissue lysates at 7 days post infection. PBS or sialidase were injected intravenously into BALB/c mice (n=4) 3 hrs before AAV4-CBA-Luciferase administration. At 7 days post infection, heart, liver, lung, leg muscle, spleen, and kidney were harvested from animals. Luciferase transgene expression in each tissue lysate was measured using a bioluminescence plate reader. Values were expressed in RLU normalized to amount of protein in each lysate sample. Data are shown as mean ± s.e.m. (n=4). **(C)** Biodistribution of AAV4 viral genomes in BALB/c mice. DNA contents in tissue lysates from experiment in **(B)** were isolated and subjected to quantitative PCR analysis, as described in Materials and Methods. Viral genome copies were normalized to copy numbers of housekeeping gene mLamin. Data are shown as mean ± s.e.m. (n=4). Statistical significance was analyzed using the one-tailed Student's *t*-test (* $p < 0.05$; ** $p < 0.01$).

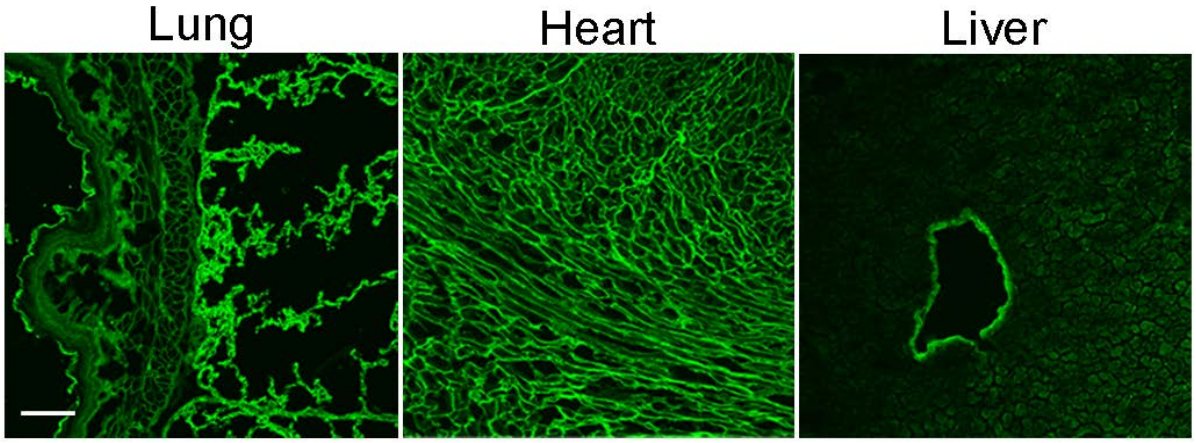

Figure 33. Expression of sialylated *O*-glycans in murine lung, heart, and liver. Lung, heart, and liver tissues harvested from BALB/c mice were stained with FITC-Jacalin, a lectin specific to *O*-linked polysaccharides. Confocal micrographs were obtained using a Zeiss 710 Confocal Laser Scanning Microscope with a 20x objective at zoom 0.6. Scale bar = 100 µm.

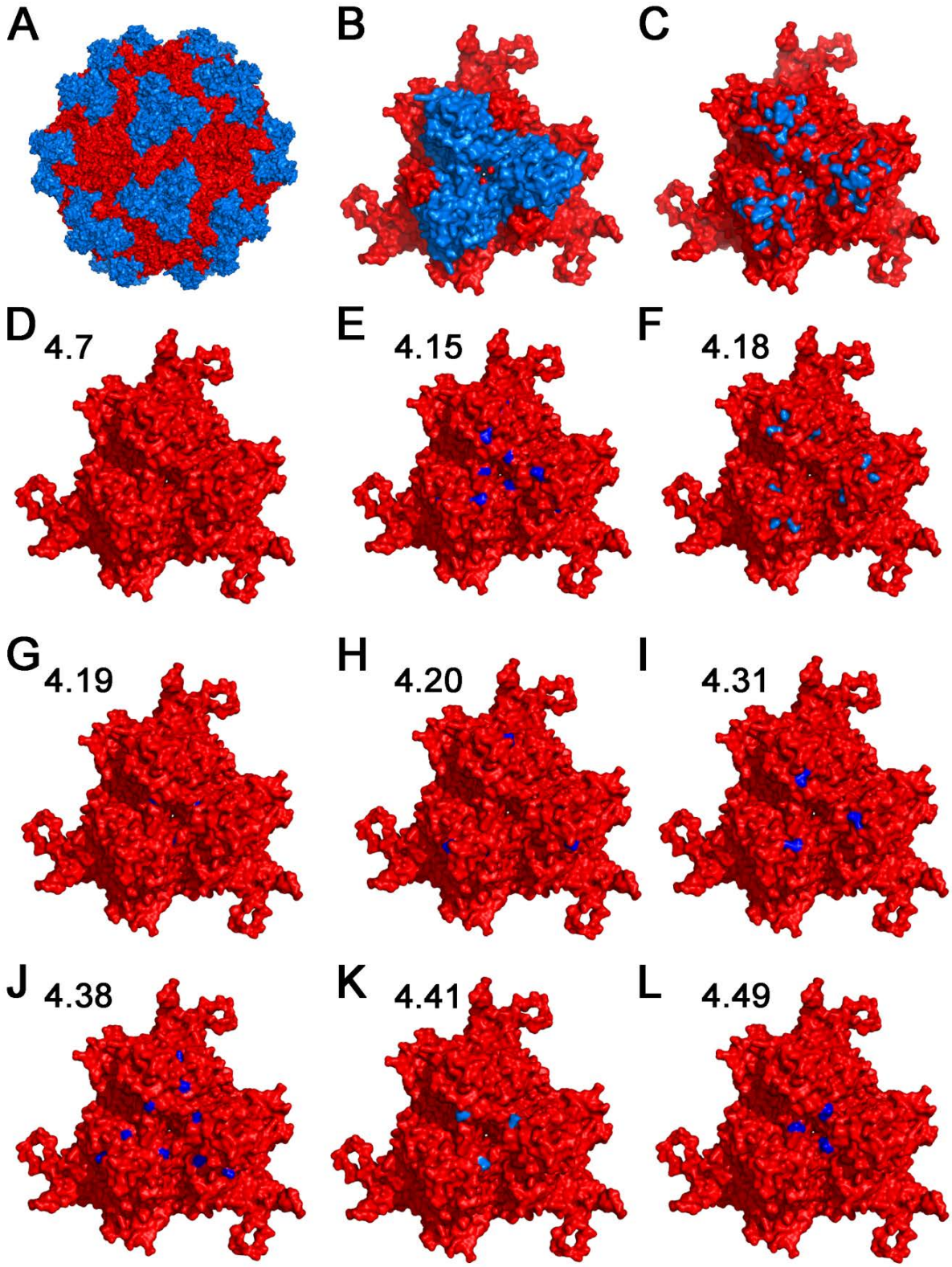

Figure 34. Three dimensional models of AAV4 capsid and trimer models of 4.18 and 4.41. (A) Three-dimensional structural model of an intact AAV4 capsid with its backbone colored in red and GH loops (VP1 numbering: 428-617) highlighted in blue, respectively. **(B**

and C) Visualization of AAV4 trimeric assembly from exterior of the capsid. The GH loops **(B)** or the cluster of point mutations **(C)** introduce by random mutagenesis are marked in blue. **(D -L)** Surface representations of the VP3 trimers of individual AAV4 mutants from the random mutagenesis library, respectively. Residues mutated from the parental AAV4 are highlighted in blue. Models were visualized and generated using PyMOL.

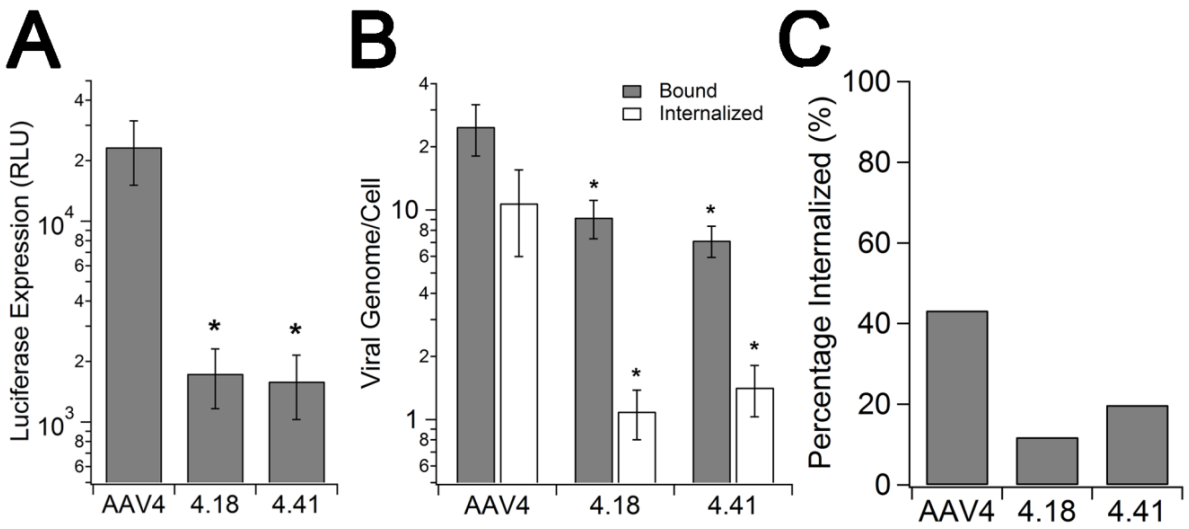

Figure 35. Infectivity, binding and internalization assays of AAV4, 4.18, and 4.41 on CV-1 cells. (A) Infectivity of AAV4, 4.18, and 4.41 on CV-1 cells. AAV4, 4.18, and 4.41-CBA-Luciferase (MOI=1000 vg/cell) were incubated on pre-chilled CV-1 cells for 1 hour before three washes of PBS to remove unbound viral particles. 18 hrs later, cells were lysed to quantify luciferase transgene expression using bioluminescence plate reader. Data were presented as mean ± s.e.m. (n=4) in RLU. **(B)** Binding (grey bars) and internalization (white bars) of AAV4, 4.18, and 4.41 on CV-1 cells. Viruses packaging CBA-Luciferase transgene were allowed to bind on prechilled CV-1 cells for 1 hour. After three washes with ice-cold PBS, cells were either lysed to measure numbers of bound virions using qPCR, or moved into 37ºC incubator to allow for internalization. 1 hour after 37ºC incubation, viral particles not internalized were removed by trypsin. Trypsin-resistant viral particles in CV-1 cells were quantified using qPCR. Data were shown as mean ± s.e.m. (n=4) in vg/cell. **(C)** Percentage of internalization of AAV4, 4.18, and 4.41. The amount of internalized virions (**B**, white bars) was normalized to the amount of bound virons (**B**, grey bars) for AAV4 and mutants to show differences in internalization efficiency on CV-1 cells. Statistical significance was analyzed using the one-tailed Student's *t*-test (* $p < 0.05$; ** $p < 0.01$).

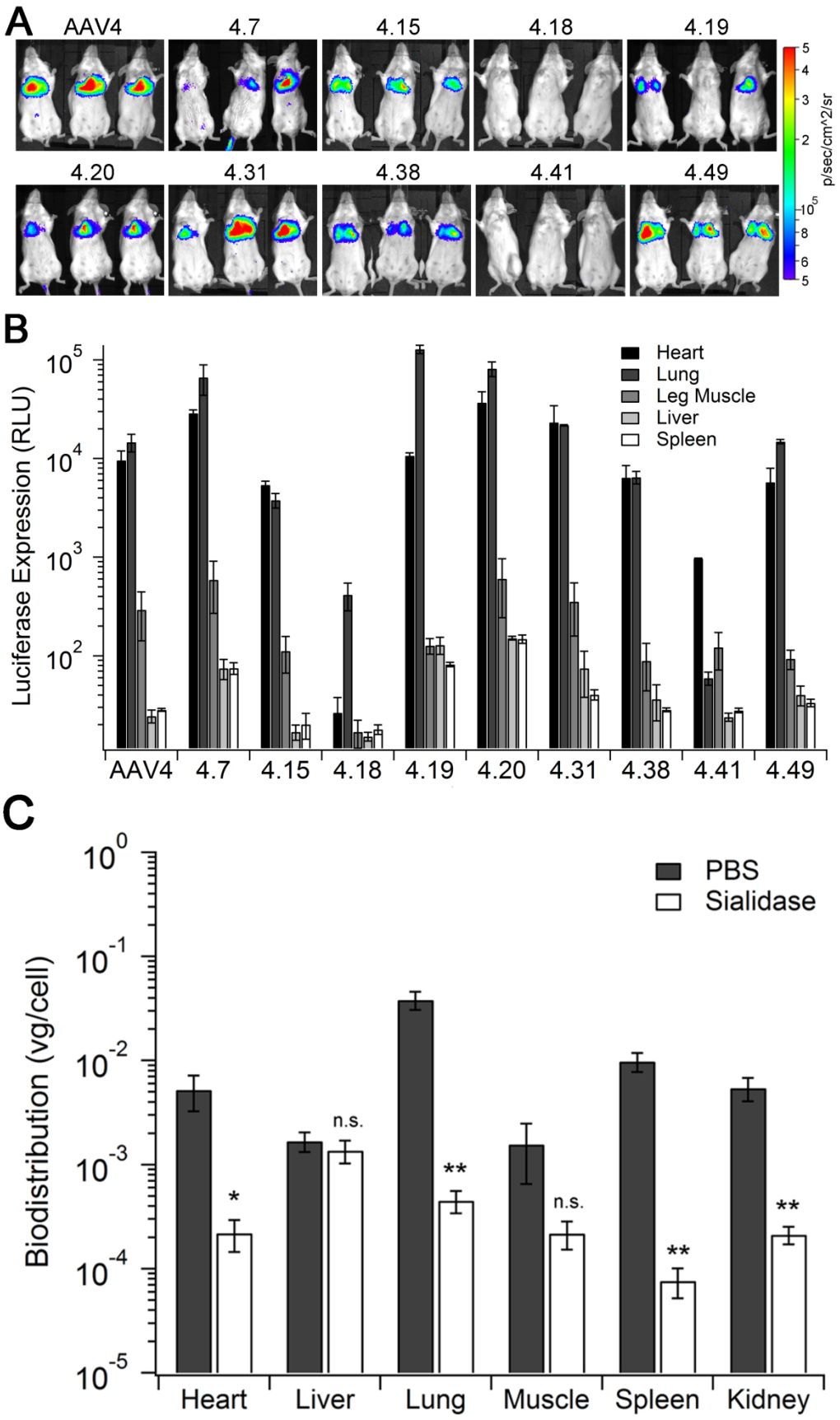
A
AAV4
4.7
4.15
4.18
4.19
4.20
4.31
4.38
4.41
4.49
p/sec/cm^2/sr
B
Luciferase Expression (RLU)
Heart
Lung
Leg Muscle
Liver
Spleen
AAV4
4.7
4.15
4.18
4.19
4.20
4.31
4.38
4.41
4.49
C
Biodistribution (vg/cell)
PBS
Sialidase
n.s.
n.s.
Heart
Liver
Lung
Muscle
Spleen
Kidney

Figure 36. Transduction profiles of 4.18 and 4.41 *in vivo*. (A) Live animal images of luciferase transgene expression mediated by AAV4, 4.18, and 4.41. AAV4 and mutants packaging CBA promoter-driven Luciferase transgene were injected into BALB/c mice (5e10 vg/mouse) through tail vein. At 1 month post infection, luciferase expression profiles in live animals were obtained using Xenogen IVIS Lumina Luciferase Imaging System (n=3). Scales represents RLU as photons per sec per cm2 per steradian. **(B)** Luciferase expression in tissue lysates. Animals shown in (A) were sacrificed at 1 month post infection to harvest heart, liver, lung, leg muscle, and spleen tissues. After lysing, luciferase expression in these tissues was recorded using bioluminescence plate reader. Data are shown as mean mean ± s.e.m. (n=4) in RLU. **(C)** Biodistribution of 4.18 and 4.41 in BALB/C mice at 3 days post infection. AAV4, 4.18, and 4.41 (5e10 vg/ul) packaging CBA-Luciferase transgene were injected into BALB/c mice intravenously. 3 days post infection, heart, liver, lung, leg muscle, and spleen were collected from animals to extract DNA. Viral genome copy numbers were obtained using qPCR, and normalized to number of mLamin genomic DNA copy numbers. Data are plotted as mean mean ± s.e.m. (n=4). Statistical significance was analyzed using the one-tailed Student's *t*-test (* $p < 0.05$; ** $p < 0.01$).

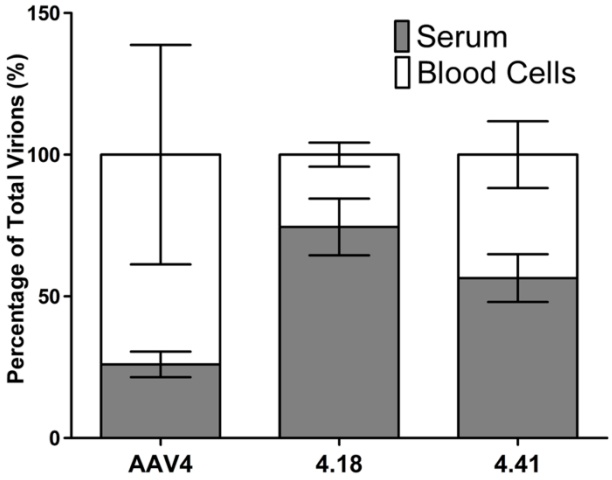

Figure 37. Distribution of AAV4 and mutants in blood *ex vivo*. 2e10 vg of AAV4, 4.18, or 4.41-CBA-Luciferase was incubated with 100 µl of 10% blood collected from BALB/c mice on ice for 1 hour. Plasma and cells were separated by centrifugation. Blood cells were washed with PBS three times to remove loosely bound virions. Amount of virus particles in plasma and cells were analyzed using qPCR after DNA extraction from both phases. Data were presented as percentage of total virions as normalized to total viral particles recovered. Error bars represent s.e.m. (n=5). Statistical significance was analyzed using the one-tailed Student's *t*-test (* $p < 0.05$; ** $p < 0.01$).

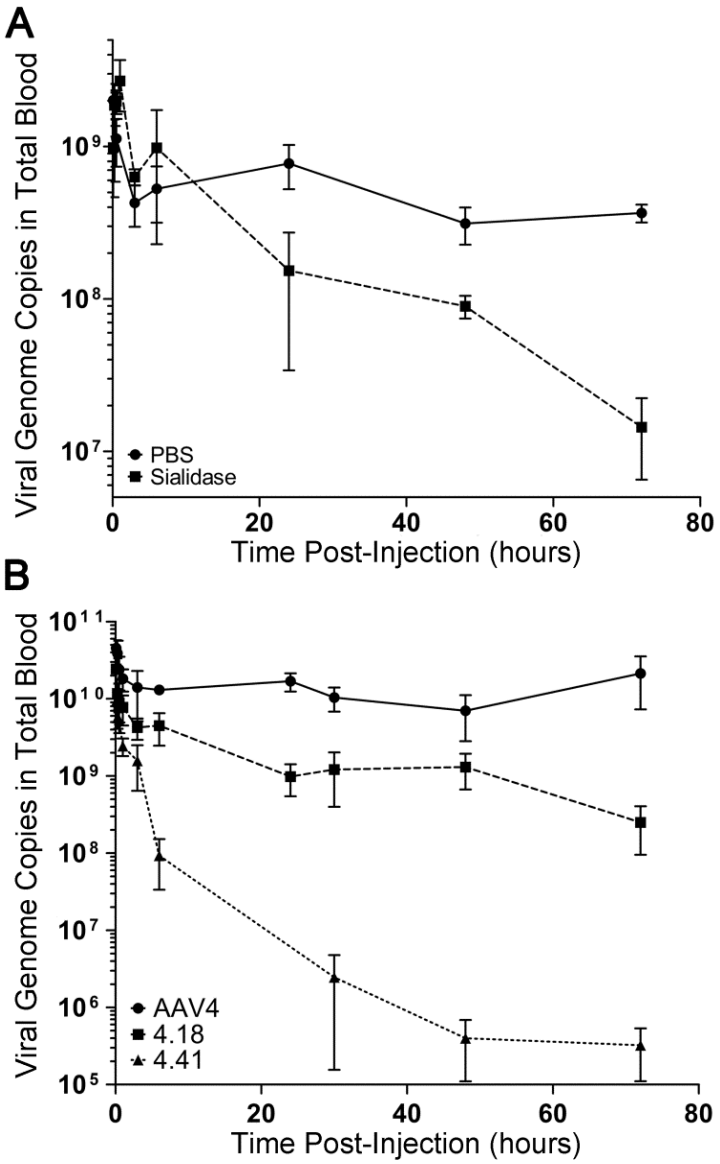

Figure 38. Sialic acid binding avidity affects the blood clearance profiles of AAV4. (A) Kinetics of AAV4 clearance in blood within 72 hrs post infection. PBS or sialidase from *Vibrio cholerae* (400 mU/mouse) was intravenously injected into BALB/c mice (n=4) 3 hrs prior to AAV4-CBA-Luciferase administration via the same route. 10 µl of blood was collected from tail vein at indicated time points post AAV4 injection: 5 min, 15 min, 30 min, 1 hr, 3 hrs, 6 hrs, 24 hrs, 48 hrs, and 72 hrs. Viral genome copies in blood samples were quantified using qPCR with primers annealing to Luciferase transgene. Data are shown as mean ± s.e.m. (n=4). **(B)** Blood clearance profiles of 4.18 and 4.41. AAV4, 4.18, and 4.41 (1e11 vg/mouse) were administered into BALB/c mice via intravenous injections. 10 µl of blood was collected through tail vein at following time points post injection: 5 min, 15 min, 30 min, 1 hr, 3 hrs, 6 hrs, 32 hrs, 48 hrs, and 72 hrs. Amount of viral genome in each blood sample was quantified using qPCR recognizing Luciferase transgene. Data are shown in vg in total blood as mean ± s.e.m. (n=4).

REFERENCES

1. Olofsson, S, Bergstrom, T (2005). Glycoconjugate glycans as viral receptors. *Ann Med* **37:** 154-172.

2. Cardin, AD, Demeter, DA, Weintraub, HJ, Jackson, RL (1991). Molecular design and modeling of protein-heparin interactions. *Methods Enzymol* **203:** 556-583.

3. Stehle, T, Yan, Y, Benjamin, TL, Harrison, SC (1994). Structure of murine polyomavirus complexed with an oligosaccharide receptor fragment. *Nature* **369:** 160-163.

4. de Paz, JL, Noti, C, Bohm, F, Werner, S, Seeberger, PH (2007). Potentiation of fibroblast growth factor activity by synthetic heparin oligosaccharide glycodendrimers. *Chem Biol* **14:** 879-887.

5. Neumann, G, Kawaoka, Y (2006). Host range restriction and pathogenicity in the context of influenza pandemic. *Emerg Infect Dis* **12:** 881-886.

6. Kuiken, T, Holmes, EC, McCauley, J, Rimmelzwaan, GF, Williams, CS, Grenfell, BT (2006). Host species barriers to influenza virus infections. *Science* **312:** 394-397.

7. Nguyen, DC, Uyeki, TM, Jadhao, S, Maines, T, Shaw, M, Matsuoka, Y, Smith, C, Rowe, T, Lu, X, Hall, H, Xu, X, Balish, A, Klimov, A, Tumpey, TM, Swayne, DE, Huynh, LP, Nghiem, HK, Nguyen, HH, Hoang, LT, Cox, NJ, Katz, JM (2005). Isolation and characterization of avian influenza viruses, including highly pathogenic H5N1, from poultry in live bird markets in Hanoi, Vietnam, in 2001. *J Virol* **79:** 4201-4212.

8. Rogers, GN, Paulson, JC (1983). Receptor determinants of human and animal influenza virus isolates: differences in receptor specificity of the H3 hemagglutinin based on species of origin. *Virology* **127:** 361-373.

9. Tumpey, TM, Maines, TR, Van Hoeven, N, Glaser, L, Solorzano, A, Pappas, C, Cox, NJ, Swayne, DE, Palese, P, Katz, JM, Garcia-Sastre, A (2007). A two-amino acid change in the hemagglutinin of the 1918 influenza virus abolishes transmission. *Science* **315:** 655-659.

10. Srinivasan, A, Viswanathan, K, Raman, R, Chandrasekaran, A, Raguram, S, Tumpey, TM, Sasisekharan, V, Sasisekharan, R (2008). Quantitative biochemical rationale for differences in transmissibility of 1918 pandemic influenza A viruses. *Proc Natl Acad Sci U S A* **105:** 2800-2805.

11. Chandrasekaran, A, Srinivasan, A, Raman, R, Viswanathan, K, Raguram, S, Tumpey, TM, Sasisekharan, V, Sasisekharan, R (2008). Glycan topology determines human adaptation of avian H5N1 virus hemagglutinin. *Nat Biotechnol* **26:** 107-113.

12. Arnberg, N, Edlund, K, Kidd, AH, Wadell, G (2000). Adenovirus type 37 uses sialic acid as a cellular receptor. *J Virol* **74:** 42-48.

13. Schwegmann-Wessels, C, Herrler, G (2006). Sialic acids as receptor determinants for coronaviruses. *Glycoconj J* **23:** 51-58.

14. Kunkel, F, Herrler, G (1993). Structural and functional analysis of the surface protein of human coronavirus OC43. *Virology* **195:** 195-202.

15. Maisner, A, Schneider-Schaulies, J, Liszewski, MK, Atkinson, JP, Herrler, G (1994). Binding of measles virus to membrane cofactor protein (CD46): importance of disulfide bonds and N-glycans for the receptor function. *J Virol* **68:** 6299-6304.

16. Markwell, MA, Svennerholm, L, Paulson, JC (1981). Specific gangliosides function as host cell receptors for Sendai virus. *Proc Natl Acad Sci U S A* **78:** 5406-5410.

17. Connaris, H, Takimoto, T, Russell, R, Crennell, S, Moustafa, I, Portner, A, Taylor, G (2002). Probing the sialic acid binding site of the hemagglutinin-neuraminidase of Newcastle disease virus: identification of key amino acids involved in cell binding, catalysis, and fusion. *J Virol* **76:** 1816-1824.

18. Nokhbeh, MR, Hazra, S, Alexander, DA, Khan, A, McAllister, M, Suuronen, EJ, Griffith, M, Dimock, K (2005). Enterovirus 70 binds to different glycoconjugates containing alpha2,3-linked sialic acid on different cell lines. *J Virol* **79:** 7087-7094.

19. Uncapher, CR, DeWitt, CM, Colonno, RJ (1991). The major and minor group receptor families contain all but one human rhinovirus serotype. *Virology* **180:** 814-817.

20. Blomqvist, S, Savolainen, C, Raman, L, Roivainen, M, Hovi, T (2002). Human rhinovirus 87 and enterovirus 68 represent a unique serotype with rhinovirus and enterovirus features. *J Clin Microbiol* **40:** 4218-4223.

21. Paul, RW, Choi, AH, Lee, PW (1989). The alpha-anomeric form of sialic acid is the minimal receptor determinant recognized by reovirus. *Virology* **172:** 382-385.

22. Gentsch, JR, Pacitti, AF (1987). Differential interaction of reovirus type 3 with sialylated receptor components on animal cells. *Virology* **161:** 245-248.

23. Pacitti, AF, Gentsch, JR (1987). Inhibition of reovirus type 3 binding to host cells by sialylated glycoproteins is mediated through the viral attachment protein. *J Virol* **61:** 1407-1415.

24. Ciarlet, M, Ludert, JE, Iturriza-Gomara, M, Liprandi, F, Gray, JJ, Desselberger, U, Estes, MK (2002). Initial interaction of rotavirus strains with N-acetylneuraminic (sialic) acid

residues on the cell surface correlates with VP4 genotype, not species of origin. *J Virol* **76:** 4087-4095.

25. Ciarlet, M, Crawford, SE, Cheng, E, Blutt, SE, Rice, DA, Bergelson, JM, Estes, MK (2002). VLA-2 (alpha2beta1) integrin promotes rotavirus entry into cells but is not necessary for rotavirus attachment. *J Virol* **76:** 1109-1123.

26. Svensson, L (1992). Group C rotavirus requires sialic acid for erythrocyte and cell receptor binding. *J Virol* **66:** 5582-5585.

27. Liu, CK, Wei, G, Atwood, WJ (1998). Infection of glial cells by the human polyomavirus JC is mediated by an N-linked glycoprotein containing terminal alpha(2-6)-linked sialic acids. *J Virol* **72:** 4643-4649.

28. Sinibaldi, L, Goldoni, P, Pietropaolo, V, Longhi, C, Orsi, N (1990). Involvement of gangliosides in the interaction between BK virus and Vero cells. *Arch Virol* **113:** 291-296.

29. Hutson, AM, Atmar, RL, Marcus, DM, Estes, MK (2003). Norwalk virus-like particle hemagglutination by binding to h histo-blood group antigens. *J Virol* **77:** 405-415.

30. Hutson, AM, Atmar, RL, Estes, MK (2004). Norovirus disease: changing epidemiology and host susceptibility factors. *Trends Microbiol* **12:** 279-287.

31. Vongchan, P, Warda, M, Toyoda, H, Toida, T, Marks, RM, Linhardt, RJ (2005). Structural characterization of human liver heparan sulfate. *Biochim Biophys Acta* **1721:** 1-8.

32. Barth, H, Schafer, C, Adah, MI, Zhang, F, Linhardt, RJ, Toyoda, H, Kinoshita-Toyoda, A, Toida, T, Van Kuppevelt, TH, Depla, E, Von Weizsacker, F, Blum, HE, Baumert, TF (2003). Cellular binding of hepatitis C virus envelope glycoprotein E2 requires cell surface heparan sulfate. *J Biol Chem* **278:** 41003-41012.

33. Bissig, KD, Wieland, SF, Tran, P, Isogawa, M, Le, TT, Chisari, FV, Verma, IM (2010). Human liver chimeric mice provide a model for hepatitis B and C virus infection and treatment. *J Clin Invest* **120:** 924-930.

34. Jiang, J, Cun, W, Wu, X, Shi, Q, Tang, H, Luo, G (2012). Hepatitis C virus attachment mediated by apolipoprotein E binding to cell surface heparan sulfate. *J Virol* **86:** 7256-7267.

35. Kalia, M, Chandra, V, Rahman, SA, Sehgal, D, Jameel, S (2009). Heparan sulfate proteoglycans are required for cellular binding of the hepatitis E virus ORF2 capsid protein and for viral infection. *J Virol* **83:** 12714-12724.

36. Vanlandschoot, P, Van Houtte, F, Serruys, B, Leroux-Roels, G (2005). The arginine-rich carboxy-terminal domain of the hepatitis B virus core protein mediates attachment of nucleocapsids to cell-surface-expressed heparan sulfate. *J Gen Virol* **86:** 75-84.

37. Berns K, Parrish CR (2007). Parvoviridae. In: Knipe DM and Howley PM (eds). *Fields Virology,* 5th edn. Lippincott Williams & Wilkins:Philadelphia, PA, pp 2437-2477.

38. Schmidt, M, Chiorini, JA (2006). Gangliosides are essential for bovine adeno-associated virus entry. *J Virol* **80:** 5516-5522.

39. Chipman, PR, Agbandje-McKenna, M, Kajigaya, S, Brown, KE, Young, NS, Baker, TS, Rossmann, MG (1996). Cryo-electron microscopy studies of empty capsids of human parvovirus B19 complexed with its cellular receptor. *Proc Natl Acad Sci U S A* **93:** 7502-7506.

40. Summerford, C, Samulski, RJ (1998). Membrane-associated heparan sulfate proteoglycan is a receptor for adeno-associated virus type 2 virions. *J Virol* **72:** 1438-1445.

41. Handa, A, Muramatsu, S, Qiu, J, Mizukami, H, Brown, KE (2000). Adeno-associated virus (AAV)-3-based vectors transduce haematopoietic cells not susceptible to transduction with AAV-2-based vectors. *J Gen Virol* **81:** 2077-2084.

42. Rabinowitz, JE, Rolling, F, Li, C, Conrath, H, Xiao, W, Xiao, X, Samulski, RJ (2002). Cross-packaging of a single adeno-associated virus (AAV) type 2 vector genome into multiple AAV serotypes enables transduction with broad specificity. *J Virol* **76:** 791-801.

43. Wu, Z, Asokan, A, Grieger, JC, Govindasamy, L, Agbandje-McKenna, M, Samulski, RJ (2006). Single amino acid changes can influence titer, heparin binding, and tissue tropism in different adeno-associated virus serotypes. *J Virol* **80:** 11393-11397.

44. Wu, Z, Miller, E, Agbandje-McKenna, M, Samulski, RJ (2006). Alpha2,3 and alpha2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *J Virol* **80:** 9093-9103.

45. Walters, RW, Yi, SM, Keshavjee, S, Brown, KE, Welsh, MJ, Chiorini, JA, Zabner, J (2001). Binding of adeno-associated virus type 5 to 2,3-linked sialic acid is required for gene transfer. *J Biol Chem* **276:** 20610-20616.

46. Kaludov, N, Brown, KE, Walters, RW, Zabner, J, Chiorini, JA (2001). Adeno-associated virus serotype 4 (AAV4) and AAV5 both require sialic acid binding for hemagglutination and efficient transduction but differ in sialic acid linkage specificity. *J Virol* **75:** 6884-6893.

47. Walters, RW, Pilewski, JM, Chiorini, JA, Zabner, J (2002). Secreted and transmembrane mucins inhibit gene transfer with AAV4 more efficiently than AAV5. *J Biol Chem* **277:** 23709-23713.

48. Nam, HJ, Gurda-Whitaker, B, Gan, WY, Ilaria, S, McKenna, R, Mehta, P, Alvarez, RA, Agbandje-McKenna, M (2006). Identification of the sialic acid structures recognized by

minute virus of mice and the role of binding affinity in virulence adaptation. *J Biol Chem* **281:** 25670-25677.

49. Lopez-Bueno, A, Rubio, MP, Bryant, N, McKenna, R, Agbandje-McKenna, M, Almendral, JM (2006). Host-selected amino acid changes at the sialic acid binding pocket of the parvovirus capsid modulate cell binding affinity and determine virulence. *J Virol* **80:** 1563-1573.

50. Kontou, M, Govindasamy, L, Nam, HJ, Bryant, N, Llamas-Saiz, AL, Foces-Foces, C, Hernando, E, Rubio, MP, McKenna, R, Almendral, JM, Agbandje-McKenna, M (2005). Structural determinants of tissue tropism and in vivo pathogenicity for the parvovirus minute virus of mice. *J Virol* **79:** 10931-10943.

51. Shen, S, Bryant, K, Sun, J, Brown, S, Troupes, A, Pulicherla, N, Asokan, A (2012). Glycan binding avidity determines the systemic fate of adeno-associated virus 9. *J Virol*

52. Shen, S, Bryant, KD, Brown, SM, Randell, SH, Asokan, A (2011). Terminal N-linked galactose is the primary receptor for adeno-associated virus 9. *J Biol Chem* **286:** 13532-13540.

53. Bell, CL, Vandenberghe, LH, Bell, P, Limberis, MP, Gao, GP, Van Vliet, K, Agbandje-McKenna, M, Wilson, JM (2011). The AAV9 receptor and its modification to improve in vivo lung gene transfer in mice. *J Clin Invest* **121:** 2427-2435.

54. Srivastava, A, Lusby, EW, Berns, KI (1983). Nucleotide sequence and organization of the adeno-associated virus 2 genome. *J Virol* **45:** 555-564.

55. Horowitz, ED, Rahman, KS, Bower, BD, Dismuke, DJ, Falvo, MR, Griffith, JD, Harvey, SC, Asokan, A (2013). Biophysical and ultrastructural characterization of adeno-associated virus capsid uncoating and genome release. *J Virol* **87:** 2994-3002.

56. Lusby, E, Fife, KH, Berns, KI (1980). Nucleotide sequence of the inverted terminal repetition in adeno-associated virus DNA. *J Virol* **34:** 402-409.

57. Wang, XS, Ponnazhagan, S, Srivastava, A (1995). Rescue and replication signals of the adeno-associated virus 2 genome. *J Mol Biol* **250:** 573-580.

58. Kyostio, SR, Owens, RA, Weitzman, MD, Antoni, BA, Chejanovsky, N, Carter, BJ (1994). Analysis of adeno-associated virus (AAV) wild-type and mutant Rep proteins for their abilities to negatively regulate AAV p5 and p19 mRNA levels. *J Virol* **68:** 2947-2957.

59. King, JA, Dubielzig, R, Grimm, D, Kleinschmidt, JA (2001). DNA helicase-mediated packaging of adeno-associated virus type 2 genomes into preformed capsids. *EMBO J* **20:** 3282-3291.

60. Pereira, DJ, McCarty, DM, Muzyczka, N (1997). The adeno-associated virus (AAV) Rep protein acts as both a repressor and an activator to regulate AAV transcription during a productive infection. *J Virol* **71:** 1079-1088.

61. Parrish, CR (2010). Structures and functions of parvovirus capsids and the process of cell infection. *Curr Top Microbiol Immunol* **343:** 149-176.

62. Stahnke, S, Lux, K, Uhrig, S, Kreppel, F, Hosel, M, Coutelle, O, Ogris, M, Hallek, M, Buning, H (2011). Intrinsic phospholipase A2 activity of adeno-associated virus is involved in endosomal escape of incoming particles. *Virology* **409:** 77-83.

63. Girod, A, Wobus, CE, Zadori, Z, Ried, M, Leike, K, Tijssen, P, Kleinschmidt, JA, Hallek, M (2002). The VP1 capsid protein of adeno-associated virus type 2 is carrying a phospholipase A2 domain required for virus infectivity. *J Gen Virol* **83:** 973-978.

64. Sonntag, F, Schmidt, K, Kleinschmidt, JA (2010). A viral assembly factor promotes AAV2 capsid formation in the nucleolus. *Proc Natl Acad Sci U S A* **107:** 10220-10225.

65. Naumer, M, Sonntag, F, Schmidt, K, Nieto, K, Panke, C, Davey, NE, Popa-Wagner, R, Kleinschmidt, JA (2012). Properties of the adeno-associated virus assembly-activating protein. *J Virol* **86:** 13038-13048.

66. Sonntag, F, Kother, K, Schmidt, K, Weghofer, M, Raupp, C, Nieto, K, Kuck, A, Gerlach, B, Bottcher, B, Muller, OJ, Lux, K, Horer, M, Kleinschmidt, JA (2011). The assembly-activating protein promotes capsid assembly of different adeno-associated virus serotypes. *J Virol* **85:** 12686-12697.

67. Harrison SC (2007). Principles of Virus Structure. In: Knipe DM and Howley PM (eds). *Fields Virology,* Fifth Eds. edn. Lippincott, Williams & Wilkins:Philadelphia, PA, pp 59-98.

68. CRICK, FH, WATSON, JD (1956). Structure of small viruses. *Nature* **177:** 473-475.

69. Crick F, Watson JD (1957) Virus Structure: General Principles. 5-18

70. CASPAR, DL (1956). Structure of bushy stunt virus. *Nature* **177:** 475-476.

71. KLUG, A, FINCH, JT, FRANKLIN, RE (1957). Structure of turnip yellow mosaic virus. *Nature* **179:** 683-684.

72. Harrison, SC, Olson, AJ, Schutt, CE, Winkler, FK, Bricogne, G (1978). Tomato bushy stunt virus at 2.9 A resolution. *Nature* **276:** 368-373.

73. Rossmann, MG, Johnson, JE (1989). Icosahedral RNA virus structure. *Annu Rev Biochem* **58:** 533-573.

74. CASPAR, DL, KLUG, A (1962). Physical principles in the construction of regular viruses. *Cold Spring Harb Symp Quant Biol* **27:** 1-24.

75. Kronenberg, S, Bottcher, B, von der Lieth, CW, Bleker, S, Kleinschmidt, JA (2005). A conformational change in the adeno-associated virus type 2 capsid leads to the exposure of hidden VP1 N termini. *J Virol* **79:** 5296-5303.

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

77. Wu, P, Xiao, W, Conlon, T, Hughes, J, Agbandje-McKenna, M, Ferkol, T, Flotte, T, Muzyczka, N (2000). 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.

78. Bleker, S, Pawlita, M, Kleinschmidt, JA (2006). Impact of capsid conformation and Rep-capsid interactions on adeno-associated virus type 2 genome packaging. *J Virol* **80:** 810-820.

79. DiPrimio, N, Asokan, A, Govindasamy, L, Agbandje-McKenna, M, Samulski, RJ (2008). Surface loop dynamics in adeno-associated virus capsid assembly. *J Virol* **82:** 5178-5189.

80. DiPrimio N (2009) Surface Loop Dynamics of a Parvovirus.

81. Xie, Q, Bu, W, Bhatia, S, Hare, J, Somasundaram, T, Azzi, A, Chapman, MS (2002). The atomic structure of adeno-associated virus (AAV-2), a vector for human gene therapy. *Proc Natl Acad Sci U S A* **99:** 10405-10410.

82. Govindasamy, L, Padron, E, McKenna, R, Muzyczka, N, Kaludov, N, Chiorini, JA, Agbandje-McKenna, M (2006). Structurally mapping the diverse phenotype of adeno-associated virus serotype 4. *J Virol* **80:** 11556-11570.

83. O'Donnell, J, Taylor, KA, Chapman, MS (2009). Adeno-associated virus-2 and its primary cellular receptor--Cryo-EM structure of a heparin complex. *Virology* **385:** 434-443.

84. Zhong, L, Li, B, Mah, CS, Govindasamy, L, Agbandje-McKenna, M, Cooper, M, Herzog, RW, Zolotukhin, I, Warrington, KH,Jr, Weigel-Van Aken, KA, Hobbs, JA, Zolotukhin, S, Muzyczka, N, Srivastava, A (2008). Next generation of adeno-associated virus 2 vectors: point mutations in tyrosines lead to high-efficiency transduction at lower doses. *Proc Natl Acad Sci U S A* **105:** 7827-7832.

85. Qiao, C, Zhang, W, Yuan, Z, Shin, JH, Li, J, Jayandharan, GR, Zhong, L, Srivastava, A, Xiao, X, Duan, D (2010). Adeno-associated virus serotype 6 capsid tyrosine-to-phenylalanine mutations improve gene transfer to skeletal muscle. *Hum Gene Ther* **21:** 1343-1348.

86. Horowitz, ED, Finn, MG, Asokan, A (2012). Tyrosine cross-linking reveals interfacial dynamics in adeno-associated viral capsids during infection. *ACS Chem Biol* **7:** 1059-1066.

87. Bowles, DE, McPhee, SW, Li, C, Gray, SJ, Samulski, JJ, Camp, AS, Li, J, Wang, B, Monahan, PE, Rabinowitz, JE, Grieger, JC, Govindasamy, L, Agbandje-McKenna, M, Xiao, X, Samulski, RJ (2012). Phase 1 gene therapy for Duchenne muscular dystrophy using a translational optimized AAV vector. *Mol Ther* **20:** 443-455.

88. Li, C, Diprimio, N, Bowles, DE, Hirsch, ML, Monahan, PE, Asokan, A, Rabinowitz, J, Agbandje-McKenna, M, Samulski, RJ (2012). Single amino acid modification of adeno-associated virus capsid changes transduction and humoral immune profiles. *J Virol* **86:** 7752-7759.

89. Kern, A, Schmidt, K, Leder, C, Muller, OJ, Wobus, CE, Bettinger, K, Von der Lieth, CW, King, JA, Kleinschmidt, JA (2003). Identification of a heparin-binding motif on adeno-associated virus type 2 capsids. *J Virol* **77:** 11072-11081.

90. Opie, SR, Warrington, KH,Jr, Agbandje-McKenna, M, Zolotukhin, S, Muzyczka, N (2003). 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.

91. Lochrie, MA, Tatsuno, GP, Christie, B, McDonnell, JW, Zhou, S, Surosky, R, Pierce, GF, Colosi, P (2006). Mutations on the external surfaces of adeno-associated virus type 2 capsids that affect transduction and neutralization. *J Virol* **80:** 821-834.

92. Levy, HC, Bowman, VD, Govindasamy, L, McKenna, R, Nash, K, Warrington, K, Chen, W, Muzyczka, N, Yan, X, Baker, TS, Agbandje-McKenna, M (2009). Heparin binding induces conformational changes in Adeno-associated virus serotype 2. *J Struct Biol* **165:** 146-156.

93. Lerch, TF, Chapman, MS (2012). Identification of the heparin binding site on adeno-associated virus serotype 3B (AAV-3B). *Virology* **423:** 6-13.

94. Bell, CL, Gurda, BL, Van Vliet, K, Agbandje-McKenna, M, Wilson, JM (2012). Identification of the galactose binding domain of the AAV9 capsid. *J Virol*

95. Pulicherla, N, Shen, S, Yadav, S, Debbink, K, Govindasamy, L, Agbandje-McKenna, M, Asokan, A (2011). Engineering liver-detargeted AAV9 vectors for cardiac and musculoskeletal gene transfer. *Mol Ther* **19:** 1070-1078.

96. Asokan, A, Hamra, JB, Govindasamy, L, Agbandje-McKenna, M, Samulski, RJ (2006). Adeno-associated virus type 2 contains an integrin alpha5beta1 binding domain essential for viral cell entry. *J Virol* **80:** 8961-8969.

97. Akache, B, Grimm, D, Pandey, K, Yant, SR, Xu, H, Kay, MA (2006). The 37/67-kilodalton laminin receptor is a receptor for adeno-associated virus serotypes 8, 2, 3, and 9. *J Virol* **80:** 9831-9836.

98. Gurda, BL, Raupp, C, Popa-Wagner, R, Naumer, M, Olson, NH, Ng, R, McKenna, R, Baker, TS, Kleinschmidt, JA, Agbandje-McKenna, M (2012). Mapping a neutralizing epitope onto the capsid of adeno-associated virus serotype 8. *J Virol* **86:** 7739-7751.

99. Qing, K, Mah, C, Hansen, J, Zhou, S, Dwarki, V, Srivastava, A (1999). Human fibroblast growth factor receptor 1 is a co-receptor for infection by adeno-associated virus 2. *Nat Med* **5:** 71-77.

100. Kashiwakura, Y, Tamayose, K, Iwabuchi, K, Hirai, Y, Shimada, T, Matsumoto, K, Nakamura, T, Watanabe, M, Oshimi, K, Daida, H (2005). Hepatocyte growth factor receptor is a coreceptor for adeno-associated virus type 2 infection. *J Virol* **79:** 609-614.

101. Ling, C, Lu, Y, Kalsi, JK, Jayandharan, GR, Li, B, Ma, W, Cheng, B, Gee, SW, McGoogan, KE, Govindasamy, L, Zhong, L, Agbandje-McKenna, M, Srivastava, A (2010). Human hepatocyte growth factor receptor is a cellular coreceptor for adeno-associated virus serotype 3. *Hum Gene Ther* **21:** 1741-1747.

102. Weller, ML, Amornphimoltham, P, Schmidt, M, Wilson, PA, Gutkind, JS, Chiorini, JA (2010). Epidermal growth factor receptor is a co-receptor for adeno-associated virus serotype 6. *Nat Med* **16:** 662-664.

103. Di Pasquale, G, Davidson, BL, Stein, CS, Martins, I, Scudiero, D, Monks, A, Chiorini, JA (2003). Identification of PDGFR as a receptor for AAV-5 transduction. *Nat Med* **9:** 1306-1312.

104. Hansen, J, Qing, K, Kwon, HJ, Mah, C, Srivastava, A (2000). Impaired intracellular trafficking of adeno-associated virus type 2 vectors limits efficient transduction of murine fibroblasts. *J Virol* **74:** 992-996.

105. Duan, D, Li, Q, Kao, AW, Yue, Y, Pessin, JE, Engelhardt, JF (1999). Dynamin is required for recombinant adeno-associated virus type 2 infection. *J Virol* **73:** 10371-10376.

106. Bartlett, JS, Wilcher, R, Samulski, RJ (2000). Infectious entry pathway of adeno-associated virus and adeno-associated virus vectors. *J Virol* **74:** 2777-2785.

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

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

109. Duan, D, Yue, Y, Yan, Z, Yang, J, Engelhardt, JF (2000). Endosomal processing limits gene transfer to polarized airway epithelia by adeno-associated virus. *J Clin Invest* **105:** 1573-1587.

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

111. Yan, Z, Lei-Butters, DC, Keiser, NW, Engelhardt, JF (2013). Distinct transduction difference between adeno-associated virus type 1 and type 6 vectors in human polarized airway epithelia. *Gene Ther* **20:** 328-337.

112. Ding, W, Zhang, LN, Yeaman, C, Engelhardt, JF (2006). rAAV2 traffics through both the late and the recycling endosomes in a dose-dependent fashion. *Mol Ther* **13:** 671-682.

113. Ding, W, Zhang, L, Yan, Z, Engelhardt, JF (2005). Intracellular trafficking of adeno-associated viral vectors. *Gene Ther* **12:** 873-880.

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

115. Sonntag, F, Bleker, S, Leuchs, B, Fischer, R, Kleinschmidt, JA (2006). Adeno-associated virus type 2 capsids with externalized VP1/VP2 trafficking domains are generated prior to passage through the cytoplasm and are maintained until uncoating occurs in the nucleus. *J Virol* **80:** 11040-11054.

116. Zhong, L, Zhao, W, Wu, J, Li, B, Zolotukhin, S, Govindasamy, L, Agbandje-McKenna, M, Srivastava, A (2007). A dual role of EGFR protein tyrosine kinase signaling in ubiquitination of AAV2 capsids and viral second-strand DNA synthesis. *Mol Ther* **15:** 1323-1330.

117. Zhong, L, Li, B, Jayandharan, G, Mah, CS, Govindasamy, L, Agbandje-McKenna, M, Herzog, RW, Weigel-Van Aken, KA, Hobbs, JA, Zolotukhin, S, Muzyczka, N, Srivastava, A (2008). Tyrosine-phosphorylation of AAV2 vectors and its consequences on viral intracellular trafficking and transgene expression. *Virology* **381:** 194-202.

118. Sanlioglu, S, Benson, PK, Yang, J, Atkinson, EM, Reynolds, T, Engelhardt, JF (2000). Endocytosis and nuclear trafficking of adeno-associated virus type 2 are controlled by rac1 and phosphatidylinositol-3 kinase activation. *J Virol* **74:** 9184-9196.

119. Zhao, W, Zhong, L, Wu, J, Chen, L, Qing, K, Weigel-Kelley, KA, Larsen, SH, Shou, W, Warrington, KH,Jr, Srivastava, A (2006). Role of cellular FKBP52 protein in intracellular trafficking of recombinant adeno-associated virus 2 vectors. *Virology* **353:** 283-293.

120. Xiao, C, Rossmann, MG (2007). Interpretation of electron density with stereographic roadmap projections. *J Struct Biol* **158:** 182-187.

121. Hirosue, S, Senn, K, Clement, N, Nonnenmacher, M, Gigout, L, Linden, RM, Weber, T (2007). Effect of inhibition of dynein function and microtubule-altering drugs on AAV2 transduction. *Virology* **367:** 10-18.

122. Kelkar, S, De, BP, Gao, G, Wilson, JM, Crystal, RG, Leopold, PL (2006). A common mechanism for cytoplasmic dynein-dependent microtubule binding shared among adeno-associated virus and adenovirus serotypes. *J Virol* **80:** 7781-7785.

123. Xiao, W, Warrington, KH,Jr, Hearing, P, Hughes, J, Muzyczka, N (2002). Adenovirus-facilitated nuclear translocation of adeno-associated virus type 2. *J Virol* **76:** 11505-11517.

124. Gao, G, Vandenberghe, LH, Alvira, MR, Lu, Y, Calcedo, R, Zhou, X, Wilson, JM (2004). Clades of Adeno-associated viruses are widely disseminated in human tissues. *J Virol* **78:** 6381-6388.

125. Zincarelli, C, Soltys, S, Rengo, G, Rabinowitz, JE (2008). Analysis of AAV serotypes 1-9 mediated gene expression and tropism in mice after systemic injection. *Mol Ther* **16:** 1073-1080.

126. Burger, C, Gorbatyuk, OS, Velardo, MJ, Peden, CS, Williams, P, Zolotukhin, S, Reier, PJ, Mandel, RJ, Muzyczka, N (2004). Recombinant AAV viral vectors pseudotyped with viral capsids from serotypes 1, 2, and 5 display differential efficiency and cell tropism after delivery to different regions of the central nervous system. *Mol Ther* **10:** 302-317.

127. Cearley, CN, Wolfe, JH (2006). Transduction characteristics of adeno-associated virus vectors expressing cap serotypes 7, 8, 9, and Rh10 in the mouse brain. *Mol Ther* **13:** 528-537.

128. Klein, RL, Dayton, RD, Tatom, JB, Henderson, KM, Henning, PP (2008). AAV8, 9, Rh10, Rh43 vector gene transfer in the rat brain: effects of serotype, promoter and purification method. *Mol Ther* **16:** 89-96.

129. White, E, Bienemann, A, Sena-Esteves, M, Taylor, H, Bunnun, C, Castrique, E, Gill, S (2011). Evaluation and optimization of the administration of recombinant adeno-associated viral vectors (serotypes 2/1, 2/2, 2/rh8, 2/9, and 2/rh10) by convection-enhanced delivery to the striatum. *Hum Gene Ther* **22:** 237-251.

130. Muller, OJ, Kaul, F, Weitzman, MD, Pasqualini, R, Arap, W, Kleinschmidt, JA, Trepel, M (2003). Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors. *Nat Biotechnol* **21:** 1040-1046.

131. Koerber, JT, Maheshri, N, Kaspar, BK, Schaffer, DV (2006). Construction of diverse adeno-associated viral libraries for directed evolution of enhanced gene delivery vehicles. *Nat Protoc* **1:** 701-706.

132. Maheshri, N, Koerber, JT, Kaspar, BK, Schaffer, DV (2006). Directed evolution of adeno-associated virus yields enhanced gene delivery vectors. *Nat Biotechnol* **24:** 198-204.

133. Kwon, I, Schaffer, DV (2008). Designer gene delivery vectors: molecular engineering and evolution of adeno-associated viral vectors for enhanced gene transfer. *Pharm Res* **25:** 489-499.

134. Excoffon, KJ, Koerber, JT, Dickey, DD, Murtha, M, Keshavjee, S, Kaspar, BK, Zabner, J, Schaffer, DV (2009). Directed evolution of adeno-associated virus to an infectious respiratory virus. *Proc Natl Acad Sci U S A* **106:** 3865-3870.

135. Dickey, DD, Excoffon, KJ, Koerber, JT, Bergen, J, Steines, B, Klesney-Tait, J, Schaffer, DV, Zabner, J (2011). Enhanced sialic acid-dependent endocytosis explains the increased efficiency of infection of airway epithelia by a novel adeno-associated virus. *J Virol* **85:** 9023-9030.

136. Koerber, JT, Klimczak, R, Jang, JH, Dalkara, D, Flannery, JG, Schaffer, DV (2009). Molecular evolution of adeno-associated virus for enhanced glial gene delivery. *Mol Ther* **17:** 2088-2095.

137. Dalkara, D, Kolstad, KD, Guerin, KI, Hoffmann, NV, Visel, M, Klimczak, RR, Schaffer, DV, Flannery, JG (2011). AAV mediated GDNF secretion from retinal glia slows down retinal degeneration in a rat model of retinitis pigmentosa. *Mol Ther* **19:** 1602-1608.

138. Klimczak, RR, Koerber, JT, Dalkara, D, Flannery, JG, Schaffer, DV (2009). A novel adeno-associated viral variant for efficient and selective intravitreal transduction of rat Muller cells. *PLoS One* **4:** e7467.

139. Dickey DD, Bergen J, Nassar B, Steines B, Excoffon KJ, Schaffer DV, Zabner J (2011) Directed Evolution of AAV in Pig Airway Epithelia *In Vivo*. *Mol. Ther.;* **19:** S251

140. Jang, JH, Koerber, JT, Kim, JS, Asuri, P, Vazin, T, Bartel, M, Keung, A, Kwon, I, Park, KI, Schaffer, DV (2011). An evolved adeno-associated viral variant enhances gene delivery and gene targeting in neural stem cells. *Mol Ther* **19:** 667-675.

141. Li, W, Zhang, L, Johnson, JS, Zhijian, W, Grieger, JC, Ping-Jie, X, Drouin, LM, Agbandje-McKenna, M, Pickles, RJ, Samulski, RJ (2009). Generation of novel AAV variants by directed evolution for improved CFTR delivery to human ciliated airway epithelium. *Mol Ther* **17:** 2067-2077.

142. Gray, SJ, Blake, BL, Criswell, HE, Nicolson, SC, Samulski, RJ, McCown, TJ, Li, W (2010). Directed evolution of a novel adeno-associated virus (AAV) vector that crosses the seizure-compromised blood-brain barrier (BBB). *Mol Ther* **18:** 570-578.

143. Asokan, A, Conway, JC, Phillips, JL, Li, C, Hegge, J, Sinnott, R, Yadav, S, DiPrimio, N, Nam, HJ, Agbandje-McKenna, M, McPhee, S, Wolff, J, Samulski, RJ (2010). Reengineering a receptor footprint of adeno-associated virus enables selective and systemic gene transfer to muscle. *Nat Biotechnol* **28:** 79-82.

144. Bowles DE, Rabinowitz JE, Samulski RJ (2006). The genus *Dependovirus*. In: Kerr JR, Cotmore SF, Bloom ME, Linden RM and Parrish CR (eds). *Parvoviruses,* Edward Arnold Ltd.:New York, pp 15-24.

145. ATCHISON, RW, CASTO, BC, HAMMON, WM (1965). Adenovirus-Associated Defective Virus Particles. *Science* **149:** 754-756.

146. Hoggan, MD, Blacklow, NR, Rowe, WP (1966). Studies of small DNA viruses found in various adenovirus preparations: physical, biological, and immunological characteristics. *Proc Natl Acad Sci U S A* **55:** 1467-1474.

147. Parks, WP, Green, M, Pina, M, Melnick, JL (1967). Physicochemical characterization of adeno-associated satellite virus type 4 and its nucleic acid. *J Virol* **1:** 980-987.

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

149. Mori, S, Wang, L, Takeuchi, T, Kanda, T (2004). Two novel adeno-associated viruses from cynomolgus monkey: pseudotyping characterization of capsid protein. *Virology* **330:** 375-383.

150. Hermonat, PL, Quirk, JG, Bishop, BM, Han, L (1997). The packaging capacity of adeno-associated virus (AAV) and the potential for wild-type-plus AAV gene therapy vectors. *FEBS Lett* **407:** 78-84.

151. Samulski, RJ, Chang, LS, Shenk, T (1989). Helper-free stocks of recombinant adeno-associated viruses: normal integration does not require viral gene expression. *J Virol* **63:** 3822-3828.

152. Asokan, A, Schaffer, DV, Jude Samulski, R (2012). The AAV Vector Toolkit: Poised at the Clinical Crossroads. *Mol Ther* **20:** 699-708.

153. Hogle, JM, Chow, M, Filman, DJ (1985). Three-dimensional structure of poliovirus at 2.9 A resolution. *Science* **229:** 1358-1365.

154. Akhtar, J, Shukla, D (2009). Viral entry mechanisms: cellular and viral mediators of herpes simplex virus entry. *FEBS J* **276:** 7228-7236.

155. Dechecchi, MC, Melotti, P, Bonizzato, A, Santacatterina, M, Chilosi, M, Cabrini, G (2001). Heparan sulfate glycosaminoglycans are receptors sufficient to mediate the initial binding of adenovirus types 2 and 5. *J Virol* **75:** 8772-8780.

156. Schmidt, M, Govindasamy, L, Afione, S, Kaludov, N, Agbandje-McKenna, M, Chiorini, JA (2008). Molecular characterization of the heparin-dependent transduction domain on the capsid of a novel adeno-associated virus isolate, AAV(VR-942). *J Virol* **82:** 8911-8916.

157. Viswanathan, K, Chandrasekaran, A, Srinivasan, A, Raman, R, Sasisekharan, V, Sasisekharan, R (2010). Glycans as receptors for influenza pathogenesis. *Glycoconj J* **27:** 561-570.

158. Guglielmi, KM, Johnson, EM, Stehle, T, Dermody, TS (2006). Attachment and cell entry of mammalian orthoreovirus. *Curr Top Microbiol Immunol* **309:** 1-38.

159. Neu, U, Stehle, T, Atwood, WJ (2009). The Polyomaviridae: Contributions of virus structure to our understanding of virus receptors and infectious entry. *Virology* **384:** 389-399.

160. Wu, Z, Miller, E, Agbandje-McKenna, M, Samulski, RJ (2006). Alpha2,3 and alpha2,6 N-linked sialic acids facilitate efficient binding and transduction by adeno-associated virus types 1 and 6. *J Virol* **80:** 9093-9103.

161. Mueller, C, Flotte, TR (2008). Clinical gene therapy using recombinant adeno-associated virus vectors. *Gene Ther* **15:** 858-863.

162. Gao, G, Vandenberghe, LH, Wilson, JM (2005). New recombinant serotypes of AAV vectors. *Curr Gene Ther* **5:** 285-297.

163. Mitchell, AM, Nicolson, SC, Warischalk, JK, Samulski, RJ (2010). AAV's Anatomy: Roadmap for Optimizing Vectors for Translational Success. *Curr Gene Ther* **10:** 319-340.

164. Cohen, M, Varki, A (2010). The sialome--far more than the sum of its parts. *OMICS* **14:** 455-464.

165. Ng, R, Govindasamy, L, Gurda, BL, McKenna, R, Kozyreva, OG, Samulski, RJ, Parent, KN, Baker, TS, Agbandje-McKenna, M (2010). Structural characterization of the dual glycan binding adeno-associated virus serotype 6. *J Virol* **84:** 12945-12957.

166. Grieger, JC, Choi, VW, Samulski, RJ (2006). Production and characterization of adeno-associated viral vectors. *Nat Protoc* **1:** 1412-1428.

167. Deutscher, SL, Nuwayhid, N, Stanley, P, Briles, EI, Hirschberg, CB (1984). Translocation across Golgi vesicle membranes: a CHO glycosylation mutant deficient in CMP-sialic acid transport. *Cell* **39:** 295-299.

168. Deutscher, SL, Hirschberg, CB (1986). Mechanism of galactosylation in the Golgi apparatus. A Chinese hamster ovary cell mutant deficient in translocation of UDP-galactose across Golgi vesicle membranes. *J Biol Chem* **261:** 96-100.

169. Stanley, P, Chaney, W (1985). Control of carbohydrate processing: the lec1A CHO mutation results in partial loss of N-acetylglucosaminyltransferase I activity. *Mol Cell Biol* **5:** 1204-1211.

170. Elbein, AD, Solf, R, Dorling, PR, Vosbeck, K (1981). Swainsonine: an inhibitor of glycoprotein processing. *Proc Natl Acad Sci U S A* **78:** 7393-7397.

171. Kuan, SF, Byrd, JC, Basbaum, C, Kim, YS (1989). Inhibition of mucin glycosylation by aryl-N-acetyl-alpha-galactosaminides in human colon cancer cells. *J Biol Chem* **264:** 19271-19277.

172. Wang, WC, Cummings, RD (1988). The immobilized leukoagglutinin from the seeds of Maackia amurensis binds with high affinity to complex-type Asn-linked oligosaccharides containing terminal sialic acid-linked alpha-2,3 to penultimate galactose residues. *J Biol Chem* **263:** 4576-4585.

173. Bai, X, Brown, JR, Varki, A, Esko, JD (2001). Enhanced 3-O-sulfation of galactose in Asn-linked glycans and Maackia amurensis lectin binding in a new Chinese hamster ovary cell line. *Glycobiology* **11:** 621-632.

174. Shibuya, N, Goldstein, IJ, Broekaert, WF, Nsimba-Lubaki, M, Peeters, B, Peumans, WJ (1987). The elderberry (Sambucus nigra L.) bark lectin recognizes the Neu5Ac(alpha 2-6)Gal/GalNAc sequence. *J Biol Chem* **262:** 1596-1601.

175. Wu, AM, Wu, JH, Tsai, MS, Yang, Z, Sharon, N, Herp, A (2007). Differential affinities of Erythrina cristagalli lectin (ECL) toward monosaccharides and polyvalent mammalian structural units. *Glycoconj J* **24:** 591-604.

176. Yamamoto, K, Tsuji, T, Matsumoto, I, Osawa, T (1981). Structural requirements for the binding of oligosaccharides and glycopeptides to immobilized wheat germ agglutinin. *Biochemistry* **20:** 5894-5899.

177. Ohyama, Y, Kasai, K, Nomoto, H, Inoue, Y (1985). Frontal affinity chromatography of ovalbumin glycoasparagines on a concanavalin A-sepharose column. A quantitative study of the binding specificity of the lectin. *J Biol Chem* **260:** 6882-6887.

178. Stults, NL, Fechheimer, M, Cummings, RD (1989). Relationship between Golgi architecture and glycoprotein biosynthesis and transport in Chinese hamster ovary cells. *J Biol Chem* **264:** 19956-19966.

179. North, SJ, Huang, HH, Sundaram, S, Jang-Lee, J, Etienne, AT, Trollope, A, Chalabi, S, Dell, A, Stanley, P, Haslam, SM (2010). Glycomics profiling of Chinese hamster ovary cell glycosylation mutants reveals N-glycans of a novel size and complexity. *J Biol Chem* **285:** 5759-5775.

180. Kornegay, JN, Li, J, Bogan, JR, Bogan, DJ, Chen, C, Zheng, H, Wang, B, Qiao, C, Howard, JF,Jr, Xiao, X (2010). Widespread muscle expression of an AAV9 human mini-dystrophin vector after intravenous injection in neonatal dystrophin-deficient dogs. *Mol Ther* **18:** 1501-1508.

181. Pacak, CA, Mah, CS, Thattaliyath, BD, Conlon, TJ, Lewis, MA, Cloutier, DE, Zolotukhin, I, Tarantal, AF, Byrne, BJ (2006). Recombinant adeno-associated virus serotype 9 leads to preferential cardiac transduction in vivo. *Circ Res* **99:** e3-9.

182. Baker, AH, Mcvey, JH, Waddington, SN, Di Paolo, NC, Shayakhmetov, DM (2007). The influence of blood on in vivo adenovirus bio-distribution and transduction. *Mol Ther* **15:** 1410-1416.

183. Nemerow, GR, Pache, L, Reddy, V, Stewart, PL (2009). Insights into adenovirus host cell interactions from structural studies. *Virology* **384:** 380-388.

184. Hennet, T, Chui, D, Paulson, JC, Marth, JD (1998). Immune regulation by the ST6Gal sialyltransferase. *Proc Natl Acad Sci U S A* **95:** 4504-4509.

185. Martin, LT, Marth, JD, Varki, A, Varki, NM (2002). Genetically altered mice with different sialyltransferase deficiencies show tissue-specific alterations in sialylation and sialic acid 9-O-acetylation. *J Biol Chem* **277:** 32930-32938.

186. Wu, Z, Asokan, A, Samulski, RJ (2006). Adeno-associated virus serotypes: vector toolkit for human gene therapy. *Mol Ther* **14:** 316-327.

187. Hundt, C, Peyrin, JM, Haik, S, Gauczynski, S, Leucht, C, Rieger, R, Riley, ML, Deslys, JP, Dormont, D, Lasmezas, CI, Weiss, S (2001). Identification of interaction domains of the prion protein with its 37-kDa/67-kDa laminin receptor. *EMBO J* **20:** 5876-5886.

188. Malakhov, MP, Aschenbrenner, LM, Smee, DF, Wandersee, MK, Sidwell, RW, Gubareva, LV, Mishin, VP, Hayden, FG, Kim, DH, Ing, A, Campbell, ER, Yu, M, Fang, F (2006). Sialidase fusion protein as a novel broad-spectrum inhibitor of influenza virus infection. *Antimicrob Agents Chemother* **50:** 1470-1479.

189. Triana-Baltzer, GB, Gubareva, LV, Nicholls, JM, Pearce, MB, Mishin, VP, Belser, JA, Chen, LM, Chan, RW, Chan, MC, Hedlund, M, Larson, JL, Moss, RB, Katz, JM, Tumpey, TM, Fang, F (2009). Novel pandemic influenza A(H1N1) viruses are potently inhibited by DAS181, a sialidase fusion protein. *PLoS One* **4:** e7788.

190. Mountney, A, Zahner, MR, Lorenzini, I, Oudega, M, Schramm, LP, Schnaar, RL (2010). Sialidase enhances recovery from spinal cord contusion injury. *Proc Natl Acad Sci U S A* **107:** 11561-11566.

191. Viswanathan, K, Koh, X, Chandrasekaran, A, Pappas, C, Raman, R, Srinivasan, A, Shriver, Z, Tumpey, TM, Sasisekharan, R (2010). Determinants of glycan receptor specificity of H2N2 influenza A virus hemagglutinin. *PLoS One* **5:** e13768.

192. Neu, U, Bauer, J, Stehle, T (2011). Viruses and sialic acids: rules of engagement. *Curr Opin Struct Biol* **21:** 610-618.

193. Agbandje-McKenna, M, Kleinschmidt, J (2011). AAV capsid structure and cell interactions. *Methods Mol Biol* **807:** 47-92.

194. Nonnenmacher, M, Weber, T (2012). Intracellular transport of recombinant adeno-associated virus vectors. *Gene Ther*

195. Mingozzi, F, High, KA (2011). Therapeutic in vivo gene transfer for genetic disease using AAV: progress and challenges. *Nat Rev Genet* **12:** 341-355.

196. Vandenberghe, LH, Breous, E, Nam, HJ, Gao, G, Xiao, R, Sandhu, A, Johnston, J, Debyser, Z, Agbandje-McKenna, M, Wilson, JM (2009). Naturally occurring singleton residues in AAV capsid impact vector performance and illustrate structural constraints. *Gene Ther* **16:** 1416-1428.

197. Lerch, TF, O'Donnell, JK, Meyer, NL, Xie, Q, Taylor, KA, Stagg, SM, Chapman, MS (2012). Structure of AAV-DJ, a Retargeted Gene Therapy Vector: Cryo-Electron Microscopy at 4.5 A Resolution. *Structure*

198. Xie, Q, Lerch, TF, Meyer, NL, Chapman, MS (2011). Structure-function analysis of receptor-binding in adeno-associated virus serotype 6 (AAV-6). *Virology* **420:** 10-19.

199. Bevan, AK, Duque, S, Foust, KD, Morales, PR, Braun, L, Schmelzer, L, Chan, CM, McCrate, M, Chicoine, LG, Coley, BD, Porensky, PN, Kolb, SJ, Mendell, JR, Burghes, AH, Kaspar, BK (2011). Systemic gene delivery in large species for targeting spinal cord, brain, and peripheral tissues for pediatric disorders. *Mol Ther* **19:** 1971-1980.

200. Ganesan, LP, Mohanty, S, Kim, J, Clark, KR, Robinson, JM, Anderson, CL (2011). Rapid and efficient clearance of blood-borne virus by liver sinusoidal endothelium. *PLoS Pathog* **7:** e1002281.

201. Perrier D, Gibaldi M (1982). Multicompartment Models. In: Anonymous *Pharmacokinetics,* 2nd edn. Marcel Dekker, INC.:New York, pp 45-111.

202. Inagaki, K, Fuess, S, Storm, TA, Gibson, GA, Mctiernan, CF, Kay, MA, Nakai, H (2006). Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Mol Ther* **14:** 45-53.

203. Kotchey, NM, Adachi, K, Zahid, M, Inagaki, K, Charan, R, Parker, RS, Nakai, H (2011). A potential role of distinctively delayed blood clearance of recombinant adeno-associated virus serotype 9 in robust cardiac transduction. *Mol Ther* **19:** 1079-1089.

204. Smedsrod, B, Pertoft, H, Gustafson, S, Laurent, TC (1990). Scavenger functions of the liver endothelial cell. *Biochem J* **266:** 313-327.

205. Seternes, T, Sorensen, K, Smedsrod, B (2002). Scavenger endothelial cells of vertebrates: a nonperipheral leukocyte system for high-capacity elimination of waste macromolecules. *Proc Natl Acad Sci U S A* **99:** 7594-7597.

206. Di Paolo, NC, Shayakhmetov, DM (2009). Adenovirus de-targeting from the liver. *Curr Opin Mol Ther* **11:** 523-531.

207. Waddington, SN, McVey, JH, Bhella, D, Parker, AL, Barker, K, Atoda, H, Pink, R, Buckley, SM, Greig, JA, Denby, L, Custers, J, Morita, T, Francischetti, IM, Monteiro, RQ, Barouch, DH, van Rooijen, N, Napoli, C, Havenga, MJ, Nicklin, SA, Baker, AH (2008). Adenovirus serotype 5 hexon mediates liver gene transfer. *Cell* **132:** 397-409.

208. Kalyuzhniy, O, Di Paolo, NC, Silvestry, M, Hofherr, SE, Barry, MA, Stewart, PL, Shayakhmetov, DM (2008). Adenovirus serotype 5 hexon is critical for virus infection of hepatocytes in vivo. *Proc Natl Acad Sci U S A* **105:** 5483-5488.

209. Varki A, Schauer R (2009). Sialic Acids. In: Varki A, Cummings RD, Esko JD, Freeze HH, Stanley P, Bertozzi CR, et al (eds). *Essentials of Glycobiology,* 2nd edn. The Consortium of Glycobiology Editors, La Jolla, California:Cold Spring Harbor (NY),

210. Altheide, TK, Hayakawa, T, Mikkelsen, TS, Diaz, S, Varki, N, Varki, A (2006). System-wide genomic and biochemical comparisons of sialic acid biology among primates and rodents: Evidence for two modes of rapid evolution. *J Biol Chem* **281:** 25689-25702.

211. Halbert, CL, Allen, JM, Miller, AD (2001). Adeno-associated virus type 6 (AAV6) vectors mediate efficient transduction of airway epithelial cells in mouse lungs compared to that of AAV2 vectors. *J Virol* **75:** 6615-6624.

212. DiMattia, MA, Nam, HJ, Van Vliet, K, Mitchell, M, Bennett, A, Gurda, BL, McKenna, R, Olson, NH, Sinkovits, RS, Potter, M, Byrne, BJ, Aslanidi, G, Zolotukhin, S, Muzyczka, N,

Baker, TS, Agbandje-McKenna, M (2012). Structural insight into the unique properties of adeno-associated virus serotype 9. *J Virol* **86:** 6947-6958.

213. Arnold, K, Bordoli, L, Kopp, J, Schwede, T (2006). The SWISS-MODEL workspace: a web-based environment for protein structure homology modelling. *Bioinformatics* **22:** 195-201.

214. Carrillo-Tripp, M, Shepherd, CM, Borelli, IA, Venkataraman, S, Lander, G, Natarajan, P, Johnson, JE, Brooks, CL,3rd, Reddy, VS (2009). VIPERdb2: an enhanced and web API enabled relational database for structural virology. *Nucleic Acids Res* **37:** D436-42.

215. Gao, GP, Alvira, MR, Wang, L, Calcedo, R, Johnston, J, Wilson, JM (2002). Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy. *Proc Natl Acad Sci U S A* **99:** 11854-11859.

216. Inagaki, K, Fuess, S, Storm, TA, Gibson, GA, Mctiernan, CF, Kay, MA, Nakai, H (2006). Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8. *Mol Ther* **14:** 45-53.

217. Wang, J, Xie, J, Lu, H, Chen, L, Hauck, B, Samulski, RJ, Xiao, W (2007). Existence of transient functional double-stranded DNA intermediates during recombinant AAV transduction. *Proc Natl Acad Sci U S A* **104:** 13104-13109.

218. Wang, L, Bell, P, Lin, J, Calcedo, R, Tarantal, AF, Wilson, JM (2011). AAV8-mediated hepatic gene transfer in infant rhesus monkeys (Macaca mulatta). *Mol Ther* **19:** 2012-2020.

219. Imai, M, Watanabe, T, Hatta, M, Das, SC, Ozawa, M, Shinya, K, Zhong, G, Hanson, A, Katsura, H, Watanabe, S, Li, C, Kawakami, E, Yamada, S, Kiso, M, Suzuki, Y, Maher, EA, Neumann, G, Kawaoka, Y (2012). Experimental adaptation of an influenza H5 HA confers respiratory droplet transmission to a reassortant H5 HA/H1N1 virus in ferrets. *Nature* **486:** 420-428.

220. Herfst, S, Schrauwen, EJ, Linster, M, Chutinimitkul, S, de Wit, E, Munster, VJ, Sorrell, EM, Bestebroer, TM, Burke, DF, Smith, DJ, Rimmelzwaan, GF, Osterhaus, AD, Fouchier, RA (2012). Airborne transmission of influenza A/H5N1 virus between ferrets. *Science* **336:** 1534-1541.

221. High, KA (2012). The gene therapy journey for hemophilia: are we there yet? *Blood*

222. McCarty, DM (2008). Self-complementary AAV vectors; advances and applications. *Mol Ther* **16:** 1648-1656.

223. Du, J, Meledeo, MA, Wang, Z, Khanna, HS, Paruchuri, VD, Yarema, KJ (2009). Metabolic glycoengineering: sialic acid and beyond. *Glycobiology* **19:** 1382-1401.

224. Bai, X, Wei, G, Sinha, A, Esko, JD (1999). Chinese hamster ovary cell mutants defective in glycosaminoglycan assembly and glucuronosyltransferase I. *J Biol Chem* **274:** 13017-13024.

225. Bame, KJ, Zhang, L, David, G, Esko, JD (1994). Sulphated and undersulphated heparan sulphate proteoglycans in a Chinese hamster ovary cell mutant defective in N-sulphotransferase. *Biochem J* **303 (Pt 1):** 81-87.

226. Lidholt, K, Weinke, JL, Kiser, CS, Lugemwa, FN, Bame, KJ, Cheifetz, S, Massague, J, Lindahl, U, Esko, JD (1992). A single mutation affects both N-acetylglucosaminyltransferase and glucuronosyltransferase activities in a Chinese hamster ovary cell mutant defective in heparan sulfate biosynthesis. *Proc Natl Acad Sci U S A* **89:** 2267-2271.

227. Zhang, L, Esko, JD (1995). Accumulation of a pentasaccharide terminating in alpha-N-acetylglucosamine in an animal cell mutant defective in heparan sulfate biosynthesis. *J Biol Chem* **270:** 12557-12562.

228. Potvin, B, Kumar, R, Howard, DR, Stanley, P (1990). Transfection of a human alpha-(1,3)fucosyltransferase gene into Chinese hamster ovary cells. Complications arise from activation of endogenous alpha-(1,3)fucosyltransferases. *J Biol Chem* **265:** 1615-1622.

229. Zuckerkandl E, Pauling LB (1962). Molecular disease, evolution, and genetic heterogeneity. In: Kasha M and Pullman B (eds). *Horizons in Biochemistry*, Academic Press:New York, pp 189-225.

230. Padron, E, Bowman, V, Kaludov, N, Govindasamy, L, Levy, H, Nick, P, McKenna, R, Muzyczka, N, Chiorini, JA, Baker, TS, Agbandje-McKenna, M (2005). Structure of adeno-associated virus type 4. *J Virol* **79:** 5047-5058.

231. Davidson, BL, Stein, CS, Heth, JA, Martins, I, Kotin, RM, Derksen, TA, Zabner, J, Ghodsi, A, Chiorini, JA (2000). Recombinant adeno-associated virus type 2, 4, and 5 vectors: transduction of variant cell types and regions in the mammalian central nervous system. *Proc Natl Acad Sci U S A* **97:** 3428-3432.

232. Dodge, JC, Treleaven, CM, Fidler, JA, Hester, M, Haidet, A, Handy, C, Rao, M, Eagle, A, Matthews, JC, Taksir, TV, Cheng, SH, Shihabuddin, LS, Kaspar, BK (2010). AAV4-mediated expression of IGF-1 and VEGF within cellular components of the ventricular system improves survival outcome in familial ALS mice. *Mol Ther* **18:** 2075-2084.

233. Bunn-Moreno, MM, Campos-Neto, A (1981). Lectin(s) extracted from seeds of Artocarpus integrifolia (jackfruit): potent and selective stimulator(s) of distinct human T and B cell functions. *J Immunol* **127:** 427-429.

234. Hortin, GL (1990). Isolation of glycopeptides containing O-linked oligosaccharides by lectin affinity chromatography on jacalin-agarose. *Anal Biochem* **191:** 262-267.

235. Lerch, TF, Xie, Q, Chapman, MS (2010). The structure of adeno-associated virus serotype 3B (AAV-3B): insights into receptor binding and immune evasion. *Virology* **403:** 26-36.

236. Walters, RW, Agbandje-McKenna, M, Bowman, VD, Moninger, TO, Olson, NH, Seiler, M, Chiorini, JA, Baker, TS, Zabner, J (2004). Structure of adeno-associated virus serotype 5. *J Virol* **78:** 3361-3371.

237. Nam, HJ, Lane, MD, Padron, E, Gurda, B, McKenna, R, Kohlbrenner, E, Aslanidi, G, Byrne, B, Muzyczka, N, Zolotukhin, S, Agbandje-McKenna, M (2007). Structure of adeno-associated virus serotype 8, a gene therapy vector. *J Virol* **81:** 12260-12271.

238. Tomassini, JE, Maxson, TR, Colonno, RJ (1989). Biochemical characterization of a glycoprotein required for rhinovirus attachment. *J Biol Chem* **264:** 1656-1662.

239. Xie, Q, Chapman, MS (1996). Canine parvovirus capsid structure, analyzed at 2.9 A resolution. *J Mol Biol* **264:** 497-520.

240. Agbandje-McKenna, M, Llamas-Saiz, AL, Wang, F, Tattersall, P, Rossmann, MG (1998). Functional implications of the structure of the murine parvovirus, minute virus of mice. *Structure* **6:** 1369-1381.