RATIONAL DESIGN OF SPLIT GENE VECTORS TO EXPAND THE PACKAGING CAPACITY OF ADENO-ASSOCIATED VIRAL VECTORS

A Dissertation Presented to the Faculty of the Graduate School

University of Missouri-Columbia

In Partial Fulfillment Of the Requirements for the Degree

Doctor of Philosophy

---

by Arkasubhra Ghosh

Dr. Dongsheng Duan, Dissertation Supervisor

December 2007

The undersigned, appointed by the dean of the Graduate School, have examined the dissertation entitled

RATIONAL DESIGN OF SPLIT GENE VECTORS TO EXPAND THE PACKAGING CAPACITY OF ADENO-ASSOCIATED VIRAL VECTORS.

presented by Arkasubhra Ghosh,

a candidate for the degree of doctor of philosophy,

and hereby certify that, in their opinion, it is worthy of acceptance.

Professor Dongsheng Duan

Professor David Pintel

Professor John Cannon

Professor Mark Hannink

Professor David Lee

DEDICATION

I dedicate this thesis to my wife, Anuprita, for putting up with my unearthly work hours and thoroughly undisciplined lifestyle and still loving me and supporting me every step of the way. I dedicate this thesis to my son, Advin for his unconditional love and being my stress-buster and making life and work so much more worthwhile.

ACKNOWLEDGEMENTS

I would like to express my sincere gratitude to Dr. Dongsheng Duan for his instruction and support that has helped me to develop as a scientist. I would also like to thank all the past and present Duan lab members for their support and friendship, especially Yongping Yue, for teaching me all the techniques and generally making our life easier in the lab. I thank Yi Lai, Brian Bostick, Dejia Li, Mingju Liu and Chun long for their help and discussion during the course of this thesis work. I thank the faculty and staff in the Department of Molecular Microbiology and Immunology for their advice and support, especially the members of my committee, Dr. David Pintel, Dr. John Cannon, Dr. David Lee and Dr. Mark Hannink from Biochemistry.

I want to thank Wendell Mitchell and late Julia Mitchell for welcoming us to this new country with open arms, being always there for us and loving us. A special thanks to all the office staff, Jana, Shelley, Karen, Leesa and Kathy for putting up with me and helping me out whenever I needed.

I want to thank all the friends that we made here in the department and in Columbia who have made every bit of our stay as enjoyable and fun as it could be.

I would especially like to thank my mother-in-law Sunanda Ringangaonkar for taking the time and trouble to come all the way here from India to take care of

us whenever we needed, a huge support for my timely completion of this thesis. I am grateful for her unconditional affection and extended discussions on the science of life.

Most importantly, my heartfelt thanks to my father Dr. Prajit Ghosh and my mother Ranjana Ghosh, for their love and sacrifice that helped me to come all this way. A huge thanks to my younger brother Rajarshi for always having faith in me and encouraging me. The long and extended discussions and arguments I had with my father, my wife Anuprita and my brother on the science we do helped me develop a more mature scientific outlook and I thank them for that.

TABLE OF CONTENTS

LIST OF ILLUSTRATIONS

Figure **Page**

LIST OF ABBREVIATIONS

AAV, adeno-associated virus

Ad, adenovirus

AP, alkaline phosphatase

BL6, C57/BL6 wild type mice

BL10, C57BL/10 wild type mice

BMD, Becker muscular dystrophy

CF, cystic fibrosis

CFTR, cystic fibrosis transmembrane regulator

CMV, cytomegalovirus

DMD, Duchenne muscular dystrophy

eGFP, enhanced green fluorescent protein

ITR, inverted terminal repeat

moi, multiplicity of infection

mdx, dystrophin null mice.

m-dko, myoD/dystrophin double knock-out

SD, splice donor

SA, splice acceptor

SV40, simian virus 40

RSV, Rous sarcoma virus

ts, trans-splicing

vg, vector genome

Abstract

Adeno associated viruses (AAV) have recently been demonstrated as a very promising gene delivery vehicle. But the limited packaging capacity of AAV vectors (~4.7kb) hinders their application for diseases involving large genes such as those responsible for Duchenne muscular dystrophy and cystic fibrosis. To overcome this hurdle, the trans-splicing and overlapping dual vector methods were developed recently to expand the packaging capacity of AAV. These methods involve the simultaneous delivery of two AAV vectors carrying different segments of a larger gene which would reconstitute the complete gene within the cell. In the “trans-splicing approach”, the gene expression cassette is split into two segments each containing a splice donor or a splice acceptor. The transgene is expressed from heterodimeric concatamers which are formed due to inverted terminal repeat (ITR) mediated recombination. In the other method, termed the “overlapping approach,” the two split parts of the transgene share a common sequence or overlap. Therefore, upon co-infection, the upstream and the downstream parts can hybridize at the homologous overlapping sequence and undergo recombination to regenerate the intact gene expression cassette.
It has been demonstrated that certain expression limiting barriers affect transduction from these dual vectors. The trans-splicing method requires an optimal gene splitting site and the overlapping method requires a highly recombinogenic domain in the middle of the gene for high levels of transduction.

Thus, due to these rate-limiting barriers, the trans-splicing or overlapping dual vector systems are applicable only to certain genes. To overcome these limitations of dual vectors, we developed a novel trans-splicing/overlapping hybrid vector system that can efficiently reconstitute any large gene. In the hybrid system, we engineered a 0.8kb highly recombinogenic alkaline phosphatase (AP) gene fragment at the end of the splice donor and beginning of the splice acceptor sequences in the trans-splicing vectors. We hypothesized that the addition of the AP gene fragment will facilitate homologous recombination in the trans-splicing vectors and increase the chance of transgene reconstitution. The experimental data demonstrate that the hybrid vector system improves gene expression compared to the traditional dual vectors. The study also demonstrates that the rationally designed trans-splicing AAV vectors can be successfully used for body-wide gene delivery. Such high levels of gene expression from these hybrid dual vectors are suitable for application to therapeutic gene transfer in disease models. Taken together, this study outlines the considerations to be taken into account for rational design of split gene vectors that would be capable of efficient transgene expression.

Chapter 1. Introduction

A. Adeno-associated virus

Adeno-associated Virus (AAV) is a human parvovirus originally isolated as a contaminant in adenoviral stocks [1]. It is classified as a dependovirus, as opposed to autonomous parvovirus because it needs a helper virus to complete a productive life cycle [2]. About 110 serotypes of AAV have been identified and they are classified into seven clades [3]. Type-2 AAV (AAV-2) was the first serotype isolated. This is also the most extensively studied prototype AAV. Wild type AAV-2 consists of a 4681nt single-stranded DNA genome and is non-pathogenic [4]. The ends of the genome are palindromic repeat sequences termed the ITRs. AAV genome consists of two open reading frames encoding structural/capsid proteins and replication/regulation proteins, respectively (Figure 1.1).

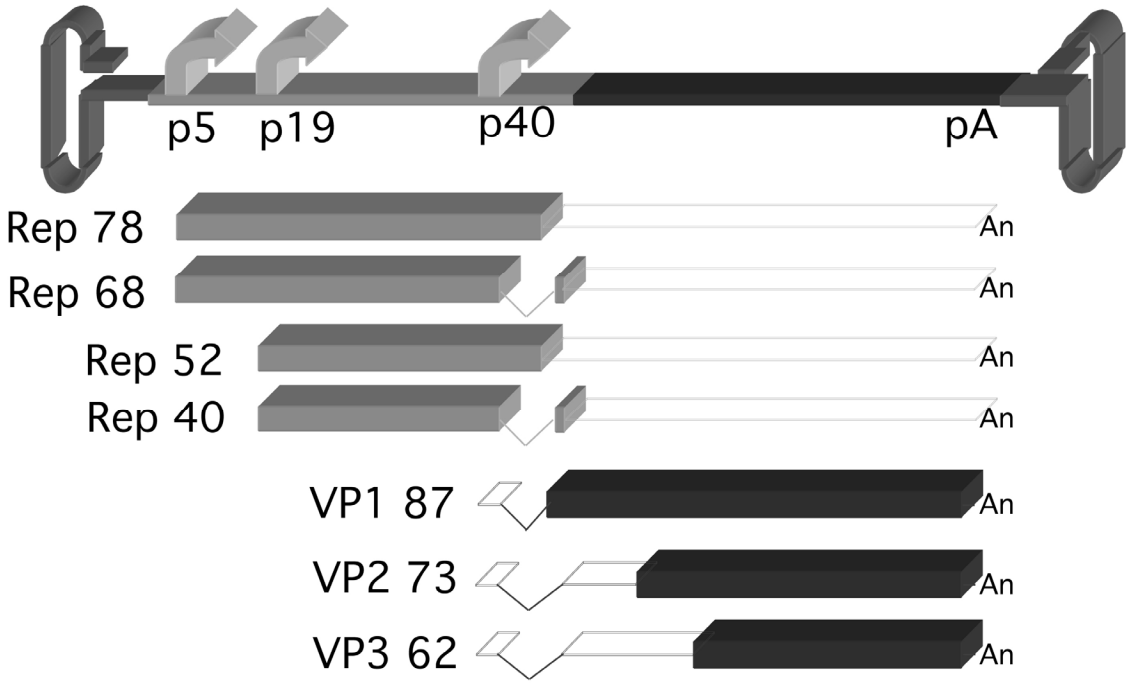


Figure 1.1. Organization of the wild-type AAV-2 genome. The wild type AAV genome has two open reading frames encoding four replication proteins (Rep78, Rep68, Rep52 and Rep40) and three structural proteins (VP1, VP2 and VP3; the numerical numbers indicate MW of each capsid protein). The replication proteins are expressed from the p5 and p19 promoters and undergo alternative splicing to generate different transcripts. The structural proteins are all expressed from the p40 promoter. They are generated by alternative splicing and/or different start codons. (This figure is contributed by Dr. Dongsheng Duan). (*Adapted from Ghosh, et al, Biotech. and Genetic Engg Rev., 2007, Vol.24)*.

AAV enters cells through receptor-mediated endocytosis followed by endosomal sorting and trafficking into the nucleus [5]. After uncoating, the AAV genome is converted to the double stranded form. In the presence of a helper virus, AAV is reproduced [6] (Figure 1.2). Without helper virus, a latent infection cycle is established and AAV is propagated as an integrated provirus [7]. In human cells, AAV-2 specifically integrates in the short arm of chromosome 19 [8].

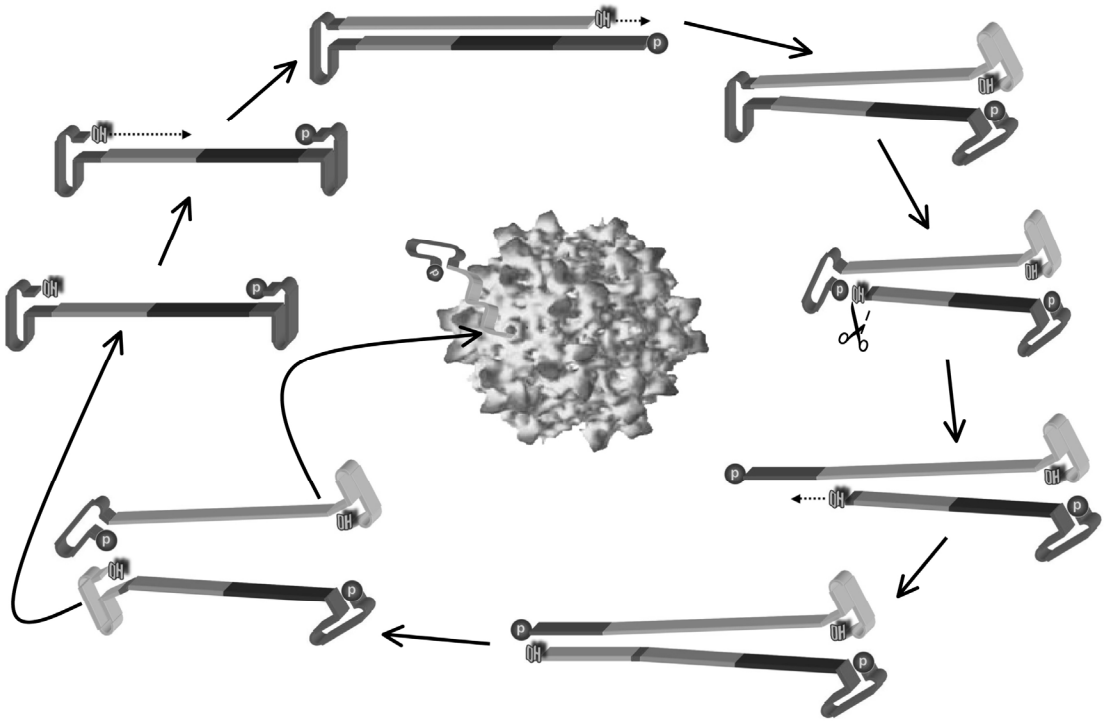


Figure 1.2. Replication and packaging of the AAV genome. AAV replication starts from a single-stranded template. Primed by the 3' hydroxy group at the end of the ITR, a second strand is synthesized. The second strand synthesis continues across the ITR on the other end, followed by resolution at the left-hand terminal resolution site (scissor sign). This allows completion of the left hand ITR replication. Finally a pair of positive and negative single-stranded AAV genomes is produced. These new genomes may now serve as templates for further rounds of amplification. Alternatively they can be packaged into pre-assembled empty capsid. (This figure is contributed by Dr. Dongsheng Duan). (*Adapted from Ghosh, et al, Biotech. and Genetic Engg Rev., 2007, Vol.24).*

B. Adeno-associated virus as a gene transfer vector

Recent advances in the field of gene therapy have led to a spurt in using recombinant AAV (rAAV) for gene delivery [9, 10]. In rAAV, the native AAV genes are replaced by the therapeutic gene(s) (Figure 1.3). After infection, rAAV mainly persists as a latent circular episomal genome in host cells [11]. These circular genomes are thought to be responsible for longterm gene expression [11, 12].

AAV vectors carry several features that are advantageous for gene therapy. These include their ability to infect both mitotic and post mitotic cells, longterm gene expression in a wide range of tissues and very low immune response. Many preclinical and clinical gene therapy studies have been performed with rAAV over the last two decades [9, 10]. Although most of the studies have used the prototype AAV-2, investigators have begun to appreciate the newly identified AAV serotypes as gene therapy vehicles over the past few years [13, 14]. These serotypes display unique transduction patterns in different tissues. One of the most striking findings is that some serotypes, in particular AAV-6, -8 and -9, can mediate efficient whole body gene transfer [15-18]. It is interesting to note that AAV-2 ITRs have been used as packaging signals in these studies. The serotype-specific differences in transduction are most likely attributed to the differences in viral uptake and/or intracellular trafficking.

Figure 1.3. Organization of a typical recombinant AAV (rAAV) genome. In a rAAV vector, an expression cassette containing a reporter or a therapeutic gene of choice replaces all the viral genes in the wild-type virus. Only the ITRs are kept to serve as the essential packaging signal. (This figure is contributed by Dr. Dongsheng Duan). (*Adapted from Ghosh, et al, Biotech. and Genetic Engg Rev., 2007, Vol.24).*

C. Strategies for expanding AAV packaging capacity

Despite its promise as an excellent therapeutic vehicle, AAV has been excluded for gene therapy of many diseases. This is primarily due to the small packaging capacity. Up to a 5 kb therapeutic expression cassette can be encapsidated in a rAAV virion, beyond which the packaging efficiency drops significantly [19]. This is not a problem for diseases involving small genes such as hemophilia B and alpha 1-antitrypsin deficiency [20, 21]. However, it cannot meet the need of a large therapeutic gene such as the dystrophin gene (the gene responsible for Duchenne and Becker muscular dystrophy, DMD and BMD) and the cystic fibrosis transmembrane conductance regulator (CFTR) gene (the gene responsible for cystic fibrosis).

To overcome this size limitation, efforts have been made to delete the less important regions in a therapeutic gene and to generate minimized genes that can fit into a single AAV capsid. However, such mini-genes may be less functional than the full-length gene [22, 23]. Alternative strategies are needed to deliver a large gene with AAV vector. It has been demonstrated that AAV genomes undergo intermolecular recombination at their ITRs to form larger concatameres [11, 24-28]. At the junction of the viral genomes is a unique structure called the double-D ITR [12]. It is further observed that the majority of the recombined genomes exist in a head-to-tail orientation [11, 12]. This feature of the AAV genome metabolism offers a great opportunity to double AAV packaging capacity with two vectors [29-31]. Several distinctive dual vector approaches have being developed since then. These include cis-activation and trans-splicing (Figure 1.4) [29, 30, 32]. In the cis-activation approach, the therapeutic expression cassette is split into two parts one containing the regulatory elements such as the promoter and the enhancer while the other carrying the protein coding sequence (Figure 1.4A). These two parts are packaged into independent AAV vectors. Co-administration will reconstitute the expression cassette. In trans-splicing approach, the gene itself is split into two AAV vectors (Figure 1.4B). The split gene fragments are engineered with splicing signals. After co-infection, a functional open reading frame is reconstituted by cellular splicing machinery from the recombined vector genome.

Besides viral ITR-mediated recombination, intermolecular recombination between AAV genomes also occurs through homologous recombination [33, 34].

This yields another option to expand AAV packaging capacity called the overlapping approach (Figure 1.4C) [35, 36]. Essentially, a large gene is split into two overlapping fragments and packaged into two AAV vectors. When these vectors enter the same cell, the full-length gene is reconstituted through homologous recombination at the overlapping region.

Cis-activation approach

The size of many therapeutic genes (such as the CFTR cDNA) is in the range of ~4.6 kb. These genes will fit into a single AAV virion themselves but there is no space for a strong promoter/enhancer to drive gene expression. The AAV ITR has weak promoter activity and it can drive minimal level expression in the absence of a strong exogenous promoter [37]. This is apparently not sufficient for gene therapy. One way to overcome this problem is to introduce strong enhancers through a second AAV vector to strengthen transcription. Duan et al tested this approach by constructing a “super-enhancer” AAV vector [29]. This vector contains two enhancer elements from the simian virus 40 (SV40) and cytomegalovirus (CMV) genomes, respectively. They also made a “promoter-less” luciferase AAV vector. Co-administration of the “super-enhancer” AAV vector resulted in 200-fold increase in transgene expression from the “promoter-less” vector. Molecular analysis of the retrieved vector genomes revealed intermolecular recombination between two incoming vector genomes and the formation of a strong expression cassette containing the ITR promoter, the luciferase gene and the super-enhancer. The cis-activation method has several

advantages for certain gene therapy applications. First, it does not entail the splitting of the gene within the coding sequence. Second, it works irrespective of the orientation of the recombined vector genome. However, there are also apparent limitations. One of the biggest limitations is that it cannot work for the genes where the coding sequence is beyond the packaging capacity of a single AAV. Safety is another important concern. Random integration of the super-enhancer elements may lead to oncogenic expression of the neighboring genes.

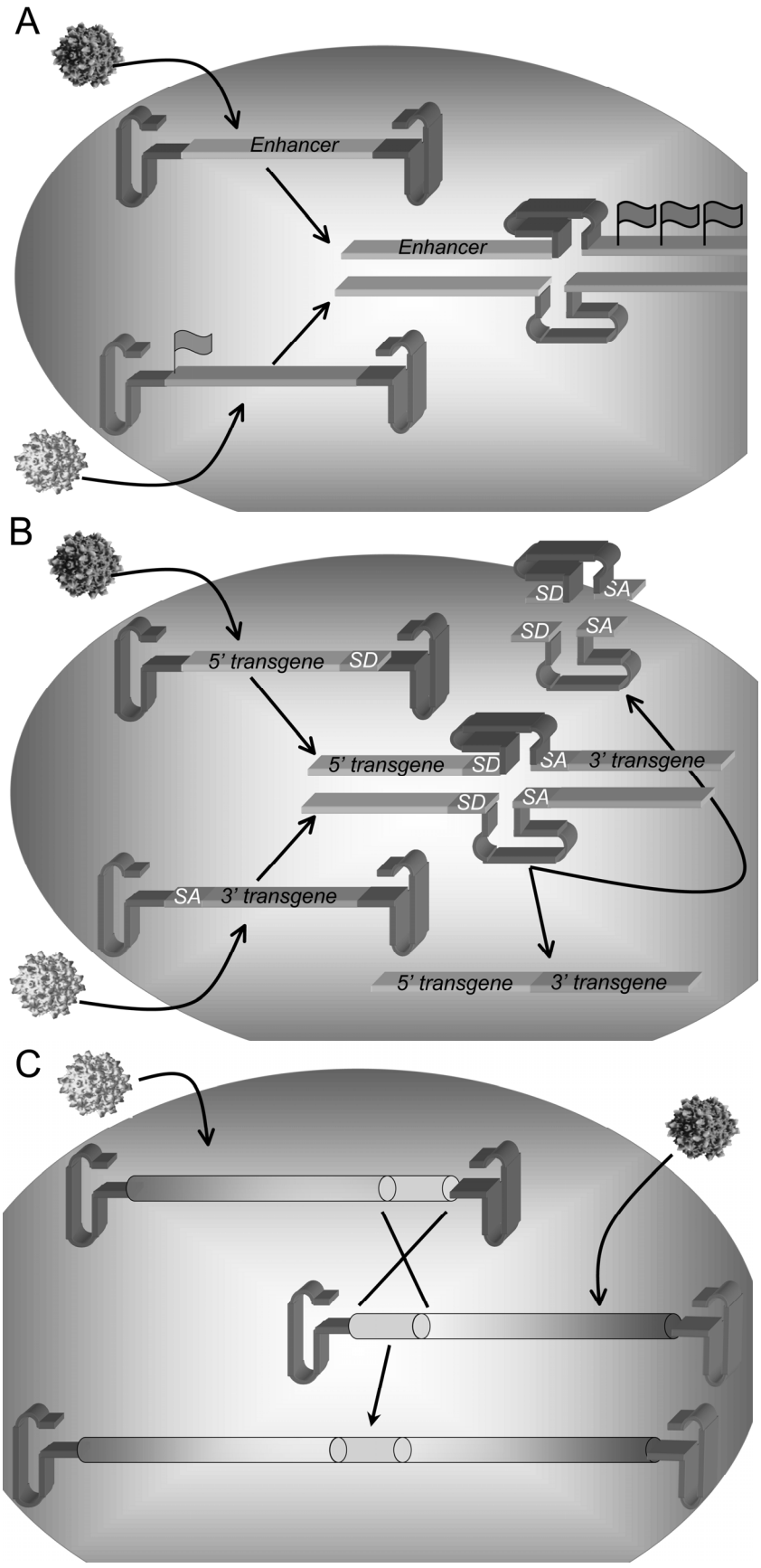

A
Enhancer
Enhancer
B
SD
SA
SD
SA
5' transgene
SD
5' transgene
SD
SA
3' transgene
SD
SA
SA
3' transgene
5' transgene
3' transgene
C

Figure 1.4. Schematic illustration of the different dual vector approaches to expand the AAV packaging capacity. A, Cis-activation. In this case, a vector carrying a promoterless transgene (or a transgene with a weak promoter) is co-infected with a vector containing the "super-enhancer". The rAAV genomes undergo ITR mediated recombination to form heterodimeric molecules. The transgene expression is highly enhanced by the presence of the "super-enhancer" elements in the concatamer. **B,** Trans-splicing. In this approach, the transgene is split into two halves in the middle of an intron. The AV.Donor molecule contains the 5' half of the expression cassette with the splice donor (SD) and the AV.Acceptor contains the 3' half of the expression cassette containing the splice acceptor (SA). Upon co-infection with both vectors, ITR mediated heterodimers are formed. The transgene is transcribed from the recombined vector genome. The splice signals from the intron will then splice out the ITR junction and generate a full length mature mRNA transcript. **C,** Overlapping. This method requires the transgene to be split in such a way that the two split parts share a sequence of homology between them. Upon co-infection, the two split parts will recombine at their homologous sequence. This recombination event will thus lead to the generation of an intact expression cassette. (This figure is contributed by Dr. Dongsheng Duan). (*Adapted from Ghosh, et al, Biotech. and Genetic Engg Rev., 2007, Vol.24).*

The trans-splicing vectors

Eukaryotic genes consist of protein coding regions called exons interspersed with non-coding regions called introns. At the junction of these introns and exons there are the splicing signals. These splicing signals direct the formation of the spliceosome complex at the appropriate locations. Subsequently, the introns are removed and mature mRNA is generated to make protein. The trans-splicing approach takes advantage of the cellular Mrna splicing process. Basically, the transgene is divided into two parts either at an endogenous intron or an engineered synthetic intron. The vector carrying the splicing donor signal is called the AV.Donor. This vector also carries the promoter and the 5' half of the transgene. The vector carrying the splicing acceptor signal is called the AV.Acceptor. This vector also carries the 3' half of the transgene and the polyadenylation signals. Upon co-infection, two vector genomes form the head-to-tail heterodimer. The RNA transcript from such a heterodimer contains both the 5' and the 3' halves of the transgene as well as splicing signals and the double-D ITR at the junction. The junction region is then spliced out from pre-mRNA to form mature mRNA for protein expression.

A number of therapeutic and/or reporter transgenes have been tested in the trans-splicing vectors [23, 30, 31, 35, 38-40]. Yan et al split the human erythopoitin gene at its 3rd intron and packaged them into the trans-splicing vectors [30]. Co-infection in mouse skeletal muscle resulted in therapeutic level of erythropoietin in the serum. Single vector administration did not lead to any erythropoietin expression. Chao et al generated the factor VIII gene trans-

splicing vectors [39]. In immune deficient mice, intra-portal vein co-administration produced 2% of normal level of factor VIII in blood. Several groups also split the LacZ reporter gene using the synthetic splicing signals [31, 35, 38, 40]. In these proof of principle studies, the trans-splicing vectors consistently regenerated the split gene and lead to the expression of functional protein. However, the actual transduction efficiency is much lower than that of a single intact AAV vector [35].

There are a number of potential expression limiting factors in the trans-splicing approach [23, 41, 42]. These include co-infection efficiency, formation of the head-to-tail heterodimer genome, transcription and the stability of the double-D ITR containing pre-mRNA and splicing across the double-D ITR structure (Figure 1.5). It is now clear that the co-infection is not a barrier. Even in a diseased tissue, co-infection efficiency can reach ≥ 90% [41]. To test whether unidirectional heterodimer formation is a rate-limiting step, Yan et al generated trans-splicing AAV vectors with AAV-2 ITR and AAV-5 ITR at the opposite ends of the genomes [42]. Recombination between AAV-2 ITR and AAV-5 ITR is very inefficient. This leads to preferential unidirectional recombination of the ITRs from the same serotype. Comparing with the traditional trans-splicing vectors which carry the same ITR at both ends, the hybrid ITR vectors resulted in 6~10-fold enhancement in gene expression. This study confirms that head-to-tail intermolecular recombination is a rate-limiting barrier.

To test whether the double-D ITR represents a critical rate-limiting factor, Xu et al generated a mini-cassette that resembles the recombined vector genome [41]. Essentially, a double-D ITR structure is inserted in the intron of a synthetic *lacZ* gene. Comparing with these of the original LacZ construct and an intron-containing LacZ construct, the total mRNA transcript level was significantly reduced in the presence of the double-D ITR structure. Pre-mRNA stability was also reduced by the double-D ITR structure. Ultimately, gene expression was reduced by 50% at the protein level. These results demonstrated that mRNA accumulation is also a rate-limiting factor.

An often neglected issue is the gene splitting site. Depending on the sequence specificity, different exon/intron/exon junctions are spliced at different efficiencies. Furthermore, surrounding exonic splicing enhancer elements may also modulate the overall level of mRNA production. To test whether the gene splitting site influences transduction efficiency of the trans-splicing vectors, Lai et al systematically screened a number of different exon/intron/exon junctions in a 6 kb mini-dystrophin gene [23]. At first, they evaluated mRNA accumulation in synthetic mini-cassettes that contain the double-D ITR structure. Two sets of the trans-splicing vectors were then generated based on the sites that yielded the highest levels of mRNA. In vivo examination in a mouse model of DMD revealed wide-spread mini-dystrophin expression from the vector set that performed the best in mRNA accumulation screening. The other vector set was developed on a different site that had ~50% mRNA level of the best site. This second vector set failed to yield therapeutic level transduction.

Taken together, tremendous progress has been made in the last few years in optimizing the trans-splicing vectors. Identification of the potential rate-limiting factors has resulted in novel approaches to bypass these barriers. Therapeutic levels of both secretary and structural proteins have been achieved. The development of the trans-splicing vectors has finally moved from the proof-of-principle stage to the gene therapy application stage.

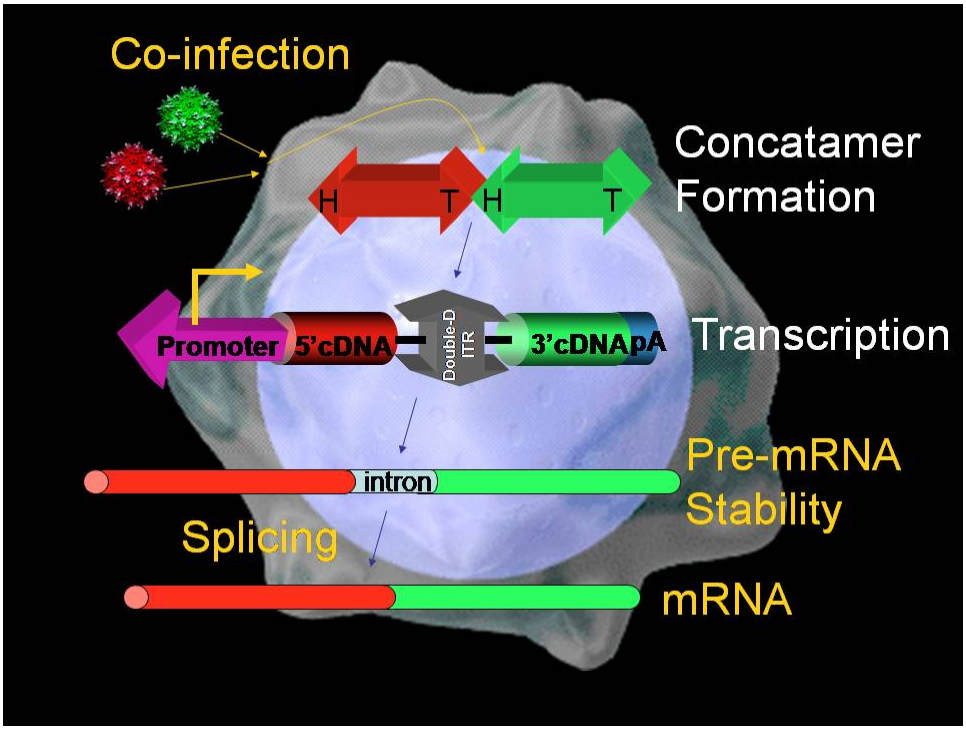


Figure 1.5. Rate-limiting steps in the trans-splicing AAV vectors. There are several potential rate-limiting barriers in the trans-splicing approach. These include co-infection efficiency, heterodimer formation, transcription across the double-D ITR, pre-mRNA stability and splicing. (This figure is contributed by Dr.

Dongsheng Duan). (*Adapted from Ghosh, et al, Biotech. and Genetic Engg Rev., 2007, Vol.24).*

The overlapping vectors

The overlapping vector approach provides an alternative to the trans-splicing approach where a homologous recombination dependent strategy is employed to reconstitute the intact transgene. Homologous recombination has been shown to regenerate the deleted regions in the ITR [33, 34]. Recent evidence suggests that homologous recombination may also occur between rAAV genome and genomic DNA [43-46]. It is therefore expected that AAV vectors would be capable of homologous recombination between each other should they share a region of homology.

The possibility of using homologous recombination-based approach to expand AAV packaging capacity was initially tested by Duan et al [35]. They generated two vectors including AV.Upstream and AV.Downstream, respectively. The AV.Upstream vector carries the promoter and the 5' two-thirds of the LacZ gene. The AAV.Downstream vector carries the 3' two-thirds of the LacZ gene followed by the polyadenylation signal. There is about a one kb LacZ sequence (the middle one-third) shared by both vectors. After co-infection in muscle, LacZ positive fibers were detected. Despite the proof-of-principle, transduction efficiency of the overlapping vectors is several logs lower than that of a single intact vector [35]. Interestingly, when Halbert et al applied the similar strategy to the alkaline phosphatase (AP) gene in the lung, they observed an extraordinarily

high level expression, a level that was compatible to a single intact AAV vector [36]. Taken together, these results demonstrate that the reasons affecting the efficacy of the overlapping AAV vectors should be elucidated in greater detail to resolve the conflicting results from different groups.

D. Summary

The demonstration that AAV can mediate safe and longterm gene expression has spawned a very deep interest in developing this virus for gene therapy. Being one of the smallest viruses, AAV faces a unique size-related challenge. It cannot deliver a large therapeutic expression cassette. This has essentially excluded AAV gene therapy for a number of common diseases such DMD/BMD, hemophilia A and cystic fibrosis. The dual vector strategies developed over the past few years have offered a real promise in overcoming the size limitation. With the increased packaging capacity, the territory of AAV gene therapy can now be expanded further. The next important step will be to move from the exciting results seen in rodents to large animal models of human diseases and eventually to human patients. Studies in several areas may likely accelerate this process. The first is to better understand the basic transduction biology of rAAV. A good example of this is the development of new vectors based on alternative AAV serotypes. These new vectors have not only enhanced transduction from a single intact AAV virion, they have also played a critical role for the success of dual AAV vectors [36, 40, 47]. Currently, only AAV-2 has been approved for patient use. Clinical application of these new

serotype vectors will certainly benefit from a clear understanding of their transduction process such as their receptor(s) and co-receptor(s), intracellular trafficking and potential integration. Secondly, it is important to evaluate systemic gene delivery with dual AAV vectors. Many diseases affect multiple tissues/organs in the body. Local gene transfer may only yield limited benefit to these patients. From this standpoint, a body-wide gene therapy will be required to cure the disease. None of the dual vector strategies has been evaluated for whole body transduction. Current successes are all based on local gene transfer. And finally, the dual vector design needs to be further streamlined and optimized. Although we have a collection of different dual vector strategies to choose for different applications, all these strategies have their inherent limitations. For example, in the trans-splicing approach, we may need to screen a large number of potential gene splitting sites before we can find one that works. On the other hand, the overlapping approach may only work for the selected genes. Creative means have to be developed for generic application of the dual vector approach in any gene in the future.

E. Hypothesis and research objectives

A major obstacle limiting adeno-associated virus (AAV) gene therapy is its small packaging capacity. The dual vector strategies such as trans-splicing and overlapping approaches have been developed to overcome this hurdle. These methods involve the simultaneous delivery of two AAV vectors carrying parts of a gene which would reconstitute the complete gene within the cell. However, the dual vector strategies have certain inherent limitations which prevent their use for delivering any gene of choice. The goal of this thesis work is two-fold. The first goal is to identify the expression limiting factors of overlapping dual vector reconstitution. This information should lead to the designing of further streamlined and optimized dual vectors that can reconstitute a large gene without the limitations of the current overlapping and trans-splicing vectors.

Many diseases affect multiple tissues/organs in the body and hence body-wide gene therapy will be required to cure the disease. None of the dual vector strategies has been evaluated for whole body transduction. Therefore, the second goal is to evaluate and optimize the dual vector strategies for systemic gene delivery. Chapter 2 details the experimental determination of expression limiting factors for gene reconstitution using the overlapping dual vectors. The methods and experimental evaluation of systemic gene transfer using trans-splicing dual vectors is detailed in the chapter 3. Chapter 4 describes the concept, design and evaluation of a new generation of dual vectors, namely, hybrid vectors, which utilize both overlapping and trans-splicing dual vector approaches in a single system. In summary, this thesis study lays the platform

for the design of optimal dual vector systems which may efficiently reconstitute and express large therapeutic transgenes in multiple tissues/organs throughout the body.

Chapter 2. Viral serotype and the transgene sequence influence overlapping vector-mediated gene transfer

A. Introduction

AAV can package vector genomes up to 5 kb [19]. However, large therapeutic gene expression cassettes (such as these carrying the full-length dystrophin gene and the cystic fibrosis transmembrane conductance regulator gene) are excluded for AAV packaging.

In order to overcome this limitation, a large gene can be split into smaller fragments followed by packaging each individual piece into separate AAV viruses. The key issue of such attempts is to re-assemble the split pieces into a functional expression unit inside cell. In the overlapping approach, a large gene is split into two overlapping fragments and the gene reconstitution is achieved through homologous recombination of the two overlapping fragments [35].

The overlapping approach does not require complicated cloning procedures to introduce splicing signals [48]. It is therefore much easier to manipulate. However, the initial characterization of the AAV-2 LacZ overlapping vectors in muscle revealed a very disappointing efficiency [35]. The absolute amount of β-galactosidase protein produced from the LacZ overlapping vectors reached only 0.37% of that from a single AAV carrying the intact LacZ gene [35]. Surprisingly, a second study by Halbert et al reported a much higher efficiency [36]. In this later study, a set of AAV-6 viruses carrying the split alkaline phosphatase (AP) gene were examined in the mouse lung [36]. The differences

in these two studies include the transgene, AAV serotype and target tissue, any one and/or the combined effect of these three factors may have contributed to the observed discrepancy in transduction efficiency. To identify the key factors underlying these results, we systematically compared the LacZ and the AP overlapping vectors in three AAV serotypes including AAV-2, -5 and –6 in mouse skeletal muscle. Our results suggest that a fairly low dose of the overlapping vectors can mediate extremely efficient transduction in skeletal muscle. In particular, ~80% myofibers in the tibialis anterior (TA) muscle were successfully transduced by the AAV-6 AP overlapping vectors (5×10^8 vg particles of each upstream and downstream vectors).

B. Materials and Methods

Proviral plasmids. Eight different *cis* proviral plasmids were used to generate rAAV stocks. These AAV vectors carry either the intact or the overlapping LacZ or AP (alkaline phosphatase) or mini-dystrophin genes, respectively. In all of these constructs, transgene expression was regulated by the Rous sarcoma virus (RSV) promoter and the Simian Virus 40(SV40) polyadenylation signal. The *cis* plasmids, pcisRSV.LacZ, pcisLacZUpstream, pcisLacZDownstream and pcisRSV.AP were described before [35]. In the LacZ overlapping vectors, the shared sequence (1093 bp) spans about one third of the central region of the intact LacZ gene [35]. In this study, we generated AP overlapping vectors that also shared the central one third of the intact AP gene. Briefly, one of the two SalI sites (at nt 3817 and nt 7127, respectively) in pcisRSV.AP was randomly

inactivated by partial digestion and blunt-end ligated to create pAG1 and pAG2. The SalI site at nt 7127 is inactivated in pAG1 and the SalI site at nt 3817 was inactivated in pAG2. To generate pcisAPUpstream, pAG2 was double digested with RsrII and SalI to remove a 1061 bp fragment containing the 3' one-third of the AP gene and the polyA sequence. To generate pcisAPDownstream, a 1377 bp fragment containing the RSV promoter and the 5' one-third of the AP gene was removed from pAG1 by AvrII and SalI double digestion. The final AP upstream construct (pcisAPUpstream) contains the RSV promoter and the 5' two-third of the AP gene. The final AP downstream construct (pcisAPDownstream) contains the 3' two-thirds of the AP gene along with the SV40 polyA signal sequence. These two constructs share 872 bp of the overlapping sequence.

The overlapping mini-dystrophin vectors (pcisRSV.mini-dystrophinUpstream and pcis.mini-dystrophinDownstream) were derived from a human mini-dystrophin minigene described previously [22]. This minigene is in turn derived from the full length human dystrophin gene by deleting the sequence between the domains hinge 2 and spectrin like repeat 19, shortening the size to approximately 6 kb. The mini-dystrophin overlapping vectors share a 1410 bp overlap sequence in the middle of the mini-dystrophin gene. Similar to the AP and LacZ overlapping constructs, the overlapping mini-dystrophin vector gene expression is regulated by the RSV promoter and SV40 polyadenylation signal.

Recombinant AAV production. AAV-2 and AAV-5 vectors were generated using an adenovirus-free system as previously described [49]. AAV-6 was produced with a packaging system generously provided by Dr. A. Dusty Miller (Fred Hutchinson Cancer Research Center, Seattle, WA) [50]. Briefly, 70% confluent 293 cells were co-transfected with a *cis* plasmid, pMT-Rep2, pCMVCap6 and an adenoviral helper plasmid (pHelper, Stratagene, Catalog #240071) at a ratio of 1:1:3:3. The viral lysate was harvested at 62 hrs post-transfection. Recombinant viral stocks were purified by two sequential rounds of isopycnic ultracentrifugation as described before [49]. Viral fractions were pooled and dialyzed through two rounds of Hepes-buffered saline. Finally, viral titer was determined by slot blot as previously described [49]. In the case of the mini-dystrophin overlapping vectors, only the AAV-6 vectors were generated.

***In vitro* studies.** Transduction efficiency of the overlapping AAV vectors was evaluated in HT1080 cells, a human fibrosarcoma cell line [51] (a gift from Dr. David J. Pintel, University of Missouri, Columbia, MO), CF 16 cells, a papillomavirus immortalized human airway epithelial cell line originated from a cystic fibrosis patient [52] (a gift from Dr. A. Dusty Miller, Fred Hutchinson Cancer Research Center, Seattle, WA) and MO59K cells (a gift from Dr. Mark Hannink, University of Missouri). The HT1080 and CF16 cell lines were used by Halbert et al to characterize AAV AP overlapping vector [36]. HT1080 cells were cultured in Dulbecco Modified Eagle's Medium supplemented with 10% fetal bovine serum (FBS) and 1% PS (penicillin 100U/ml and streptomycin 100µg/ml). CF 16 cells

were cultured in Keratinocyte- serum free media (Gibco-BRL, Catalog#10724-011) supplemented with pituitary extracts and recombinant epidermal growth factor. MO59K cells were cultured in DMEM/F12 mixture supplemented with 1% non-essential amino acids, 1% sodium pyruvate (100X working stock), 10% fetal bovine serum and 1% PS (penicillin 100U/ml and streptomycin 100U/ml).

For infection with AV.LacZ and AV.AP, 80% confluent cells were infected at a multiplicity of infection (moi) of 1,000 vg particles/cell. For co-infection experiments, each of the overlapping vectors was applied at an moi of 1,000 vg particles/cell. All infections were performed in serum free medium. In HT1080 cells, FBS concentration was brought to 10% by addition of 20%FBS/2%PS at one hr post-infection. Transgene expression was analyzed at 48 hrs post-infection by standard cytochemical staining as described before [53].

Control studies. In order to confirm that the single overlapping split vectors do not express any functional protein, complete expression constructs were generated for the AP and LacZ overlapping parts. To the pcisRSV.APUpstream and pcisRSV.LacZUpstream, a SV40 polyadenylation sequence was added at the end of the AP and LacZ sequences respectively. The RSV promoter sequence was introduced 5' to the AP and LacZ coding sequences of pcis.APDownstream and pcis.LacZDownstream respectively. These constructs were then co-transfected into 293 cells and gene expression was evaluated for each construct with the intact pcisRSV.AP and pcisRSV.LacZ as positive controls respectively. A plasmid expressing eGFP (enhanced green fluorescent protein)

was co-transfected with each AP and LacZ construct as a transfection control. Transfections were done using 1µg of each of the AP and LacZ expression plasmids and 0.5µg of the eGFP plasmid per well in 6-well plates. Transfections were done using Lipofectamine (Invitrogen, Catalog # 18324-12) and Plus (Invitrogen, Catalog #11514-015) reagents as per manufacturer's protocols. The transfected cells were analyzed for gene expression from each of the plasmid constructs at 48 hrs post-transfection by standard cytochemical staining [53] for AP and LacZ. Expression from the eGFP control plasmid was analysed by fluorescence microscopy.

Low molecular weight Hirt DNA preparation and PCR. Low molecular weight DNA was isolated from AAV infected HT 1080 and MO59K cells using a modified Hirt DNA isolation protocol. AV6.RSV.AP, co-infection with AV6.RSV.APUpstream and AV6.RSV.APDownstream, AV6.RSV.APUpstream single and AV6.RSV.APDownstream single infections respectively were carried out as described above. The infected cells were harvested after 45 hrs using a cell scraper in 1X phosphate buffered saline (PBS). After centrifugation for 1 min at 3000 rpm, the cell pellet was resuspended in 300 µl lysis buffer containing 1% SDS, 50 mM EDTA, and 100 mM Tris pH 8.0. Cell lysate was then digested with proteinase K (1µg/µl final concentration) and pronase (0.5µg/µl final concentration) for 30 min. To precipitate low molecular weight DNA, 75 µl 5M NaCl was added to the cell lysate. After incubating for 2 hrs at 4ºC, 250 µl TE (10 mM Tris, pH 8.0, 1 mM EDTA) and 100 µl Promega neutralizing solution

(Wizard Plus SV minipreps, Promega, Madison, WI) was added to the tube. The cell lysate was mixed by inverting the tube several times. Cellular debris (including chromosome DNA) was then removed by a 35 min centrifugation at 13,000 rpm. Finally, low molecular weight DNA was purified using a DNA miniprep column (Wizard Plus SV minipreps, Promega, Madison, WI) and eluted in DNase free water. 100 ng of this DNA was used for a PCR reaction using forward primer DL373 (CATTGGTGTGCACCTCCAAG) and reverse primer DL316 (GGGACGGGACACTCAGGGAGCAGTG). The forward primer is located in the RSV promoter sequence while the reverse primer is located downstream of the overlapping sequence. Stringent conditions were used for the PCR reaction including high annealing temperature (t_m = (T-2) ^oC; t_m is the calculated annealing temperature; T is the primer melting temperature) short annealing time (20 sec) and lower primer concentration (50 nMol).

Animal studies. All the animal experiments were carried out in accordance with NIH and institutional guidelines of the University of Missouri. Seven-week-old C57Bl/6 (BL6) male mice were purchased from Jackson Laboratory (Bar Harbor, Maine). Recombinant AAV was delivered to the TA muscle according to a previously described protocol [11]. Briefly, a 2 mm incision was made on the skin covering the proximal end of the TA muscle. 20 μl of AAV viruses were directly injected into the muscle with a 33G gas-tight Hamilton syringe. In the case of AV.LacZ or AV.AP infection, 5 x 10^8 vg particles were delivered to the muscle. In overlapping virus co-infection studies, a total of 1 x 10^9 vg particles (5 x 10^8 vg

particles of each virus) were delivered. To evaluate whether there was leaky expression from upstream or downstream viruses, we examined transgene expression in individual upstream or downstream vector infected muscles, respectively. Muscles injected with 20 μl Hepes-buffered saline served as negative controls. To examine transduction efficiency in the lung, same doses of viruses (5×10^8 vg particles for each virus) were diluted in 50 μl Hepes-buffered saline and directly administrated to the trachea by supraglottic injection with a 24 G bent feeding syringe. Pilot studies with food dye and adenovirus (Ad.CMV.AP) revealed very efficient delivery to the entire lung with this method (data not shown). Mice were sacrificed at six weeks post-infection and the TA muscles and the lungs were harvested according to published protocols [53]. To examine transgene expression, 10 μm cryosections were stained for LacZ or AP as described before [53]. Specifically, LacZ staining was carried on for 90 min at 37ºC and AP staining was 10 min for muscle and 20 min for the lung sections at 37ºC. To determine the level of transgene expression in muscle, the number of transgene positive myofibers (irrespective of staining intensity) and the number of total myofibers in an entire muscle section were manually counted from digitized images. At least four sections (two in the middle belly, two in the peripheral) were quantified for each muscle sample. Transduction efficiency was expressed as the percentage of staining positive myofibers in the entire muscle section. For histochemical staining in the lung, we used Ad.CMV.LacZ and Ad.CMV.AP infected lung samples as the positive controls.

In the case of the mini-dystrophin overlapping vectors, co-infections were done in *mdx* (dystrophin null) mice and *mdko* (dystrophin/myoD double knock-out) mice. 25 µL of a mixture of AV6.RSV. mini-DystrophinUpstream and AV6.RSV. mini-DystrophinDownstream at a dosage of 2 x 10^{10} vg (1 x 10^{10} vg particles of each virus particles of each virus) was injected in each tibialis anterior muscle (TA). 10 µl of the same mixture was injected into the extensor digitorum longus muscle (EDL). Single vector infections were done as controls. AV6.CMV.micro-Dystrophin was used as a positive control for dystrophin expression from AAV6. Single vector injections were also done in wild type C57BL/10 mice at a dosage of 1 x 10^{10} vg to account for any adverse effects of the virus infection or transgene expression. TA injections were done as described above. The mice were sacrificed after 6 weeks and TA muscles harvested for analysis of mini-dystrophin expression using indirect immunofluorescence as previously published [23, 54]. Since the mini-dystrophin gene is derived from the human dystrophin gene, 2 different antibodies were used to evaluate gene expression. The two antibodies detecting different regions of the mini-dystrophin gene are the human dystrophin N-terminus specific monoclonal antibody Dys-3 and the dystrophin C-terminus specific monoclonal antibody Dys-2 (Novocastra Laboratories, Newcastle, UK). The EDL muscles were harvested and muscle function analysis was performed as described before [54].

Statistical analysis. Data are presented as mean ± standard error of mean. Statistical analysis was performed with the SPSS software. At first, a Normality

test was performed to evaluate data distribution. The majority of the data follow normal distribution except for two groups in which transduction efficiency in each muscle sample is 0~1%. These two groups are muscles infected by AAV-2 AV.LacZ and muscles co-infected by AAV-2 AV.APUpstream and AV.APDownstream viruses. For normally distributed data, a parametric statistical analysis was performed. Specifically, statistical significance was determined by oneway ANOVA followed by Bonferroni *post hoc* analysis among different groups. For data that do not follow a normal distribution, a non-parametric Kruskal-Wallis test was performed. Comparison between the intact gene vector and the overlapping vector in the same group (same serotype) was carried out with paired student *t* test. Difference was considered significant when $P < 0.05$. (The statistical analysis was performed by Dr. Mingju Liu.)

C. Results

Overlapping vector construction and cross-packaging into different AAV serotypes. A critical consideration in homologous recombination is the length of the homology region between two substrates. In *E. coli*, the minimal length for RecBC pathway is 23-27 bp and for RecF pathway it is 44-90 bp [55]. In yeast, a 60 bp minimal length has been proposed [56]. The minimal length of homology region for overlapping AAV vector-mediated gene transfer has not been determined. In our previous study with the AAV-2 LacZ overlapping vectors, we have arbitrarily selected an overlapping length of one-third of the LacZ gene (1093 bp). Halbert et al has shown that recombination efficiency is not affected

between a 440 bp and a 1000 bp overlap in the AP gene [36]. To be consistent with our design in the LacZ overlapping vectors, we generated a pair of the AP overlapping vectors that also shared one-third of the transgene (872 bp). Similarly, the mini-dystrophin overlapping vectors were generated with an overlap of 1410bp in the middle of the gene. Since the mini-dystrophin gene is much larger (>6kb) than the reporter genes, the overlap of 1410 bp represents the maximum achievable overlap for the mini-dystrophin overlapping vectors.

AAV-2 is the prototype AAV vector used in many gene therapy studies. Recently AAV vectors based on several new viral serotypes were developed [57]. Cross-packaging of the AAV-2 genome with different viral capsids has a profound effect on transduction. In skeletal muscle, pseudopackaged AAV-5 and AAV-6 were found to be more efficient than AAV-2 [15, 16, 58, 59]. In airway epithelia, both AAV-5 and AAV-6 were considered to be superior than AAV-2 [60]. To clarify the impact of viral capsid on the overlapping vector-mediated gene transfer, we packaged the LacZ and the AP vectors in AAV-2, -5 and –6 capsids. The mini-dystrophin overlapping vectors were packaged exclusively into AAV-6 vectors based on the results observed from the AP and LacZ overlapping viruses.

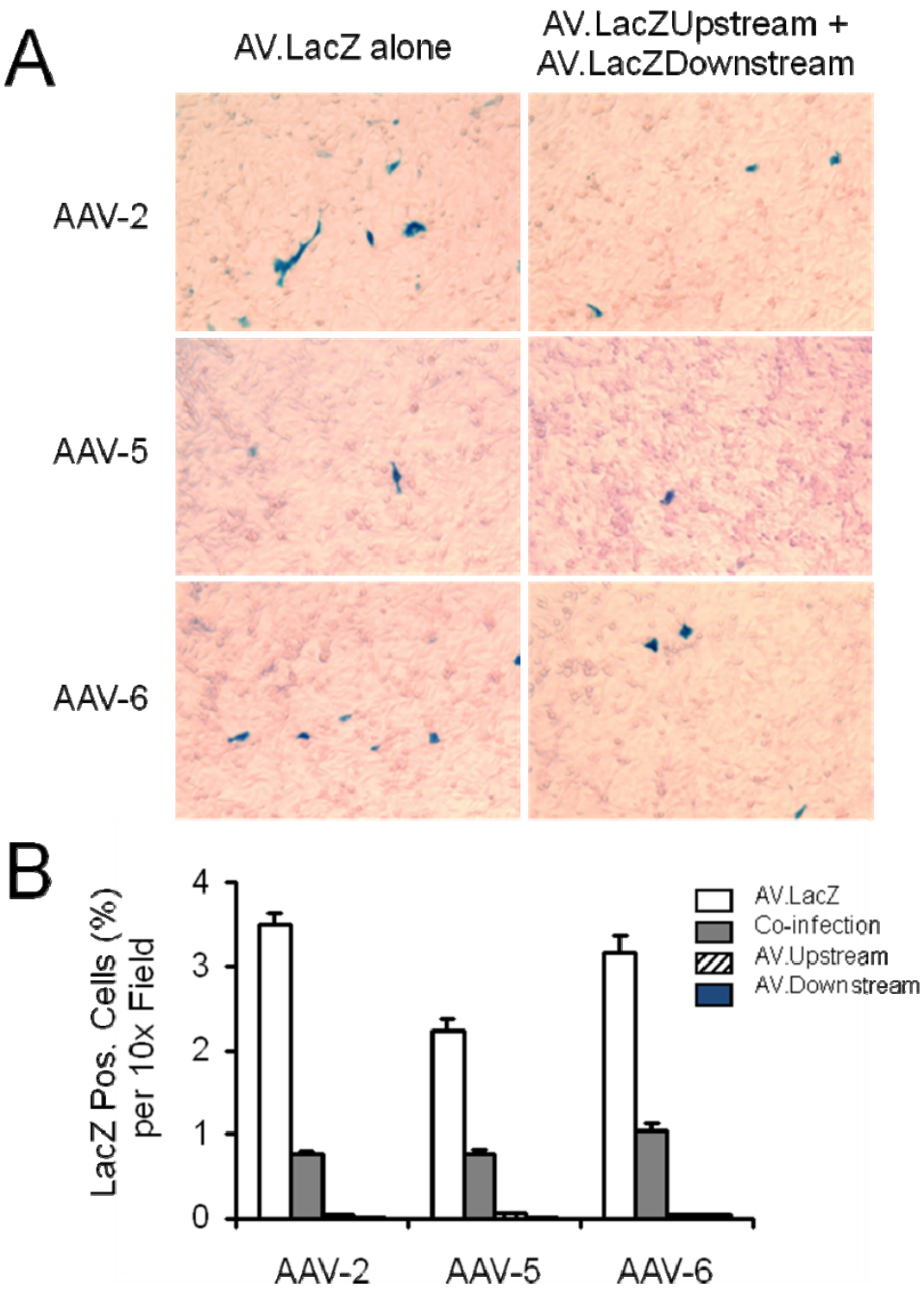
***In vitro* evaluation of the overlapping vectors**. Transduction efficiency of the intact gene vector and the overlapping vectors were first examined in HT1080 cells, a cell line that has been previously used to characterize AP overlapping vector [36]. As shown in Figure 2.1 A-B, very few cells were transduced by

AV.LacZ (intact single vector). At a dosage of 1,000 moi, AAV-2 and AAV-6 were slightly better than AAV-5, but all were lower than 4 %. To exclude the possibility of poor viral preparation, we applied the same viral stocks to 293 cells *in vitro* and mouse skeletal muscle *in vivo*. These studies revealed the expected transduction efficiency (data not shown). We also achieved similar results with new AAV preparations (data not shown). The transduction profile in the MO59K cells was marginally better (3-6%), but essentially a trend similar to HT1080 was observed (Figure 2.1 C-D).

We next examined transduction efficiency of the LacZ overlapping vectors (Figure 2.1 A-B). Surprisingly, the number of LacZ positive cells following overlapping vector infection was fairly high comparing with that of the intact gene vector infection. Based on the percentage of LacZ positive cells, transduction efficiency of the overlapping vectors reached one-fourth to one-third of that of the intact gene vector (Figure 2.1 B, 2.1 D). Intriguingly, we observed a substantially different transduction profile when the AP gene was used as the reporter gene (Figure 2.2 A-B). At the dosage as we used in the LacZ virus experiment, the intact AP vector resulted in 98% transduction in all three AAV serotypes. However, very few AP positive cells were detected in overlapping vector co-infected cells. AAV-6 overlapping vectors transduced approximately 3% cells while AAV-2 and AAV-5 were slightly less efficient and they transduced approximately 1% cells. Although the absolute number of the positive cells were higher with the AP overlapping vectors than with the LacZ overlapping vectors,

relative efficiency of the overlapping vectors to that of the intact gene vector was greatly reduced in the AP overlapping vectors.

The infection profile in the MO59K cells markedly differed from that of the other cell lines (Figure 2.2 C-D). Importantly, the expression from the AP overlapping vector co-infection sets was much higher than those from corresponding HT1080 and CF16 cells. At a dosage of 1000vg per cell, the number of AP positive cells from the AAV-2 and AAV-5 overlapping vectors reached 16.7% and 17.3% of the respective intact AP vectors. Remarkably, the number of positive cells from the AAV-6 overlapping set reached 70.7% of the intact vector (Figure 2.2 D). Thus, in these cells, the difference in efficiency of the AP and LacZ overlapping vectors is not as pronounced as with the HT1080 cells. Similar experiments were also performed in CF16 cells, another cell line that had been used to study AAV overlapping vector [36]. The infection profiles in these cells were similar to that observed for the HT1080 cells, but the overall level of transduction for all vectors was much lower (data not shown).


A
AV.LacZ alone
AV.LacZUpstream +
AV.LacZDownstream
AAV-2
AAV-5
AAV-6
B
LacZ Pos. Cells (%)
per 10x Field
0
1
2
3
4
AV.LacZ
Co-infection
AV.Upstream
AV.Downstream
AAV-2
AAV-5
AAV-6

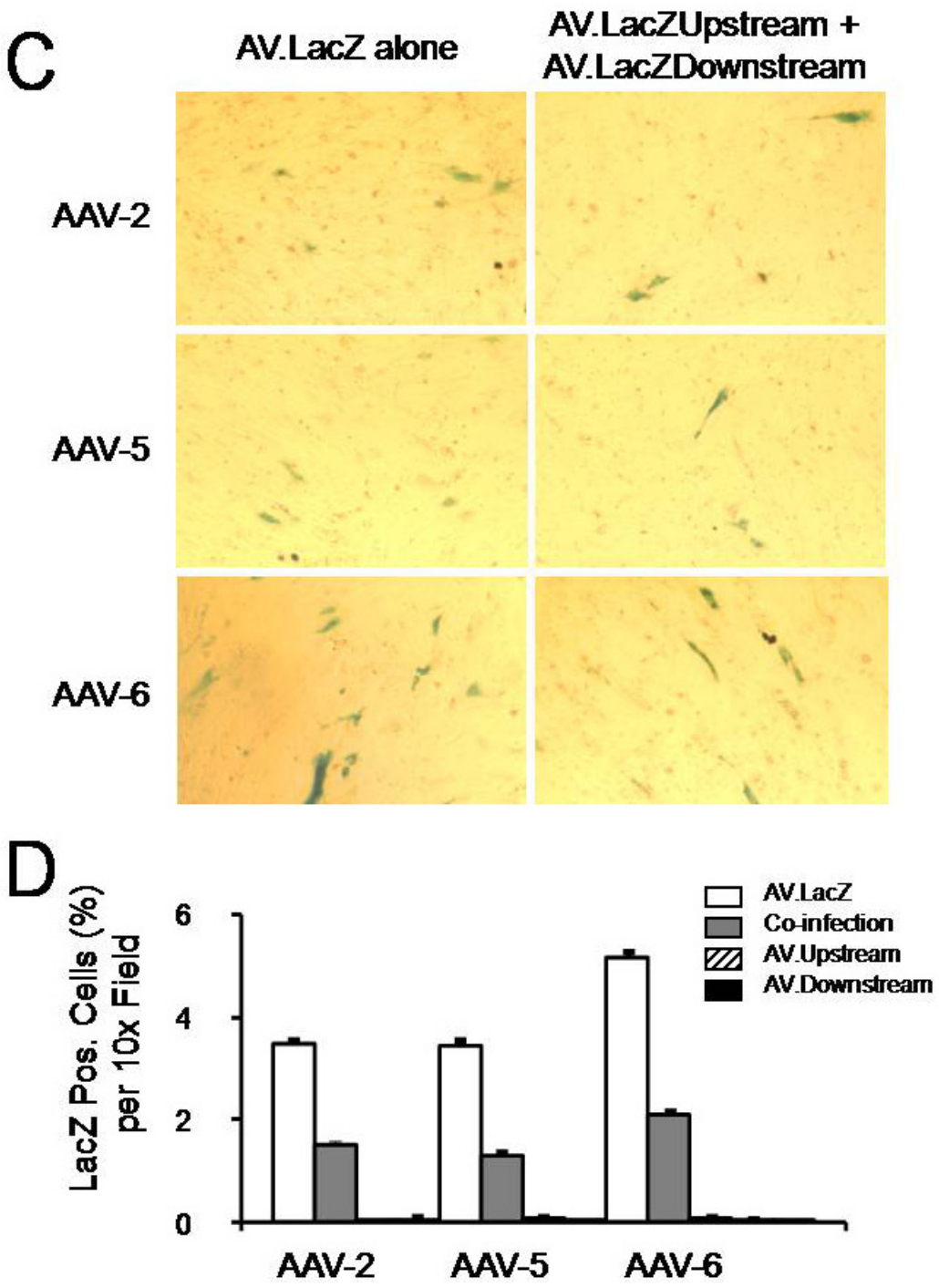
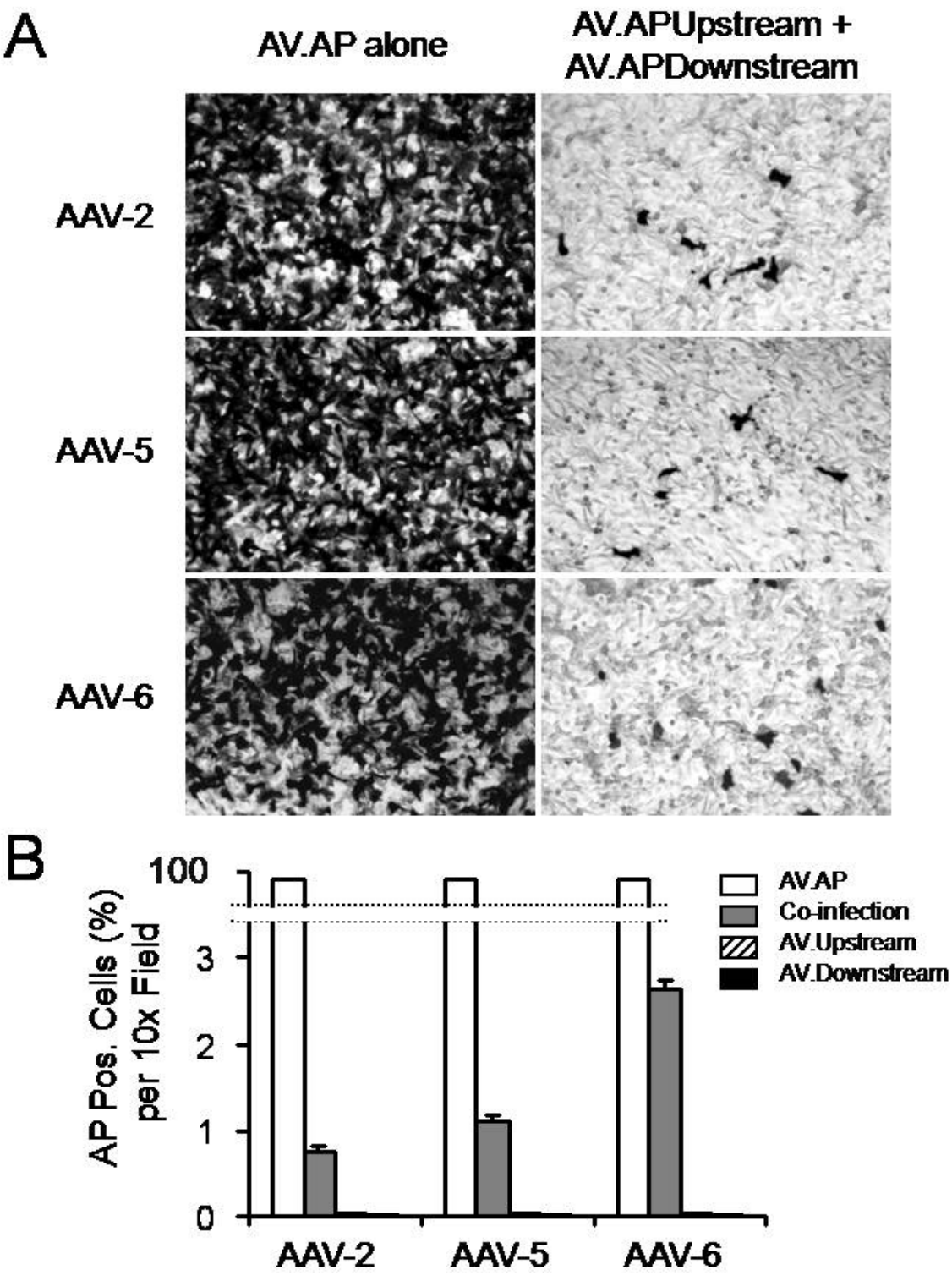


Figure 2.1. *In vitro* evaluation of the LacZ overlapping vectors in HT1080 and MO59K cells. 80% confluent cells were infected at an moi of 1,000 vg particles/cell for each virus and LacZ expression was examined by cytochemical

staining 48 hrs later. **A**, Representative photomicrographs from cells infected with AV.LacZ alone or co-infected with both AV.LacZ.Upstream and AV.LacZ.Downstream in HT1080 cells. **B**, Morphometric quantification of LacZ positive cells following intact or overlapping vector infection, respectively in HT1080 cells. N = 5 for each group. **C**, Representative photomicrographs from cells infected with AV.LacZ alone or co-infected with both AV.LacZ.Upstream and AV.LacZ.Downstream in MO59K cells. **D**, Morphometric quantification of LacZ positive cells following intact or overlapping vector infection, respectively in MO59K cells. N = 5 for each group. (*Adapted from Ghosh, et al, J. Gene Med., 2006, 8(3):298-305).*


A
AV.AP alone
AV.APUpstream +
AV.APDownstream
AAV-2
AAV-5
AAV-6
B
AP Pos. Cells (%)
per 10x Field
100
3
2
1
0
AV.AP
Co-infection
AV.Upstream
AV.Downstream
AAV-2
AAV-5
AAV-6

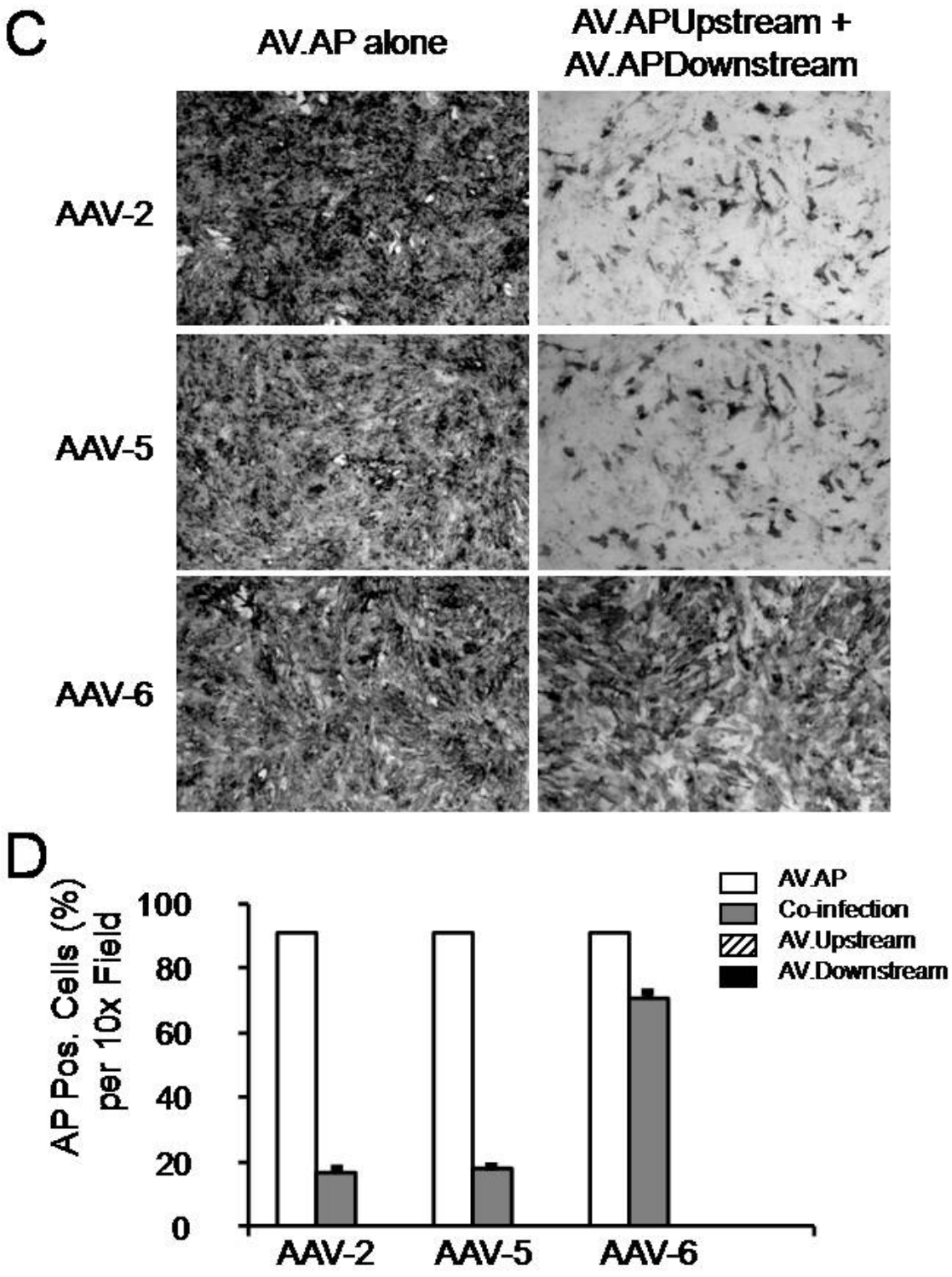


Figure 2.2. *In vitro* evaluation of the AP overlapping vectors in HT1080 and MO59K cells. 80% confluent cells were infected at an moi of 1,000 vg

particles/cell for each virus and AP expression was examined by cytochemical staining 48 hrs later. **A**, Representative photomicrographs from the intact gene vector (AV.AP) infection or the overlapping vectors (AV.AP.Upstream and AV.AP.Downstream) co-infection, respectively in HT1080 cells. **B**, Morphometric quantification of transduction efficiency in HT1080 cells. N = 5 for each group. **C**, Representative photomicrographs from the intact gene vector (AV.AP) infection or the overlapping vectors (AV.AP.Upstream and AV.AP.Downstream) co-infection, respectively in MO59K cells. **D**, Morphometric quantification of transduction efficiency in MO59K cells. N = 5 for each group. (*Adapted from Ghosh, et al, J. Gene Med., 2006, 8(3):298-305).*

Control studies. In order to determine if there was any leaky expression from the individual overlapping vectors, complete expression cassettes for each individual overlapping vector were generated. The transfection studies indicate that these smaller parts of the AP and LacZ genes (the individual overlapping parts) were not capable of making a functional protein as indicated by the cytochemical staining (Figure 2.3).

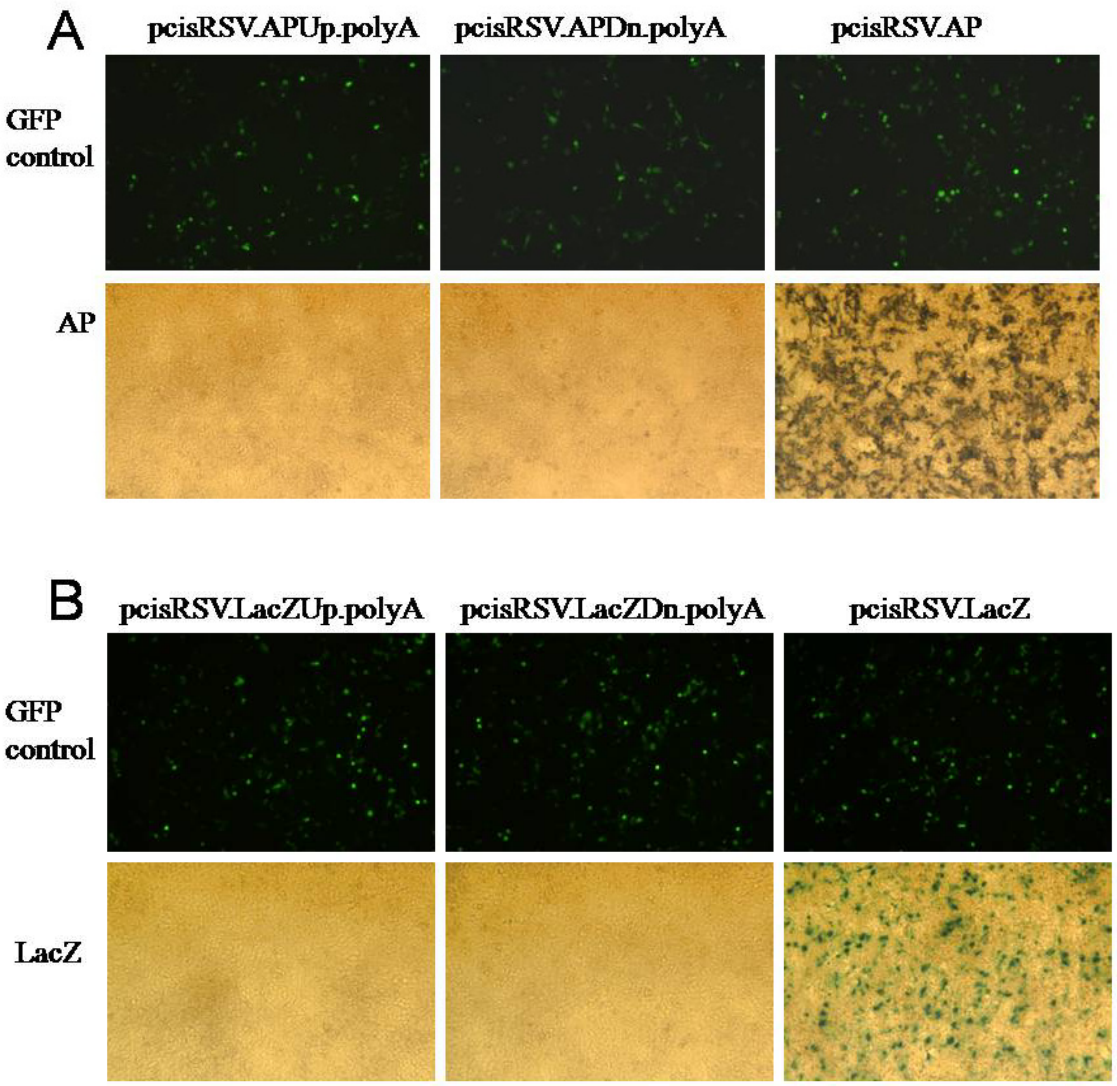


Figure 2.3. Control experiments to evaluate gene expression from overlapping vector parts. Cytochemical staining for gene expression analysis of 293 cells transfected with different plasmid constructs after 48hrs. The top set of photomicrographs depict GFP expression analysed by fluorescence microscopy. **A**. Cytochemical staining for AP expression with GFP as transfection control. **B**. Cytochemical staining for LacZ expression with GFP as transfection control.

Detection of recombined genomes from the overlapping vector co-infection: To detect the recombined expression cassettes generated within cells co-infected with the overlapping AAV vectors, low molecular weight DNA was isolated from HT1080 and MO59K cells. The Hirt DNA was then amplified by PCR using forward primer based on the RSV promoter sequence and the reverse primer based on the AP coding sequence 3' of the shared overlap region of the AP overlapping vector (Figure 2.4 A). These primers were designed such that they would selectively amplify a 1.72kb fragment from the reconstituted AP gene expression cassette (Figure 2.4 A) but not from the single upstream or downstream vectors (Figure 2.4 B, Lanes 5 and 6 respectively). In figure 2.4 B, lane 2-3, the expected amplified product for a reconstituted genome is observed at 1.72 kb. A similar product is seen in the AV.RSV.AP infected positive control lanes 4-5. To control for in-reaction recombination events during the PCR protocol, a mixture of AV.RSV.APUpstream and AV.RSV.APDownstream viral DNA was used for amplification (Figure 2.4 B, lane 8) as well as a mixture of Hirt DNA from AV.RSV.APUpstream and AV.RSV.APDownstream single vectors infections (Figure 2.4 B, lane 9). No definitive bands were observed in any of the control lanes (Figure 2.4 B, lanes 5, 6, 7, 8 and 9). This experiment demonstrates that the overlapping vectors physically recombine to generate intact expression cassettes. The same pattern of amplified products was observed for PCR of Hirt preps from MO59K cells (data not shown).

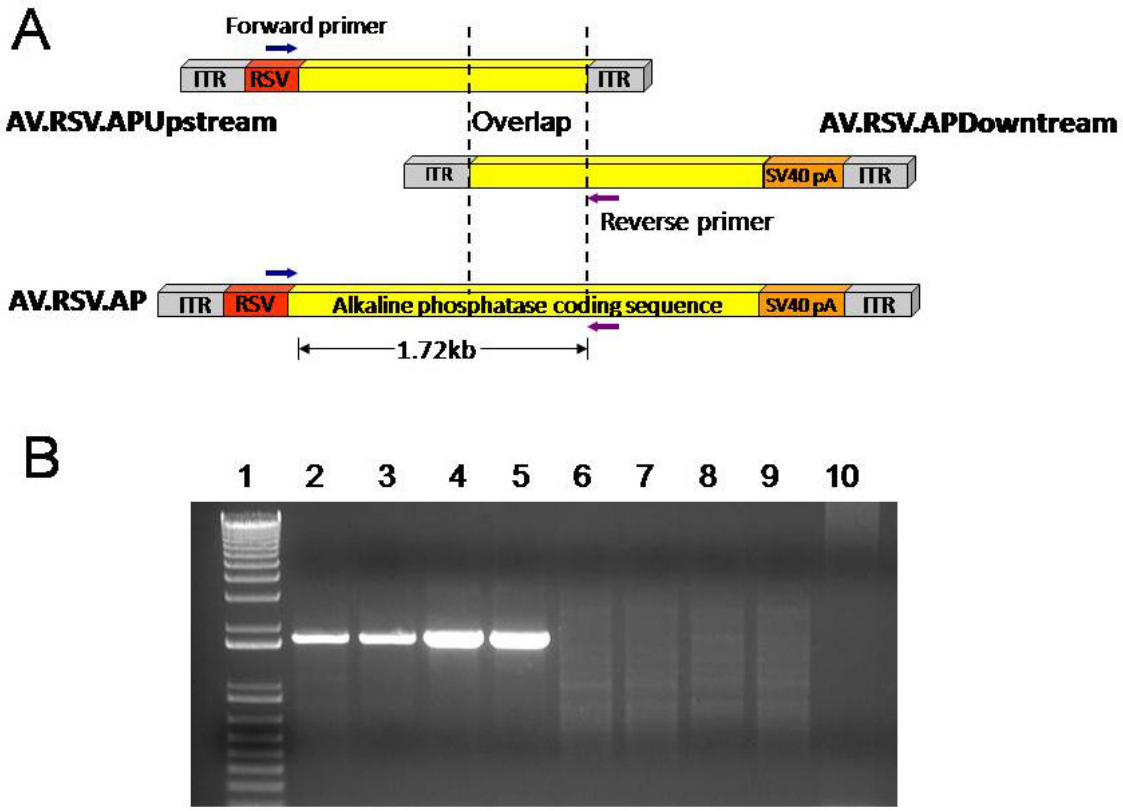


Figure 2.4. Detection of reconstituted transgene expression cassettes from AAV overlapping vector infected cells. Hirt DNA was prepared from HT1080 cells infected with AP overlapping, AP intact and AP overlapping single vectors followed by PCR. **A**. Location of PCR primers and expected size of amplified fragment. **B**. Lane 1. Molecular weight marker. Lanes 2-3. AV.RSV.APUpstream and AV.APDownstream co-infection. Lanes 4-5 AV.RSV.AP infection. Lane 6. AV.RSV.APUpstream single vector infection. Lane 7 AV.APDownstream single vector infection. Lane 8. AV.RSV.APUpstream and AV.APDownstream DNA purified directly from the viral capsids. Lane 9. Mixture of AV.RSV.APUpstream and AV.APDownstream single vector Hirt DNA. Lane10. AV.RSV.AP without *Taq* polymerase as negative control.

***In vivo* comparison of the intact gene vector and the overlapping vectors**: To determine whether viral capsid and the nature of the transgene sequence influence transduction efficiency of the overlapping vectors in muscle, we performed a cohort study in the TA muscles of BL6 mice. Transduction efficiency was compared among four groups including muscles infected with the intact gene vector only, the upstream vector only, the downstream vector only and muscles that were co-infected with both the upstream and the downstream vectors. In each case, 5×10^8 vg particles of each virus were directly injected into the TA muscle. We have intentionally chosen this low dosage to avoid potential saturation effect in AAV-5 and/or AAV-6 infection. Transduction efficiency was examined six weeks later by scoring positive cells. Although this quantification procedure did not distinguish the staining intensity, we chose it because it allowed us to compare relative efficiency between the LacZ gene and the AP gene systems.

Figure 2.5 depicts the transduction profile with the intact gene and the overlapping LacZ vectors. As expected, low dosage infection resulted in poor transduction with AAV-2 vectors and only few LacZ positive myofibers were observed in AV.LacZ infected muscles. Consistent with previous reports, AAV-5 capsid pseudotyped virus led to a much higher transduction and ~11% myofibers were positive following AV.LacZ infection [49]. However, no LacZ positive fibers were found in AAV-5 overlapping vector co-infected muscles. AAV-6 yielded the highest transduction efficiency. Approximately 80% myofibers were positive in

intact gene vector infected muscles. Furthermore, the LacZ staining intensity was much stronger than that in AAV-5 infected muscles. Encouragingly, about 35% myofibers were LacZ positive in AAV-6 overlapping vector co-infected muscles. This corresponds to approximately 42% of the efficiency of that of the intact gene vector. Surprisingly, transduction efficiency of the AAV-6 overlapping LacZ vectors was higher than that of the AAV-5 vector carrying the intact LacZ gene. In both cases, positive myofibers demonstrated similar staining intensity. However, more myofibers were transduced by the AAV-6 overlapping vectors than by the AAV-5 intact gene vector.

Results from Halbert et al [36] and our *in vitro* data (Figures 2.1 and 2.2) suggest that the nature of the transgene sequence may also influence overall transduction efficiency. To investigate whether overlapping vector-mediated gene transfer in muscle was effected by the transgene sequence, we performed a similar study as described in figure 2.5, except that this time we used the AP gene, instead of the LacZ gene. Transduction efficiency of AV.AP reached 25% and 40% in AAV-2 and AAV-5, respectively (Figure 2.6). These levels were substantially higher than that of the respective efficiency in AV.LacZ study (Figure 2.5). Consistent with AV.LacZ infection results (Figure 2.5), the highest transduction was observed in AAV-6. Interestingly, AAV-6 led to quite compatible levels of transduction between AV.AP and AV.LacZ. It is possible that the levels of expression from the AAV-6 intact LacZ and/or AP gene vectors might have reached the plateau. Surprisingly, the percentage of AP positive myofibers reached the similar level between the intact gene vector and the

overlapping AP vectors in AAV-5 and AAV-6 (Figure 2.6 B). Nevertheless, as reflected in the staining intensity, the absolute amount of AP protein produced from the overlapping vectors seemed to be significantly lower than that from the intact gene vector.

In the LacZ gene study (Figure 2.5), transduction efficiency of the AAV-6 overlapping vectors was higher than that of the AAV-5 intact gene vector. Interestingly, a similar result was observed in the AP gene study. Transduction efficiency was ~ 40% for the AAV-5 intact gene vector. However, it reached ~80% in the AAV-6 overlapping vectors.

Consistent with previous publications [35], infection with one of the two overlapping vectors (upstream vector alone or downstream vector alone) did not yield transgene expression (data not shown). Despite impressive transduction with AAV-6 overlapping vector in muscle, we did not detect any positive cells in the lung from either the intact gene vector or the overlapping vectors (data not shown). It is very likely that the dosage we used in this study (5×10^8 vg particles per lung infection) was too low for lung gene transfer.

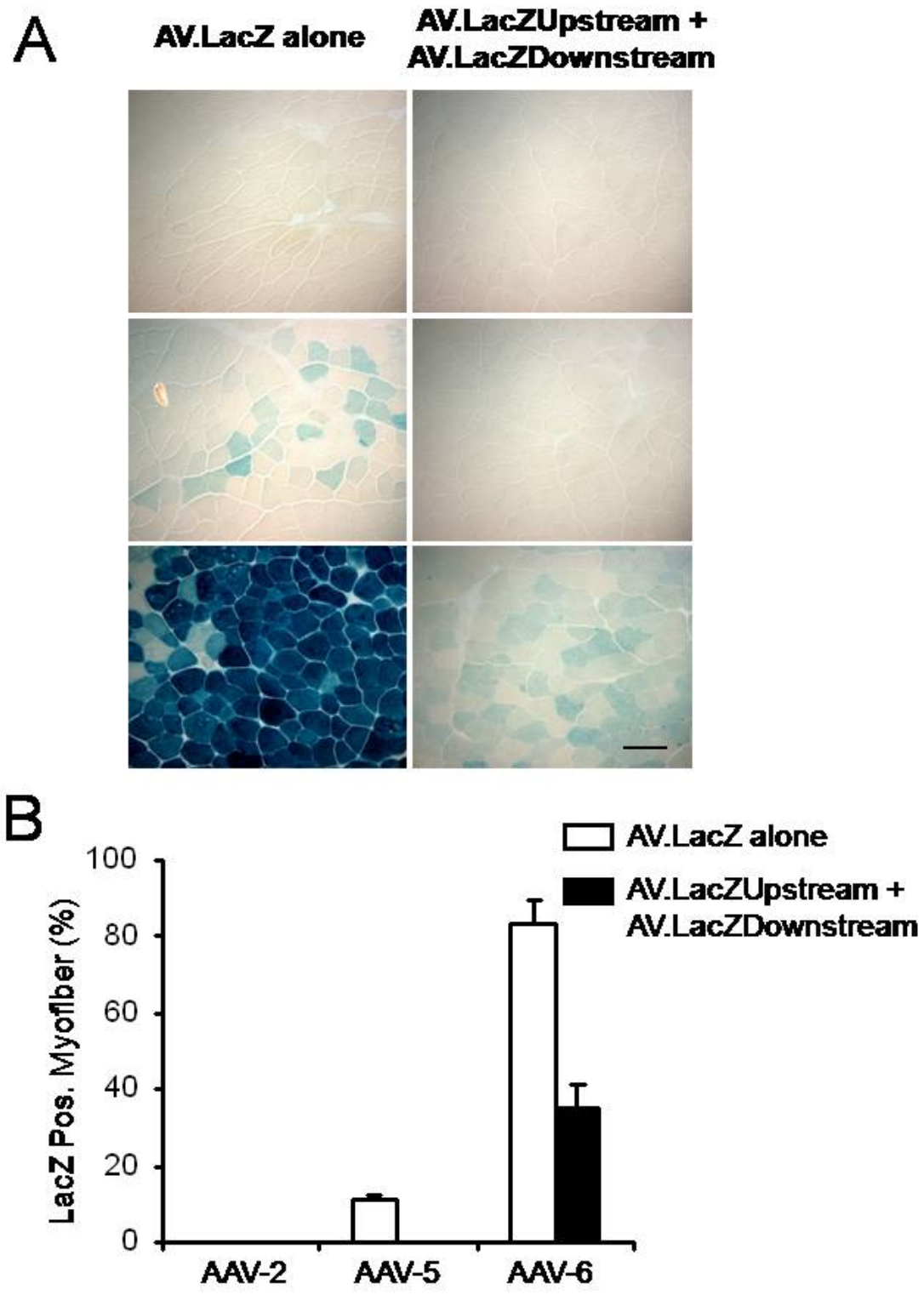

A
AV.LacZ alone
AV.LacZUpstream +
AV.LacZDownstream
B
AV.LacZ alone
AV.LacZUpstream +
AV.LacZDownstream
LacZ Pos. Myofiber (%)
100
80
60
40
20
0
AAV-2
AAV-5
AAV-6

Figure 2.5. *In vivo* evaluation of the LacZ overlapping vectors in skeletal muscle. The TA muscles of BL6 male mice were infected with 5 x 10^8 vg particles of each indicated AAV viruses. LacZ expression was examined at six weeks later by histochemical staining. **A**, Representative photomicrographs following AV.LacZ infection or AV.LacZ.Upstream and AV.LacZ.Downstream co-infection, respectively. Scale bar, 100μm. **B**, Morphometric quantification of transduction efficiency. Positive myofibers were not scored for the staining intensity. Asterisk, transduction efficiency of AV.LacZ is significantly higher than that of the LacZ overlapping vectors of the same AAV serotype. Cross, transduction efficiency of the AAV-6 overlapping vectors is significantly higher than that of the AAV-5 intact gene vector. N = 4 for each group. (*Adapted from Ghosh, et al, J. Gene Med., 2006, 8(3):298-305).*

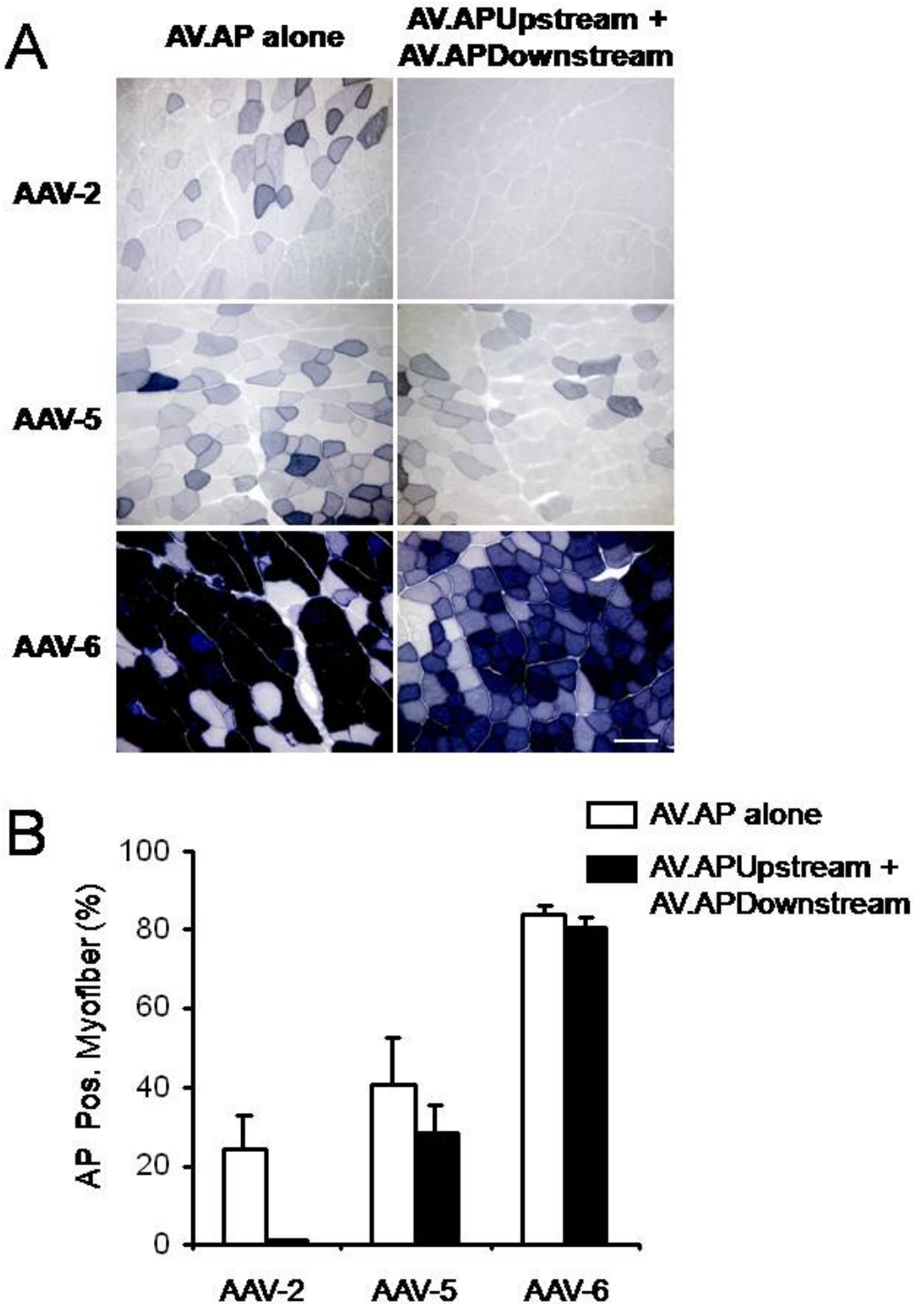

A
AV.AP alone
AV.APUpstream +
AV.APDownstream
AAV-2
AAV-5
AAV-6
B
AV.AP alone
AV.APUpstream +
AV.APDownstream
AP Pos. Myofiber (%)
100
80
60
40
20
0
AAV-2
AAV-5
AAV-6

Figure 2.6. *In vivo* evaluation of the AP overlapping vectors in skeletal muscle. The TA muscles of BL6 male mice were infected with 5 x 10^8 vg particles of each indicated AAV viruses. AP expression was examined at six weeks later by histochemical staining. **A**, Representative photomicrographs from AV.AP infection or AP overlapping vector co-infection, respectively. Scale bar, 100μm. **B**, Morphometric quantification of AP positive myofibers. The staining intensity was not scored. N = 4 for each group. There was no significant difference between AV.AP infection and the overlapping AP vector co-infection in AAV-5 and AAV-6. Asterisk, transduction efficiency of the AAV-2 intact gene vector is significantly higher than that of the AAV-2 AP overlapping vectors. Cross, transduction efficiency of AAV-6 is significantly higher than that of AAV-2 and AAV-5. (*Adapted from Ghosh, et al, J. Gene Med., 2006, 8(3):298-305).*

***In vivo* analysis of the mini-dystrophin overlapping vectors.** The observations that the AAV-6 vectors provided the maximum gene expression from dual vectors prompted us to develop the AAV-6 mini-dystrophin overlapping vectors. We evaluated the efficacy of the mini-dystrophin overlapping vectors in a cohort study in *mdx* and *m-dko* mice. The TA muscles were injected with 1 x 10^{10} vg particles per virus (1 x 10^{10} vg particles for co-infections) whereas EDL muscles were injected at a dosage of 2.5 x 10^9 vg particles per virus (5 x10^9 vg particles for co-infections). The transduction efficiency was evaluated by immunofluorescence analysis of mini-dystrophin expression in the infected TA muscles. We observed very few fibers that were positive for the mini-dystrophin

transgene expression (Figure 2.7A, positive fibers indicated by arrows). There was no expression from the AV6.RSV.mini-dystrophinUpstream and AV6.RSV.mini-dystrophinDownstream single vector infections. The muscle function assay correspondingly indicated that the specific force and eccentric contraction parameters for the mini-dystrophin overlapping co-infected EDL were not improved over the uninfected controls (Figure 2.7 B-C). The cohort study simultaneously done in the C57/BL10 mice indicated no adverse effect of the injections in normal mice on the basis of morphological analysis of TA and EDL muscles (data not shown).

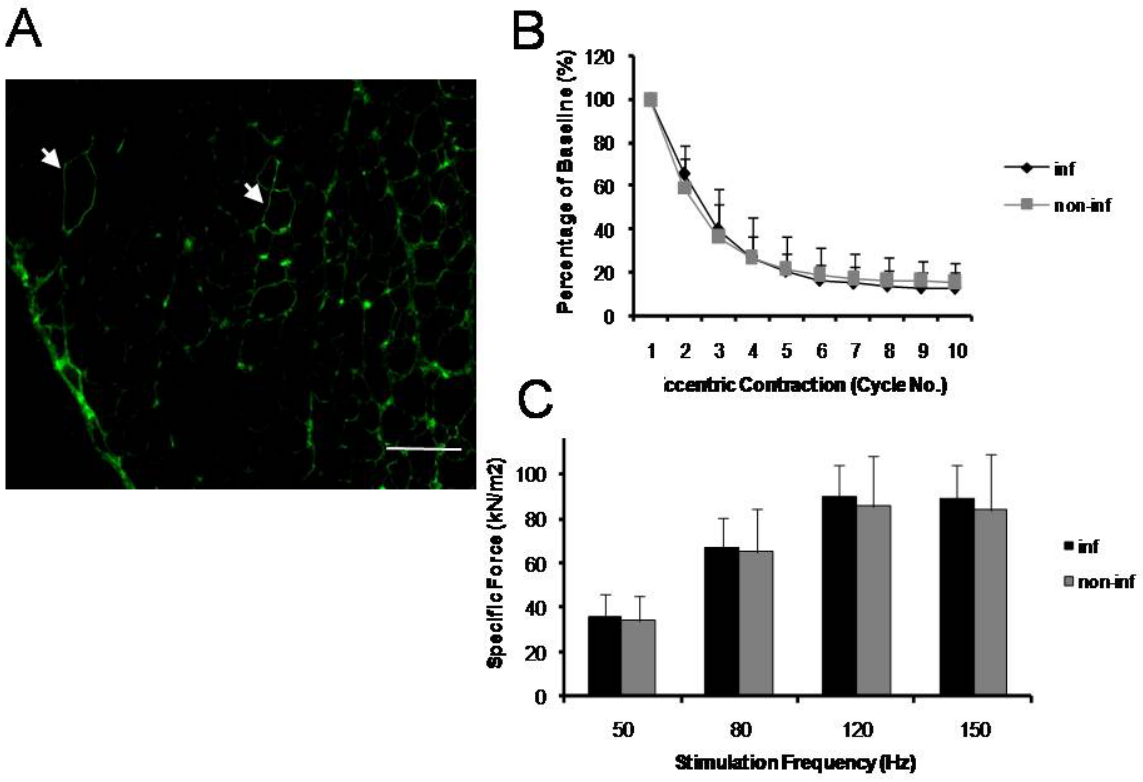


Figure 2.7. *In vivo* evaluation of the mini-dystrophin overlapping vector in *mdx* skeletal muscle. **A**. Detection of dystrophin expression by immunofluorescence staining in *mdx* TA muscle injected with 2 x 10^{10} vg of a mixture of AV6.RSV.mini-dystrophinUpstream and AV6.minidystrophinDownstream. The dystrophin N-terminal specific antibody was used for staining Scale bar, 100μm. Arrows indicate myofibers expressing mini-dystrophin. This experiment was performed with the help of Yongping Yue. **B**. Muscle function analysis in *mdx* mice; protection from eccentric contraction induced injury in EDL muscles. Black line indicates mini-dystrophin overlapping infected EDL muscles and gray line indicated uninfected control. **C**. Muscle function analysis in *mdx* mice; tetanic force generation in EDL muscles upon electrical stimulation. Black bars indicate mini-dystrophin overlapping vector infected EDL muscles and gray bars indicate uninfected control. (These experiments were performed by Dr. Mingju Liu and Yongping Yue.)

To control for mouse strain differences between this study in *mdx* mice and the C57/BL10 mice used for the reporter gene study, we infected *mdx* mice with the AP and LacZ overlapping vectors. The transduction profiles of the overlapping LacZ and AP as well as their corresponding LacZ and AP intact viruses was similar to that observed in BL10 mice (Figure 2.8). In the absence of any dramatic differences in transgene expression (compare figures 2.5, 2.6 and 2.8), quantification of transgene positive myofibers for the *mdx* set is not shown. These controls indicate that instead of the mouse strain difference, it is the

transgene sequence itself that is the most important determinant of efficiency in the overlapping approach.

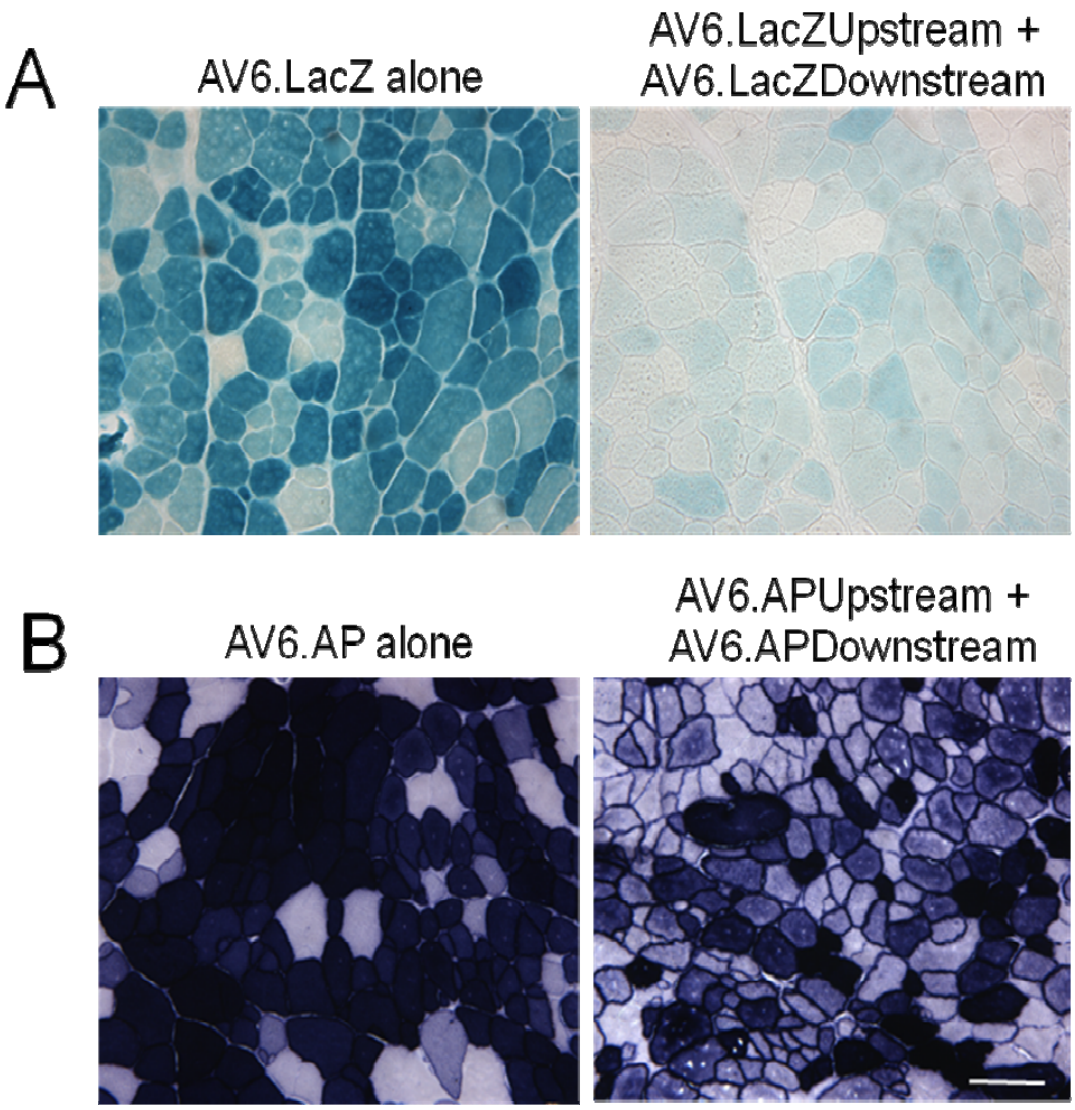


Figure 2.8. Control experiments to exclude the effect of mouse strain on efficiency of overlapping vectors. The TA muscles of *mdx* male mice were infected with 5 x 10^8 vg particles of each indicated AAV viruses. LacZ and AP expression was examined at six weeks later by histochemical staining. **A.** AAV-6

LacZ overlapping and intact infections in *mdx* TA muscles. **B.** AAV-6 AP overlapping and intact infections in *mdx* TA muscles.

D. Discussion

To overcome the size limitation hurdle of AAV vector, several novel approaches have been developed including the *trans*-splicing, *cis*-activation and overlapping methods [31]. Each approach has its unique advantages and disadvantages [32]. In this part of the study we examined the potential rate-limiting factors in the overlapping approach. The beauty of this method lies in its simplicity of its genetic engineering and the independence of the viral ITR-mediated recombination. Interestingly, two published studies revealed contradictory results, from a very poor transduction in muscle with the AAV-2 LacZ overlapping vectors to an extremely high efficiency in the lung with the AAV-6 AP overlapping vectors [35, 36]. A better understanding of the mechanisms underlying these observations is critical to improve overlapping vector-mediated gene therapy.

The transgene, viral serotype and target tissue are the three major differences between two published studies [35, 36]. To determine the relative contribution of each of these factors, we generated a series of AAV-2, -5 and –6 viruses carrying the intact and the overlapping LacZ or AP genes, respectively. We first examined transduction efficiency in cell lines. Previous studies suggest that the relative transduction efficiency of the overlapping vectors is 1-2.4 % of that of the intact gene vector in cell lines [35]. However, the relative transduction

efficiency was surprisingly much higher with the LacZ vectors in both HT1080 and CF16 cells. Since transduction efficiency with the intact gene vector (AV.LacZ) was also very low, we initially suspected that there might exist a subset population of cells that were highly susceptible to AAV infection and these cells might also carry an inherent property for efficient homologous recombination. However, additional studies with the AP vectors suggest that this may not be the case. A 98% transduction efficiency was achieved with AV.AP in all three AAV serotypes at the same viral dosage as was used in the LacZ virus experiments. Consistent with previous studies [36], the relative transduction efficiency of the AP overlapping vectors is 1-3%. Since the LacZ gene has been widely used in many different cell types for many years, it is unlikely that the LacZ protein be toxic to HT1080 and CF16 cells. The cytochemical methods for detecting the LacZ and the AP proteins are based on the enzymatic activity. Since LacZ is a tetrameric enzyme, while AP functions as a dimer [61], it is possible that more LacZ monomers will be needed to achieve the same enzymatic activity as compared to the AP protein. In support of this notion, it has been shown that the enzymatic staining assay may underestimate LacZ gene expression [62]. Although the difference in assay method may have partially contributed to the observed difference between the LacZ and the AP experiments *in vitro*, the exact mechanisms remain to be explored.

However, the transduction profile of the overlapping AP vectors in the MO59K cells was dramatically different. The number of positive fibers from the AAV-6 AP overlapping virus infections reached 70.7% of the intact while the

numbers for AAV-2 and AAV-5 reached 16.7% and 17.3%. Thus the relative levels of expression of the overlapping AP vectors in this cell line follows a similar trend to the LacZ set. This high level of transduction in the MO59K cells indicates that the repair and recombination pathways in this cell line are perhaps not impaired as compared to the HT1080 and CF16 cells. Furthermore, the transduction profile of the different viruses in this cell line closely resembles that from the *in vivo* results. Therefore it suggests that the MO59K cells are more suitable for evaluation of dual vectors *in vitro.*

AAV serotypes are known to affect viral uptake, intracellular processing and transgene expression. To determine the impact of viral serotype on overlapping vector-mediated gene transfer *in vivo*, we compared the AAV-2, -5 and –6 overlapping vectors. AAV-5 and AAV-6 are more efficient than AAV-2 for muscle gene transfer [15, 16, 49, 58, 59]. Since the effect of viral capsids on transduction may not be evident at high viral doses, we deliberately used a low viral dose in our *in vivo* studies. Consistent with previous publications [49], AAV-5 intact gene vector was much more efficient than that of AAV-2. The highest muscle transduction was achieved with AAV-6 [15, 16, 59].

One of the most important findings of this study is efficient muscle transduction with the AAV-6 overlapping vectors. Approximately 35% and 80% myofibers were transduced by the AAV-6 LacZ overlapping vectors and the AAV-6 AP overlapping vectors, respectively (Figure 2.5 and 2.6). This result strongly argues against the tissue tropism of the overlapping vectors. Our results suggest that in addition to the lung [36], skeletal muscle is also an excellent target tissue

for overlapping vector-mediated gene therapy. This finding holds practical implication for DMD gene therapy. To overcome the size limitation, 3.8-4.2 kb micro-dystrophin genes have been engineered for AAV-mediated DMD therapy [63]. However, these microgenes are functionally less competent than that of the 6 kb mini-dystrophin gene [22]. Since 20% level of dystrophin expression is sufficient to correct the skeletal muscle pathology in DMD [64], overlapping vector-mediated mini-dystrophin expression may represent a viable approach for DMD gene therapy. To this end we developed and tested the mini-dystrophin overlapping vectors pseudopackaged in AAV-6 capsids in the disease models for DMD including *mdx* (dystrophin knock out) and *mdko* (dystrophin/myoD double knockout) mice. But this set of vectors provided very poor transduction, the level of which was far lower to be of any meaningful therapeutic effect (Figure 2.7). To rule out the mouse strain differences, we duplicated the LacZ and AP overlapping infections in *mdx* mice using the same batch of AAV-6 virus used previously. We observed similar levels of transduction in the *mdx* mice by the overlapping LacZ and AP vectors.

The PCR studies of Hirt DNA isolated from overlapping vector infected cells indicate the presence of detectable levels of reconstituted transgenes (Figure 2.4). Homologous recombination has been proposed as the underlying mechanism for the overlapping approach. However, the detailed recombination process remains to be defined. One interesting aspect is the contribution of the transgene sequence to recombination efficiency. The level of GC content, the diversity of nucleotide distribution, the presence of the secondary structure and

the existence of certain sequence motifs may ultimately influence the kinetics of the recombination reaction [65-67]. We observed a dramatic difference between the LacZ gene and the AP gene overlapping vectors. The relative efficiency of the LacZ overlapping vectors was higher than that of the AP overlapping vectors *in vitro*, although the level of difference was dependent on the cell line used (Figures 2.1 and 2.2). However, in mouse muscle, the relative efficiency of the LacZ overlapping vectors (~42%) was lower than that of the AP overlapping vectors (~96%). It was not immediately clear whether this is due to the source of the transgene (bacterial for the LacZ gene and eukaryotic for the AP gene), or the sequence itself. Since recombination hot spots have been identified in bacterial as well as mammalian genes [65-67], it is possible that the organization and/or composition of the middle one third of the AP gene may be more prone for homologous recombination *in vivo*. The difference in recombination capacity of different overlapping sequence/genes is further exemplified by the case of the mini-dystrophin overlapping vectors which apparently failed to reconstitute to therapeutic levels *in vivo* (Figure 2.7). Furthermore, the strain of mouse used in these experiments did not influence the recombination efficiencies (Figure 2.8). These results reiterate our observation that the transgene sequence itself is indeed the most important factor influencing the success of overlapping dual vectors.

It should be noted that the GC content of the mini-dystrophin overlapping sequence is 47.05%, which is lower compared to LacZ (57.81%) and AP (65.03%) overlapping sequences which may be a consideration for efficiency of

recombination [65, 66]. Despite the difference in GC content, further experiments will be required to determine whether the GC content is an important factor affecting the transduction efficiency of the different overlapping vector sets. Nevertheless, transduction efficiency achieved with the AAV-6 overlapping vectors may meet the therapeutic needs for certain muscle gene therapy applications. Although the overlapping approach did not work for the mini-dystrophin gene, it may still work for other genes like dysferlin or CFTR, depending on the recombinational qualities of the chosen overlapping sequences.

E. Acknowledgements

We thank Dr. A. Dusty Miller for providing the AAV-6 packaging plasmids and CF16 cells. We thank Dr. David J. Pintel for providing HT1080 cell line. I thank Yongping Yue for assistance in harvesting and sectioning of muscle samples. I thank Dr. Mingju Liu for the muscle function and statistic analysis and Dr. Dongsheng Duan for assistance in writing. This work was supported by grants from the National Institutes of Health (AR-49419, DD) and the Muscular Dystrophy Association (DD).

Chapter 3. Whole body gene transfer using dual vectors

A. Introduction

For genetic diseases that affect multiple tissues and/or organs, the primary challenge in gene therapy for these diseases is delivering the therapeutic gene to all the affected tissues/organs. Systemic gene transfer has been a daunting task for many years until recent success with adeno-associated virus (AAV) [15-17, 59, 68-74]. To this end, earlier interest in serotype-2 AAV (AAV-2) has now expanded to a large number of newly isolated serotypes [13]. The diverse capsid sequence and/or structure of new serotypes have resulted in striking differences in the transduction profiles. Although the initial comparisons of different AAV serotypes were made in a single tissue by local delivery, it was soon evident that many new serotypes could reach remote tissues after intra-vascular injection. The biological activity of these novel AAV serotypes remained largely unaffected by their presence in the bloodstream or capacity for capillary escape. Impressive body-wide transduction has been reported for AAV-6, -8 and -9 [15-17, 59, 70, 73]. Several groups have capitalized on this unique property and achieved whole body amelioration in the mouse models of several inherited diseases including DMD, congenital muscular dystrophy, limb-girdle muscular dystrophy, Pompe disease and type II glycogen storage disease [15, 68, 69, 71-74].

We have recently demonstrated efficient transduction of tsAAV vectors in a mouse model of DMD by local injection [23]. More than 90% myofibers in a single muscle were transduced by tsAAV. This efficiency is quite comparable to that of a single intact AAV vector [23, 54, 75]. These studies established that the expression limiting factors influencing gene expression from tsAAV vectors are the choice of splice junction used to split the gene, mRNA accumulation [23] and to an extent, the intron sequence [76].

Despite the success in local gene transfer, it is not clear whether one can achieve efficient whole body transduction with tsAAV vectors. A cure for many genetic diseases requires global transduction of all the affected tissues and/or organs. In the case of DMD, morbidity and mortality are the results of severe pathology not only in skeletal muscle, but also in the heart. These observations laid the foundation for experimentally determining the efficiency of dual vectors for body-wide gene transfer. Achieving systemic gene transfer will bring tsAAV a step further to clinical application. In this section of the study, we explored the potential of systemic gene transfer with AAV serotype 9 (AAV-9) trans-splicing vectors. Using a pair of optimized alkaline phosphatase (AP) gene tsAAV vectors, we detected therapeutic level expression in all body skeletal muscle and the heart. We also observed high-level transduction in the lung. This is the first demonstration of successful whole body transduction with tsAAV. These results provide the foundation for applying this promising vector technology to the disease models for future human gene therapy. We further evaluated if the age

of mice affects AAV-9 transduction. We observed certain interesting differences in gene expression in certain tissues depending on age.

B. Materials and Methods

Proviral plasmids. The cis plasmid for the single AP intact vector (pcisAV.RSV.AP) has been reported before [53, 54]. The cis plasmids for AP tsAAV were generated by a multi-step cloning procedure. We first inserted a synthetic intron from pCI (Promega, Madison, WI) at the junction of exons 8 and 9 in pcisAV.RSV.AP. Next, we created a unique KpnI site at the center of the synthetic intron in this intermediate construct by inactivating a second KpnI site at the backbone. There are two SalI sites in the original pcisAV.RSV.AP plasmid. One is located right before the RSV promoter and another is located right after the polyA. We then generated two additional intermediate constructs by selectively inactivating one of the SalI sites. Finally, pcisAV.RSV.AP.Donor was generated by KpnI/SalI double digestion from one intermediate of the last step to remove the second half of the intron, the 3' end of the AP gene and polyA. pcisAV.RSV.AP.Acceptor was generated by SalI/KpnI digestion from the second intermediate from the last step to remove the RSV promoter, the 5' portion of the AP gene and the first half of the synthetic intron.

AAV production. AAV-6 vectors were produced as described in chapter 2. AAV-9 vectors were produced by triple plasmid transfection using a pRep2/Cap9

helper plasmid generously provided by Drs Gao and Wilson [3, 77]. Recombinant viral stocks were purified through two rounds of isopycnic $CsCl_2$ ultracentrifugation as we have described before in chapter 2. Viral titration and quality control were performed according to our previous publications [23, 41].

Cell culture. MO59K cells were cultured as described in chapter 2. Viral infections were done at an moi of 500 vg particles per virus per cell (1000 vg particles for co-infections). These sets of infections were done at a low dose to properly differentiate between the transduction efficiencies of the AP intact, overlapping and trans-splicing vectors. All cell culture experiments were performed using AAV-6 vectors.

***In vivo* gene delivery.** All animal experiments were approved by the Animal Care and Use Committee at the University of Missouri and were in accordance with NIH guidelines. In this study, we used newborn C57Bl/10 (BL10) mice. Adult breeders were purchased from The Jackson Laboratory (Bar Harbor, Maine). Breeding was carried out in a specific-pathogen-free facility. We first optimized the injection conditions with tattoo dye (time, route and volume). Best results were achieved at 48 hrs after mice were born by facial vein injection. The maximal deliverable volume was 70 µl. For AV.RSV.AP infection, 3.5×10^{11} vg AAV-9 particles were delivered in 70 µl HEPES buffer with a custom-designed 33-gauge gas-tight Hamilton syringe (Hamilton Co, Reno, NV). For tsAAV infection, 3.5×10^{11} vg particles (1.2×10^{11} vg/g body weight, 50 µl/g body

weight) of each vector (AV.RSV.AP.Donor and AV.AP.Acceptor) were mixed thoroughly in 70 µl HEPES buffer and then delivered using similar method. In mock treated animals, we injected 70 µl HEPES buffer. After injection, mice were returned to the cage for recovery. More than 95% neonatal mice survived the procedure.

In young adult BL10 mice (7-wk-old), 7×10^{10} vg/g body weight (1×10^{12} vg/mouse) of AAV were delivered via tail vein injection (adult venous). AP expression was examined 12 weeks later by histochemical staining of tissue sections and AP activity assay of tissue lysates, according to our published protocols.

Gene expression evaluation. At six weeks post-infection, mice were euthanized and various tissues/organs were examined for AP expression using our published histochemical staining protocol [53] and an AP enzyme activity assay (StemTAG AP Activity Assay kit; Cell biolabs Inc. *http://www.cellbiolabs.com ;* Cat # CBA-301). To compare transduction efficiency among experimental groups, we used two different quantification methods. We first performed morphometric analyses on digitized tissue section images. In skeletal and cardiac muscle, the liver and the kidney, we used NIH Image J software (version 1.36b) to quantify the percentage of AP positive area (AP positive area/total tissue area) and relative AP staining intensity (the mean intensity of AP positive area in AAV infected tissue normalized by the average mean intensity in mock-infected tissue) according to previous publications [78-

81]. In morphometric study, three to seven representative cross-section images were quantified for each tissue sample. Images from three to eight independent samples were examined for each tissue. In lung gene therapy studies, transduction efficiency is usually quantified by counting the number of positive cells in conducting airway and alveoli, respectively [36, 53, 82]. We have used similar manual quantification method for the lung. Besides morphometric quantification, we also measured the total AP activity in tissue lysate using the StemTAG AP kit. The assay was performed essentially as described in the manufacturer's manual except that we first incubated tissue lysate at 65ºC for 1 hr to inactivate endogenous heat-labile AP. In lysate assay, the AP activity is defined as μM of para-nitrophenol being generated per μg of total protein in the tissue lysate (one molecule of para-nitrophenol is generated when one molecule of phosphate is released from the substrate).

Statistical analysis. Data are presented as mean ± standard error of mean. Statistical analysis was performed with the SPSS software (SPSS, Chicago, IL). In two-group comparison (such as these in morphometric quantification studies), we used two tailed independent student t test. In three-group comparison (such as tissue lysate AP activity assay), we used one-way analysis of variance followed by Tukey post hoc analysis. Difference was considered significant when $P < 0.05$.

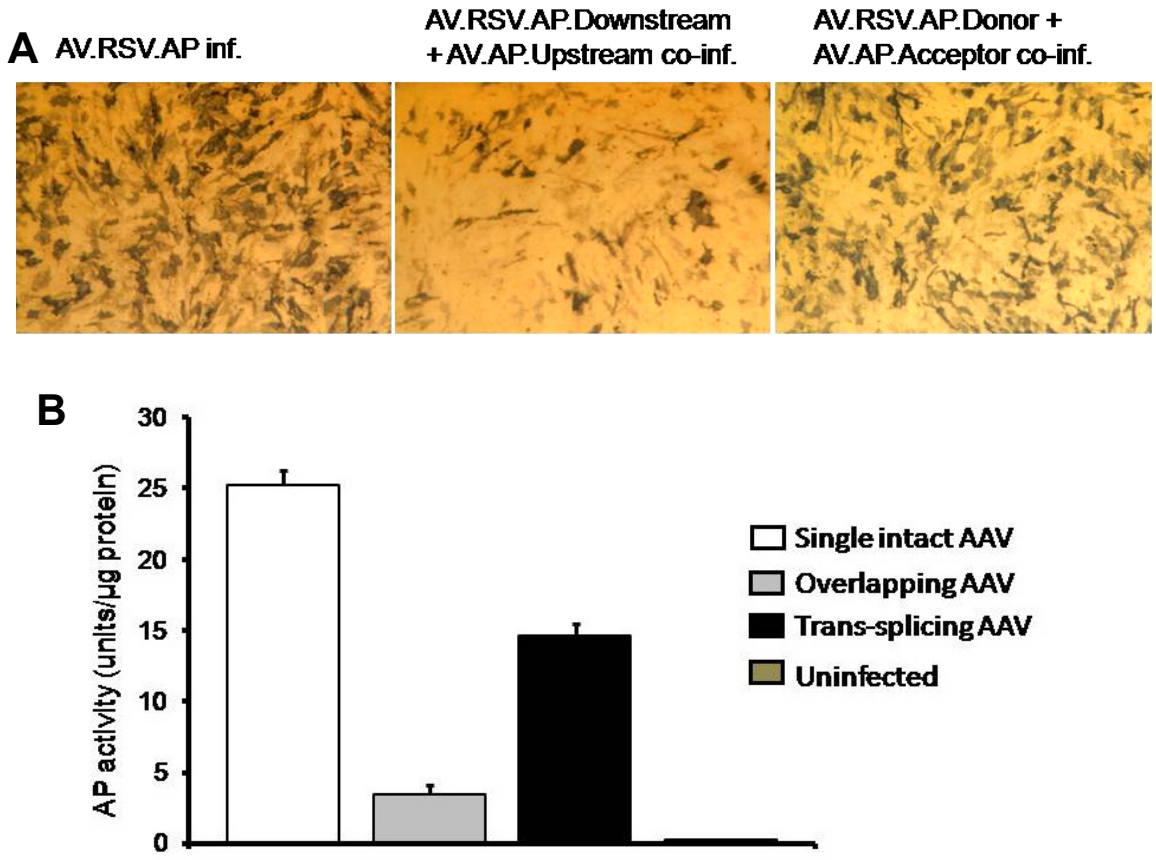


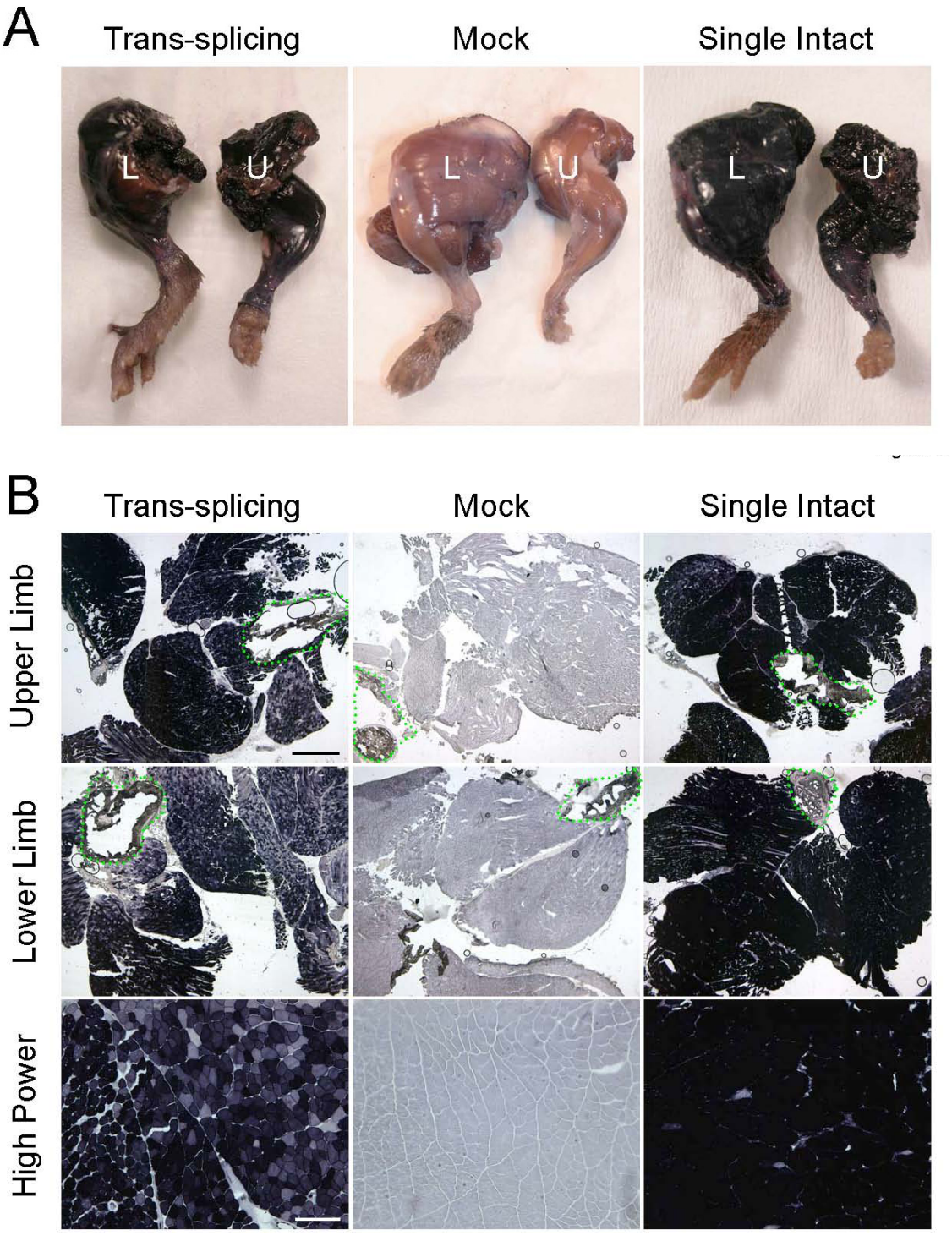
Figure 3.1 Comparison of tsAAV AP vectors with overlappingAAV AP vectors and single intact vectors AP *in vitro*. MO59K cells were infected with indicated virus sets at 500 vg particles per virus per cell (1000 vg particles for dual vectors). **A.** Cytochemical staining for alkaline phosphatase expression was done 45 hrs. post infection. **B.** Quantitation of transgene expression levels by alkaline phosphatase assay done for the indicated virus infections and reported as µMols of assay substrate utilized/µg total protein in lysate.

C. Results

Developing optimized reporter gene tsAAV vectors. Several sets of reporter/therapeutic gene tsAAV vectors have been published to express the LacZ gene, the erythropoietin gene, the factor VIII gene and the mini-dystrophin gene, respectively [23, 30, 31, 35, 39, 40]. Besides our recently reported mini-dystrophin tsAAV, none of the published tsAAV vectors achieved the same transduction efficiency as that of a single AAV. To evaluate systemic transduction of the tsAAV vectors, we first constructed a pair of rationally designed reporter gene tsAAV to express the heat-resistant human placental AP gene (AV.RSV.AP.Donor and AV.AP.Acceptor). Using mRNA accumulation as the guide, we split the AP gene at the junction of exons 8 and 9 [23, 41]. The splicing signals are obtained from a synthetic intron as we have described [41, 83]. As a control, we also included a single intact AAV vector (AV.RSV.AP) in our study [53, 54]. Gene expression in these vectors (both tsAAV and single AAV) is controlled by the identical transcription regulatory elements (the RSV promoter and the SV40 pA signal).

We first examined the transduction efficiency of AP tsAAV vectors in cell culture (Figure 3.1). Efficient AP expression was observed only in AV.RSV.AP.Donor and AV.AP.Acceptor co-infected cells. There was no leaky expression from singly infected vector (AV.RSV.AP.Donor alone or AV.AP.Acceptor alone) (data not shown). The AP lysate assay demonstrated the efficient performance of the tsAAV AP vectors over the overlapping AP vectors

even at a low moi of 500 vg particles per virus per cell (1000 vg particles for co-infections).



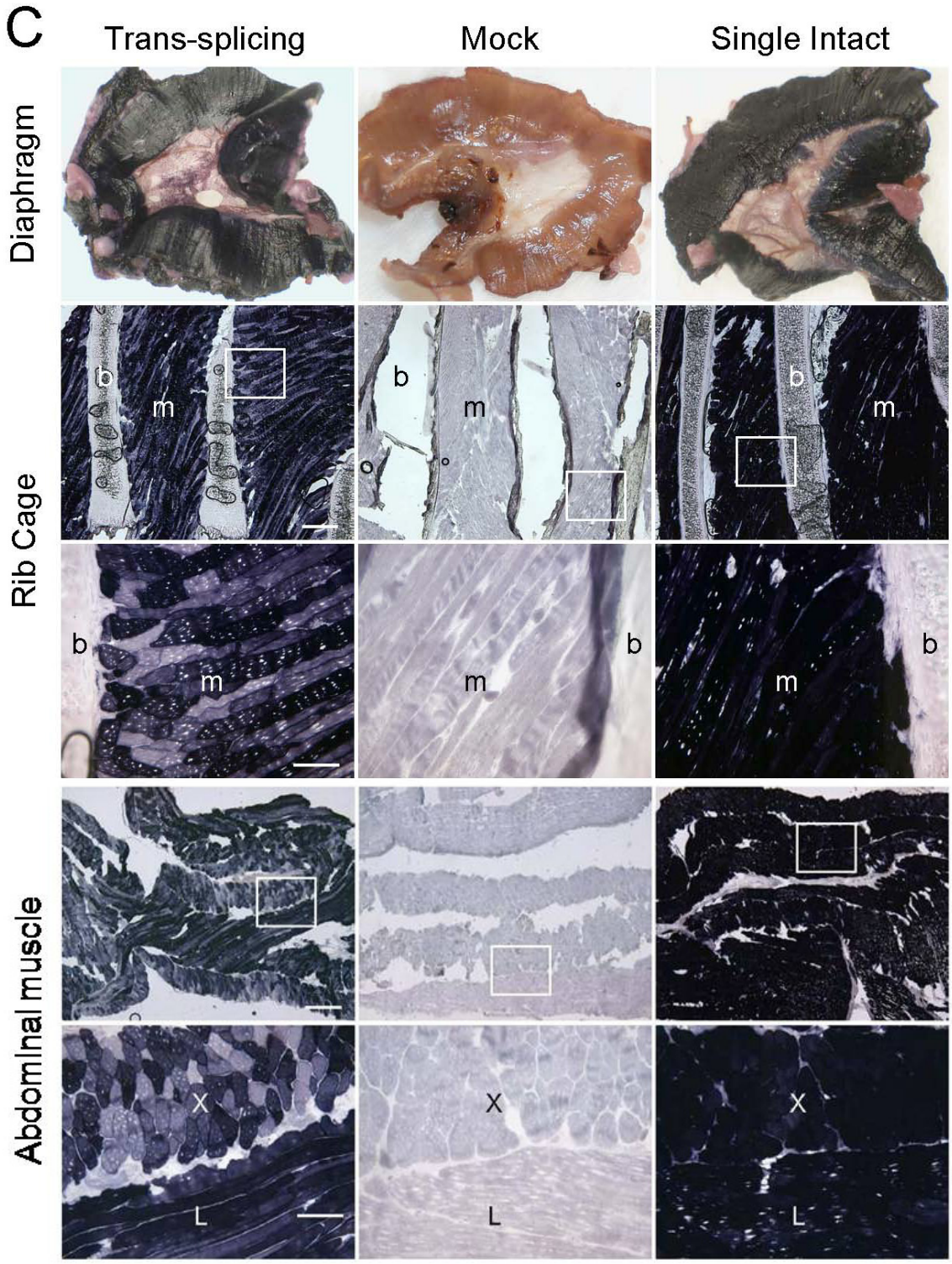


Figure 3.2. Wide-spread skeletal muscle transduction following intra-venous delivery of the trans-splicing AAV vectors in newborn mice.

Representative histochemical staining photomicrographs from each group (N = 3 to 5 mice/group) are shown. **A,** Full-view of the whole mount limb muscle (L, lower limb; U, upper limb); **B,** Cross-sections of the entire limb (top two rows; scale bar, 1 mm) or high magnification view of representative areas of muscles in the lower limb (bottom row; scale bar, 200 μm), respectively. Dotted lines in low magnification photomicrographs mark the location of bone; **C,** AP expression in respiratory muscles including the diaphragm (top row, whole mount images), inter-costal muscles (middle two rows, cross sections), and abdominal muscles (bottom two rows, cross sections), respectively. Photomicrographs in high magnification (scale bar, 100 μm) are the closer view of the boxed areas in the respective low magnification (scale bar, 400 μm) photomicrographs. b, bone (rib); m, muscle; X, regions where myofibers are in cross-section orientation; L, regions where myofibers are in longitudinal orientation. (*Adapted from Ghosh, et al, Molecular Therapy., 2007, (4):750-5).*

Systemic delivery of AP tsAAV resulted in bodywide transduction of striated muscle but not smooth muscle. Since our goal is to achieve whole body transduction, we delivered AP tsAAV vectors through facial vein in newborn mice. Control groups included mock-infection (negative control) and single intact AV.RSV.AP infection (positive control). The AV.RSV.AP group represented the maximal achievable level of expression. Transgene expression was examined at six weeks later.

In single intact vector infected mice, we obtained widespread high-level expression in all striated muscles including all skeletal muscle of the body and the heart (Figures 3.2, 3.3 and 3.5A). When AP trans-splicing vector infected mice were evaluated, we also observed broad AP expression in all skeletal muscles throughout the entire body (Figures 3.2, 3.3 and 3.5A). Although tsAAV group showed a reduced staining intensity (Figure 3.5A), the total numbers of transduced skeletal myofibers were quite comparable between the tsAAV and single AAV groups. Overall, we achieved ≥ 90% transduction in both groups (Figure 3.5A). In tissue lysate assay, the total AP activity was not significantly lower in tsAAV infected skeletal muscle but there was a trend towards reduced activity (Figure 3.5A). High-level AP expression was also observed in the hearts in tsAAV infected mice (Figures 3.3, 3.5A). In morphometric quantification, more than 50% cardiomyocytes were transduced. In tissue lysate, the AP activity in tsAAV infected heart reached ~60% of that in single AAV infected heart.

Despite robust transduction in skeletal and cardiac muscle, we observed limited AP expression in smooth muscle (Figure 3.6). Since surrounding cells (such as skeletal muscle myofibers, cardiomyocytes and endothelial cells) were efficiently transduced, it is very likely that the poor transduction in smooth muscle is due to certain intrinsic biological properties of these cells rather than vector spreading/distribution.

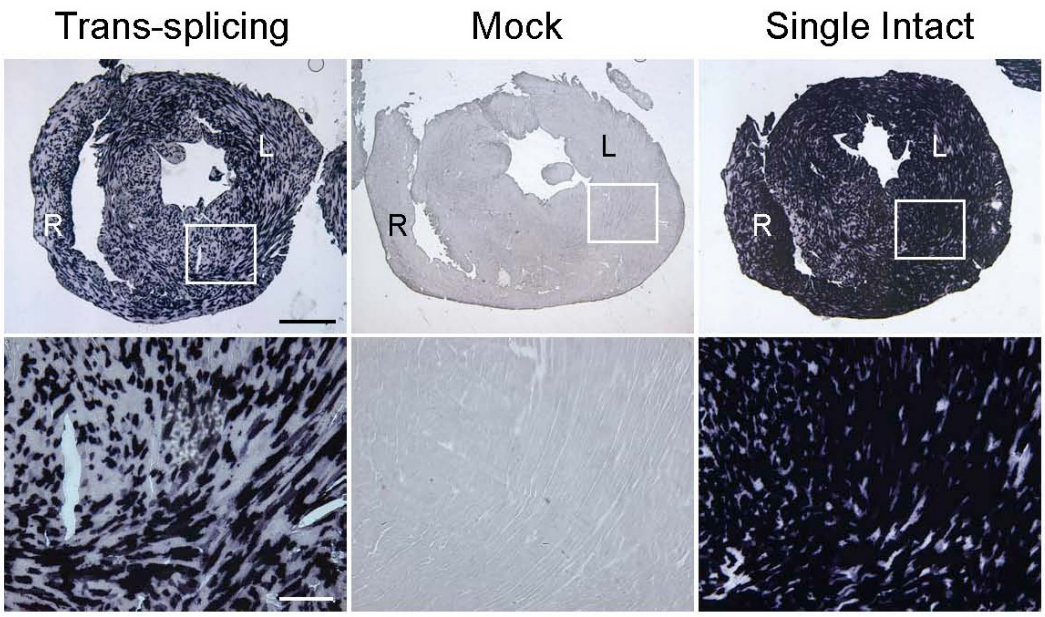
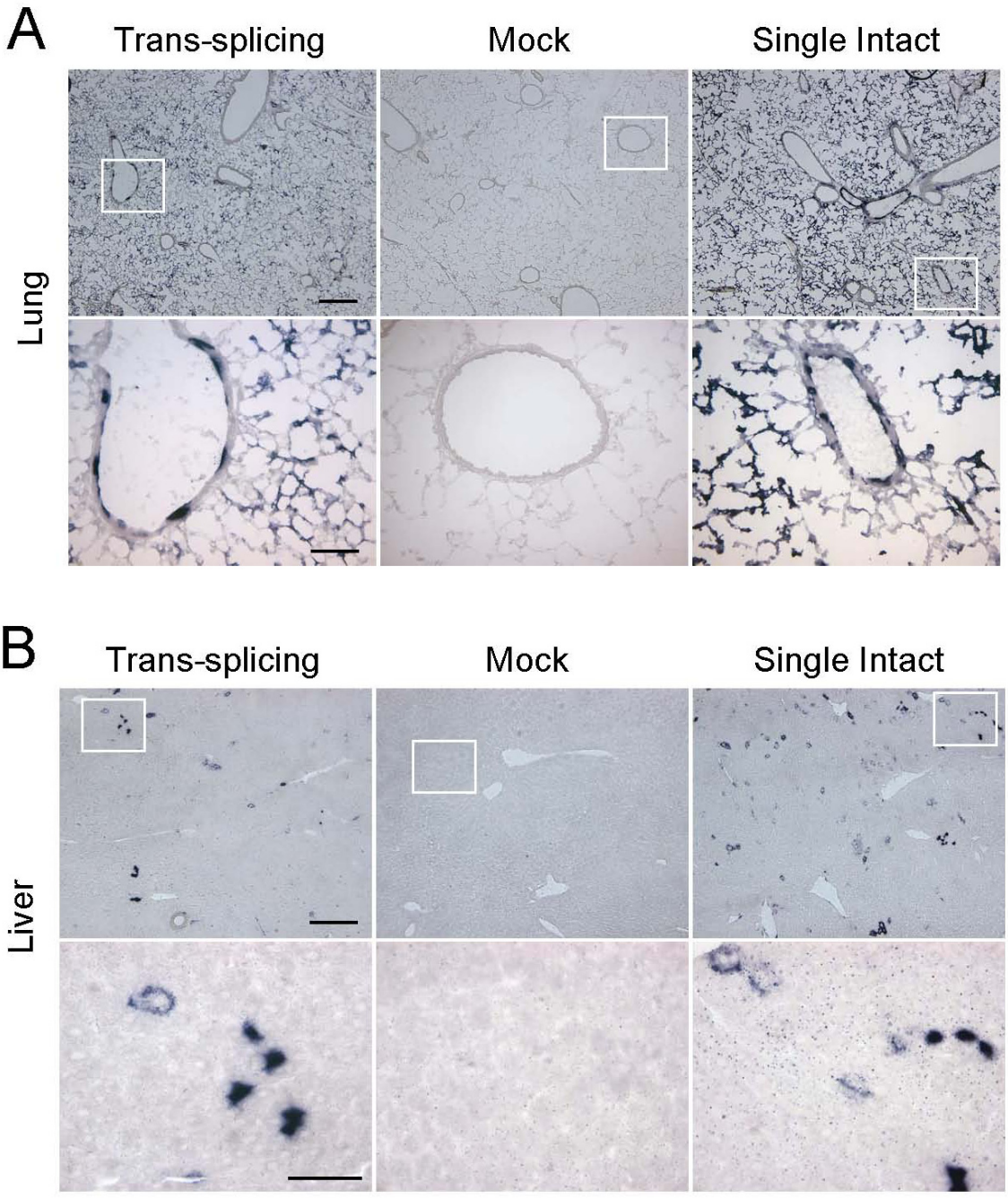


Figure 3.3. The trans-splicing AAV vectors efficiently transduce cardiac muscle after neonatal systemic delivery. Photomicrographs are representative images of tsAAV, mock and single AAV infected hearts, respectively. Top panels, full view images of the entire heart sections. R, right ventricular wall; L, left ventricular wall. Scale bar, 1 mm (low magnification); Bottom panels, closer view of the boxed areas in the top panels. Scale bar, 200 μm. (*Adapted from Ghosh, et al, Molecular Therapy., 2007, (4):750-5).*

Systemic delivery of AP tsAAV resulted in therapeutic level transduction in the lung but not in the liver and the kidney. We next evaluated transduction efficiency in several internal organs including the lung, the liver and the kidney (Figures 3.4 and 3.5B, 3.5C). In the lung, both alveolar cells and bronchiolar epithelial cells were transduced by tsAAV (Figures 3.4A and 3.44B). But the

overall AP activity in the tsAAV-infected lung was lower than that of the AV.RSV.AP infected lung (Figure 3.5B). Despite a high level transduction in the lung, we detected only limited transduction in the liver and the kidney with single intact AAV (Figures 3.4B and 3.4C). Interestingly, when we delivered the exact same batch of the AV.RSV.AP vector to adult mice via tail vein injection, we observed high-level transduction in the adult liver and kidney (Figure 3.7). Consistent with the poor transduction of AV.RSV.AP, tsAAV did not yield strong transduction either in the liver or the kidney after neonatal systemic injection (Figures 3.4B and 3.4C).


A
Trans-splicing
Mock
Single Intact
Lung
B
Trans-splicing
Mock
Single Intact
Liver

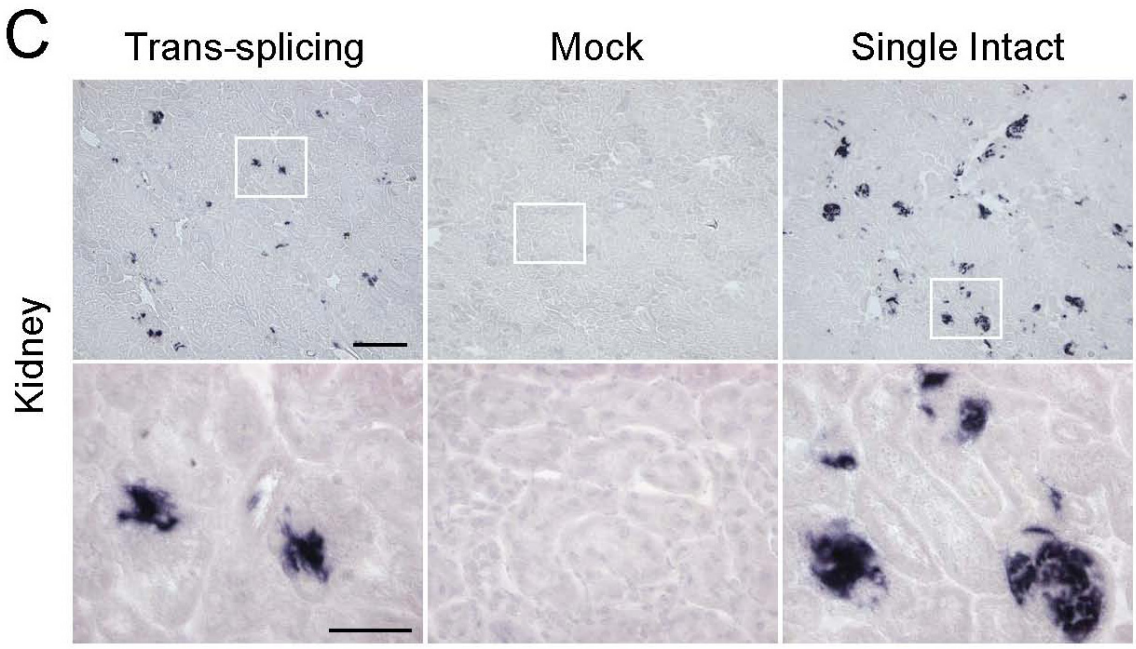


Figure 3.4. Systemic delivery of a single intact AAV or a pair of trans-splicing AAV vectors results in efficient transduction in the lung but only minimal transduction in the liver and the kidney. Photomicrographs are representative sections in **A,** the lung; **B,** the liver; and **C,** the kidney. Higher power magnification of the boxed areas in the top row is shown in the corresponding panels in the bottom row. Scale bar in panel A, 400 μm (top panel, low magnification) and 100 μm (bottom panel, high magnification), respectively; Scale bar in panel B, 200 μm (top panel, low magnification) and 50 μm (bottom panel, high magnification), respectively; Scale bar in panel C, 200 μm (top panel, low magnification) and 50 μm (bottom panel, high magnification), respectively. (*Adapted from Ghosh, et al, Molecular Therapy., 2007, (4):750-5).*

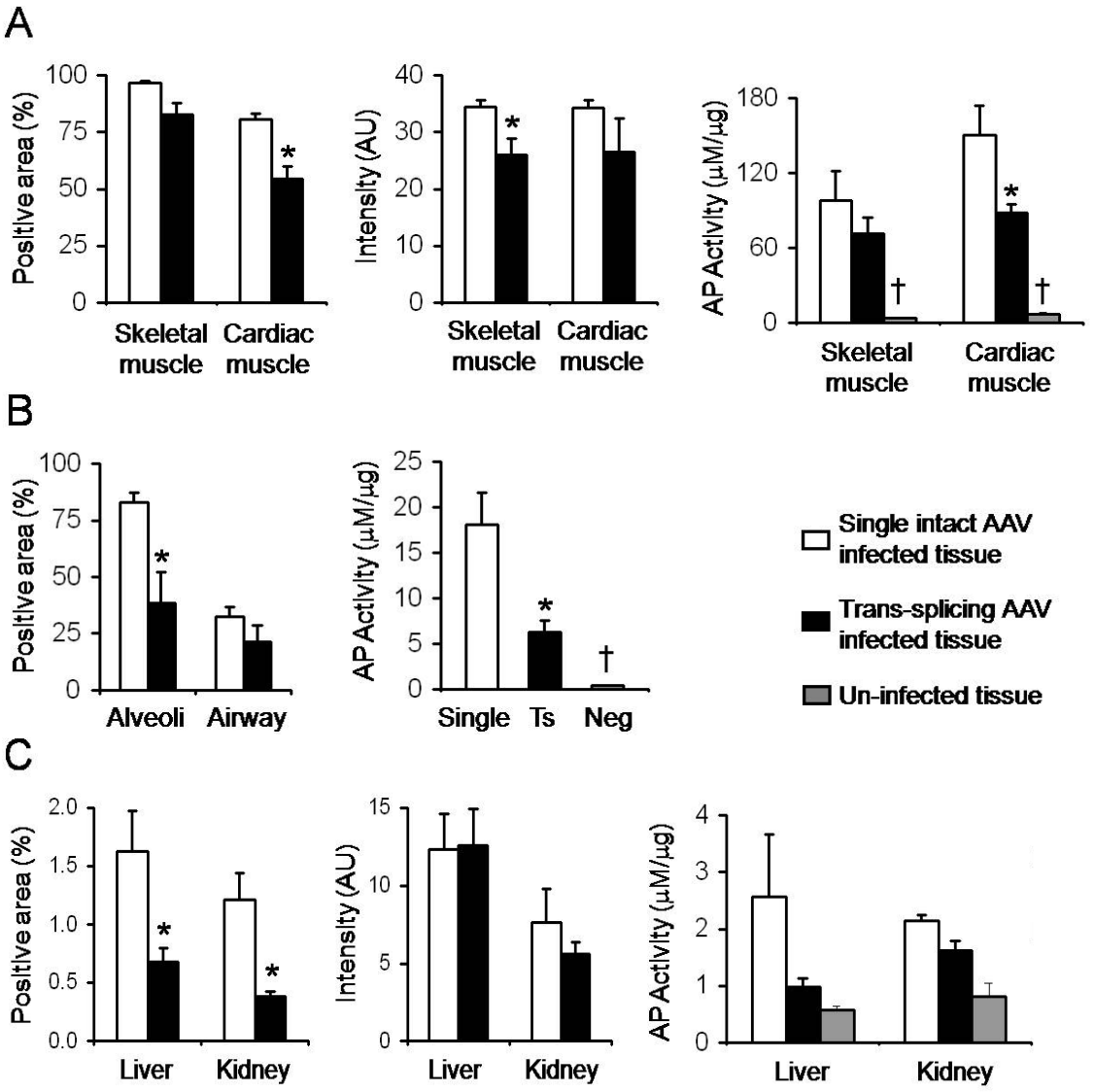
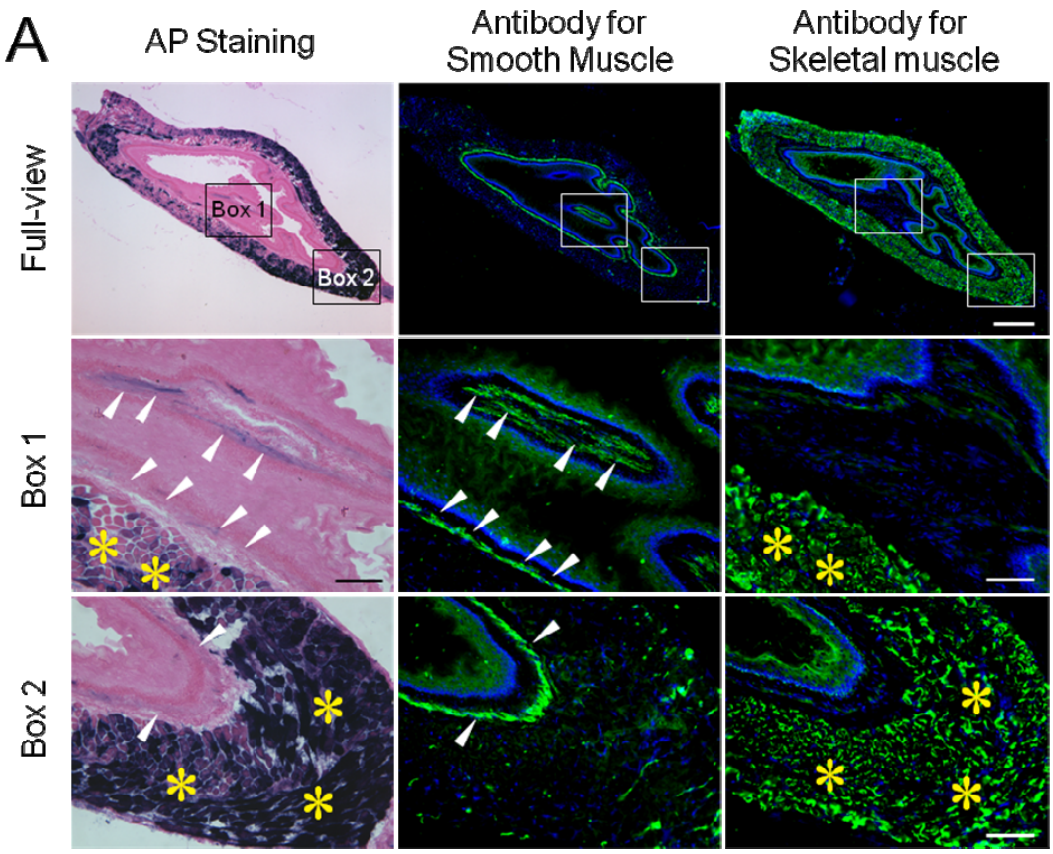
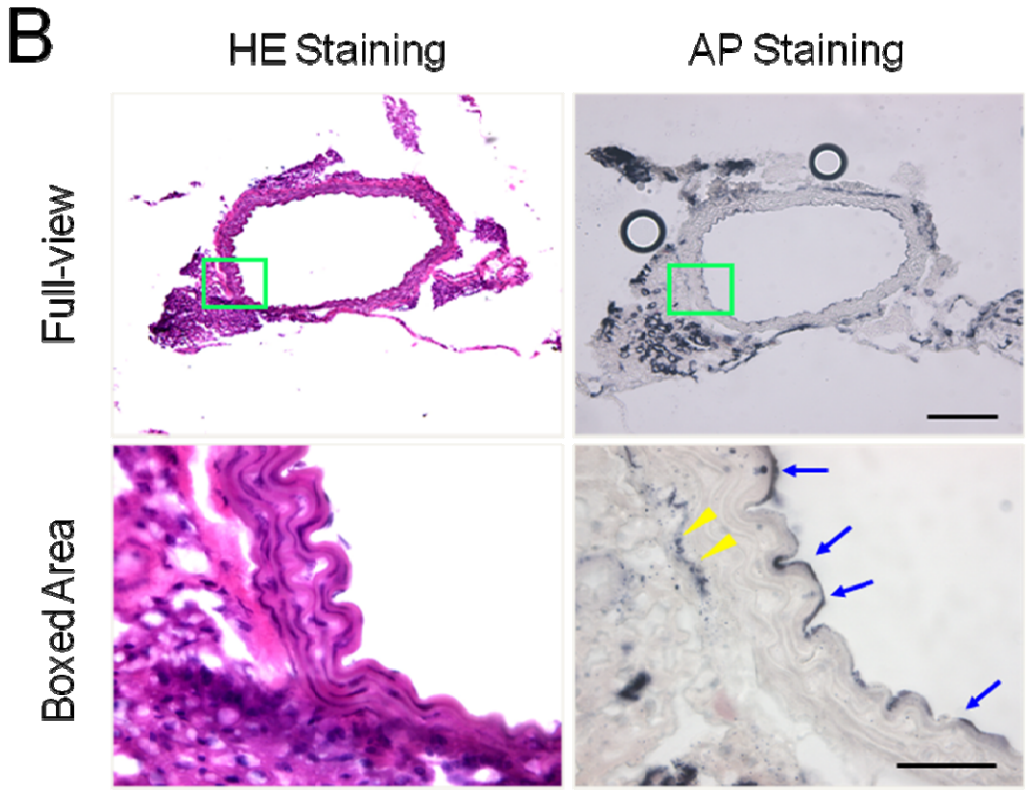


Figure 3.5. Quantitative comparison of the transduction efficiency of the trans-splicing vectors and that of the single intact vector. A, AP positive area, relative intensity and total tissue lysate activity in skeletal muscle and the heart. **B,** Percentage of AP positive cells in alveoli and conducting airway, respectively (quantified by manual counting). Also shown is the total tissue lysate AP activity. Single, AV.RSV.AP infected lung; Ts, tsAAV infected lung; Neg, mock-treated lung. **C,** AP positive area, relative intensity and total tissue lysate activity in the liver and the kidney. AU, artificial unit. Asterisk, statistically

different between the trans-splicing vector group and the single intact vector group. Cross, the AP activity in uninfected tissues is significantly lower than those of AAV infected tissues. (*Adapted from Ghosh, et al, Molecular Therapy., 2007, (4):750-5).*

Age of the mice an important aspect affecting gene expression in kidney, liver and eye. Another relevant issue in clinical gene therapy is patient age. Many diseases affect both newborn and adult patients. While some diseases are diagnosed late in life, a number of inherited diseases can be identified through neonatal screening. AAV gene therapies that are developed for adult patients (such as the timing, the route and the dose of vector delivery) may not suit well for neonatal patients and vice versa. Exploring the age-associated differences in AAV-9 transduction may have significant clinical implications. In contrast to the lung, systemic AAV-9 transduction of the liver and the kidney was affected by the age (Figure 3.7). Efficient transduction was achieved only in adult mice [84]. In most tissues/organs, we observed equal (skeletal muscle, the heart and the lung, data not shown) or better (the liver, the kidney and the aorta; Figure 3.7, 3.8) transduction in adult mice. In sharp contrast, the inner layer of the retina showed a different pattern (Figure 3.9). Robust AP staining was found only in mice infected at the neonatal stage. Very few cells in the inner retinal layer were transduced when AAV-9 was delivered to adult mice (Figure 3.9). This suggests that certain tissues may have developmental “windows” that facilitate AAV-9 transduction.


A
AP Staining
Antibody for Smooth Muscle
Antibody for Skeletal muscle
Full-view
Box 1
Box 2
Box 1
Box 2


B
HE Staining
AP Staining
Full-view
Boxed Area

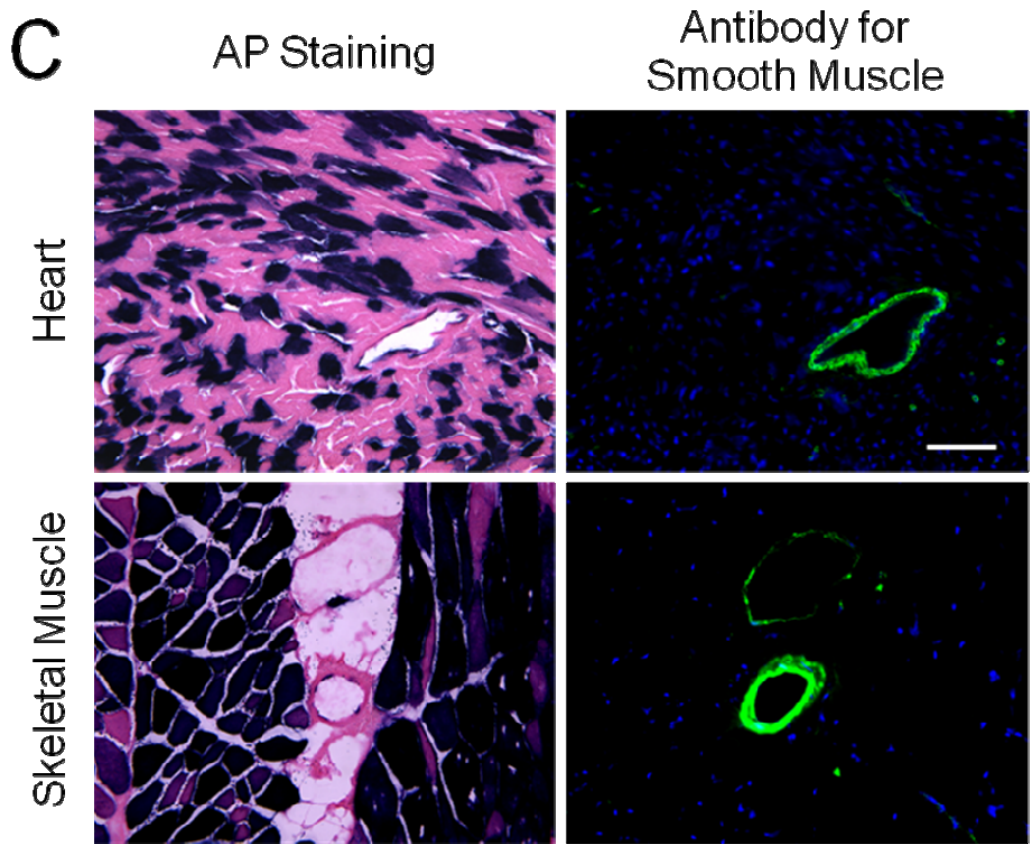


Figure 3.6. Systemic delivery of AAV-9 in neonatal mice does not lead to efficient smooth muscle transduction. 3.5 x 10^{11} vg particles of AV.AP (pseudopackaged in AAV-9 capsid) were delivered to newborn mice via facial vein. AP expression in smooth muscle was evaluated at six weeks post-infection. **A,** AP expression in the esophagus is robust in the skeletal muscle layer but poor in the smooth muscle layer. Serial sections of the middle one third of the esophagus was examined by AP staining (counter-stained with eosin) and immunostaining with antibodies specific for smooth muscle (monoclonal anti-α smooth muscle actin; catalog # A 2547; Sigma, St. Louis, MO) and skeletal muscle (polyclonal anti-skeletal muscle myosin; catalog # M 7523; Sigma, St.

Louis, MO), respectively. Top panel, full view photomicrographs of the entire esophagus cross-section; middle and bottom panels, high power photomicrographs of the boxed areas (box 1 and box 2, respectively) in the top panel. Arrowheads mark smooth muscle in the middle and bottom panels; Asterisks mark skeletal muscle in the middle and bottom panels. Scale bar in the top panel, 400 μm; scale bar in the middle and bottom panel, 100 μm. **B,** AP expression in the aorta. Some expression was seen in endothelial cells (arrows) but smooth muscle was scarcely transduced (arrowheads). Scale bar in the top panel, 200 μm; scale bar in the bottom panel, 50 μm. **C,** AP expression in small vessels in the heart (top panel) and skeletal muscle (bottom panel). Expression was detected in skeletal muscle myofibers, cardiomyocytes and endothelial cells but not in smooth muscle. Serial tissue sections were stained by AP (counter-stained with eosin) and smooth muscle α-actin antibody, respectively. Scale bar, 100 μm. (The immunofluorescence and histochemical staining were performed by Yongping Yue.). (*Adapted from Ghosh, et al, Molecular Therapy., 2007, (4):750-5).*

D. Discussion

The trans-splicing AAV vectors have given an important option for expanding the small packaging capacity of the AAV vector (reviewed in [85]). Therefore, to further the application of tsAAV vectors for actual therapy, it is important to evaluate their performance in body-wide gene delivery. In this

chapter, we investigate the efficacy of tsAAV vectors by intravenous delivery of AAV-9 pseudotyped tsAAV AP vectors.

AAV-9 was recently reported to mediate strong expression in both skeletal and cardiac muscle [17, 73]. We observed excellent transduction in skeletal muscle from AV.RSV.AP and tsAAV vectors. In AP activity assay from tissue lysate and positive area quantification, tsAAV transduction is comparable to that of the single intact AAV (Figure 3.5A). Comparatively, the tsAAV transduction in heart was not as high as the skeletal muscle and was much lower than the AV.RSV.AP in heart. Nevertheless, it has reached the 50% threshold needed for treating cardiomyopathy in diseases such as DMD [86]. We previously packaged the exact same expression cassette of AV.RSV.AP in AAV-2 and AAV-5 and achieved almost no (for AAV-2) or very limited expression in selected regions in myocardium (AAV-5) after direct intra-cavity injection in newborn mice [87]. AAV-9 apparently worked much better than AAV-2 and AAV-5 (Figure 3.3). The cardiac tropism of AAV-9 may have contributed to the high-level tsAAV transduction in the heart.

AAV-9 has been shown as one of the best serotypes for lung gene transfer after direct intra-tracheal injection [3, 82]. It also resulted in moderate transduction in lung after systemic delivery in mice [17, 73]. Consistent with these reports, we also detected numerous AP positive cells in the lungs from AV.RSV.AP infected mice. Strikingly, tsAAV also resulted in broad transduction in both alveolar cells and conducting airway epithelial cells (Figure 3.4A and 3.4B). Although the single vector was significantly more efficient than the tsAAV

vectors in alveolar cells, intriguingly, both groups showed similar efficiency in the conducting airway. It is possible that the difference may have resulted from different transduction properties in these two cell types. Importantly, levels of gene transfer in both alveoli and airways were higher than 10%, a level that was thought to be sufficient for CF gene therapy [88].

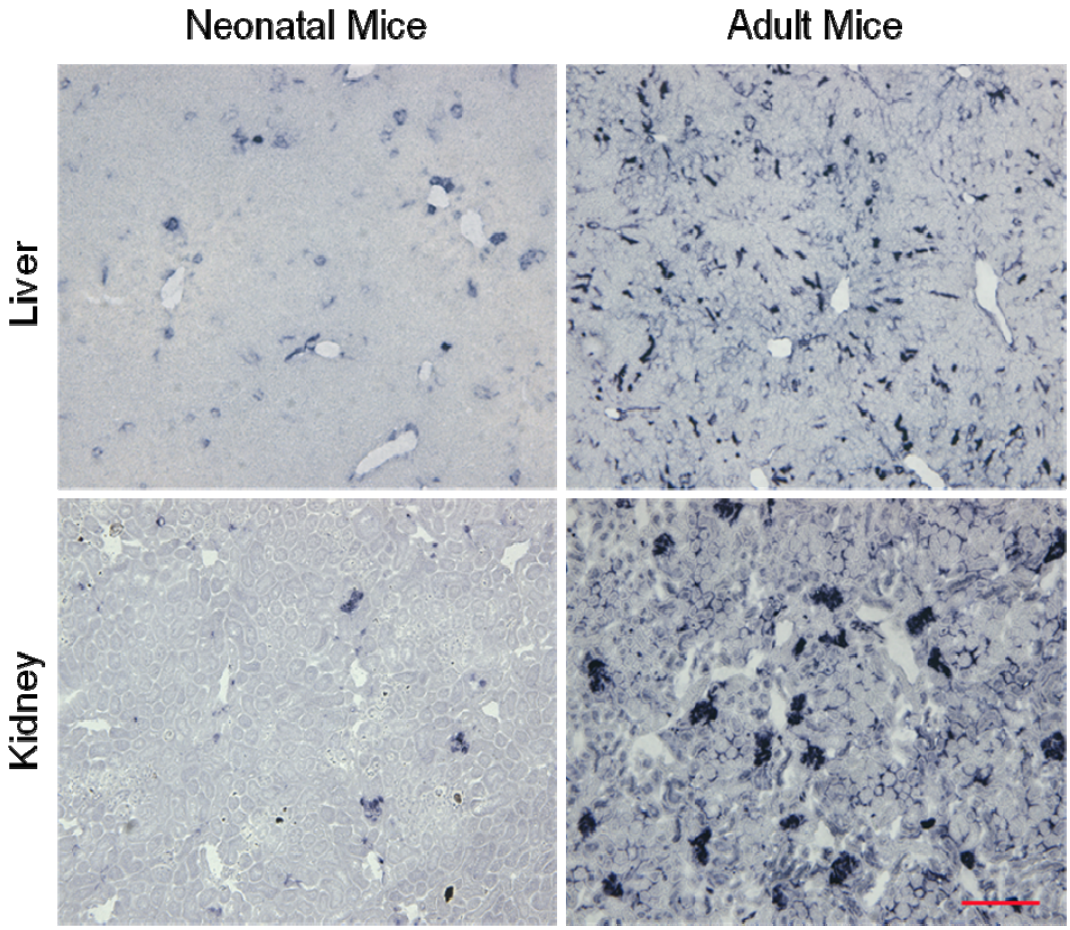


Figure 3.7 AAV-9 efficiently transduces the liver and the kidney in adult mice but not in neonatal mice. To compare the relative transduction efficiency, the same batch of AAV-9 virus (AV.AP) was delivered to neonatal mice (via facial vein; 3.5 x 10^{11} vg particles/mouse = 2.33 x10^{11} vg particles/gram body weight)

and 6-week-old adult mice (via tail vein; 1 x10^{12} vg particles/mouse = 5 x 10^{10} vg particles/gram body weight), respectively. Transgene expression was examined at six weeks later by histochemical staining of the tissue sections. The liver and the kidney in adult mice were transduced much more efficiently than the corresponding organs in neonatal mice, respectively. Scale bar, 200 μm. (*Adapted from Ghosh, et al, Molecular Therapy., 2007, (4):750-5).*

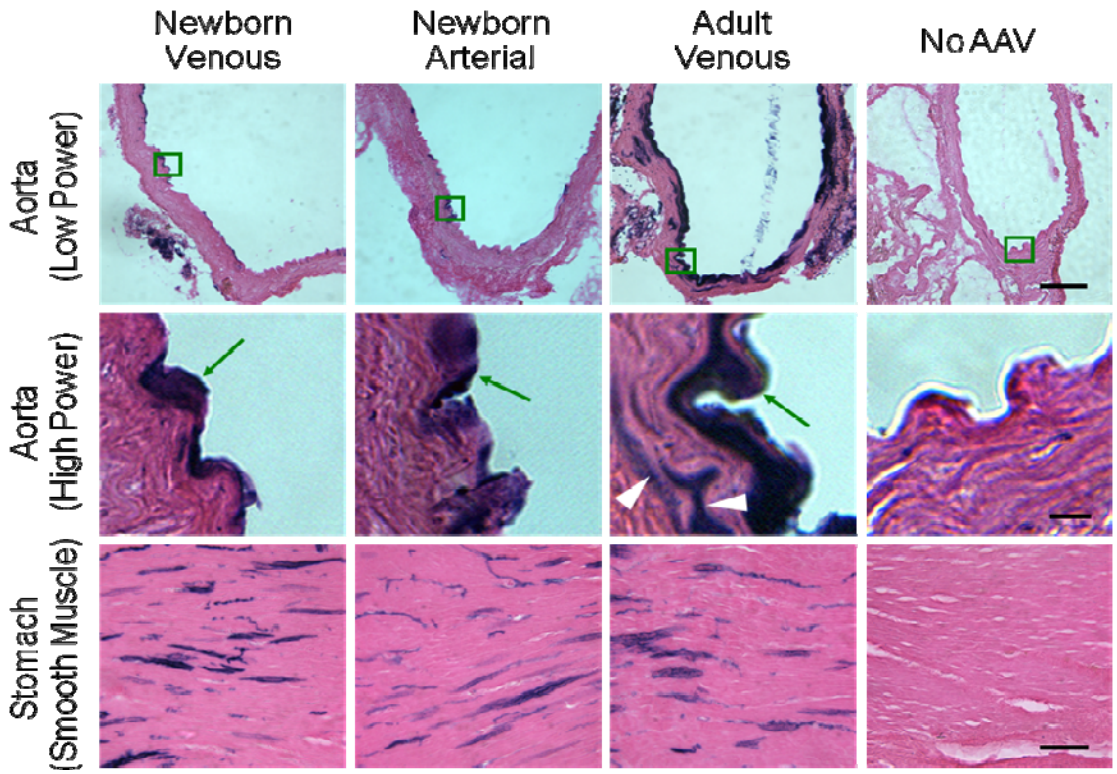


Figure 3.8. Aortic transduction is improved in adult mice following systemic AAV-9 infection but stomach smooth muscle transduction is not influenced by age. Representative photomicrographs depict AP expression in the aorta and stomach smooth muscle from the same cohort study described in Figure 1. N = 3 or 4 for each group. Boxed areas in low power aorta photomicrographs are enlarged in the corresponding high power

photomicrographs. Arrow, AAV-9 transduced endothelial layer; arrowhead, AAV-9 transduced smooth muscle in the aorta of adult mice. Scale bar, 100 μm for low power photomicrographs of the aorta; 10 μm for high power photomicrographs of the aorta; 50 μm for stomach smooth muscle. (The AP staining was done by Brian Bostick.). (*Adapted from Bostick, et al, Gene Therapy., in press).*

Animal age seems to play a minimal role in striated muscles and the lung transduction. These tissues/organs were efficiently transduced in both newborn and adult mice. This allows maximal clinical flexibility for muscle or lung-specific gene therapies such as those directed towards Duchenne muscular dystrophy. In adult mice, AAV-9 has been shown to effectively transduce the liver either by portal vein or tail vein injection [3, 17]. We have obtained similar results in adult mice (Figure 3.7). However, in neonatal mice, both the liver and the kidney were poorly transduced (Figure 3.4 B and C, 3.5C). This observation is reminiscent of a recent finding in AAV-8 [16]. AAV-8 is considered the prime serotype for liver gene transfer [89]. Systemic delivery of AAV-8 in adult mice resulted in strong expression in the liver [70]. However, systemic delivery of AAV-8 in newborn mice failed to maintain persistent transduction [16]. After two weeks, liver transduction drops to undetectable levels [16]. It is currently thought that the loss of expression after neonatal infection is mainly due to the loss of vector genome during post-natal development in these organs [16].

Age-associated differences in the aorta and retina are also quite interesting. They highlight the importance of designing age-specific gene therapy applications for these organs. They also raise some intriguing perspectives on the underlying mechanism(s) and/or the rate-limiting barriers in AAV-9 mediated systemic gene transfer.

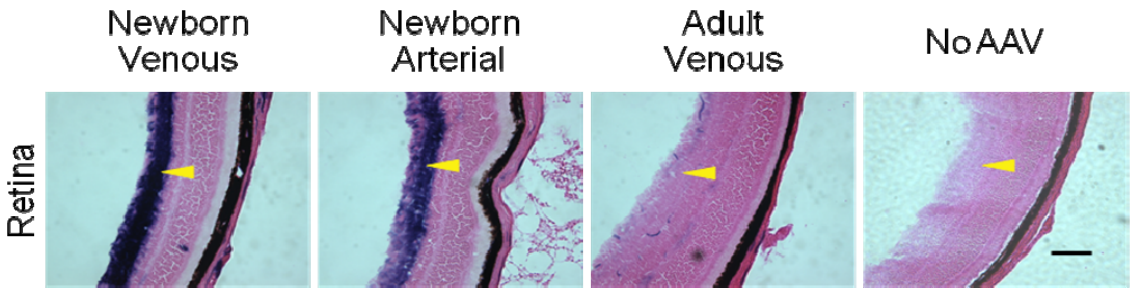


Figure 3.9. The age of mice and route of administration differentially affects systemic AAV-9 transduction in the retina. Representative AP stained photomicrographs from the retina following systemic AAV-9 infection (newborn venous, newborn arterial and adult venous) or without AAV-9 infection. N = 3 or 4 for each group. Yellow arrowhead, the inner layer of the retina. Scale bar, 100 μm. (The AP staining was done by Brian Bostick). (*Adapted from Bostick, et al, Gene Therapy., in press).*

In summary, we have demonstrated for the first time that trans-splicing AAV can be used to for whole body gene transfer. Importantly the level of transduction in several tissues is quite suitable for therapeutic applications. CF and DMD are two lethal systemic diseases affecting mainly the lung and muscle, respectively. Traditionally, the large size of the disease genes has excluded the

possibility of delivering a full-length cDNA expression cassette with a single intact AAV vector. The development of the trans-splicing vectors has raised the hope of solving the size issue. The demonstration of efficient gene expression in all body muscle as well as in the lung has finally made it possible to capitalize on the large packaging capacity of tsAAV to treat these diseases. One remaining issue is the unwanted expression in tissues that are not affected by the disease. This represents an important safety concern for any systemic gene delivery approach. It is quite possible that the expanded packaging capacity in tsAAV may allow inclusion of large tissue-specific regulatory elements (such as the tissue specific promoter) to minimize the risk associated with systemic delivery.

E. Acknowledgements

We thank Drs. Guangping Gao and Jim Wilson for providing the AAV-9 packaging plasmid pRep2/Cap9. We thank Dr. Gang Yao for the help with image quantification and we thank Dr. Hyun-Hee Lee for sharing facial vein injection technique. I thank Yongping Yue, Brian Bostick and Chun Long for assistance in tissue harvesting and sectioning. I thank Dr. Dongsheng Duan for assistance in writing this section. This work is supported by grants from the National Institutes of Health (AR-49419, DD), the Muscular Dystrophy Association (DD) and the Cystic Fibrosis Foundation (DD).

Chapter 4. Hybrid dual vectors, a new generation of split gene vectors reconstitutes cargo genes in a transgene independent manner

A. Introduction

The dual vector strategies involve simultaneous delivery of two AAV vectors carrying different segments of a larger expression cassette. Subsequent reconstitution of these segments leads to transgene expression. The trans-splicing (ts) and overlapping vectors are mechanistically distinctive dual vector strategies. Since the initial publication of these dual vector strategies [29-31, 35, 38], a number of improvements have been made to increase the transduction efficiency from these vectors. Of notice is the use of novel AAV serotypes to enhance expression [36, 40, 47]. Other advances include engineering the ITR sequences to improve tsAAV transduction [42, 90]. However, we demonstrated that there are certain limitations inherent within the target gene itself. In an effort to explore the rate-limiting steps in the tsAAV vectors, Xu et al identified mRNA accumulation as a barrier [41]. Further studies revealed the gene splitting site as a key rate-limiting factor for the tsAAV vectors [23, 91]. Highly efficient transduction can be achieved with the tsAAV vectors if there exists an optimal gene splitting site in the target gene and if this optimal site is used to generate the tsAAV vectors [23, 84, 91]. The transgene specific limitations have have been discussed for the overlapping vectors in chapter 2. In this case, transduction efficiency is largely dependent on the recombination efficiency of the

overlapping sequence [47]. A generic dual vector strategy that is applicable to any gene is currently unavailable.

To bypass the inherent limitations associated with the traditional dual AAV vectors, in this study, we have explored a novel approach by combining the features of the ts and the overlapping vectors into a single hybrid system. We hypothesized that engineering a highly recombinogenic foreign DNA sequence into the tsAAV vectors should provide a transgene independent pathway for reconstitution. To test this hypothesis, we generated a series of AAV vectors to express a common LacZ expressing cassette. These include the single intact, the ts, the overlapping and the hybrid AAV vectors. In the single intact vector, the entire LacZ expression cassette is carried by one AAV virus. In the tsAAV vectors, the LacZ gene was split into two fragments and the synthetic splicing signals were used for transgene reconstitution. In the overlapping vectors, the split LacZ gene fragments shared a 1 kb overlap. In the hybrid vectors, an 872 bp sequence from the middle of the human placental alkaline phosphatase (AP) gene was incorporated into the tsAAV vectors. The transduction efficiency of these vectors was then examined in cell line *in vitro* and adult mouse muscle *in vivo*. The performance of the hybrid vectors significantly surpassed these of the ts and overlapping vectors. In hybrid vector infected cells and muscles, transduction efficiency reached 100% and 80% of these from the single intact vector, respectively. Additional Southern analysis provided molecular evidence for AP sequence-mediated homologous recombination and reconstitution of the LacZ expression cassette. We further tested this hybrid vector strategy in the

therapeutic context of the mini-dystrophin gene. The ability of the hybrid mini-dystrophin vectors to significantly improve gene expression compared to corresponding tsAAV vectors has further validated the usefulness of this strategy. Thereafter, we further streamlined the hybrid vectors by reducing the size of the AP sequence overlap in order to maximize the capacity of the hybrid vectors. We further examined the influence of the AP overlap sequence on the function of hybrid vectors and discovered an interesting dependence on the overlapping sequence for reconstitution based on context.

B. Materials and Methods

Proviral plasmids. All proviral *cis* plasmids used to generate AAV vectors used in this study were flanked with AAV serotype-2 ITR. In the expression cassette, transcriptional regulation is controlled by the CMV promoter and the SV40 polyadenylation signal. pcisCMV.LacZ is a previously published plasmid and it was used to generate the single intact AAV vector (AV.CMV.LacZ) [32]. pcisCMV.LacZ.Upstream and pcisCMV.LacZ.Downstream were used to make the overlapping vectors. pcisCMV.LacZ.Upstream was generated by replacing the RSV promoter in the previously described pcisRSV.LacZ.Upstream with the CMV promoter from pcisCMV.LacZ [49]. pcisCMV.LacZ.Downstream is identical to the previously described pcisRSV.LacZ.Downstream [49].

We previously published a set of RSV promoter-based LacZ trans-splicing vectors [49]. In the current study, we generated a new set of *cis* plasmids to

make the CMV promoter-based LacZ trans-splicing vectors. In the new LacZ ts vectors, the LacZ gene was split at a putative splicing junction (AAG/G) we described before [41]. To generate pcisCMV.LacZ.Donor (also called pDD535), a 1687 bp DNA fragment was amplified using a previously reported pZX18 plasmid as the template [41]. The PCR primers were DL301 (forward primer) 5'GCGCGCTAGCGGGATCGAAAGAGCCTGCTAAAGC (the underlined nucleotides represent the Nhe I site) and DL246 (reverse primer) 5'GCGCGGATCCATGCGGTACCTCAGAAACGC (the BamH I site is underlined). This PCR fragment contains the 5' end of the LacZ gene and a splicing donor signal derived from the first intron of the human β-globulin gene. The Nhe I/BamH I double digested fragment was then cloned between the CMV promoter and the 3' ITR in pDD472, an intermediate *cis* plasmid carrying the CMV promoter. The final pcisCMV.LacZ.Donor plasmid contains the CMV promoter, 5' end of the LacZ gene and the splicing donor signal. To generate pcisCMV.LacZ.Acceptor (also called pDD536), a 1735 bp DNA fragment was amplified using pZX18 as the template and DL302 (5'GAAGACTCTTGCGTTTCTG) as the forward primer and DL303 (5'GCGCTCTAGACGGGCAGACATGGCCTGCCCGG, the Xba I site is underlined) as the reverse primer. The PCR product was digested with Kpn I and Xba I and then cloned into pDD295, an intermediate *cis* plasmid carrying the SV40 polyadenylation signal. The final pcisCMV.LacZ.Acceptor plasmid contains the splicing acceptor signal from the human immunoglobulin heavy chain gene, 3' end of the LacZ gene and the pA signal.

The LacZ hybrid vectors were generated by engineering an 872 bp AP gene fragment into the CMV LacZ tsAAV vectors. The AP gene fragment comes from the middle one-third of the AP cDNA. Our previous studies in the AP overlapping vectors suggest that this 872 bp fragment is highly recombinogenic [47]. We first generated an intermediate plasmid (pAG15). Briefly, we amplified the 872 bp AP gene fragment using the following set of the primers, DL412 (forward primer) 5'AGGTCTGCATGCGTGATCCTAGGTGG (underlined nucleotides represent the Sph I site) and DL413 (reverse primer) 5'ATATTAGCATGCACCACGGACCGCCC(underlined nucleotides represent the Sph I site). There is a unique Avr II site and a unique Rsr II site at the 5' and 3' ends of the AP gene fragment, respectively. The PCR product was cloned into the Sph I site in the middle of the synthetic intron in pZX18 to make pAG15 [41]. To generate the pcisCMV.LacZ.Hybrid5' (also called pAG27), pAG15 was first digested with Avr II. After blunting, the plasmid was digested with Not I. The Avr II (blunted)/Not I fragment was then cloned between the BamH I (blunted) and Not I sites in pDD2, a previously reported *cis* plasmid [32]. The resulting pcisCMV.LacZ.Hybrid5' plasmid contains the CMV promoter, the 5' half of the LacZ gene, the splicing donor signal from the first intron of the human β-globulin gene, and the 872 bp AP sequence (Figure 1). To generate pcisCMV.LacZ.Hybrid3' (also called pAG28), pAG15 was digested with Rsr II, blunted, then digested with Not I. The Rsr II (blunted)/Not I fragment was then cloned between the Sal I (blunted) and Not I sites in pAG1, an intermediate construct. The resulting pcisCMV.LacZ.Hybrid3'plasmid contains the same 872

bp AP sequence, the splicing acceptor from the human immunoglobulin heavy chain gene, the 3'half of the LacZ gene and the SV40 polyadenylation signal (Figure 1).

The mini-dystrophin hybrid vectors have been derived from the human mini-dystrophin c-DNA construct [22]. Briefly, the pCI intron containing the AP sequence from pAG15 described above was cloned at the exon55-56 junction of the mini-dystrophin construct using a 3-step PCR protocol. In the first PCR set, primers DL382 (GCCAGAGCCAAGCTTGAGTCATGG) and DL383 (GATACTTACTTGCCATTGTTTCATCAG) were used to amplify a 420 bp fragment on the 5' side of the mini-dystrophin exon 55/56. A second set of PCR was done with primers DL384 (TCTCCACAGGACCTCCAAGGTGAAATTG) and DL385 (CCACCTGCAGAAGCTTCCATCTGG) to amplify a 1 kb fragment 3' from the exon 55/56 junction. Finally using pAG15 as template, a third PCR with primers DL386 (CAATGGCAAGTAAGTATCAAGGTTACAAG) and DL387 (TTGGAGGTCCTGTGGAGAGAAAGGCAAAG) amplified the 1 kb pCI intron/AP overlapping sequence. Each of the first two products had overlap tails with the 5' and 3' ends of the third PCR product owing to the sequence being in the primer tails. A final PCR was done using DL382 and DL385, using all three PCR fragments as template to generate a 2.5kb fragment including the pCI intron and AP overlapping sequence within at the junction of exon 55/56. This fragment had HinDIII sites flanking it, which were used to replace the corresponding 1.45 kb HinDIII fragment in mini-dystrophin to generate pAG21. The pcisMini-Dystrophin.Hybrid5' was constructed by cloning the mini-dystrophin sequence

upto exon55 along with the pCI donor and 872bp AP sequences from pAG21 into a AAV proviral construct driven by the CMV promoter. For this the pYL8 construct was digested with SalI, blunted and then digested with NsiI to obtain a 5184bp fragment. The construct pAG21 was digested with AvrII, blunted and digested with NsiI to obtain the 3325bp insert. These two fragments were ligated to generate the hybrid mini-dystrophin 5' construct pAG25. The pcisMini-Dystrophin.Hybrid3' was generated by cloning the KpnI-HpaI fragment from pAG21 containing the same 872bp AP sequence followed by the pCI donor and the mini-dystrophin sequence from exon56 onwards into an AAV proviral plasmid containing the SV40 polyadenylation signal. For this, pDD295 was digested with XbaI, blunted and digested with KpnI to obtain a 4316bp vector fragment. Then pAG21 was sequentially digested with HpaI and KpnI to obtain a 4 kb insert fragment. The two fragments were ligated to generate the hybrid mini-dystrophin 3' construct pAG26. To generate the pcisMini-dystrophin.Donor, the pAG21 construct was digested with RsrII, Blunted and digested with NsiI to obtain a 2453bp fragment. This fragment was ligated to the 5184 bp vector fragment generated previously from pYL8 construct (for the pcisMini-Dys.Hybrid 5' construction). The resulting construct was pAG39, pcisMini-dystrophin.Donor. The pcisMini-dystrophin.Acceptor was generated from the pAG26 construct described previously. The pAG26 plasmid was double digested with AvrII and RsrII. The 7.4 kb band was the blunted and ligated to obtain pAG40, pcisMini-dystrophin.Acceptor.

In order to determine which part of the AP gene sequence is essential and sufficient for recombination and gene expression from overlapping vector co-infection, different regions of the 872 bp overlapping region were used to make pairs of AP overlapping viruses. To construct the overlapping sets sharing either the 5' or the 3' half of the AP overlap sequence, the pcisRSV.AP (previously described in chapter 2) was partially digested with SalI and a 7.4 kb product was gel purified. The fragment was then blunted and digested with BbsI. This produced two sets of bands, one resulting from the 5'SalI and BbsI cut (5.8 kb and 1.6 kb) and another resulting from the 3'SalI and BbsI cut (5.4 kb and 2.0 kb). The 5.8 kb and 5.4 kb bands were then gel purified and ligated to obtain the pcisRSV.APupstream-5' half and the pcisRSV.APdownstream-3' half respectively. These two vectors, therefore, share the 5' half and the 3' half overlaps when paired with the original pcisADownstream and pcisAPUpstream (described previously in chapter 2) respectively. To generate the one third AP overlapping sets, the AP sequences upto the one third sequences were PCR amplified using primers with SalI and AvrII (for upstream vectors) or SalI and RsrII (for downstream vectors) adapters. The different AP overlapping upstream vectors were generated by cloning the respective PCR fragments into pAG2, also digested with SalI and AvrII. The AP overlapping downstream vectors were generated by cloning the respective PCR fragments into pAG1, digested with SalI and RsrII. These different overlapping vectors were then paired correspondingly to generate five different sets of overlapping AP vectors to compare with the original AP overlapping vector set described in chapter 2.

For mapping of the AP sequence to determine which part(s) of the AP overlap sequence would be essential and sufficient for the function of the hybrid vectors, different regions of the AP sequence were PCR amplified and cloned into the AvrII-RsrII sites of the pAG15 plasmid. Thereafter, from each such construct, AAV.Hybrid 5' and AAV.Hybrid 3' vectors were constructed by splitting the full length construct and cloning it into within the AAV ITRs as described above for pAG27 and pAG28. Five different sets of such hybrid vectors were constructed that shared either the 5' or 3' half of the AP overlap sequence or the 5' one-third, middle or 3' one thirds of the AP sequence.

Recombinant AAV production. All viral stocks were made of serotype-6 recombinant AAV (AAV-6) and produced as described in Chapter 2.

***In vitro* studies.** Transduction efficiency of the AAV vectors was examined in MO59K cells, a human glioblastoma cell line. MO59K cells were cultured in a mixture (1:1) of Dulbecco's Modified Eagle's medium and Ham's F12 medium containing 2.5 mM L-glutamine and 1.2 g/L sodium bicarbonate supplemented with 0.05 mM non-essential amino acids, 0.5 mM sodium pyruvate and 10% fetal bovine serum (FBS) and 1% penicillin and streptomycin (PS) (penicillin 100 U/ml and streptomycin 100 μg /ml). 80% confluent cells were infected at an moi of 10,000 vg particles per cell for each virus (a total of 20,000 vg particles per cell for the dual AAV vector infection). Infection was carried out in serum-free, antibiotic-free medium. Two hours after AAV inoculation, FBS and PS

concentration was brought to 10% and 1% respectively by adding equal volume medium containing 20% FBS and 2% PS. LacZ expression was analyzed at 45 hrs post infection by cytological staining and the β-galactosidase assay (Applied Biosystems, Bedford, MA) as described before [49].

For the *in vitro* viral competition assays, all infections were done at moi of 2500 vg particles per cell per virus and at a ratio of 1:1:1 (for 3 vectors) or 1:1 (for 2 vectors). Analysis of gene expression was done as described above.

Southern blot. Low molecular weight DNA was isolated from AAV infected MO59K cells using a modified Hirt DNA isolation protocol. Briefly, at 45 hours after infection cells were harvested from 6-well plates using a cell scraper in phosphate buffered saline (PBS). After one min centrifugation at 800 x g, cell pellet was resuspended in 300 μl lysis buffer containing 1% SDS, 50 mM EDTA, and 100 mM Tris pH 8.0. Cell lysate was then digested with proteinase K (1 μg/μl final concentration) and pronase (0.5 μg/μl final concentration) for 30 min. To precipitate low molecular weight DNA, 75 μl 5 M NaCl was added to the cell lysate. After two hours incubation at 4ºC, 250 μl TE (10 mM Tris, pH 8.0, 1 mM EDTA) and 100 μl Promega neutralizing solution (Wizard Plus SV minipreps, Promega, Madison, WI) was added to the tube. Cell lysate was mixed by inverting the tube several times. Cellular debris (including chromosome DNA) was then removed by a 35 min centrifugation at 15,700 x g. Finally, low molecular weight DNA was purified using a DNA miniprep column (Wizard Plus SV minipreps, Promega, Madison, WI) and eluted in DNase free water.

Southern blot was performed as previously described using radiolabeled LacZ, AP and ITR probes, respectively [26]. For the LacZ probe, a 780 bp Mlu I fragment from the LacZ gene was used. This probe spans both the 5' and 3' parts of the split LacZ gene. For the AP probe, a 359 bp Sac I fragment from the middle of the AP gene was used. For the ITR probe, a 266 bp Sph I fragment from a previously reported pZX19 plasmid was used [41]. This fragment contains the double-D ITR sequence. To compare results from different probes, we first hybridized the membrane with the LacZ probe. After stripping, the same membrane was re-hybridized with the AP probe. The membrane was then stripped again and re-hybridized with the ITR probe. There was no leftover radioactivity in the membrane between different hybridizations.

Animal studies. Experiments in animals were performed in accordance with NIH and institutional guidelines of the University of Missouri. Six to eight-week-old BL10 male mice were purchased from Harlan (Indianapolis, IN). Recombinant AAV was delivered to the TA muscle according to a previously described protocol [11]. For single intact vector infection, we delivered 1×10^{10} vg particles of AV.CMV.LacZ to the muscle. For dual vector co-infection, a total of 2×10^{10} vg particles (1×10^{10} vg particles for each virus) were delivered. To determine whether there was leaky expression from a single virus in each set of dual AAV vectors, we also performed single vector infection with viruses carrying either the 5' or the 3' LacZ gene. In negative control (mock infection), we injected 20 μl HEPES-buffered saline to the TA muscle. Mice were sacrificed at

six weeks post-infection and TA muscles were harvested according to our published protocol [11]. LacZ expression was analyzed by histochemical staining and β-galactosidase assay (Applied Biosystems, Bedford, MA) as described before in Chapter 2, [47, 49]. Mini-dystrophin vector injections were done in *mdx* mice at the same dosage of 1×10^{10} vg particles per virus per TA muscle. Immunohistochemical detection of dystrophin expression in tissue sections was done as described previously in chapter 2.

For the set of experiments done to compare the hybrid vector sets with different lengths of AP overlap, infections were done at a total dosage of 2×10^{9} vg particles (1×10^{9} vg particles for each virus)

Statistical analysis. Data are presented as mean ± standard error of mean. Statistical analysis was performed with the SPSS software (SPSS, Chicago, IL). Statistical significance among all experimental groups was determined using one-way ANOVA followed by Tukey *post hoc* analysis. Difference was considered significant when $P < 0.05$.

C. Results

Design and construction of the hybrid LacZ vectors. To accurately compare different vector systems, we packaged the same LacZ gene expression cassette into the single intact vector (AV.CMV.LacZ), the traditional dual AAV vectors and the hybrid vectors. This design will eliminate potential influence of the

transcription regulatory elements in different systems. Figure 1 illustrates the design of the hybrid vectors. Similar to the traditional ts and overlapping vectors, the LacZ gene is divided into two parts and separately packed into two independent AAV virions called AV.CMV.LacZ.Hybrid5' and AV.CMV.LacZ.Hybrid3', respectively. In the previously published dual AAV vectors, transgene reconstitution is either dependent on the gene splitting site (for the ts vectors) or the transgene sequence (for the overlapping vectors) [23, 47, 91]. The success of these approaches hinges on the transgene itself. The key innovation in the hybrid vectors is the inclusion of a highly recombinogenic foreign DNA sequence to allow for transgene-independent reconstitution. Essentially, AV.CMV.LacZ.Hybrid5' contains the CMV promoter, the 5' end of the LacZ gene, the splicing donor signal and an 872 bp AP gene fragment. AV.CMV.LacZ.Hybrid3' contains the same 872 bp AP gene fragment, then the splicing acceptor signal, the 3' end of the LacZ gene and the poly A signal. The 872 bp AP gene fragment represents the middle one third of the human placental AP cDNA. We have recently shown that this fragment is highly recombinogenic in the context of overlapping AAV vectors [47]. To minimize the influence of the gene splitting site, the LacZ gene was split at exactly the same location in both the ts and hybrid vectors. We hypothesize that the AP gene fragment will facilitate homologous recombination between AV.CMV.LacZ.Hybrid5'and AV.CMV.LacZ.Hybrid3' in co-infected cells. Ultimately, the full-length LacZ expression cassette will be regenerated through a transgene independent pathway. Since the hybrid vectors are built on the tsAAV vectors, transgene

reconstitution can also undergo ITR-mediated recombination as in the traditional tsAAV vectors (Figure 4.1).

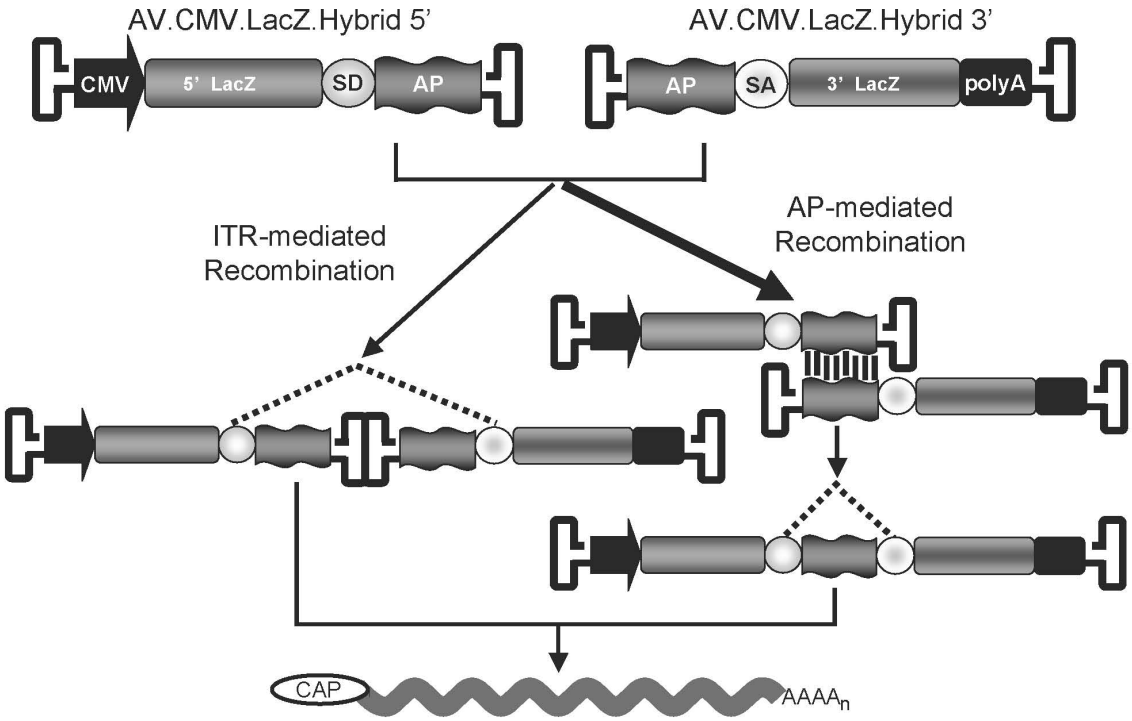


Figure 4.1. Schematic outline of hybrid vector-mediated transgene reconstitution. The AV.CMV.LacZ.Hybrid5' vector contains the CMV promoter, the 5' half of the LacZ gene, the splicing donor (SD) and the 872 bp alkaline phosphatase (AP) sequence. The AV.CMV.LacZ.Hybrid 3' vector contains the 872 bp AP sequence followed by the splicing acceptor (SA), the 3' half of the LacZ gene and the SV40 polyadenylation signal (polyA). Vector genomes are flanked by the AAV-2 ITRs. Upon co-infection, the complete LacZ expression cassette can be reconstituted via AP sequence-mediated homologous recombination. Alternatively, it can be generated via viral ITR-mediated head-to-

tail recombination. LacZ expression is achieved after the junction sequences are removed by cellular splicing machinery (dotted lines). (*Adapted from Ghosh, et al, Molecular Therapy., in press).*

LacZ expression from the hybrid vectors reached that of the single intact vector in cell line *in vitro*. In all the dual AAV vector sets, each individual vector only carries part of the LacZ gene. We first confirmed that individual infection with these vectors did not yield LacZ expression (data not shown). To evaluate the relative transduction efficiency of different dual vector strategies, we infected MO59K cells with the ts, overlapping and hybrid LacZ vectors at an moi of 20,000 vg particles per cell (10, 000 vg particles/cell for each vector). As a control, we also infected MO59K cells with AV.CMV.LacZ at an moi of 10,000 vg particles/cell. This infection should illustrate the maximal achievable level. At 45 hours post infection, we performed LacZ staining. We observed highly efficient transduction from AV.CMV.LacZ and the hybrid vectors, but not from the ts and overlapping vectors (data not shown). We next quantified the β-galactosidase activity in cell lysate. AV.CMV.LacZ infection yielded an activity of 1,307 ± 22 units/μg protein. As expected from our previous studies, the activity levels from overlapping and ts vector infections were significantly lower than that from AV.CMV.LacZ infection [47, 49]. Furthermore, the transduction efficiency of the overlapping vectors (389 ± 20 units/μg protein) was lower than that of the ts vectors (654 ± 40 units/μg protein) (Figure 4.2). Consistent with *in situ* LacZ staining results, the activity in hybrid vector infected cells reached 1,137 ± 82

units/μg protein. This level was significantly higher than those of the ts and overlapping vector infections, and it was statistically comparable to that of AV.CMV.LacZ infection (Figure 4.2).

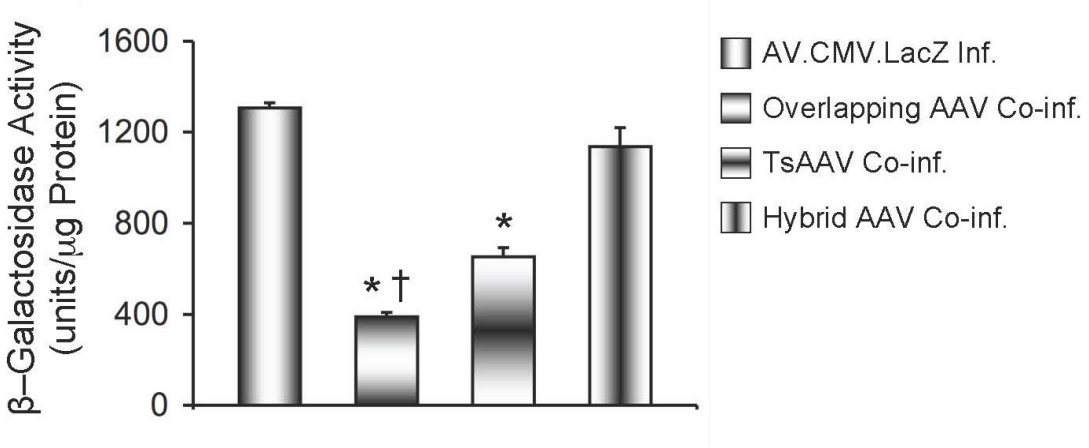
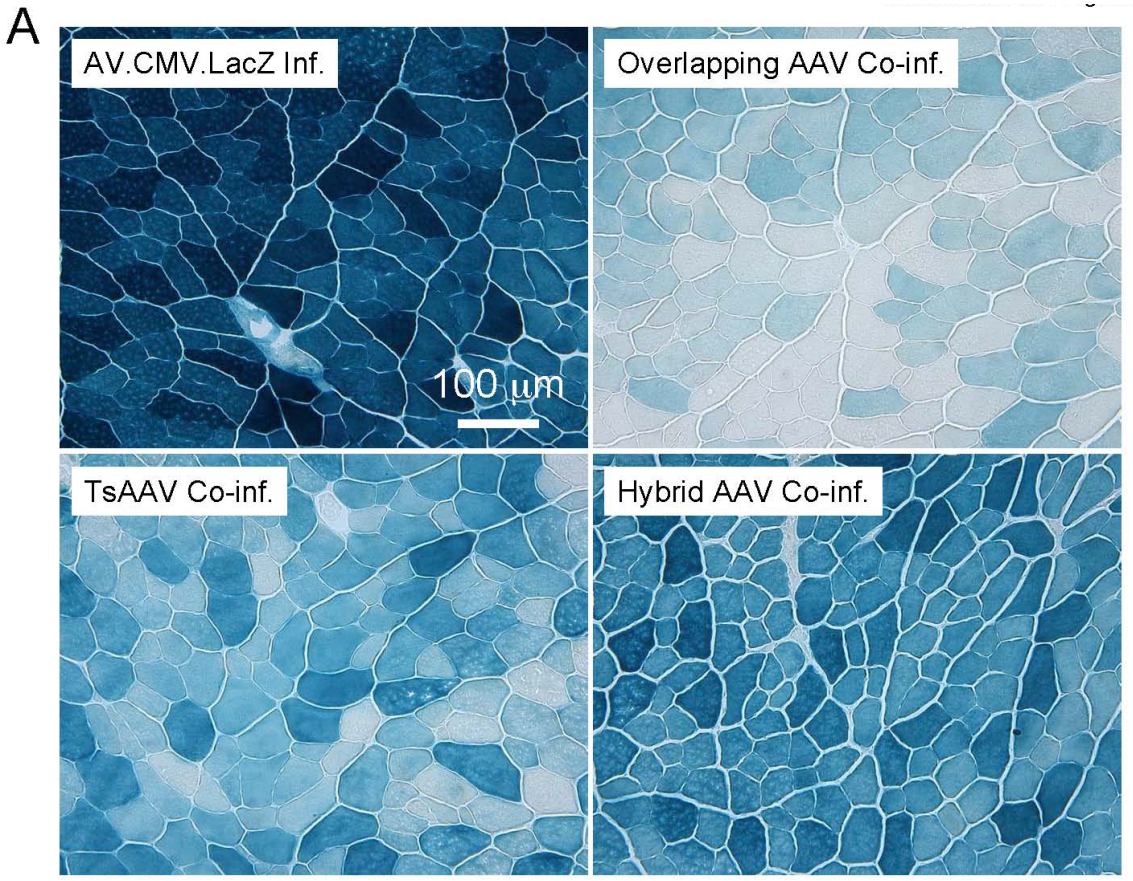


Figure 4.2. Comparative analysis of LacZ expression from different AAV vectors in MO59K cells *in vitro*. AV.CMV.LacZ inf, single intact AAV infected cells; Overlapping AAV co-inf, AV.CMV.LacZ.Upstream and AV.CMV.LacZ.Downstream co-infected cells; TsAAV co-inf, AV.CMV.LacZ.Donor and AV.CMV.LacZ.Acceptor co-infected cells; Hybrid AAV co-inf, AV.CMV.LacZ.Hybrid5' and AV. CMV.LacZ.Hybrid3' co-infected cells. Asterisk, significantly lower than that from AV.CMV.LacZ infection and hybrid AAV co-infection; Cross, significantly lower than that from the tsAAV co-infection. There is no significant difference between AV.CMV.LacZ infection and hybrid AAV co-infection. N = 4 for each group. (*Adapted from Ghosh, et al, Molecular Therapy., 2007).*

Transduction efficiency of the hybrid vectors was significantly higher than those from the traditional dual vectors in mouse muscle *in vivo*. We next tested whether the hybrid vectors could efficiently reconstitute LacZ expression in skeletal muscle. We injected AV.CMV.LacZ (1×10^{10} vg particles) and different sets of the dual AAV vectors (1×10^{10} vg particles of each vector and a total of 2×10^{10} vg particles per muscle) into the tibialis anterior (TA) muscles of 6 to 8-week-old BL10 mice. We examined LacZ expression six weeks later (Figure 4.3). In histochemical staining, we observed wide-spread expression in AV.CMV.LacZ infected muscle and hybrid vector co-infected muscle. The ts and overlapping vectors showed patchy expression with lower intensity (Figure 4.3A). To quantify the relative efficiency among different vector sets, we measured the β-galactosidase activity (Figure 4.3B). Similar to our *in vitro* results in MO59K cells, the highest activity was achieved with AV.CMV.LacZ (2,071 ± 103 units/μg protein). The lowest activity was seen with the overlapping vectors (711 ± 69 units/μg protein). The ts vectors resulted in an activity (1,288 ± 40 units/μg protein) higher than that of the overlapping vectors, but significantly lower than that of the single intact vector. Among three dual vector systems, the hybrid vectors yielded the highest β-galactosidase activity (1,674 ± 86 units/μg protein). This level reached 80.8 % of that from the single intact vector and was significantly higher than that from the other dual vector systems.


A
AV.CMV.LacZ Inf.
Overlapping AAV Co-inf.
100 μm
TsAAV Co-inf.
Hybrid AAV Co-inf.

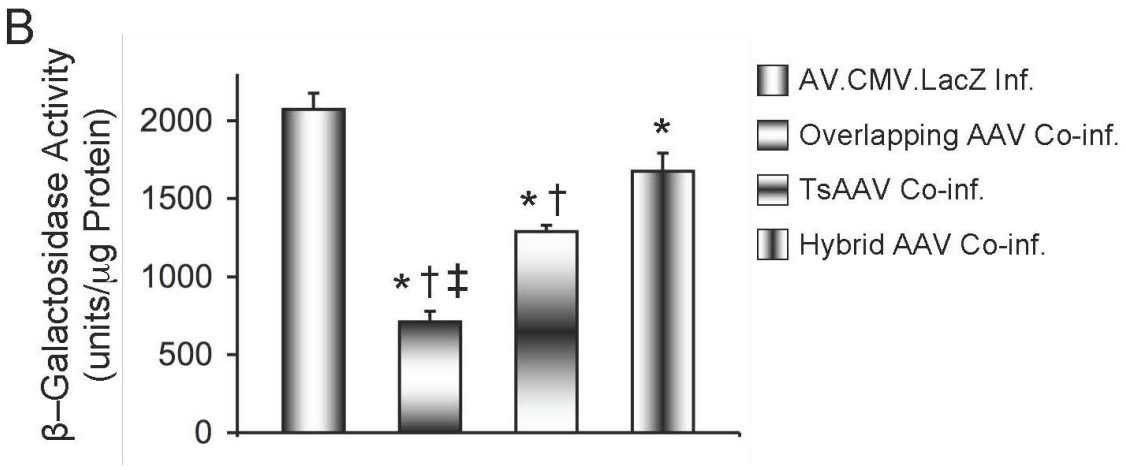

B
β–Galactosidase Activity
(units/μg Protein)
2000
1500
1000
500
0
*†‡
*†
*
AV.CMV.LacZ Inf.
Overlapping AAV Co-inf.
TsAAV Co-inf.
Hybrid AAV Co-inf.

Figure 4.3. The hybrid vectors enhance transduction efficiency of the traditional dual AAV vectors in mouse muscle *in vivo*. A, Representative photomicrographs of LacZ staining from different experiment groups (N = 4 for each group). Scale bar applies to all images. **B,** Quantitative evaluation of β-galactosidase activity in AAV infected muscles (N = 4 for each group). AV.CMV.LacZ inf, single intact AAV infected muscle; Overlapping AAV co-inf, AV.CMV.LacZ.Upstream and AV.CMV.LacZ.Downstream co-infected muscle; TsAAV co-inf, AV.CMV.LacZ.Donor and AV.CMV.LacZ.Acceptor co-infected muscle; Hybrid AAV co-inf, AV.CMV.LacZ.Hybrid5' and AV. CMV.LacZ.Hybrid3' co-infected muscle. Asterisk, significantly lower than that of AV.CMV.LacZ infection; Cross, significantly lower than that of the hybrid vector co-infection; Double cross, significantly lower than that of the tsAAV vector co-infection. (*Adapted from Ghosh, et al, Molecular Therapy., 2007).*

Molecular evidence for transgene-independent reconstitution in the hybrid vectors. Based on our vector design, we hypothesized that the LacZ expression cassette reconstitution in the hybrid vectors should occur in a LacZ gene independent manner. Specifically, the engineered AP-sequence should mediate homologous recombination between the vector genomes in AV.CMV.LacZ.Hybrid5' and AV.CMV.LacZ.Hybrid3' viruses. Thus, the complete expression cassette is regenerated irrespective of the gene splitting site and the recombiogenic potential of the LacZ gene sequence. In addition, the hybrid vectors can still undergo ITR-mediated ts pathway.

To test this hypothesis, we examined the vector genome in AAV infected cells (Figure 4.4). Low molecular weight Hirt DNA was digested with Not I (for AV.CMV.LacZ, overlapping vector and hybrid vector infection) or Not I/Spe I (for ts vector infection). In AV.CMV.LacZ infected cells, Not I released a 3.47 kb band which was recognized only by the LacZ probe but not by the AP or ITR probes (Figure 4.4B lanes 1, 12 and 23). Both AV.CMV.Upstream and AV.CMV.Downstream have only one Not I site. In cells individually infected with these vectors, we observed 3.37 kb (for AV.CMV.Upstream) and 3.02 kb (for AV.CMV.Downstream) diagnostic bands for head-to-tail monomers/concatamers by the LacZ and ITR probes but not by the AP probe (Figure 4.4B lanes 3, 4, 14, 15, 25 and 26). Using the LacZ probe, we also observed a 5.06 kb tail-to-tail concatamer band for AV.CMV.Upstream infection and a 4.82 kb head-to-head concatamer band for AV.CMV.Downstream infection (Figure 4.4B lanes 3 and 4). These bands were weaker than the head-to-tail bands. We did not observe these bands in the ITR sequentially probed membranes probably because of multiple stripping prior to ITR hybridization had weakened the signal. In overlapping vector co-infected cells, we detected the 3.47 kb diagnostic band representing the reconstituted genome (Figure 4B lane 2). As expected, this band was not seen by the AP or ITR probes (Figure 4.4B lanes 13 and 24). Interestingly, in overlapping vector co-infected cells, we also detected a ~ 5 kb band by the LacZ and ITR probes (Figure 4.4B lanes 2 and 24). Should AV.CMV.Upstream and AV.CMV.Downstream form heterodimer through ITR recombination, we would expect a 4.94 kb band. Since the size of this band is

very close to homodimer bands (5.06 kb tail-to-tail dimer of AV.CMV.Upstream or 4.82 kb head-to-head dimer of AV.CMV.Downstream), this ~ 5 kb band may likely represent a mixture of all three bands (4.82 kb, 4.94 kb and 5.06 kb). Nevertheless, this suggests the existence of ITR-mediated recombination in cells co-infected with the overlapping vectors.

In AV.CMV.Donor, the Not I site was inactivated during cloning. So we used Spe I/Not I double digestion. Spe I cuts AV.CMV.Donor once but does not cut AV.CMV.Acceptor. In cells individually infected with AV.CMV.Donor or AV.CMV.Acceptor, we observed head-to-tail bands with the LacZ and ITR probes (2.90 kb for AV.CMV.Donor and 2.38 kb for AV.CMV.Acceptor; Figure 4.4B lanes 6, 7, 28 and 29). We also observed the 5.12 kb tail-to-tail band (for AV.CMV.Donor; Figure 4.4B lane 6) and the 3.9 kb head-to-head band (for AV.CMV.Acceptor; Figure 4.4B lane 7) with the LacZ probe. Interestingly, in AV.CMV.Acceptor infected cells, we detected a prominent ~ 2 kb band (Figure 4.4B lanes 7 and 29, Asterisk). AV.CMV.Acceptor has a vector genome size of 2.38 kb, a perfect size for self-complementary AAV packaging [92, 93]. In the case of a self-annealed linear double stranded vector genome, we would expect to see a 1.95 kb band. Thus, this ~ 2 kb band may represent the self-annealed vector genome. In AV.CMV.Donor and AV.CMV.Acceptor co-infected cells, we observed all the bands seen in individual vector infected cells. Importantly, we also observed the 4.51 kb diagnostic band representing the reconstituted vector genome (Figure 4.4B lanes 5 and 27).

Individual infection with each one of the hybrid vectors yielded similar band patterns as we saw with single infection of other dual vectors. We detected head-to-tail monomer/concatamer bands (a 3.83 kb band for AV.CMV.Hybrid 5' infection and a 3.59 kb band for AV.CMV.Hybrid 5' infection; Figure 4.4B lanes 9, 10, 20, 21, 31 and 32) with all three probes (LacZ, AP and ITR, respectively). Using the LacZ and AP probes, we also detected the 5.48 kb tail-to-tail concatamer band in AV.CMV.Hybrid 5' infected cells (Figure 4.4B lanes 9 and 20) and the 5.98 kb head-to-head concatamer band in AV.CMV.Hybrid 5' infected cells (Figure 4.4B lanes 10 and 21). For hybrid vector co-infection, we expect a 4.52 kb band (detectable by the LacZ and AP probes but not by the ITR probe) should vector genomes be reconstituted through the AP sequence-mediated homologous recombination (the overlapping pathway). As shown in Figure 4.4B lanes 8, 19 and 30, this is indeed the case. On the other hand, if reconstitution of the LacZ expression cassette occurs through ITR-mediated recombination (the tsAAV pathway), we should see a 5.73 kb band that can be detected by all three probes. We were able to detect a faint band at this location. However, a definitive band has been elusive because of the close migration of the AV.CMV.Hybrid 5' head-to-head homodimer band (5.96 kb) and the AV.CMV.Hybrid 3' tail-to-tail homodimer band (5.48 kb).

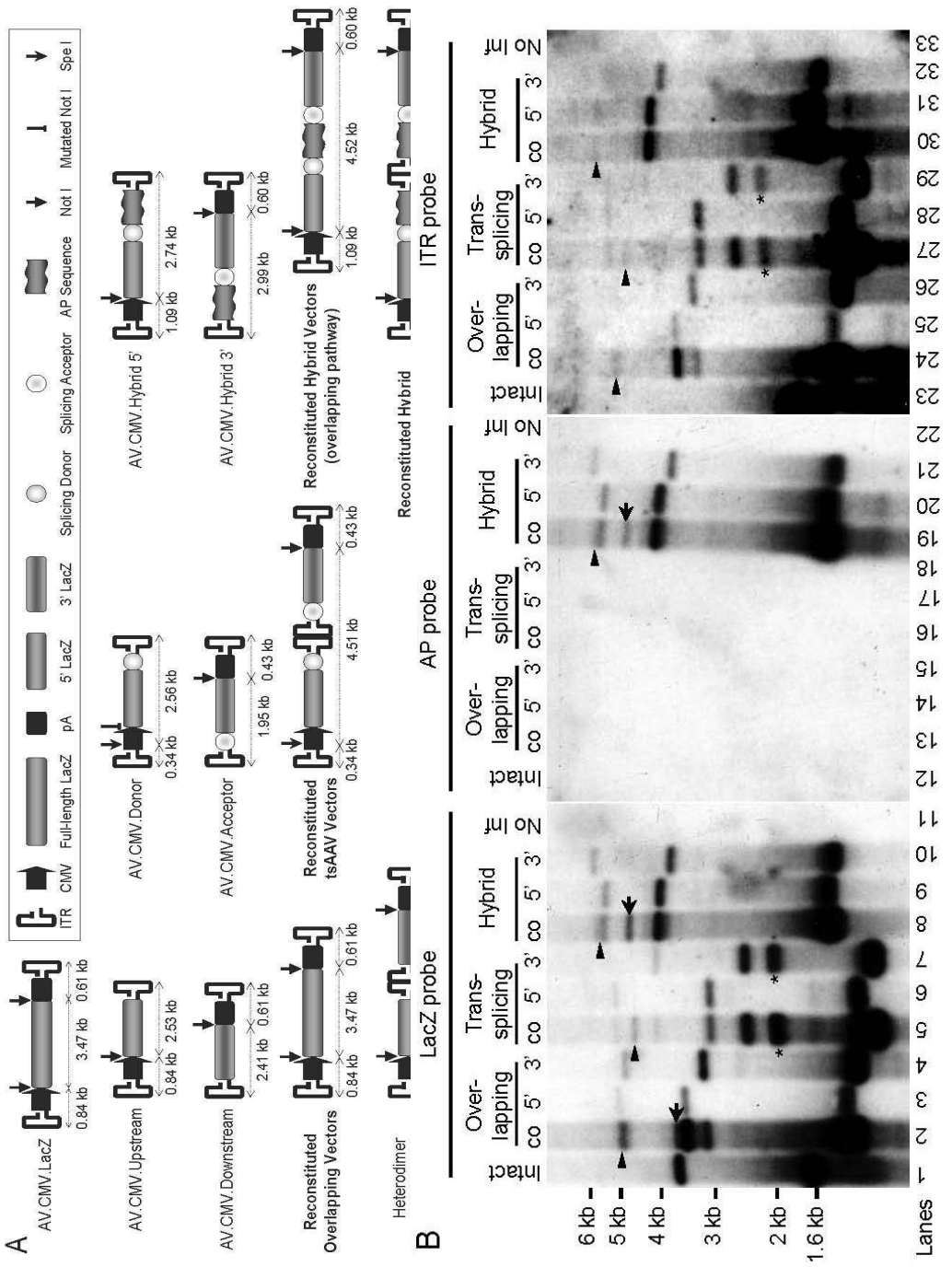


Fig.4.4 Southern blot analysis of AAV genome reconstitution. A, Schematic outline of the vector genome structure and the expected restriction pattern in individually infected or co-infected cells. Box, cartoon of the individual components in the drawing. Note, it is expected that the majority of the vector

genomes should be circular forms. For the presentation purpose, we have drawn all the vector genomes in the linear form. Also, we only show representative heterodimer structures from dual vector co-infected cells. Homodimers of each individual vectors are not presented. The predicted structures of the reconstituted vector genomes (from dual AAV vector co-infection) are highlighted in bold. **B,** Representative Southern blot of Hirt DNA from AAV infected cells. The size mark applies to all three blots. Intact, AV.CMV.LacZ infection; co, two vector co-infection; 5', individual infection with the vector carrying the 5' LacZ gene; 3', individual infection with the vector carrying the 3' LacZ gene; No Inf, Hirt DNA from cells not infected by AAV. Bands that are less than 2 kb represent single stranded AAV genomes (these bands are not discussed in the text). Arrow, the locations of the bands derived from homologous recombination in dual vector co-infected cells. Arrowhead, the locations of the bands derived from ITR head-to-tail recombination in dual vector co-infected cells. Asterisk, diagnostic band for self-complementary vector (AV.CMV.Acceptor). (*Adapted from Ghosh, et al, Molecular Therapy, 2007).*

The efficiency of the hybrid vectors is dependent on overlapping pathway more than the ts pathway. Although figure 4.4 suggests that the overlapping pathway may be the dominant pathway for hybrid vectors reconstitution, we understand that quantitative estimation based on Southern blot data could be inaccurate. Therefore, to further address the question of which pathway is more dominant, we performed competition experiments (Figure 4.5). In these

experiments, we used the AV.APDownstream (AP overlapping virus downstream (as described in Chapter 2)) or the AV.mini-dystrophinHybrid 3' to compete with the AV.LacZ hybrid 3' for recombining via the AP overlap sequence to the AV.LacZ hybrid 5'at a ratio of 1:1:1. On the other hand, we used the AV.mini-dystrophinAcceptor to compete with the AV.LacZHybrid 3' for recombining with AV.LacZHybrid 5' via the ITR mediated pathway at a ratio of 1:1:1. Control infections were done at 1:1 ratios. All infections were done in MO59K cells at an moi of 2500 vg particles per virus per cell (7500 vg for the 1:1:1 infections and 5000 vg for 1:1 infections). The levels of LacZ expression from these sets of infections are indicative of which pathway is more critical for hybrid vector recombination. The AV.APDownstream and AV.mini-dystrophinHybrid 3' competing for the AP overlapping mediated recombination reduced the LacZ expression levels to 175.21$\pm$9.74 and 181.58$\pm$20.14. The AV.mini-dystrophinAcceptor co-infection set reduced LacZ expression to 240.71$\pm$15.59. In the case of overlapping vector competition, LacZ expression levels dropped by 42-44% compared to the LacZ hybrid co-infection by itself. Comparatively, the ts competition by AV.minidystrophinAcceptor reduced LacZ expression by only 22.6%. The data suggests that the overlapping vectors are more efficient at competition as indicated by the lower levels of LacZ expression, thereby indicating the relative importance of AP sequence mediated recombination over ITR mediated recombination.

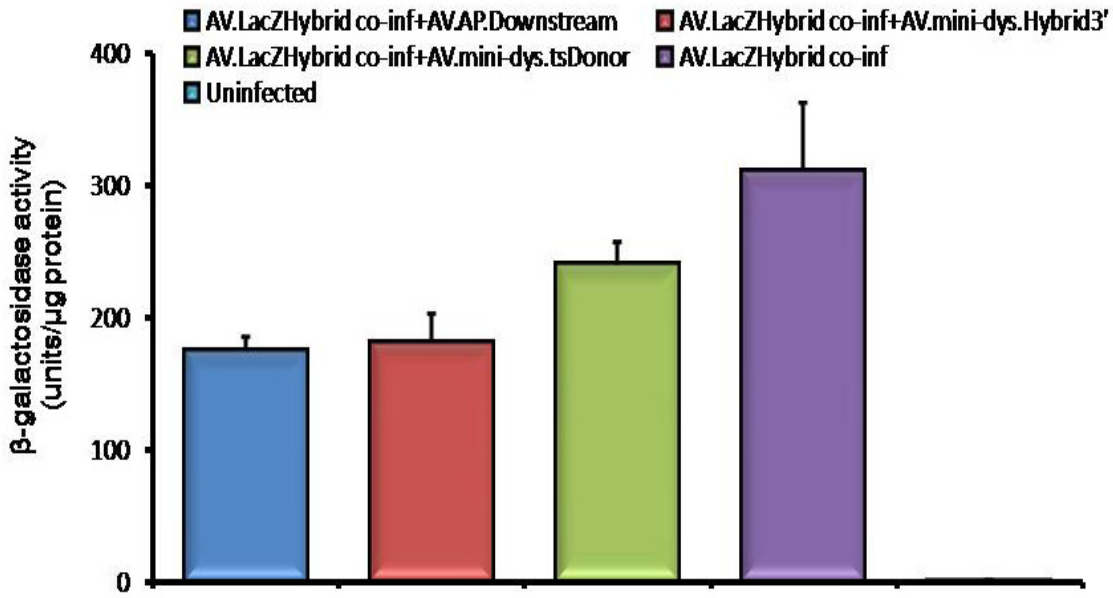
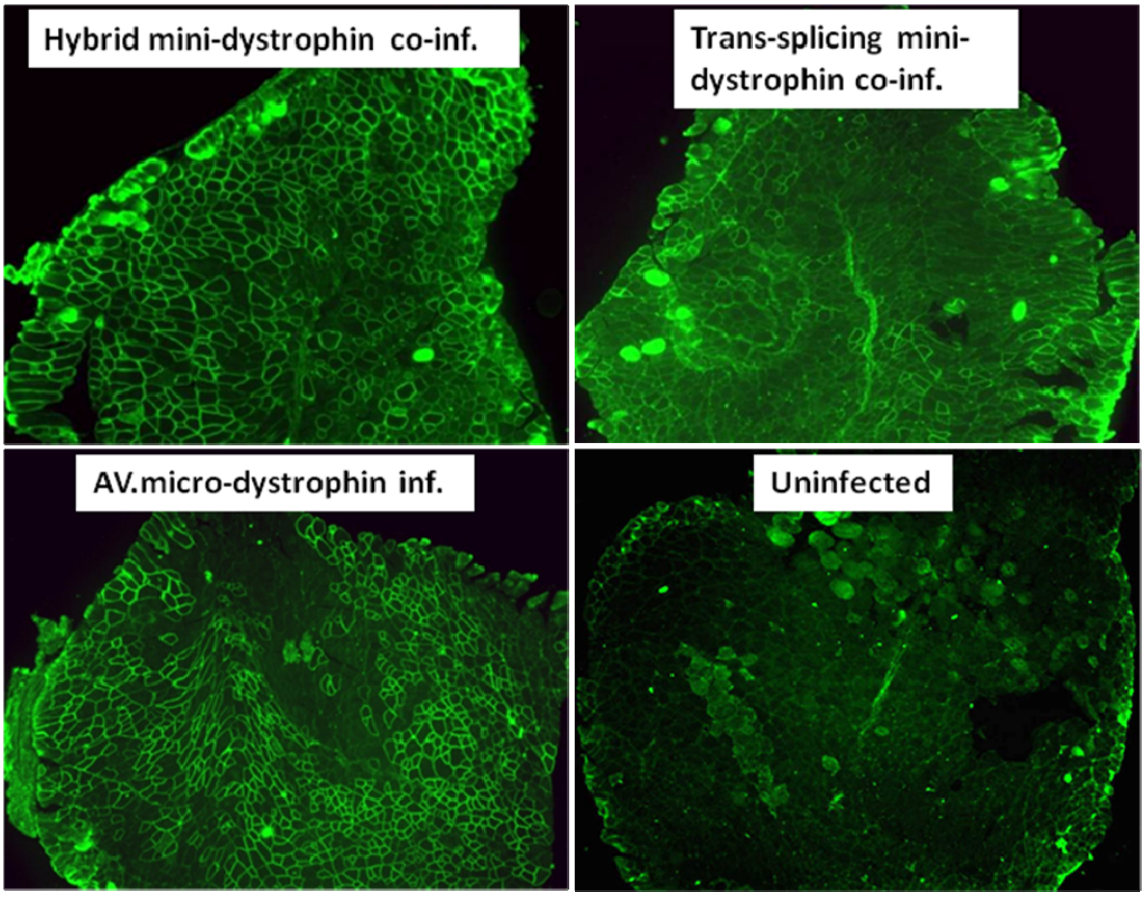


Figure 4.5. Competition assay comparing the overlapping and ts pathways in hybrid vectors recombination. MO59K cells were infected at an moi of 2500 vg particles per cell per virus (7500 vg for 1:1:1 and 5000vg for 1:1 infections). β-galactosidase levels were measured in total protein lysate at 45 hrs. post infection. Blue bar, LacZ hybrid 5', LacZ hybrid 3' and AP overlapping downstream vector co-infection. Pink bar, LacZ hybrid 5', LacZ hybrid 3' and mini-dystrophin hybrid 3' vector (based exon 55/56) co-infection.

***In vivo* gene expression from the hybrid mini-dystrophin vectors is higher than the corresponding trans-splicing vectors.** To further clarify how the choice of the gene splitting site would affect the hybrid vector system, we generated the mini-dystrophin hybrid vectors. The human mini-dystrophin vector was split at the exon 55/56 junction and the exact same pCI intron containing the AP overlap sequence (as described for the LacZ hybrid vectors) was inserted at

that junction. The exon55/56 junction was selected randomly and is placed in the center of the mini-dystrophin gene. Splicing efficiency from this junction was not experimentally determined previously.

The pcisMini-dystrophin.Hybrid5' contained the CMV promoter followed by the 5' mini-dystrophin sequence upto exon55 along with the pCI donor and 872bp AP overlap sequence. The pcisMini-dystrophin.Hybrid3' contains the same 872bp AP sequence followed by the pCI intron acceptor and the 3'mini-dystrophin sequence from exon56 onwards into an AAV proviral plasmid containing the SV40 polyadenylation signal. The corresponding mini-dystrophin tsAAV vectors were generated such that they lacked the AP overlap sequence flanking the pCI intron sequences. TA muscles of *mdx* mice were co-infected with the AAV-6 mini-dystrophin hybrid vectors at a dosage of 2×10^{10} vg particles (1×10^{10} vg particles for each virus). Uninfected *mdx* mice and AV.CMV.microdystrophin infected *mdx* mice at a dosage of 1×10^{10} vg particles served as negative and positive controls respectively. The mice were sacrificed six weeks later and TA muscles harvested. Mini-dystrophin expression was analysed by immunofluorescence staining (Figure 4.6 A) and the number of positive fibers was quantified (Figure 4.6 B) in tissue sections. The results of this experiment indicate that although this specific set of mini-dystrophin tsAAV vectors (based on exon 55/56) was poor in transgene expression (112.25$\pm$13.6), the corresponding hybrid vectors increased gene expression by 5.6-fold at an average of (635.25$\pm$146.4).


Hybrid mini-dystrophin co-inf.
Trans-splicing mini-dystrophin co-inf.
AV.micro-dystrophin inf.
Uninfected

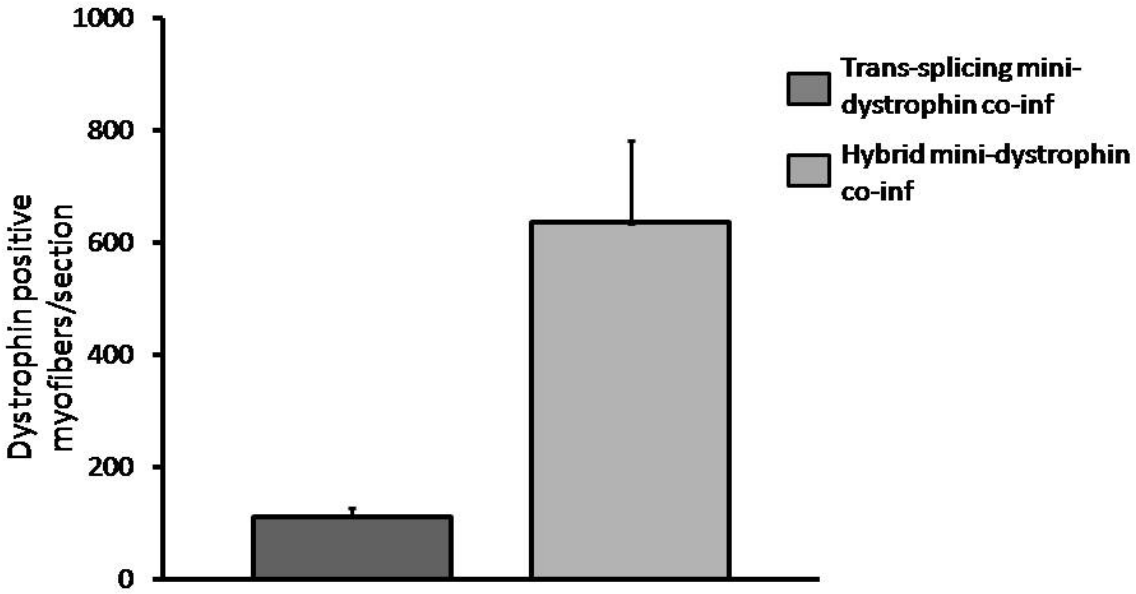

1000
800
600
400
200
0
Dystrophin positive myofibers/section
Trans-splicing mini-dystrophin co-inf
Hybrid mini-dystrophin co-inf

Figure 4.6: Hybrid mini-dystrophin vectors improve gene expression from a poor splice junction. TA muscles of 2-month old *mdx* mice were infected with 1 x 10^{10} vg particles (2 x 10^{10} for co-infections) in 20 µL volume. **A.** Immunofluorescence staining detection of mini-dystrophin expression using human dystrophin N-terminal specific antibody (as described in chapter 2) in sections from TA muscles. **B**. Relative numbers of myofibers positive for mini-dystrophin expression (N=4).

***In vivo* comparison between AP overlapping vector sets sharing different regions of overlap.** The hybrid vectors expand the AAV packaging capacity in an efficient manner, but the presence of the AP overlapping region takes away about 1.8 kb of packaging space from the hybrid vectors. Therefore, in order to use the best possible part of the AP sequence for using in the hybrid vectors system, we developed a series of overlapping AP vector sets as indicated in Figure 4.7. These vectors shared either the 5' half, the 3' half, the 5'one third, the middle one third or the 3' one third of the original 872 bp overlap sequence. These vectors were used for infecting MO59K cells at an moi of 500 vg particles per virus per cell (1000 vg for co-infections). The results demonstrate that while almost all the overlapping AP vector sets showed similar efficiency compared to the original overlapping vector set, the overlapping set for the middle one third shared sequence was much lower (Figure 4.7). This set of experiments indicates that the different parts of the AP sequence may have different efficiencies of recombination. It is therefore, imperative to test these sequences in the context

of the LacZ gene in the hybrid vector system to determine which sequence(s) may be the most efficient for hybrid vector reconstitution.

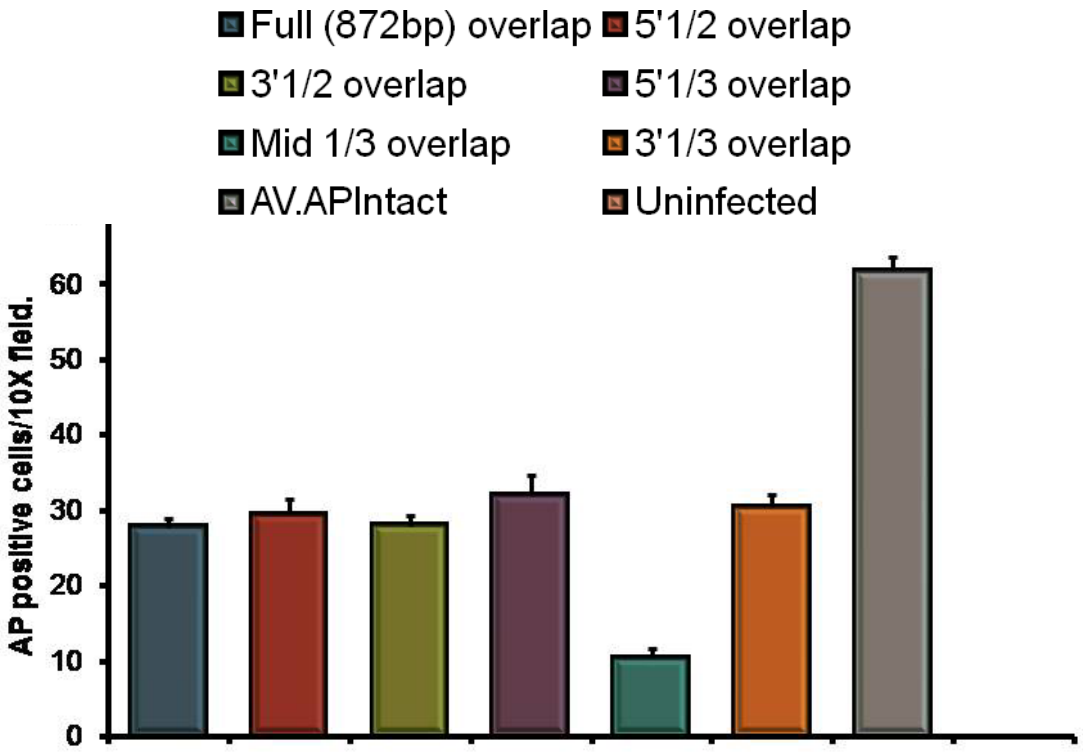
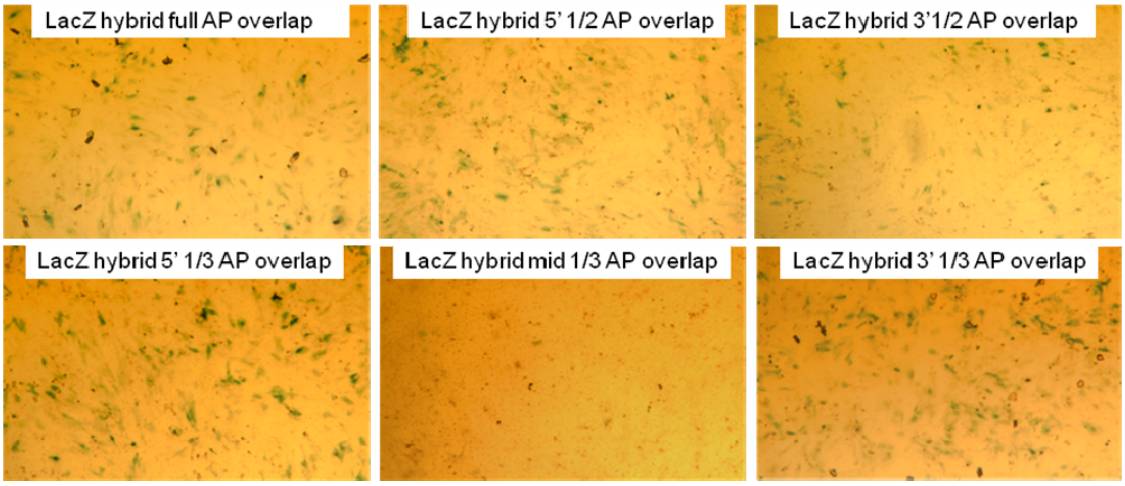


Figure 4.7. Overlapping AP vectors sharing different overlapping parts show differential transduction. MO59K cells were infected at an moi of 500 vg particles per virus per cell (1000 vg particles for co-infection) and AP expression was examined by cytochemical staining 48 hrs later. Morphometric quantification of transduction efficiency in MO59K cells. N = 5 for each group.

Different parts of the AP overlap sequence exhibit different efficiencies of gene expression in the hybrid vectors *in vitro.* Since the ultimate aim of the dual vector systems is to increase the packaging capacity of rAAV, we shortened the length of the alkaline phosphatase overlap sequence used in the construction

of the hybrid vectors. This would provide more space for packaging a therapeutic gene. Therefore, to identify the minimal sequence(s) of the alkaline phosphatase overlap that can be sufficient and equally efficient as the full length (872bp) sequence, a series of hybrid LacZ vector sets were constructed in a manner similar to the different AP overlapping vector sets described above. Five different hybrid vector sets were constructed containing the 5' half, the 3' half, the 5' one-third, middle one-third and the 3' one-third of the original AP overlap sequence. Gene expression from each of these vector sets was then evaluated in comparison to the original hybrid vector set in vitro in MO59K cells (Figure 4.8). Surprisingly, the 5' and 3' one third sequences turned out to be efficient in gene expression with the β-galactosidase activity levels reaching 1002.8±187.2 and 1178.64±169.6 units/μg protein respectively, which is comparable to the original hybrid set (1162.36±88.6 units/μg protein). But, LacZ expression from the middle one-third sequence was significantly lower than that from all the other sets at 58.31±11.8 units/μg protein (Figure 4.8). The 5' and 3' half AP overlap sets yielded 613.05±64.61 and 583.6±27.12 units/μg protein respectively.



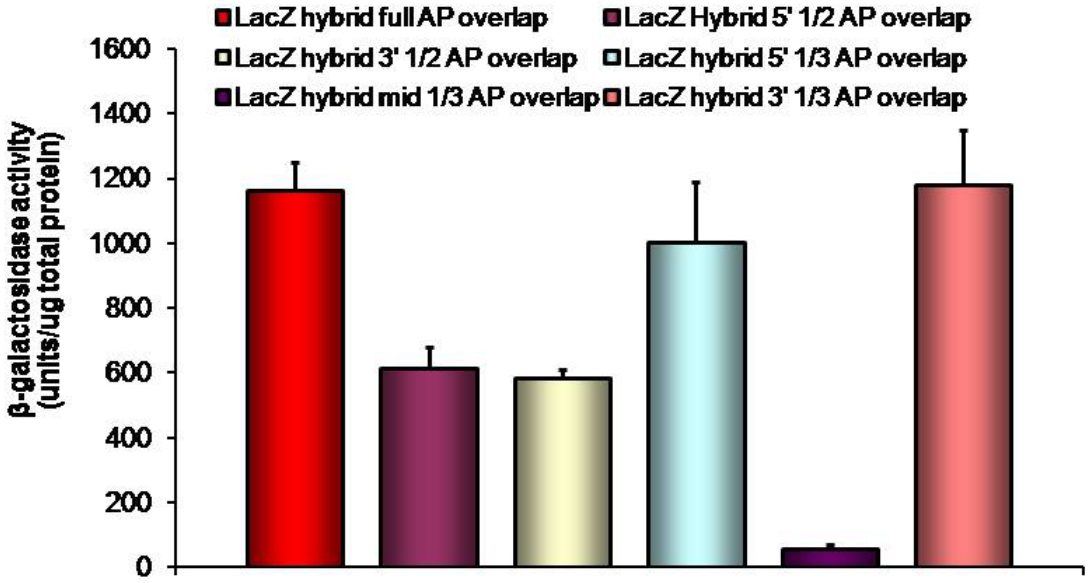


Figure.4.8. Smaller segments of the AP overlap sequence demonstrates variable efficiency of LacZ gene reconstitution in the context of the hybrid vectors *in vitro*. Analysis of LacZ expression from hybrid vector sets containing different regions of the AP sequence in MO59K cells infected at an moi of 20,000 vg particles per cell (10, 000 vg particles/cell for each vector). **A.** Cytological staining for LacZ expression in MO59K cells infected with different sets of LacZ

hybrid viruses 45 hrs post infection. **B.** Quantitative comparison of β-galactosidase activity from the different LacZ hybrid vector sets.

***In vivo* gene expression from small hybrid vectors demonstrates the variation in recombinational properties of different parts of the AP overlap sequence.**

All these different hybrid vector sets were then compared to each other *in vivo* in the TA muscles of BL10 mice. The infections were done at a dosage of 2×10^9 vg particles per muscle for each set of co-infections (1×10^9 vg per virus). Similar to the cell line experiments, the same expression profile was observed *in vivo* with the 5' and 3' one third sequences being the most efficient. Quantification of LacZ positive myofibers show that the 5' and 3' half AP overlap sets had 77.4% and 60.9% LacZ positive myofibers. Again, the middle one third AP sequence was lower at 40.4% while the full length AP overlap set reached 86.7%. The 5' and 3' AP overlap sets were comparable to the full length set at 83.5% and 95.4% respectively. This set of experiments points out that the recombination efficiencies of the overlap sequence used for hybrid vectors is a very crucial determinant of transduction efficiency.

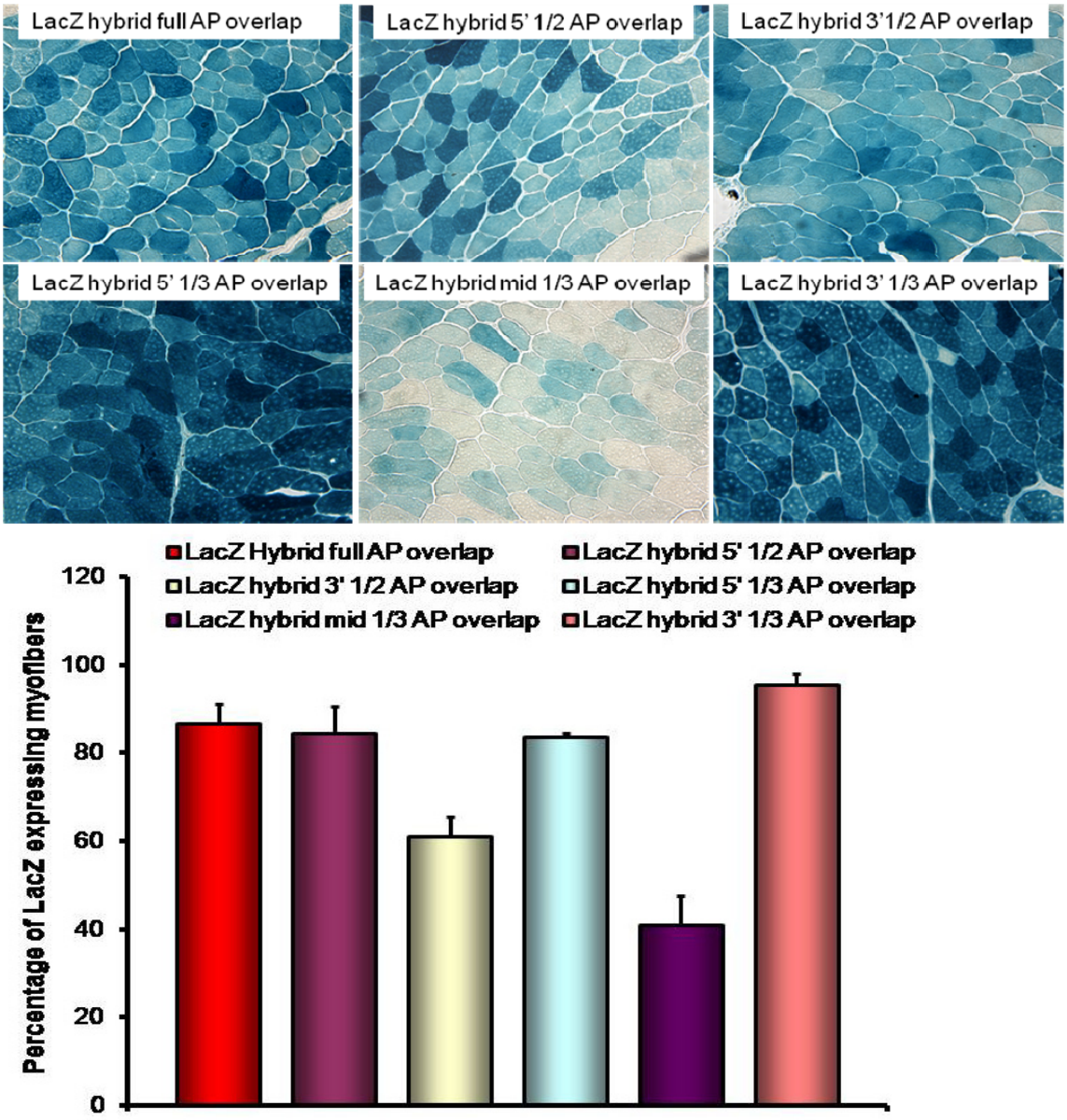


Figure 4.9. Efficient transgene reconstitution mediated by small fragments of the AP overlap sequence in the context of hybrid LacZ vectors in mouse muscle *in vivo*. TA muscles of 2-month old BL10 mice were infected with 2 x 10^9 vg particles for co-infections in 20 µL volume. **A,** Representative photomicrographs of LacZ staining from different experiment groups. **B,** Morphometric quantification of transduction efficiency (N = 3 for each group).

D. Discussion

The small AAV genome has been a longstanding hurdle in AAV gene therapy. The recent development of ts and overlapping AAV vectors has raised the hope of expanding the territory of AAV gene therapy to many diseases that need a large therapeutic gene (reviewed in [94]). However, the inherent limitation of these approaches has failed to completely solve the size problem. The small packaging capacity is still being considered as a limitation in AAV gene therapy [95]

The major problem in the current dual vector approaches is that they all depend on the molecular property of the target gene itself. Essentially, an optimal gene splitting site is required to efficiently express a gene with the tsAAV vectors. On the other hand, the overlapping vectors can only be used to deliver genes that carry a highly recombinogenic sequence. The goal of this study is to develop a generic system that is independent of the target gene. Such system will be applicable to essentially any large therapeutic gene.

We have previously shown that the middle one third of the AP gene mediates efficient homologous recombination [47]. Transduction efficiency of the AP overlapping vectors reached that of a single intact vector in muscle and the lung [36, 47]. We decided to incorporate this unique DNA sequence into both the donor and acceptor vectors in the traditional tsAAV system. We hypothesized that this AP gene fragment would guide transgene reconstitution through homologous recombination as in the traditional overlapping vectors. Since we

know this DNA fragment will recombine efficiently through homologous recombination, we expect to see a high transduction from the hybrid vectors. In addition, since the vector genomes can also undergo ITR-mediated recombination, we expect that there may be additional transgene reconstitution through the traditional tsAAV pathway. Taken together, we expect the hybrid vectors to exceed the traditional dual AAV vectors.

To test our hypothesis, we built a series of vectors using the LacZ gene, a relatively poor candidate gene for the traditional dual vector approaches [47, 49]. In both *in vitro* study in cell culture and *in vivo* study in mouse muscle, the hybrid LacZ vectors showed superior performance, approaching that of the single intact vector (Figures 4.2 and 4.3). In quantitative comparison, transduction efficiency of the hybrid vectors was significantly higher than these from the ts and overlapping LacZ vectors.

We next explored the molecular mechanisms of transgene reconstitution in the hybrid vectors. Consistent with our hypothesis, we indeed observed a strong diagnostic band derived from homologous recombination of the AP sequence (Figure 4.4, lanes 8 and 19, arrow). Interestingly, we only saw a faint band that might have come from ITR-mediated recombination (Figure 4.4, lanes 8, 19 and 30, arrowhead). This is surprising since in the absence of the AP sequence, this LacZ gene splitting site has resulted in decent transgene reconstitution in the tsAAV vectors. Since splicing is influenced by the neighboring enhancers and silencers, one would expect some positive or negative effect from the AP sequence on the overall splicing efficiency in the

hybrid vectors. However, this should not alter ITR-mediated recombination. Our results suggest that there may be a competition between ITR-driven recombination and AP sequence-driven recombination in the hybrid vectors. In the case of the LacZ hybrid vectors, AP sequence-mediated recombination seems to represent the dominant pathway.

To further elucidate such DNA level competition between the two recombination pathways, we performed a series of competition assays. In the case of overlapping vector competition, LacZ expression levels dropped by 42-44% compared to the LacZ hybrid co-infection by itself. Comparatively, the ts completion by AV.minidystrophin.Acceptor reduced LacZ expression by only 22.6%. The data suggests that the overlapping vectors are more efficient at competition as indicated by the lower levels of LacZ expression, thereby indicating the relative importance of AP sequence mediated recombination over ITR mediated recombination.

We then generated mini-dystrophin hybrid vectors and compared them to their corresponding mini-dystrophin ts vectors. The results demonstrated that this set of hybrid vectors was performed much better than the corresponding ts vectors (5.6-fold higher). This underscores the basic hypothesis of this work that the hybrid vectors may overcome the limitations of traditional dual vectors. It should also be noted that we have previously reported even higher levels of transgene expression from a set of optimized mini-dystrophin tsAAV vectors [23, 76] (based on exon 63/64 junction). It should be noted that in the hybrid vectors, even if the overlapping AP sequences mediate efficient transgene reconstitution,

there remains a splicing step prior to gene expression. A poor splice junction will therefore still be a hindrance to maximal gene expression. We speculate that this may be reason that the hybrid vectors based on the exon 55/56 junction is unable to fully complement the mini-dystrophin expression to the levels reported earlier [23, 76].

It is evident from the competition experiments (Figure 4.5) and the Southern blot data (figure 4.4) that the overlapping pathway is the important factor determining the higher levels of gene expression from hybrid vectors compared to traditional dual vectors. Furthermore, the mini-dystrophin hybrid vector experiments (Figure 4.6) and other LacZ hybrid vector experiments suggest that the relative importance of the overlapping pathway depends on the strength of the splice junction used for splitting the gene. In the case of tsAAV vectors, the problem of a poor splice junction is further confounded by the presence of the double-D ITR in the intron of the reconstituted expression cassette, thereby dramatically reducing gene expression [23, 76]. For example, in the case of the mini-dystrophin vectors, when a poor splice junction is used, then the corresponding hybrid vector improved gene expression by many fold, indicating that the overlapping pathway is of far greater importance. On the other hand, in the case of LacZ tsAAV and LacZ hybrid vectors, the improvement due to the overlapping pathway was not as large as in case of the mini-dystrophin hybrid vectors. Therefore, if the splice junction is already optimum, then the degree of improvement due to the overlapping pathway may not be significant. Overall, in terms of actual physical reconstitution of the

expression cassette, the AP overlapping sequence mediated recombination may be more efficient than the ITR mediated recombination.

To further streamline the hybrid system, we further investigated the minimum size of the overlapping sequence required for the hybrid system to function optimally. To address this, we first constructed sets of AP overlapping vectors which shared different parts of the original overlapping AP vectors. The results from this set of experiments demonstrates that except for the middle one third part of the overlap, all other smaller parts were as efficient as the original AP overlapping vector in gene expression (Figure 4.7). These experiments were therefore followed up in the context of the LacZ gene by generating hybrid LacZ vector sets containing different regions of the 872bp AP sequence. We determined that ~270bp sequences at the 5' and the 3' ends of the original 872 bp AP sequence are essential and sufficient for hybrid vector function. Further, their *in vitro* as as well as *in vivo* transduction efficiencies reached up to 100% or more when compared with the original LacZ hybrid set (Figure 4.8 and 4.9). The results from the different LacZ hybrid vectors show an exaggerated trend for what was observed for the comparison between the different AP overlapping vectors (Figure 4.7). The most intriguing aspect of this set of experiments is the extremely low efficiency of transgene reconstitution of the middle one-third of the AP overlap sequence (Figure 4.7, 4.8 and 4.9) which is more pronounced in the context of the hybrid vectors. This not only points out the importance of the overlapping sequence for hybrid vector function, but also the possibility that certain elements within the overlap sequence used may have an effect on the

splicing efficiency across that intron . It is also possible that the AP sequence might stabilize the transgene mRNA or reduce the rate of its degradation, thereby leading to higher gene expression. In this context it should be noted that the AP sequence is derived from the cDNA of alkaline phosphatase and could therefore contain exonic splice enhancers (ESE) or splice inhibitory sequences. Such ESE sequences may help recruit the splicing machinery more efficiently assuming that the splicing complexes are sterically flexible [96, 97]. Therefore it is plausible to argue that the overlapping sequence within the hybrid vectors may be further modified (or taken from a different source) to improve the expression from hybrid vector co-infection. We also calculated the GC content of the five different AP sequences and found no correlation between the GC content and the recombination efficiency (data not shown).

In summary, we have shown the proof-of-principle evidence that essentially any large gene can be efficiently expressed with the hybrid AAV vectors. Future studies with the therapeutic candidate genes (such as the cystic fibrosis transmembrane conductance gene, the full length dystrophin gene and the dysferlin gene) will further reveal the utility and therapeutic potential the hybrid AAV vectors.

E. Acknowledgements . This work was supported by grants from the National Institutes of Health (AR3 49419 and DK-76552 to DD), and research grants from the Muscular Dystrophy 4 Association (DD) and the Cystic Fibrosis Foundation (DD). I also thank Matthew Lamar, Kimberly Gillock and Nate Marschalk for assistance during cloning of vector constructs.

Chapter 5. Summary and future directions

Over the past decade, adeno-associated virus has emerged as one of the major gene delivery vehicles of choice for a wide array of gene therapy applications. While the low immunoreactivity and long term gene expression from recombinant AAV is a positive feature, its chief disadvantage is the small packaging capacity of 5 kb. This disadvantage precludes the use of AAV for treating diseases like DMD or CF. Therefore, innovative techniques were developed to expand the AAV packaging capacity, namely the trans-splicing and overlapping vectors. Initial proof-of -principle experiments demonstrated that these dual vectors are capable of transduction, but their levels of gene expression are much lower than what can be achieved from the single intact vectors. We hypothesized that if we can delineate the various expression limiting steps in the reconstitution process for dual vector co-infection, the dual vector design could be improved for better transgene expression. Additionally, various factors like dosage and viral serotype that affect the viral uptake and subsequent intracellular processing may also prove crucial for downstream gene expression.

This study first examines the expression limiting factors involved in overlapping dual vector reconstitution. The data revealed that the primary determinant of efficiency for these vectors is the recombination potential of the overlap sequence used for designing the overlapping vectors. We also demonstrated the effect of viral serotype on gene expression. In parallel to this work, it was also demonstrated that for tsAAV vectors, the transduction efficiency

is dependent on the splicing efficiency and mRNA accumulation. The revelation of these expression limiting factors enabled better design of dual vectors that led to gene expression levels required for successful gene therapy, as was shown for the mini-dystrophin gene.

Furthermore, the rationally designed tsAAV vectors were successfully applied for systemic gene transfer to multiple tissues/organs throughout the body. This observation significantly raised the hope of using the dual vector strategy for DMD gene therapy.

Using the knowledge from our previous observations and the literature, we were able to design a novel dual vector system. This was called the hybrid vector system and was capable of utilizing both the overlapping and the trans-splicing dual vector pathways for transgene reconstitution. This system benefits from the presence of a sequence within the intron shared by the two split vectors capable of mediating very efficient transgene reconstitution via homologous recombination. The hybrid vector system realizes the full potential of dual vector systems by overcoming the transgene dependent expression limiting barriers for gene expression. These vectors are therefore suitable for application in any therapeutic gene of choice that exceeds the packaging limits of AAV. We have further streamlined and optimized the hybrid vectors to shorten the overlap sequence required for hybrid vector reconstitution, thereby maximizing their packaging capacity. Therefore, the next step is to evaluate the hybrid vectors in the therapeutic context. To this end we are applying the new hybrid vector technology to the CFTR and the full length human dystrophin genes.

Cystic fibrosis (CF) is characterized by a solute transport defect in epithelial tissues. In the lungs, this defect culminates in the dehydration of the airway surface and mucus accumulation, ultimately leading to chronic bacterial infection. The current therapeutic approaches used to treat CF primarily focus on the secondary manifestations of the disease (e.g. bacterial infection, viscous mucus). However, new therapeutic approaches are targeting the underlying ion transport defect in cystic fibrosis, with the aim of restoring the function of the cystic fibrosis transmembrane conductance regulator, stimulating alternative chloride channels, inhibiting sodium absorption, and utilizing hyperosmotic agents to rehydrate the airway surface. Gene therapy for cystic fibrosis is an attractive prospect, if efficient gene delivery and long term, stable gene expression can be achieved. Thus, the infection and long term gene expression properties of AAV present an important asset for utilizing it to deliver the CFTR gene. The length of the CFTR cDNA is about 4.6kb and addition of transcription regulatory components pushes the size to about 5.5-6.0 kb thereby rendering it too large to fit into AAV. Therefore we are generating the CFTR hybrid vectors based on the design of the previously described LacZ hybrid vectors with the small 3'AP overlap sequence. To generate these hybrid CFTR vectors, we are using a biologically active CFTR-GFP transgene [98]. The expression from this transgene can be tracked by GFP expression as well as chloride channel activity. As a primer to the CFTR gene delivery, we delivered the AAV-9 intact AP and tsAAV AP vectors to the mouse lung (Figure 5.1). These vectors were delivered by spuraglottic injection using a feeding tube. We delivered 2.5×10^{11} vg

particles per lung in a total volume of 50µl to BL10 mice. The data demonstrates that the tsAAV AP infection successfully transduced the alveolar and epithelial cells as well as the vasculature. This demonstrates that the dual vectors are suited for direct lung gene transfer.

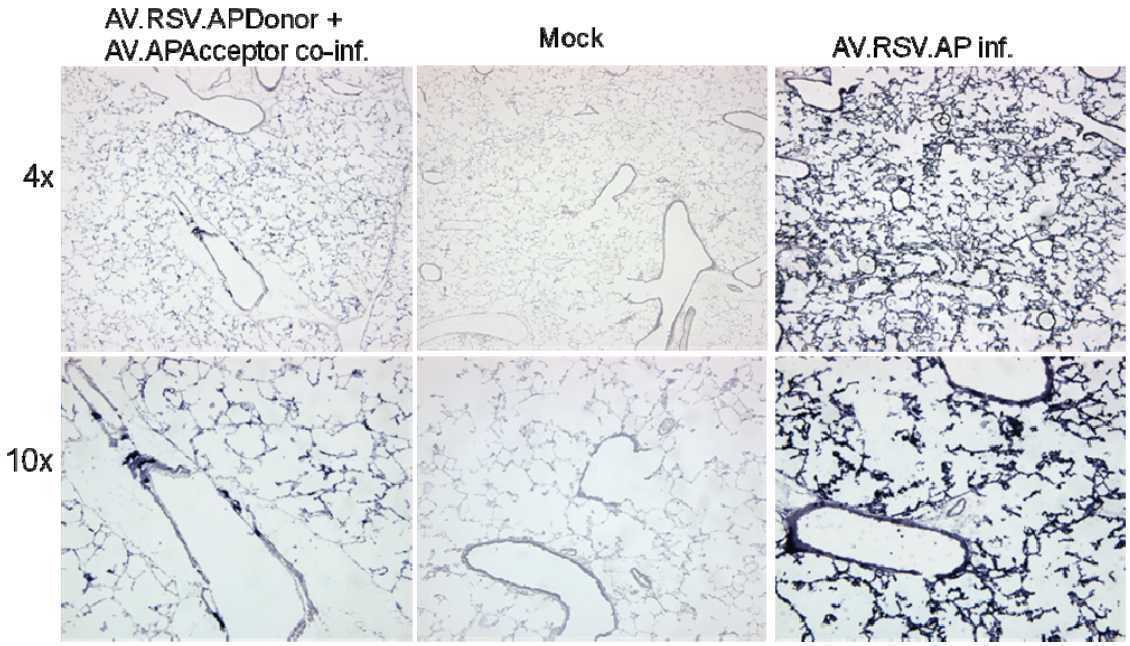


Figure 5.1. Supraglottic delivery of AP tsAAV vectors leads to widespread lung transduction. BL10 mice were infected with the indicated viruses at a dosage of 2.5 x 10^{11} vg particles and sacrificed 6 weeks later. Panel shows AV.AP infected and AP tsAAV co-infected lung sections stained for AP expression.

The second goal would be the application of hybrid vectors to the full length dystrophin gene. Although a variety of dystrophin minigenes and microgenes are available, they are still functionally less efficient than the full length dystrophin protein. To accommodate the full length dystrophin gene, it

would need to be split into three parts (Figure 5.2). In order to drive directional recombination, we will take advantage of the two different small AP overlap sequences identified previously. One of the small AP overlap sequences will be shared by the 5' one third hybrid dystrophin vector and the 5' end of the middle one third hybrid dystrophin vector. The second small overlap sequence will be shared between the 3' end of the middle one third hybrid dystrophin vectors and the 3' hybrid dystropnin vector. Furthermore, it has been demonstrated that directional recombination by using "hybrid ITR" AAV vectors can improve transgene expression. Thus, the 5' and 3' hybrid vectors in this case will have the "hybrid ITR" flanking them in reverse order to each other as shown in Figure 5.2 (rounded ITR, AAV-5; boxed ITR, AAV-2). Thus, the hybrid ITR will prove to be an additional driving force for directional recombination for the trans-splicing mediated recombination. Furthermore, the gene will be split in a way such that the 5' part and the 3' parts can also form a functional minigene in the absence of the middle one third parts as shown in figure 5.2.

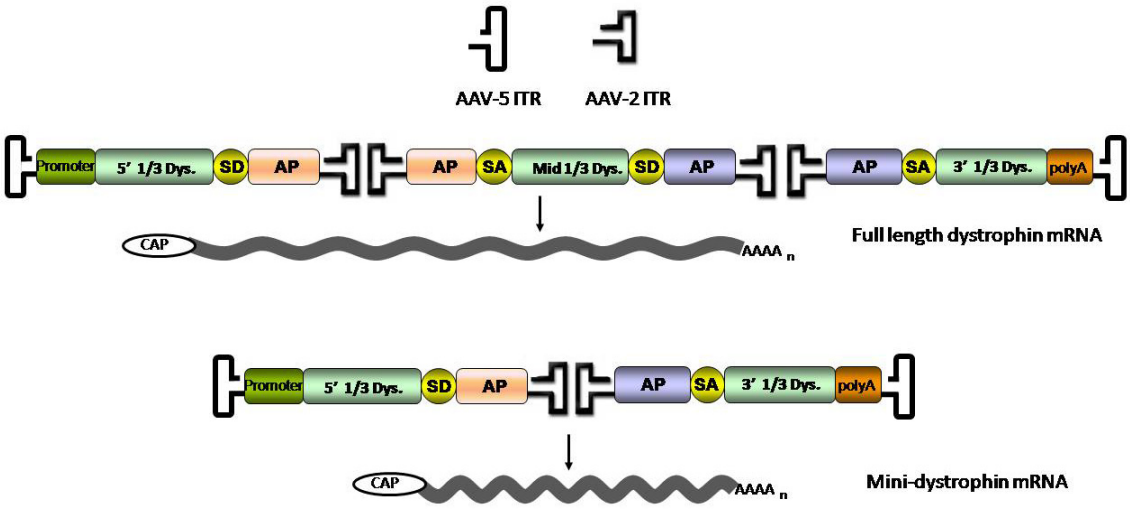


Figure 5.2. Schematic representation of full length dystrophin three vector design. The full length human dystrophin gene is split into three parts within the three hybrid vectors designed as shown. The rounded ITR represents AAV-5 ITR sequence and the boxed ITR represents AAV-2 ITR sequence. The panel also shows two possible sets of vector recombination that can lead to functional dystrophin protein production.

This design should allow for successful full-length dystrophin or mini-dystrophin expression. This unique vector design should provide relatively higher amounts of correctly recombined expression cassettes and consequently, higher levels of gene expression.

Thus, in summary, this study has laid the groundwork for rational design of split gene vectors capable of high levels of gene expression for application in actual gene therapy.

References

1. Atchison, R.W., B.C. Casto, and W.M. Hammon, *Adenovirus-Associated Defective Virus Particles.* Science, 1965. **149**: p. 754-6.
2. Flotte, T.R. and K.I. Berns, *Adeno-associated virus: a ubiquitous commensal of mammals.* Hum Gene Ther, 2005. **16**(4): p. 401-7.
3. Gao, G., et al., *Clades of Adeno-associated viruses are widely disseminated in human tissues.* J Virol, 2004. **78**(12): p. 6381-8.
4. Srivastava, A., E.W. Lusby, and K.I. Berns, *Nucleotide sequence and organization of the adeno-associated virus 2 genome.* J Virol, 1983. **45**(2): p. 555-64.
5. Ding, W., et al., *Intracellular trafficking of adeno-associated viral vectors.* Gene Ther, 2005. **12**(11): p. 873-80.
6. Berns, K.I. and R.A. Bohenzky, *Adeno-associated viruses: an update.* Adv Virus Res, 1987. **32**: p. 243-306.
7. Berns, K.I. and R.M. Linden, *The cryptic life style of adeno-associated virus.* Bioessays, 1995. **17**(3): p. 237-45.
8. Kotin, R.M., et al., *Site-specific integration by adeno-associated virus.* Proc Natl Acad Sci U S A, 1990. **87**(6): p. 2211-5.
9. Carter, B.J., *Adeno-associated virus and the development of adeno-associated virus vectors: a historical perspective.* Mol Ther, 2004. **10**(6): p. 981-9.
10. Carter, B.J., *Adeno-associated virus vectors in clinical trials.* Hum Gene Ther, 2005. **16**(5): p. 541-50.
11. Duan, D., et al., *Circular Intermediates of Recombinant Adeno-Associated Virus have Defined Structural Characteristics Responsible for Long Term Episomal Persistence In Muscle.* J. Virol., 1998. **72**(11): p. 8568-8577.
12. Duan, D., et al., *Structural analysis of adeno-associated virus transduction intermediates.* Virology, 1999. **261**(1): p. 8-14.
13. Gao, G., L.H. Vandenberghe, and J.M. Wilson, *New recombinant serotypes of AAV vectors.* Curr Gene Ther, 2005. **5**(3): p. 285-97.
14. Wu, Z., A. Asokan, and R.J. Samulski, *Adeno-associated Virus Serotypes: Vector Toolkit for Human Gene Therapy.* Mol Ther, 2006. **14**(3): p. 316-27.
15. Gregorevic, P., et al., *Systemic delivery of genes to striated muscles using adeno-associated viral vectors.* Nat Med, 2004. **10**(8): p. 828-34.
16. Wang, Z., et al., *Adeno-associated virus serotype 8 efficiently delivers genes to muscle and heart.* Nat Biotechnol, 2005. **23**(3): p. 321-8.
17. Inagaki, K., et al., *Robust systemic transduction with AAV9 vectors in mice: efficient global cardiac gene transfer superior to that of AAV8.* Mol Ther, 2006. **14**(1): p. 45-53.
18. Pacak, C.A., et al., *Recombinant Adeno-Associated Virus Serotype 9 Leads to Preferential Cardiac Transduction In Vivo.* Circ Res, 2006.
19. Dong, J.Y., P.D. Fan, and R.A. Frizzell, *Quantitative analysis of the packaging capacity of recombinant adeno-associated virus.* Hum Gene Ther, 1996. **7**(17): p. 2101-12.

20. Flotte, T.R., et al., *Phase I trial of intramuscular injection of a recombinant adeno-associated virus alpha 1-antitrypsin (rAAV2-CB-hAAT) gene vector to AAT-deficient adults.* Hum Gene Ther, 2004. **15**(1): p. 93-128.
21. Manno, C.S., et al., *Successful transduction of liver in hemophilia by AAV-Factor IX and limitations imposed by the host immune response.* Nat Med, 2006. **12**(3): p. 342-347.
22. Harper, S.Q., et al., *Modular flexibility of dystrophin: implications for gene therapy of Duchenne muscular dystrophy.* Nat Med, 2002. **8**(3): p. 253-61.
23. Lai, Y., et al., *Efficient in vivo gene expression by trans-splicing adeno-associated viral vectors.* Nat Biotechnol, 2005. **23**(11): p. 1435-9.
24. Duan, D., et al., *Formation of adeno-associated virus circular genomes is differentially regulated by adenovirus E4 ORF6 and E2a gene expression.* J Virol, 1999. **73**(1): p. 161-9.
25. Song, S., et al., *Effect of DNA-dependent protein kinase on the molecular fate of the rAAV2 genome in skeletal muscle.* Proc. Natl. Acad. Sci. U S A, 2001. **98**(7): p. 4084-8.
26. Duan, D., Y. Yue, and J.F. Engelhardt, *Consequences of DNA-Dependent Protein Kinase Catalytic Subunit Deficiency On Recombinant Adeno-associated Virus Genome Circularization and Heterodimerization in Muscle Tissue.* J Virol, 2003. **77**(8): p. 4751-9.
27. Yue, Y. and D. Duan, *Double strand interaction is the predominant pathway for intermolecular recombination of adeno-associated viral genomes.* Virology, 2003. **313**(1): p. 1-7.
28. Song, S., et al., *DNA-dependent PK inhibits adeno-associated virus DNA integration.* Proc Natl Acad Sci U S A, 2004. **101**(7): p. 2112-6.
29. Duan, D., et al., *A new dual-vector approach to enhance recombinant adeno-associated virus-mediated gene expression through intermolecular cis activation.* Nat. Med., 2000. **6**(5): p. 595-8.
30. Yan, Z., et al., *From the Cover: Trans-splicing vectors expand the utility of adeno-associated virus for gene therapy.* Proc. Natl. Acad. Sci. U S A, 2000. **97**(12): p. 6716-6721.
31. Sun, L., J. Li, and X. Xiao, *Overcoming adeno-associated virus vector size limitation through viral DNA heterodimerization.* Nat. Med., 2000. **6**(5): p. 599-602.
32. Duan, D., Y. Yue, and J.F. Engelhardt, *Dual vector expansion of the recombinant AAV packaging capacity*, in *Methods in Molecular Medicine, Cardiac Gene Transfer: Principles, Protocols and Applications*, J.M. Metzger, Editor. 2002, Humana Press Inc.: Totowa, NJ. p. 29-51.
33. Senapathy, P., J.D. Tratschin, and B.J. Carter, *Replication of adeno-associated virus DNA. Complementation of naturally occurring rep- mutants by a wild-type genome or an ori- mutant and correction of terminal palindrome deletions.* J Mol Biol, 1984. **179**(1): p. 1-20.
34. Wang, X.S., S. Ponnazhagan, and A. Srivastava, *Rescue and replication signals of the adeno-associated virus 2 genome.* J Mol Biol, 1995. **250**(5): p. 573-80.

35. Duan, D., Y. Yue, and J.F. Engelhardt, *Expanding AAV Packaging Capacity With Trans-splicing Or Overlapping Vectors: A Quantitative Comparison.* Mol Ther, 2001. **4**(4): p. 383-91.
36. Halbert, C.L., J.M. Allen, and A.D. Miller, *Efficient mouse airway transduction following recombination between AAV vectors carrying parts of a larger gene.* Nat Biotechnol, 2002. **20**(7): p. 697-701.
37. Flotte, T.R., et al., *Expression of the cystic fibrosis transmembrane conductance regulator from a novel adeno-associated virus promoter.* J Biol Chem, 1993. **268**(5): p. 3781-90.
38. Nakai, H., T.A. Storm, and M.A. Kay, *Increasing the size of rAAV-mediated expression cassettes in vivo by intermolecular joining of two complementary vectors [see comments].* Nat. Biotechnol., 2000. **18**(5): p. 527-32.
39. Chao, H., et al., *Expression of Human Factor VIII by Splicing between Dimerized AAV Vectors.* Mol Ther, 2002. **5**(6): p. 716-22.
40. Reich, S.J., et al., *Efficient trans-splicing in the retina expands the utility of adeno-associated virus as a vector for gene therapy.* Hum Gene Ther, 2003. **14**(1): p. 37-44.
41. Xu, Z., et al., *Trans-splicing adeno-associated viral vector-mediated gene therapy is limited by the accumulation of spliced mRNA but not by dual vector coinfection efficiency.* Hum Gene Ther, 2004. **15**(9): p. 896-905.
42. Yan, Z., et al., *Inverted terminal repeat sequences are important for intermolecular recombination and circularization of adeno-associated virus genomes.* J Virol, 2005. **79**(1): p. 364-79.
43. Russell, D.W. and R.K. Hirata, *Human gene targeting by viral vectors.* Nat. Genet., 1998. **18**(4): p. 325-30.
44. Liu, X., et al., *Targeted Correction of Single-Base-Pair Mutations with Adeno-Associated Virus Vectors under Nonselective Conditions.* J Virol, 2004. **78**(8): p. 4165-75.
45. Hendrie, P.C. and D.W. Russell, *Gene targeting with viral vectors.* Mol Ther, 2005. **12**(1): p. 9-17.
46. Miller, D.G., et al., *Gene targeting in vivo by adeno-associated virus vectors.* Nat Biotechnol, 2006. **24**(8): p. 1022-6.
47. Ghosh, A., Y. Yue, and D. Duan, *Viral serotype and the transgene sequence influence overlapping adeno-associated viral (AAV) vector-mediated gene transfer in skeletal muscle.* J. Gene Med., 2006. **8**(3): p. 298-305.
48. Duan, D., Y. Yue, and J.F. Engelhardt, *Dual vector expansion of the recombinant AAV packaging capacity.* Methods Mol Biol, 2003. **219**: p. 29-51.
49. Duan, D., et al., *Enhancement Of Muscle Gene Delivery With Pseudotyped AAV-5 Correlates With Myoblast Differentiation.* J Virol, 2001. **75**(16): p. 7662-71.
50. Allen, J.M., C.L. Halbert, and A.D. Miller, *Improved adeno-associated virus vector production with transfection of a single helper adenovirus gene, E4orf6.* Mol Ther, 2000. **1**(1): p. 88-95.
51. Rasheed, S., et al., *Characterization of a newly derived human sarcoma cell line (HT-1080).* Cancer, 1974. **33**(4): p. 1027-33.

52. Halbert, C.L., J.M. Allen, and A.D. Miller, *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, 2001. **75**(14): p. 6615-24.
53. Duan, D., et al., *Endosomal processing limits gene transfer to polarized airway epithelia by adeno-associated virus.* J. Clin. Invest., 2000. **105**(11): p. 1573-1587.
54. Liu, M., et al., *Adeno-associated Virus-mediated Micro-dystrophin Expression Protects Young Mdx Muscle from Contraction-induced Injury.* Mol Ther, 2005. **11**(2): p. 245-56.
55. Shen, P. and H.V. Huang, *Homologous recombination in Escherichia coli: dependence on substrate length and homology.* Genetics, 1986. **112**(3): p. 441-57.
56. Hua, S.B., et al., *Minimum length of sequence homology required for in vivo cloning by homologous recombination in yeast.* Plasmid, 1997. **38**(2): p. 91-6.
57. Rutledge, E.A., C.L. Halbert, and D.W. Russell, *Infectious clones and vectors derived from adeno-associated virus (AAV) serotypes other than AAV type 2.* J Virol, 1998. **72**(1): p. 309-19.
58. Hildinger, M., et al., *Hybrid vectors based on adeno-associated virus serotypes 2 and 5 for muscle-directed gene transfer.* J Virol, 2001. **75**(13): p. 6199-203.
59. Blankinship, M.J., et al., *Efficient transduction of skeletal muscle using vectors based on adeno-associated virus serotype 6.* Mol Ther, 2004. **10**(4): p. 671-8.
60. Halbert, C.L., et al., *Repeat transduction in the mouse lung by using adeno-associated virus vectors with different serotypes.* J Virol, 2000. **74**(3): p. 1524-32.
61. Li, F.Q., A. Coonrod, and M. Horwitz, *Preferential MyoD homodimer formation demonstrated by a general method of dominant negative mutation employing fusion with a lysosomal protease.* J Cell Biol, 1996. **135**(4): p. 1043-57.
62. Couffinhal, T., et al., *Histochemical staining following LacZ gene transfer underestimates transfection efficiency.* Hum Gene Ther, 1997. **8**(8): p. 929-34.
63. Wang, B., J. Li, and X. Xiao, *Adeno-associated virus vector carrying human minidystrophin genes effectively ameliorates muscular dystrophy in mdx mouse model.* Proc Natl Acad Sci U S A, 2000. **97**(25): p. 13714-9.
64. Phelps, S.F., et al., *Expression of full-length and truncated dystrophin mini-genes in transgenic mdx mice.* Hum Mol Genet, 1995. **4**(8): p. 1251-8.
65. Petes, T.D., *Meiotic recombination hot spots and cold spots.* Nat Rev Genet, 2001. **2**(5): p. 360-9.
66. Ptak, S.E., et al., *Absence of the TAP2 human recombination hotspot in chimpanzees.* PLoS Biol, 2004. **2**(6): p. e155.
67. Crawford, D.C., et al., *Evidence for substantial fine-scale variation in recombination rates across the human genome.* Nat Genet, 2004. **36**(7): p. 700-6.
68. Zhu, T., et al., *Sustained whole-body functional rescue in congestive heart failure and muscular dystrophy hamsters by systemic gene transfer.* Circulation, 2005. **112**(17): p. 2650-9.
69. Qiao, C., et al., *Amelioration of laminin-alpha2-deficient congenital muscular dystrophy by somatic gene transfer of miniagrin.* Proc Natl Acad Sci U S A, 2005. **102**(34): p. 11999-2004.
70. Nakai, H., et al., *Unrestricted hepatocyte transduction with adeno-associated virus serotype 8 vectors in mice.* J Virol, 2005. **79**(1): p. 214-24.

71. Mah, C., et al., *Sustained correction of glycogen storage disease type II using adeno-associated virus serotype 1 vectors.* Gene Ther, 2005. **12**(18): p. 1405-9.
72. Denti, M.A., et al., *Body-wide gene therapy of Duchenne muscular dystrophy in the mdx mouse model.* Proc Natl Acad Sci U S A, 2006. **103**(10): p. 3758-63.
73. Pacak, C.A., et al., *Recombinant Adeno-Associated Virus Serotype 9 Leads to Preferential Cardiac Transduction In Vivo.* Circ Res, 2006. **99**(4): p. e3-9.
74. Gregorevic, P., et al., *rAAV6-microdystrophin preserves muscle function and extends lifespan in severely dystrophic mice.* Nat Med, 2006. **12**(7): p. 787-9.
75. Yue, Y., M. Liu, and D. Duan, *C-terminal truncated microdystrophin recruits dystrobrevin and syntrophin to the dystrophin-associated glycoprotein complex and reduces muscular dystrophy in symptomatic utrophin/dystrophin double knock-out mice.* Mol Ther, 2006. **14**(1): p. 79-87.
76. Lai, Y., et al., *Synthetic intron improves transduction efficiency of trans-splicing adeno-associated viral vectors.* Hum Gene Ther, 2006. **17**(10): p. 1036-42.
77. Gao, G., et al., *Adeno-associated viruses undergo substantial evolution in primates during natural infections.* Proc Natl Acad Sci U S A, 2003. **100**(10): p. 6081-6.
78. Yan, Z., et al., *Distinct classes of proteasome-modulating agents cooperatively augment recombinant adeno-associated virus type 2 and type 5-mediated transduction from the apical surfaces of human airway epithelia.* J Virol, 2004. **78**(6): p. 2863-74.
79. Grider, M.H., et al., *In situ expression of brain-derived neurotrophic factor or neurotrophin-3 promotes sprouting of cortical serotonergic axons following a neurotoxic lesion.* J Neurosci Res, 2005. **82**(3): p. 404-12.
80. Ficklin, M.B., S. Zhao, and G. Feng, *Ubiquilin-1 regulates nicotine-induced up-regulation of neuronal nicotinic acetylcholine receptors.* J Biol Chem, 2005. **280**(40): p. 34088-95.
81. Chan, P., et al., *Method for multiplex cellular detection of mRNAs using quantum dot fluorescent in situ hybridization.* Nucleic Acids Res, 2005. **33**(18): p. e161.
82. Limberis, M.P. and J.M. Wilson, *Adeno-associated virus serotype 9 vectors transduce murine alveolar and nasal epithelia and can be readministered.* Proc Natl Acad Sci U S A, 2006. **103**(35): p. 12993-8.
83. Lai, Y., et al., *Synthetic intron improves transduction efficiency of the trans-splicing adeno-associated viral vectors.* Hum Gene Ther, 2006. **17**(10): p. 1036-42.
84. Ghosh, A., et al., *Efficient Whole-body Transduction with Trans-splicing Adeno-associated Viral Vectors.* Mol Ther, 2007. **15**(4): p. 750-5.
85. Duan, D., Z. Yan, and J.F. Engelhardt, *Expanding the capacity of AAV vectors*, in *Parvoviruses*, M.E. Bloom, et al., Editors. 2006, Hodder Arnold; Distributed in the U.S.A. by Oxford University Press: London
New York. p. pp525-32.
86. Yue, Y., et al., *Full-length dystrophin expression in half of the heart cells ameliorates beta-isoproterenol-induced cardiomyopathy in mdx mice.* Hum Mol Genet, 2004. **13**(15): p. 1669-75.

87. Yue, Y., et al., *Microdystrophin Gene Therapy of Cardiomyopathy Restores Dystrophin-Glycoprotein Complex and Improves Sarcolemma Integrity in the Mdx Mouse Heart.* Circulation, 2003. **108**(13): p. 1626-32.
88. Johnson, L.G., et al., *Efficiency of gene transfer for restoration of normal airway epithelial function in cystic fibrosis.* Nat Genet, 1992. **2**(1): p. 21-5.
89. Gao, G.P., et al., *Novel adeno-associated viruses from rhesus monkeys as vectors for human gene therapy.* Proc Natl Acad Sci U S A, 2002. **99**(18): p. 11854-9.
90. Yan, Z., et al., *Hybrid adeno-associated virus bearing nonhomologous inverted terminal repeats enhances dual-vector reconstruction of minigenes in vivo.* Hum Gene Ther, 2007. **18**(1): p. 81-7.
91. Blankinship, M.J., P. Gregorevic, and J.S. Chamberlain, *Gene therapy strategies for Duchenne muscular dystrophy utilizing recombinant adeno-associated virus vectors.* Mol Ther, 2006. **13**(2): p. 241-9.
92. Wang, Z., et al., *Rapid and highly efficient transduction by double-stranded adeno-associated virus vectors in vitro and in vivo.* Gene Ther, 2003. **10**(26): p. 2105-11.
93. Fu, H., et al., *Self-complementary adeno-associated virus serotype 2 vector: global distribution and broad dispersion of AAV-mediated transgene expression in mouse brain.* Mol Ther, 2003. **8**(6): p. 911-7.
94. Duan, D., *From the smallest virus to the biggest gene: marching towards gene therapy for duchenne muscular dystrophy.* Discov Med, 2006. **6**(33): p. 103-8.
95. Fischer, A.C., et al., *Expression of a Truncated Cystic Fibrosis Transmembrane Conductance Regulator with an AAV5-pseudotyped Vector in Primates.* Mol Ther, 2007. **15**(4): p. 756-63.
96. DiDonato, C.J., et al., *Regulation of murine survival motor neuron (Smn) protein levels by modifying Smn exon 7 splicing.* Hum Mol Genet, 2001. **10**(23): p. 2727-36.
97. Lorson, C.L. and E.J. Androphy, *An exonic enhancer is required for inclusion of an essential exon in the SMA-determining gene SMN.* Hum Mol Genet, 2000. **9**(2): p. 259-65.
98. Granio, O., et al., *Cellular Localization and Activity of Ad-delivered GFP-CFTR in Airway Epithelial and Tracheal Cells.* Am J Respir Cell Mol Biol, 2007.

VITA

Arkasubhra Ghosh was born July 20th, 1977 in Kolkata, India. After completing high school from Hindi High in Kolkata, he enrolled at the Nagpur University, India and earned a Bachelor of Science degree in Microbiology, Chemistry and Zoology in 1998. He then continued at the Nagpur University to earn a Master of Science degree in Microbiology in 2000. He then joined the department of Molecular Microbiology and Immunology at the University of Missouri-Columbia as a graduate student in August 2002. Arkasubhra earned his PhD on development of viral vectors for gene therapy applications in September 2007. His future interests are to remain in the field of gene therapy and disease research and plans to consider independent research positions in the near future.